UNIVERSITÉ PARIS SUD
Comprendre le monde, construire l'avenir®

Irfu

Université Paris-Sud

École Doctorale Particule, Noyaux et Cosmos
Laboratoire: CEA / Irfu - Service de Physique Nucléaire

Discipline : Physique des Particules

Thèse de doctorat

Soutenue le 11 Decembre 2014 par

Quiela Marina Curiel Garcia

Measurement of Hadron Multiplicities in Deep Inelastic Scattering and Extraction of Quark Fragmentation Functions

Composition du jury :

Président :	Dr. Elias Khan	IPN Orsay, France.
Rapporteurs :	Dr. Andrea Bressan	Universita' degli Studi di Trieste, Italy.
	Dr. Eva-Maria Kabuss	Universitaet Mainz, Germany.
Examinateurs:	Dr. Yann Bedfer	CEA Saclay IRFU, France.
	Dr. Efraín Chávez	IFUNAM, México.
Directeur de thèse :	Dr. Fabienne Kunne	CEA Saclay IRFU, France.

Abstract

One of the goals of the COMPASS collaboration is the study of the nucleon spin structure. Data were taken from a polarized muon beam (160 GeV/c) scattering off a polarized target (^{6}LiD or NH_3). In this context, the need of a precise knowledge of quark Fragmentation Functions (final-state hadronisation of quarks q into hadrons h, FFs) was raised. The FFs can be extracted from hadron multiplicities produced in Semi-Inclusive Deep Inelastic Scattering (SIDIS). This thesis presents the measurement of charged hadrons (pions and kaons) multiplicities from SIDIS data collected in 2006. The data cover a large kinematical range: $Q^2 > 1$ (GeV/c)^2, $y \in [0.1, 0.9]$, $x \in [0.004, 0.7]$ and $W \in [5, 17]$ GeV. These multiplicities provide an important input for global QCD analyses of world data at NLO, aiming at the FFs determination.

Acknowledgments

This thesis would not have been possible unless the support of the COMPASS collaboration based at Saclay.

I would like to thank first my advisor Fabienne Kunne for giving me the opportunity to become a Ph.D student, she spent a lot of time to make possible for me to come to France. Her commitment and interest for COMPASS and science had a major influence on this thesis.

I would like to express my gratitude to the members of my dissertation committee. I want to thanks Elias Khan for agreeing to head a committee. I appreciate the feedback offered by Andrea Bressan to improve the content of this work. Advice and comments given by Eva Kabuss have been a great help in my physics formation. I have greatly benefited from physics knowledge and computational skills from Yann Bedfer. I would like to express the deepest appreciation to Efraín Chávez who have been by my side all the way long trough my development as a scientist, always with warm encouragement and confidence.

I would like to express my gratitude to COMPASS-Saclay members Damien, Nicole and Andrea, they always ,made me feel welcome and supported. I am grateful for the assistance given by Stephane to the writing of this thesis.

Special thanks for all the COMPASS collaboration members. Thanks to my colleagues Nick and Daniel, without their work this would not have been possible.

I would like to offer my special thanks to Luigi Capozza and Federica Sozzi for their valuable contributions in different subjects and mostly for their support and encouragement.

My heartfelt appreciation goes to Diane Doré her sincere, helpful and kind advice, her comprehension and support in the final stage of the writing of this thesis were crucial.

To my laboratory friends Florian, Lucile, Maxime and Vincent. Thanks for the great company, the funny moments, your expertise, for the patience to show me the french way, for appreciating my research strengths and your support through hard times. I greatly look forward to being in touch with all of you.

I owe a very important debt to Erin, even if she arrived at the end of the journey, her ideas and fast work were of huge relevance.

My thanks to Stefano Panebianco for his most valuable knowledge of RICH detector and also for his important contribution to the dissertation.

My thanks to Michel Garçon, Héloïse Goutte, Françoise Auger et Jacques Ball for welcoming me in their lab SPhN.

My thanks to Danielle Coret and Isabelle Richard for helping me on all the administrative matters. To Gilles Tricoche for his helping on technical problems and to Valérie Poyeton for being a good librarian.

I wish to thank my french family, Lucile, Anne-So, Aurore, Michel, Claire-Lise et Mutty who make me feel at home.

From the bottom of my heart, I want to thank Delphine for be very suporting with this proyect and for helping me to improve my french.

To my lovely english family, my dear brother Diego, her strong and sweet wife Lucinda and my favourite little person Estelita, your confidence and support were always a motivation to pursuit this goal.

To my mom and dad, even if you are far away you never leave me alone. Thanks to everything you give me and teach me I am what I am.

I truly thank Emi for being by my side in this experience for the good and for the bad. You are the best.

Contents

Introduction

The first observation of the nucleon internal structure was performed at the Stanford Linear Accelerator Center (SLAC) in late 60's. They performed the first *Deep Inelastic Scattering* (DIS) measurements using a high energy electron beam and proton target giving the first evidence for the existence of quarks. During this time Feynman introduced the *Quark Parton Model* (QPM) which gave a good description of the DIS cross section measurements. After the SLAC results, new experiments were set up with higher energy and different beam particles (*e.g.* muons and neutrinos) which showed that in addition to quarks there must be other constituents in the proton. On the theory side, *Quantum Chromodynamics* (QCD) was developed and introduced gluons as the bosons of the strong interaction and quark confinement which explains why isolated quarks are never observed.

The final state of a DIS process is composed by the scattered particle (electron, muon or neutrino) and a hadronic system. If the scattered lepton and a produced hadron are detected (Semi-Inclusive DIS, SIDIS), the corresponding cross section can be described by a hard scattering cross section (*e.g.* electron-quark interaction calculable in perturbative QCD) and two functions: the *Parton Distribution Functions* PDFs and the *Fragmentation Functions* FFs. Both are universal, in the sense that they do not depend on the type of process from which they are measured. The PDFs parametrize the partonic structure of the target before the interaction and the FFs parametrize the hadronization process (formation of hadrons out of quarks (gluons) produced in DIS). Currently the PDFs of u, $\bar{u}$, d and $\bar{d}$ quarks are well constrained with high precision by experimental measurements, while the ones for the s and $\bar{s}$ quarks are still affected by large uncertainties. Most of their uncertainties came from the poor knowledge of the FFs, particularly those of the strange quarks. The FFs are determined from measurements of electron-positron annihilation, proton-antiproton collisions and semi-inclusive DIS. Several parameterizations of FFs exist. There are difference among, specially for what concerns the strange quark sector.

One of the goals of the COMPASS experiment is the study of the nucleon spin structure. For this purpose, SIDIS data were taken using a polarized 160 GeV muon beam and a polarized target (^{6}LiD or NH_3) for several years. The quark polarization can be accessed by measuring double spin asymmetries in polarized SIDIS. In the determination of the quark polarization for a full flavour separation the precise knowledge of quark FFs is essential. In particular, it was found that the strange quark FF into kaons is the dominant source of uncertainty [1]. In order to contribute with new accurate measurements to constrain the FFs better (through QCD fits of world data), it was decided to extract the hadron

multiplicities (average number of hadrons produced per DIS event) from COMPASS data. This observable can be used to extract the quark FFs.

The work done for this thesis concerns the extraction of charged unidentified and identified hadron (pion and kaon) multiplicities. The thesis is structured as follows: in chapter 1 the theoretical framework is given. The kinematic variables used to describe the DIS process and the corresponding cross sections are presented. The Quark Parton model and its improved version using QCD are briefly described. The PDFs and FFs are introduced and an overview on the current knowledge of FFs is provided.

The chapter 2 gives a description of the COMPASS experiment. The main components of COMPASS apparatus is briefly described. Since the Ring Imaging CHerenkov (RICH) detector is an important item for this analysis it is described in chapter 3. The determination of RICH performances and the algorithm followed to perform the particle identification are described in chapter 4.

The extraction of the unidentified hadron (h^+ and h^-), π^+, π^-, K^+ and K^- multiplicities from COMPASS data is explained in detail in chapters 5 and 6. The final results are presented and discussed in chapter 7. A fit at leading order of the pion and kaon multiplicities is performed and compared to other existent parameterizations in this same chapter.

Chapter 1

Nucleon Structure and Quark Fragmentation Functions

1.1 Motivation

One of the present goals of hadronic physics is the study of the inner structure of the nucleon, and in particular of its spin structure. Momentum distributions of partons (quarks and gluons, polarized or not) in the nucleon are described via Parton Distribution Functions (unpolarized and polarized PDFs). At COMPASS, deep inelastic scattering of longitudinally polarized muon on longitudinally polarized nucleons was used to measure the spin structure function of the nucleon, and also to determine which fraction of the nucleon spin is carried by the spin of the quarks of various flavours and of the gluons. For this purpose, polarized SIDIS (see below) data with the detection of a hadron (pion or kaon) tagging the struck quark flavour were taken. It appears that the quark fragmentation functions (FFs) into hadrons are key ingredients for the determination of the polarized quark distributions through SIDIS. The FFs describe the hadronization of partons into various hadrons. They are rather well known for the u and d quarks, but still not well determined, up to a factor of 3, for the strange quarks. Thus they constitute the largest uncertainty for the determination of the strange quark polarization from SIDIS. In the COMPASS kinematic range, the strange quark unpolarized PDF is also affected by large uncertainties.

In order to improve the situation, it was decided to use the existing abundant COMPASS data to study hadron production (hadron multiplicities), and in particular the multiplicities in unpolarized SIDIS. With the RICH detector kaons are well identified in COMPASS. The multiplicity is sensitive to both PDFs and FFs. They will serve as input to global analyses to determine the FFs. Since FFs are universal objects, they are also used for the description of various processes.

Deep Inelastic Scattering (DIS) has proven to be an effective tool to study FFs and PDFs. A DIS process occurs when a high-energy lepton l interacts with a nucleon N and new particles (X: hadrons) are produced ($l + N \rightarrow l' + X$). If only the scattered lepton

l' is detected in the final state, the measurement is called *inclusive*. If in addition to l', at least a hadron is also detected in the final state ($l + N \rightarrow l' + h + X$) it is classified as a *semi-inclusive* reaction (SIDIS). When all particles produced and l' are detected the measurement is called *exclusive*. The FFs are accessible by measuring the hadron multiplicities in SIDIS.

In order to establish an approach to the subject of the present work (the FFs) the theoretical framework to study the DIS and SIDIS processes is provided. The DIS and SIDIS kinematics and cross section are briefly introduced. The model used to interpret the DIS (SIDIS) results and to describe the nucleon structure (*Quark Parton Model* QPM) as well as its improvement version with Quantum Chromodynamics are described. The PDF and FF definitions are given. A brief description of how the PDFs are extracted is provided. Some fragmentation function characteristics are briefly discussed. The experimental measurements through which FFs are determined are presented. A brief explanation on how FFs are accessed in SIDIS data is discussed. A summary of the existent FFs parameterization is given at the end of this chapter.

1.2 Deep Inelastic Scattering

The DIS process is schematically illustrated in fig. 1.1. The incoming lepton l exchanges a single virtual photon $\gamma*$ with the nucleon N. The nucleon absorbs the $\gamma*$ energy and breaks up into a final state X. The final state of the lepton is represented by l'. This DIS description is known as the one photon exchange approximation.

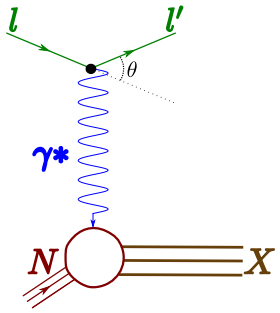


Figure 1.1: *Deep inelastic scattering diagram.*

The kinematics of a deep inelastic scattering event is determined by the 4-momentum vector of l ($\mathbf{k} = (E, \vec{k})$), l' ($\mathbf{k'} = (E', \vec{k'})$) and $N(\mathbf{P} = (M, \vec{p} = \vec{0}))$. The 4-momentum vector for the exchanged photon is given by $\mathbf{q} = \mathbf{k} - \mathbf{k'} = (\nu = E - E', \vec{q} = \vec{k} - \vec{k'})$. For a lepton l, there are only two Lorentz invariant independent variables that describe *inclusive* DIS [3]. One of them is given by the invariant mass of the virtual photon Q^2 defined in eq. (1.1). The Q^2 expression in the laboratory frame holds if the lepton l mass is neglected. Q^2 defines the scale at which the nucleon structure is being studied. If Q^2 is large enough

the probe of the nucleon components is possible.

$$Q^2 = -\mathbf{q}^2 \overset{\text{lab}}{\simeq} 4EE' \sin^2\left(\frac{\theta}{2}\right) \tag{1.1}$$

The second variable is x (eq. (1.2)). It measures the property of elasticity of the interaction: if $x = 1$ the interaction is elastic, if $x < 1$ the interaction is inelastic. This variable was first introduced by Bjorken and in the *quark parton model* [4] interpretation it is defined as the fraction of the nucleon 4-momentum $\mathbf{P}$ carried by the quark which interacts with the virtual photon.

$$x = \frac{Q^2}{2\mathbf{P}\cdot\mathbf{q}} \overset{\text{lab}}{=} \frac{Q^2}{2M\nu} \tag{1.2}$$

In addition to x and Q^2, others dimensionless Lorentz invariant variables can be used (table 1.1).

Table 1.1: *Kinematic variables used for DIS. The expressions in the laboratory frame hold in case of a fixed target and neglecting the lepton mass.*

Variable	Description
$Q^2 = -\mathbf{q}^2 \overset{\text{lab}}{\simeq} 4EE' \sin^2\left(\frac{\theta}{2}\right)$	Negative square 4-momentum transfer (Interaction scale)
$x = \frac{Q^2}{2\mathbf{P}\cdot\mathbf{q}} \overset{\text{lab}}{=} \frac{Q^2}{2M\nu}$	Fraction of the nucleon momentum $\mathbf{P}$ carried by the quark which interacts with $\gamma*$
$\nu = \frac{\mathbf{P}\cdot\mathbf{q}}{M} \overset{\text{lab}}{=} E - E'$	Energy transfer from the lepton l to $\gamma*$
$y = \frac{\mathbf{P}\cdot\mathbf{q}}{\mathbf{P}\cdot\mathbf{k}} \overset{\text{lab}}{=} \frac{\nu}{E}$	Fractional energy transfer to $\gamma*$
$s = (\mathbf{P}+\mathbf{k})^2 \overset{\text{lab}}{\simeq} M^2 + 2ME$	Center-of-mass energy squared
$W^2 = (\mathbf{P}+\mathbf{q})^2 \overset{\text{lab}}{=} M^2 + 2M\nu - Q^2$	Invariant mass of the hadronic system (X)

Cross section for the inclusive DIS process: $l + N \to l' + X$

The deep inelastic scattering cross section can be written in the one photon exchange approximation (fig. 1.1) in terms of the lepton-photon coupling ($L_{\mu\nu}$) and the hadron-photon ($W^{\mu\nu}$) coupling tensors [5]:

$$\frac{d\sigma}{dE'd\Omega} = \frac{\alpha^2}{2Mq^4}\frac{E'}{E} L_{\mu\nu} W^{\mu\nu} \tag{1.3}$$

where α is the fine structure constant. The leptonic and hadronic tensor are [6]:

$$L_{\mu\nu}(k, s; k') = 2\left\{L^{(S)}_{\mu\nu}(k; k') + iL^{(A)}_{\mu\nu}(k, s; k')\right\} \tag{1.4}$$

with

$$L^{(S)}_{\mu\nu} = k'_\mu k_\nu + k'_\nu k_\mu - g_{\mu\nu}\left(\vec{k'}\vec{k} - m^2\right) \tag{1.5a}$$

$$L^{(A)}_{\mu\nu} = -m\varepsilon_{\mu\nu\alpha\beta}s^\alpha q^\beta \tag{1.5b}$$

and

$$W^{\mu\nu}(q;P,s) = W^{\mu\nu}_{(S)}(q;P) + iW^{\mu\nu}_{(A)}(q;P,S) \tag{1.6}$$

with

$$\begin{aligned}\frac{1}{2M}W^{\mu\nu\ (S)}(q;P) \quad &= W_1(P\cdot q, q^2)\left(-g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\right) \\ &+ \frac{W_2(P\cdot q, q^2)}{M^2}\left(P^\mu - \frac{P\cdot q}{q^2}q^\mu\right)\left(P^\nu + \frac{P\cdot q}{q^2}q^\nu\right)\end{aligned} \tag{1.7a}$$

$$\begin{aligned}\frac{1}{2M}W^{\mu\nu\ (A)}(q;P,S) = \varepsilon_{\mu\nu\alpha\beta}\ q^\alpha\ \Big\{&G_1(P\cdot q, q^2)MS^\beta \\ &+ \frac{G_2(P\cdot q, q^2)}{M}(P\cdot q)S^\beta - (S.q)P^\beta\Big\}\end{aligned} \tag{1.7b}$$

where the lepton and nucleon polarizations[1] are given by s and S, respectively; $g_{\mu\nu}$ is the Minkowski metric[2] and m is the lepton mass. The functions $W_1(P\cdot q, q^2)$, $W_2(P\cdot q, q^2)$, $G_1(P\cdot q, q^2)$ and $G_2(P\cdot q, q^2)$ are the spin averaged and spin dependent *structure functions* and they parametrize the internal structure of the nucleon. They are expressed by the dimensionless functions defined in eqs. (1.8a) to (1.8d).

$$MW_1(P.q, Q^2) = F_1(x, Q^2) \tag{1.8a}$$

$$\nu W_2(P.q, Q^2) = F_2(x, Q^2) \tag{1.8b}$$

$$\frac{(P\cdot q)^2}{\nu}G_1(P.q, Q^2) = g_1(x, Q^2) \tag{1.8c}$$

$$\nu(P\cdot q)G_2(P.q, Q^2) = g_2(x, Q^2) \tag{1.8d}$$

Inserting eqs. (1.5a), (1.5b), (1.7a) and (1.7b) in eq. (1.3) gives:

$$\frac{d\sigma}{dE'd\Omega} = \frac{\alpha^2}{2Mq^4}\frac{E'}{E}\left[L_{\mu\nu(S)}W^{\mu\nu(S)} - L_{\mu\nu(A)}W^{\mu\nu(A)}\right] \tag{1.9}$$

When summing over all spin configurations of the initial and final lepton and nucleon, and neglecting the lepton mass, eq. (1.9) reads:

$$\begin{aligned}\frac{d^2\sigma^{unp}}{dxdQ^2} &= \frac{\alpha^2}{2Mq^4}\frac{E'}{E}2\left[L_{\mu\nu(S)}W^{\mu\nu(S)}\right] \\ &= \frac{4\pi\alpha}{Q^4}\left[y^2F_1(x,Q^2) + \left(\frac{1-y}{x} - \frac{My}{2E}\right)F_2(x,Q^2)\right]\end{aligned} \tag{1.10}$$

This represents the *unpolarized* DIS cross section. Its measurement gives the access to the structure functions F_1 and F_2.

[1] Covariant spin 4-vector proprieties: $s\cdot k = 0$ and $s\cdot s = -1$. Similar for S.

[2] Diagonal matrix where $g_{00} = -1$, $g_{11} = g_{22} = g_{33} = 1$.

1.3 Quark Parton Model

The *Quark Parton Model* (QPM) [4, 7] is developed in the *infinite momentum frame*, where the nucleon has a very large momentum along a certain direction and is composed by massless point-like particles called partons. In particular, the transverse momentum component of the partons can be neglected. The partons are considered to carry a fraction ξ of the nucleon 4-momentum $\mathbf{P}$. In DIS, the virtual photon interacts with the parton and the invariant mass of the initial and final states are $(\xi\mathbf{P}+\mathbf{q})^2$ and 0 respectively, leading to

$$(\xi\mathbf{P}+\mathbf{q})^2 = 0 \Rightarrow 2\xi\mathbf{P}\cdot\mathbf{q}+\mathbf{q}^2 = 0 \Rightarrow \xi = \frac{Q^2}{2\mathbf{P}\cdot\mathbf{q}} \tag{1.11}$$

which is the x definition given in table 1.1.

Within this model, the DIS interaction only occurs with quarks (since the gluons do not have electric charge). The quarks located near the interacting quark are not affected by the interaction. The cross section is the sum over all the individual lepton-parton interactions weighted by the parton charge and momentum. The hadronic tensor defined in eq. (1.6) in this model is [5]:

$$W^{\mu\nu} = \sum_{q,s} e_q^2 n_q(x,s;S)\frac{1}{\mathbf{P}\cdot\mathbf{q}}\left[2x\mathbf{P}^\mu\mathbf{P}^\nu + \mathbf{P}^\nu\mathbf{q}^\mu + \mathbf{P}^\mu\mathbf{q}^\nu - g^{\mu\nu}\mathbf{P}\cdot\mathbf{q}\right] \tag{1.12}$$

where $n_q(x,s;S)$ is the density of quarks q with charge e_q and spin s. The nucleon spin is given by S. The expression for the structure functions in this model are [5]:

$$F_1(x) = \frac{1}{2}\sum_q e_q^2 q(x) \tag{1.13a}$$

$$F_2(x) = x\sum_q e_q^2 q(x) \tag{1.13b}$$

where $q(x)$ are the *unpolarized Parton Distributions Functions* (PDFs). The sum in eqs. (1.13a) and (1.13b) is over all quark flavors. In the parton model, since $Q^2 \to \infty$ and $\nu \to \infty$ the Q^2 dependence of F_1 and F_2 is lost (eqs. (1.13a) to (1.13b)). This is called the *Bjorken scaling*. Another important result of the parton model is the Callan-Gross relation [8]:

$$F_1(x) = \frac{1}{2x}F_2(x) \tag{1.14}$$

The Bjorken scaling and the Callan-Gross relation showed that the partons were indeed point-like particles with spin 1/2. The unpolarized DIS cross section in the QPM is [9]:

$$\frac{d^2\sigma^{unp}}{dxdy} \overset{\text{QPM}}{=} \frac{8\pi\alpha^2 ME}{Q^2}\left[\frac{1}{2}y^2 + \left(1-y-\frac{y^2\gamma^2}{4}\right)\right] x\sum_q e_q^2 q(x) \tag{1.15}$$

Scaling violation

The structure function F_2 has been measured by different collaborations in a wide $x - Q^2$ kinematic range [2]. The measured values are plotted as a function of Q^2 in different x bins in fig. 1.2. The Bjorken scaling is only observed in a small x region (0.1 to 0.4). Outside this region the structure function F_2 has a logarithmic dependence on Q^2. At small x, F_2 increases with Q^2, while at large x, F_2 decreases for large Q^2 values. One conclusion

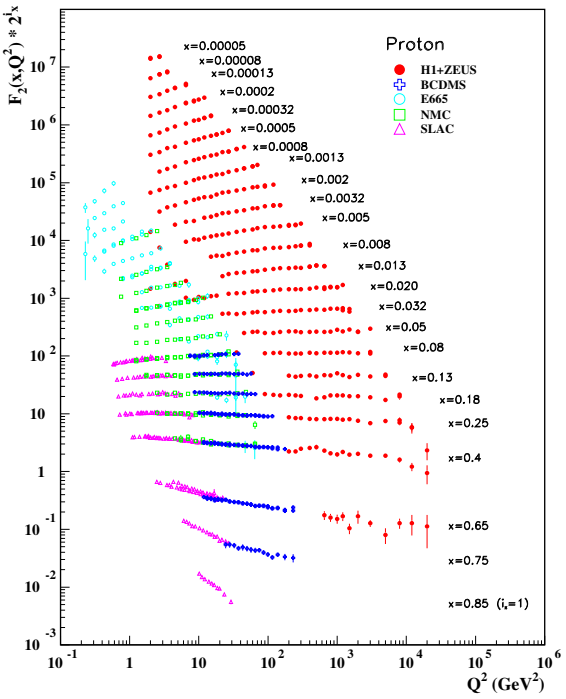


Figure 1.2: *The proton structure function F_2 as a function of Q^2 for different x values. Figure taken from [2].*

derived from the observation of the F_2 scaling violation was that something was missing in the QPM interpretation: the gluon contribution. To include the gluon contribution, the parton model was developed within the Quantum Chromodynamics frame (QCD).

Improved QPM with QCD

The Q^2 dependence observed in fig. 1.2 can be calculated by introducing the quark interaction described in QCD [3, 10]. Quantum Chromodynamics is a gauge theory based on a non-abelian group SU(3), which describes the strong interaction. In this theory, the quarks carry a charge color in addition to the electric charge. The color interactions are mediated by the gluons which are also colored particles. The internal nucleon dynamic is dominated by the gluon emission and absorption by the 3 valence quarks. The gluons

can produce a pair of quark-antiquark or emit gluons. This creates a cloud of gluons and virtual $q\bar{q}$ pairs known as *sea quarks*.

The QCD coupling constant α_s depends on the scale of the interaction: at low energies (long distances) quarks or gluons are always grouped as colorless particles (hadrons) (confinement). At high energies (short distances) the quarks and gluons behave as free particles (asymptotic freedom).

Depending on the energy, a physical process can be cataloged as a *hard* process (large energies, $\alpha_s \sim 0$) or a *soft* one (low energies, $\alpha_s \approx 1$). A hard process can be explained within perturbative QCD (pQCD) while a soft one can only be parametrized from experimental data. This two regimes differ by the factorization scale Λ. In DIS the scale of the interaction is given by Q^2. At high scales ($Q^2 > 1$ GeV^2) the DIS cross section can be *factorized* [11] in terms of soft and hard processes. The hard process involved in DIS is the lepton-quark (gluon) interaction ($\hat{\sigma}_q$) and the soft one is a non-calculable universal quantity: the PDFs.

The resolution of the virtual photon probe is illustrated in fig. 1.3; it is $1/Q^2$. At low energies $Q^2 \approx 0$, the quasi real photon sees the nucleon as a point like particle. While Q^2 increases, the virtual photon starts to be sensitive to the nucleon constituents and interacts with them. For larger Q^2, the virtual photon is sensitive to smaller distances and it is able "to see" the gluons. The first QCD correction to the quark parton model (QPM) concerns the gluon emission by the initial or final quark [10]. The *splitting functions* $P_{ij}(x/\xi)$ are the probability that a quark or gluon of type j and momentum fraction ξ is the parent of i with momentum fraction x [12].

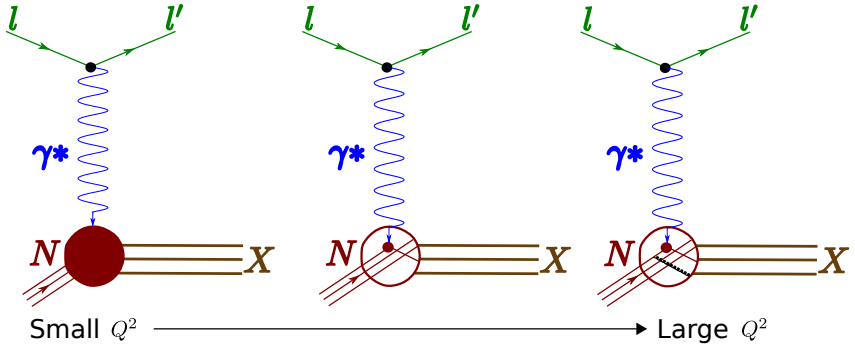


Figure 1.3: *Resolution of the photon probe versus Q^2. Figure taken from [10].*

The Q^2 dependence shown in fig. 1.2 is calculated using the Dokshiter-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [13–16]:

$$\frac{dq_i(x,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\sum_j \int_x^1 \frac{d\xi}{\xi} P_{ij}(x/\xi, \alpha_s(Q^2)) q_i(\xi, Q^2) \tag{1.16}$$

If the PDFs are known at a certain scale of Q^2 (Q_0^2), they can be calculated at any other scale.

1.4 Determination of Parton Distribution Functions

The PDFs are non-perturbative objects and therefore can only be parametrized from experimental data. A global fit to world data is performed. This is possible because PDFs are *universal quantities*, *i.e* they are hard process independent. The current fits include data coming from: lepton-nucleon DIS, collider experiments (pp or $p\bar{p}$), neutrino scattering, *etc* [2]. The different kinematic domains covered by all the experiments permit determining the PDFs in a wide x and Q^2 range.

For the fit it is also necessary to provide the PDF functional forms at a starting scale Q_0^2. The most used functional form is [2] $xq_i(x, Q_0^2) = x^a(...)(1-x)^b$ and the number of free parameters varies from 10 to 25. The DGLAP equations are used to evolve the PDFs functional forms to the Q^2 of the experimental data.The PDFs resulting from a global fit performed by the group MSTW [17] are shown at Leading Order (LO) for two different Q^2 values in fig. 1.4.

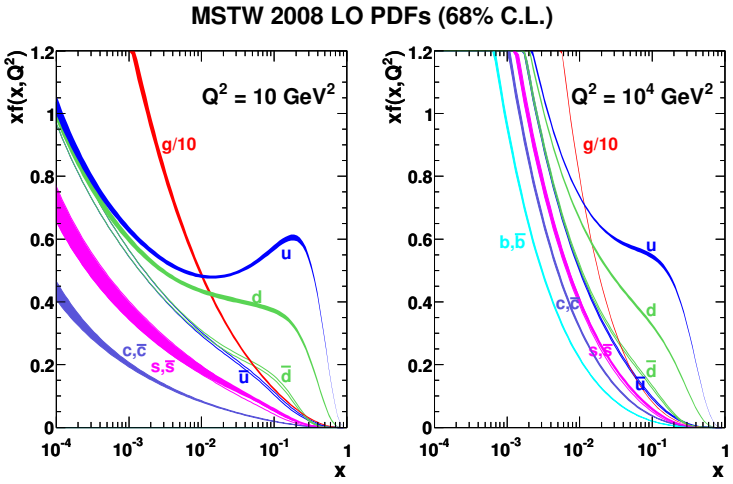


Figure 1.4: *Unpolarized PDFs at leading order (LO) (MSTW) at $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right). Figure taken from [17].*

1.5 Semi-Inclusive Deep Inelastic Scattering

In a *Semi-Inclusive Deep Inelastic Scattering Process* (SIDIS), a hadron is detected in coincidence with the scattered lepton ($l + N \to l' + h + X$). New variables are associated to the hadron (see table 1.2). A invariant one is z, the energy fraction of the virtual photon transferred to the produced hadron h (eq. (1.17)):

$$z = \frac{\mathbf{P} \cdot \mathbf{p_h}}{\mathbf{P} \cdot \mathbf{q}} \overset{\text{lab}}{=} \frac{E_h}{\nu} \tag{1.17}$$

Table 1.2: *Kinematic variables realated to the hadron in SIDIS.*

Variable	Description
$\mathbf{p} = (E_h, \vec{p_h})$	Hadron 4-momentum vector
$p_{h_\parallel} = p_h$	Component of $\vec{p_h}$ along $\vec{q}$
$p_{h_\perp}$	Transverse component of $\vec{p_h}$ with respect to q
θ_h	Angle between $\vec{q}$ and $\vec{p_h}$
ϕ_h	Angle between the scattering plane and the hadron production plane.
z	Energy fraction of the virtual photon transferred to the produced hadron h

The semi-inclusive cross section reads [18]:

$$\frac{d\sigma}{d^x dy dz} \overset{\text{QPM}}{=} \frac{8\pi\alpha^2 ME}{Q^4}\left[xy^2 H_1(x, Q^2, z) + (1-y)H_2(x, Q^2, z)\right] \tag{1.18}$$

where H_1 and H_2 are structure functions related to F_1 and F_2 [9, 18]:

$$\sum_h \int_0^1 dz z H_i(x, Q^2, z) = F_i(x, Q^2), \quad i = 1, 2 \tag{1.19}$$

SIDIS in QPM

The production of hadrons in DIS in the QCD improved quark parton model can be seen as two independent sub-processes (thanks to the *factorization*): first a quark q absorbs the virtual photon $\gamma*$ (hard process $\hat{\sigma}_q$), then the quark q fragments into a hadron h (soft process) (fig. 1.5).

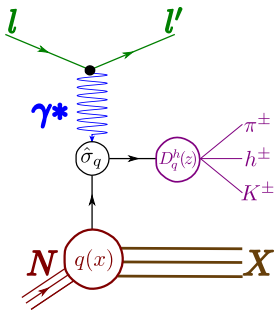


Figure 1.5: *Factorization in semi-inclusive deep inelastic scattering.*

The structure functions $H_i(x, Q^2, z)$ contain the information on what happens to the struck quark after the interaction with the virtual photon, the *quark fragmentation*. The *fragmentation function* $D_q^h(z, Q^2)$ is defined as the probability for a quark of flavor q to fragment into a hadron h with a fraction of energy z. The expression of the semi-inclusive

DIS cross section can be written in terms of the fragmentation functions (FFs) and the PDFs [9, 19]:

$$\frac{d^3\sigma^{unp}}{dxdydz} \overset{\text{LO}}{=} \frac{8\pi\alpha^2 ME}{Q^2}\left[\frac{1}{2}y^2 + \left(1 - y - \frac{y^2\gamma^2}{4}\right)\right] x\sum_q e_q^2 q(x) D_q^h(z) \tag{1.20}$$

Scaling and Q^2 evolution

A scaling violation is observed for the FFs measured in e^+e^- annihilation (fig. 1.6) [2]. The scaling is observed in the $x = 2p_h/\sqrt{s}$ bin 0.1-0.2. At low x ($x < 0.1$) the FFs increases with $\sqrt{s}$, while at large x, the FFs are shifted towards lower values for large Q^2 .

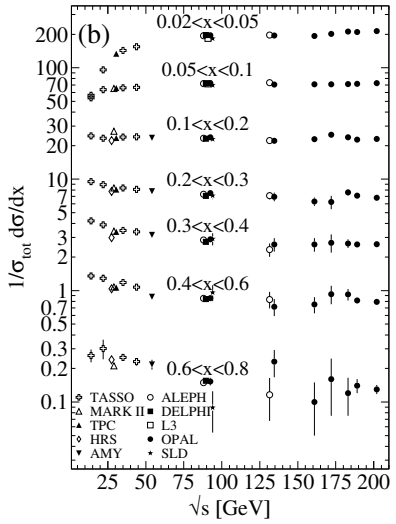


Figure 1.6: *The e^+e^- FFsn for all charged particles as a function of $\sqrt{s}$ for different $x = 2p_h/\sqrt{s}$ values. Figure taken from [2]*

The $D_q^h(z, Q^2)$ fragmentation functions are evolved via the DGLAP equations [13–16]:

$$\frac{dD_q^h(z,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\sum_j \int_x^1 \frac{d\xi}{\xi} P_{qj}(z/\xi, \alpha_s(Q^2)) D_q^h(\xi, Q^2) \tag{1.21}$$

The quark and gluon fragmentation is illustrated in fig. 1.7. The fragmentation of a quark q_i via the decay of the same quark q_i after emitting a gluon G is represented in fig. 1.7a ($P_{qq}D_{q_i}^h$). The fragmentation is also possible through the decay of a gluon G ($P_{Gq}D_G^h$)(fig. 1.7b). In fig. 1.7c, the fragmentation of a gluon caused by splitting into a quark-antiquark and subsequent fragmentation of the quark into hadron is illustrated ($P_{qG}D_{q_i}^h$). The last figure, fig. 1.7d, represents the gluon fragmentation via the three-gluon self-interaction ($P_{GG}D_G^h$).

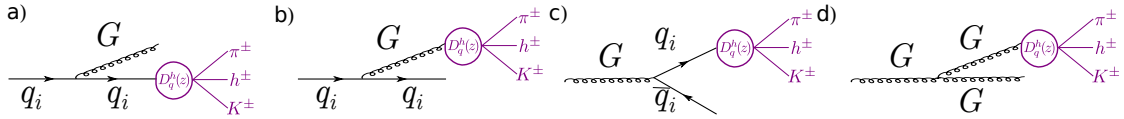


Figure 1.7: *a) The fragmentation of a quark q_i emitting a gluon G via the decay of q_i into a hadron h ($P_{qq}D^h_{q_i}$). b) The fragmentation of q_i through gluon G ($P_{Gq}D^h_G$). c) The gluon fragmentation caused by a $q\bar{q}_i$ pair creation and the subsequent decay of q_i ($P_{qG}D^h_{q_i}$). d) The gluon fragmentation by a three-gluon vertex ($P_{GG}D^h_G$). Figure taken from [20].*

1.6 Fragmentation Functions

The FFs describe the quark fragmentation into a hadron (also known as *hadronization*). They are involved in processes such as SIDIS where at least one hadron is found in the final state. The quark fragmentation into a hadron is soft process and cannot be described using pQCD. Different models have been developed to describe how quarks confine together to make a hadron [21].

1.6.1 Lund String Fragmentation Model

In the Lund string model [22], the hadron production is explained by the creation of quark-antiquark pairs ($q\bar{q}$). The interaction between partons (color force) is represented by a string (fig. 1.8a). The energy inside the string increases linearly as the distance between the constituting partons increases. When the stored energy is large enough to create a new $q\bar{q}$ pair the original string breaks into smaller strings. For each new string the same process is repeated until finally only hadrons remain. The hadronization scheme within the Lund model is illustrated in the center of mass frame in fig. 1.8b. The virtual photon is absorbed by the u quark and in consequence the u quark is ejected from the bound state. In this case, the u quark bounds with the $\bar{d}$ quark to form a π^+ which carries a fraction z of the available energy ν (see table 1.2 and table 1.1, respectively). The remaining system (in this case the d quark and ud quark pair) repeats the fragmentation process until the energy is smaller than the available energy (ν).

1.6.2 Quark Fragmentation Regions

Until now, only the hadrons produced from the struck quark were considered. However, it is also possible that a quark which is not involved in the scattering process fragments into hadrons. In order to study these phenomena two regions are defined by the magnitude of p_h [9]: the *target fragmentation* region where the final hadron h moves with small momentum in the rest frame of the target and the *current fragmentation* one, where the product $\mathbf{P}\cdot\mathbf{p_h}$ grows withQ^2. The study of the contamination by the hadrons produced in the target fragmentation region was performed by Berger in [9] using the rapidity of the final state η (eq. (1.22)). The η sign is related to the target and current regions. If $\eta > 0$

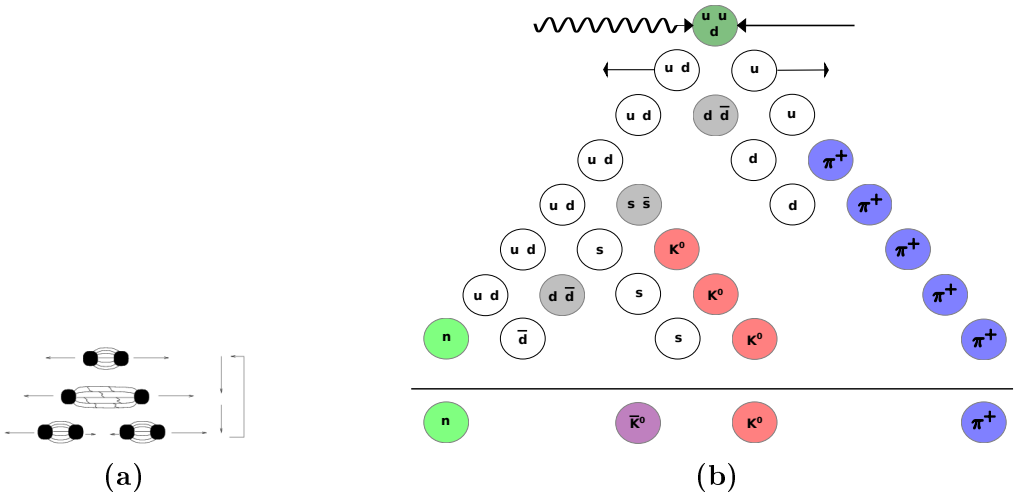


Figure 1.8: *The fragmentation process in the Lund model. Figure taken from [19].*

the hadron moves in the direction of the virtual it is a *current hadron*, while if $\eta < 0$ the hadron is considered as a *target remnant* (fig. 1.9).

$$\eta = \frac{1}{2} \ln \frac{E_h + p_h}{E_h - p_h} \tag{1.22}$$

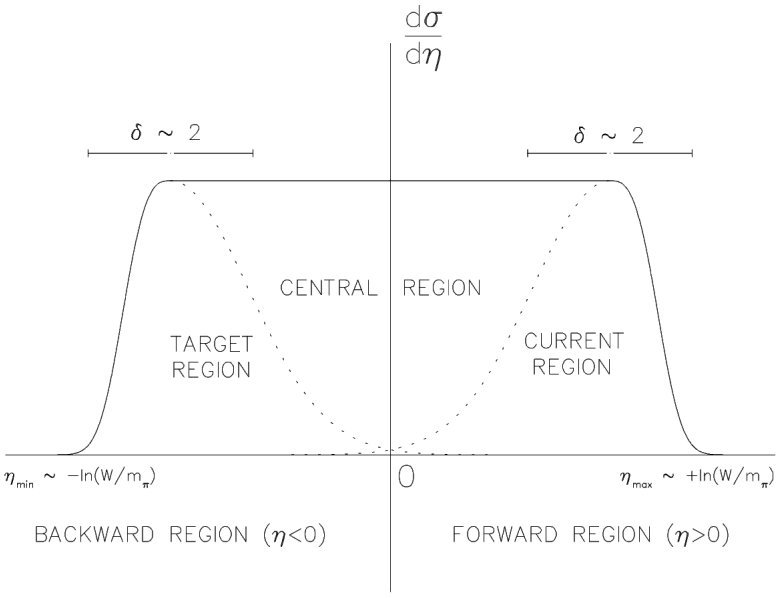


Figure 1.9: *Illustration of typical hadronic rapidity distribution (η). Figure taken from [23]*

The typical hadronic correlation length in rapidity is $\delta \sim 2$. Therefore to separate the

current and target regions it is necessary[3] that $\Delta\eta = \eta_{max} - \eta_{min} \geq 2\delta$ or $W \gtrsim 7.4\text{GeV}$. This phenomenological requirement is known as the *Berger criterion*. It is important to be able to differentiate between the current and target fragmentation to ensure the quark fragmentation into hadrons is independent of the production of the quark.

1.6.3 Fragmentation Function Symmetries

The FFs $D_q^h(z, Q^2)$ are defined for any q flavor and h hadron. Within the QCD improved quark parton model, some symmetries such as the charge-conjugation or isospin are inherited by the FFs. For the present work only the FFs from the light quark (u, $\bar{u}$, d, $\bar{d}$, s and $\bar{s}$) and gluons are considered because the mass threshold for heavy quarks is higher than the covered kinematic domain.

There are two main groups of FFs: if the q quark which fragments into the hadron h is a valence quark of h the FFs is said to be *favored* (D_{fav}^h), if the q quark belongs to the sea quark the fragmentation function is *unfavored* (D_{unf}^h).

Quark FFs into pions There are twelve quark fragmentation function for pions, half of them correspond to π^+ and the other one to π^-. Using the charge-conjugation symmetry, the number of independent FFs reduces to six. If the isospin asymmetry is considered (π^+ case $D_u^{\pi^+} = D_{\bar{d}}^{\pi^+}$) only three FFs stay independent.

$$D_{fav}^h : D_u^{\pi^+} = D^{\pi_{\bar{u}}^-}, D_{\bar{d}}^{\pi^+} = D_d^{\pi^-} \tag{1.23a}$$

$$D_{unf}^h : D_{\bar{u}}^{\pi^+} = D_u^{\pi^-}, \quad D_d^{\pi^+} = D_{\bar{d}}^{\pi^-}, D_s^{\pi^+} = D_{\bar{s}}^{\pi^-}, \quad D_s^{\pi^-} = D_{\bar{s}}^{\pi^+} \tag{1.23b}$$

Additional hypotheses can be applied. For example, the favored FFs can be considered to be the same as well as the unfavored ones (leading to two FFs).

Quark FFs into kaons In the kaon case, only the charge-conjugation symmetry is applied, leaving six independent FFs.

$$D_u^{K^+} = D_{\bar{u}}^{K^-}, \quad D_u^{K^-} = D_{\bar{u}}^{K^+} \tag{1.24a}$$

$$D_d^{K^+} = D_{\bar{d}}^{K^-}, \quad D_d^{K^-} = D_{\bar{d}}^{K^+} \tag{1.24b}$$

$$D_{\bar{s}}^{K+} = D_s^{K^-}, \quad D_{\bar{s}}^{K^-} = D_s^{K^+} \tag{1.24c}$$

Further assumptions can be made. The most commonly used are to group the quark FFs into kaons in three categories: the favored ($D_u^{K^+} = D_{\bar{u}}^{K^-}$), the strange ($D_{\bar{s}}^{K+} = D_s^{K^-}$) and the unfavored ones ($D_s^{K^+}, D_{\bar{u}}^{K^+}, D_{\bar{d}}^{K^+}, D_d^{K^+}$) (leading to three FFs).

[3] The maximum rapidity is reached when the hadron momentum $p_h = \frac{W}{2} \to \eta_{max} \sim \ln \frac{W}{M_h}$, while for the minimum $\eta_{min} \sim -\ln \frac{W}{M_h}$.

1.7 Quark Fragmentation Functions: current knowledge

1.7.1 Measurements

The same non-perturbative functions D_q^h are used to describe the quark hadronization in all processes. The *universality* of the FFs has been experimentally tested [2, 24]. Several scattering processes such as electron-positron annihilation, lepton-nucleon SIDIS and hadron-hadron collisions ($pp(\bar{p})$ for example) are used to access the quark FFs. The e^+e^- data (LEP [25], SLAC [26], BaBar [27] and BELLE [28])provide a "clean" access to the FFs since the $\sigma^{e^+e^-\to q+\bar{q}\to hX}$ cross sections do not depend on the PDFs. They do not allow the separating D_q^h and $D_{\bar{q}}^h$. The data from hadron-hadron collisions data (UA5 [29], UA1 [30], ALICE [31], CMS [32], RHIC [33], ATLAS [34]) give access to D_q^h and $D_{\bar{q}}^h$. The data are sensitive to the gluon fragmentation function D_g^h.

The quark fragmentation information obtained from a SIDIS process can be compared to the two previous ones if the data belong to the current fragmentation region (section 1.6.2). There are several collaborations contributing to enlarge the data base from SIDIS interactions: EMC [35] (muon beam, proton and deuteron target), E00-108 [36] (electron beam, proton and deuteron target), HERMES [37] (electron beam, proton and deuteron target) and COMPASS (only deuteron data analyzed so far). A preliminary version of COMPASS results has been presented in [38, 39]. In 2012, a SIDIS data taking with a proton target was performed. One of the advantages of SIDIS is that factorization has been proven to all orders in α_s [11]. In contrast to e^+e^- annihilation, the SIDIS measurements are performed at much lower scales providing a wider coverage in Q^2.

A summary of the different aspects from e^+e^- annihilation, hadron-hadron collision and semi-inclusive DIS process regarding the extraction of the fragmentation function is given in table 1.3.

1.7.2 Fragmentation Functions access in SIDIS

The relevant observable in SIDIS to access D_q^h is the hadron multiplicities, which is defined as the average number of hadrons produced per DIS interaction defined:

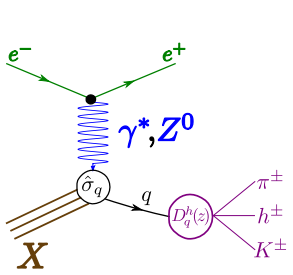
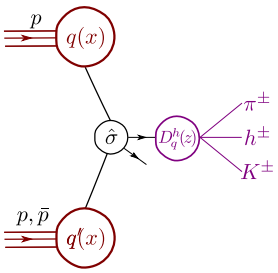
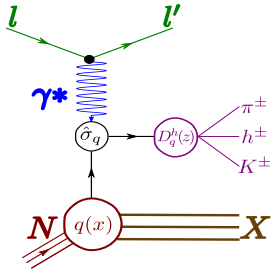
$$M(x,y,z) = \frac{1}{d\sigma^{lN\to l'X}} \frac{d\sigma^{lN\to lh'X}}{dz} \tag{1.25}$$

Using the expression for the DIS and SIDIS cross sections given in eqs. (1.15) and (1.20), eq. (1.25) becomes:

$$M(x,Q^2,z) = \frac{\sum_q e_q^2 q(x,Q^2) \otimes D_q^h(z,Q^2)}{\sum_q e_q^2 q(x,Q^2)} \tag{1.26}$$

The fact PDFs and FFs depend on two different variables x and z is a great advantage since at LO the convolution in eq. (1.26) becomes a simple product. If hadron multiplicities are measured separately for positive and negative charge, from eq. (1.26) it is possible to differentiate between D_q^h and $D_{\bar{q}}^h$. The details on how the FFs are extracted from COMPASS data set are given in section 7.4.

Table 1.3: *Fragmentation functions access in different processes.*

	e^+e^- annihilation	hadron-hadron collision	SIDIS
			
High precision data	✓	✓	✓
PDF Indep.	✓	✗	✗
Access D^h_{fav} & D^h_{unfav}	✗	✓	✓
Access parton kinematics	✓	✗	✓
Theoretical Calculation	LO, NLO, NNLO	LO, NLO	LO, NLO

1.7.3 Global Fits of multiplicities data and Parameterizations of FFs

In order to provide a general parameterization of the FFs, the *universality* of FFs is exploited to perform a global QCD analysis (fit) of available data on multiplicities from various processes. Most parameterizations are based on information from e^+e^-: Kniehl-Kramer-Pöter (KKP) [24], Kretzer (KRE) [40] and Hirai-Kumano-Nagai-Sudoh (HKNS) [41]. The Albino-Kniehl-Kramer (AKK) [42] parameterization uses some hadron-hadron scattering data in addition. The de Florian-Sassot-Stratmann (DSS) [43,44] collaboration uses the three type of data available and Leader-Sidorov-Stamenov (LSS) [45,46] uses only the semi-inclusive data from [37,47] and [38]. In addition, different assumptions in the symmetries are considered and different functional forms are used for D^h_q. A summary

of these parameterization is provided in table 1.4 [48]. Only HKNS, AKK and DSS are described in the following because they reflect all the parameterizations types.

Parametrization	Year	Data			No. FFs fitted	
		SIDIS	pp	e^+e^-	π	K
KKP [24]	2000	✗	✗	✓	5	5
KRE [40]	2001	✓	✗	✓	2	3
HKNS [41]	2007	✗	✗	✓	2	2
AKK [42]	2008	✗	✓	✓	3	5
DSS [43, 44]	2014	✓	✓	✓	4	3
LSS [45, 46]	2014	✓	✗	✗	3	–

Table 1.4: *Parameterization of FFs for pions and kaons.*

HKNS parameterization . The functional form for the FFs is given by:

$$D_i^h(z, Q^0) = N_i^h z^{\alpha_i^h} (1-z)^{\beta_i^h} \tag{1.27}$$

where N_i^h, α_i^h and β_i^h are parameters determined by the fit. The different assumptions for the FFs for pions and kaons are summarized in table 1.5. Without these assumptions, it is not possible to differentiate between the valence quarks (favored) from a sea quark (unfavored) because the e^+e^- cross section allows only to extract $D_{q+\bar{q}}^{h^+ + h^-}$. At LO for pions, all favored FFs are assumed to be equal; all the unfavored ones are assumed equal between themselves. In the kaon case, the same assumptions are used for the unfavored FFs. A difference is made between the two favored functions $D_u^{K^+} \neq D_{\bar{s}}^{K^+}$. The charm and bottom quark FFs are also fitted.

	Pions
Favored	$D_u^{\pi^+} = D_{\bar{d}}^{\pi^+} = D_d^{\pi^-} = D_{\bar{u}}^{\pi^-}$
Unfavored	$D_{\bar{u}}^{\pi^+} = D_d^{\pi^+} = D_{\bar{d}}^{\pi^-} = D_u^{\pi^-} = D_s^{\pi^+} = D_{\bar{s}}^{\pi^+} = D_s^{\pi^-} = D_{\bar{s}}^{\pi^+}$
Gluons	$D_g^{\pi^+} = D_g^{\pi^-}$

	Kaons
Favored	$D_u^{K^+} = D_{\bar{u}}^{K^-}$
Unfavored	$D_{\bar{u}}^{K^+} = D_d^{K^+} = D_{\bar{d}}^{K^+} = D_s^{K^+} = D_{\bar{s}}^{K^-} = D_u^{K^-} = D_d^{K^-} = D_{\bar{d}}^{K^-}$
Strange	$D_{\bar{s}}^{K^+} = D_s^{K^-}$
Gluons	$D_g^{K^+} = D_g^{K^-}$

Table 1.5: *Assumptions in the HKNS parameterization.*

AKK parameterization . This parameterization uses data from e^+e^- annihilation and pp collision. The FFs are parametrized with the functional form:

$$D_i^h(z,Q^0) = N_i^h z^{\alpha_i^h}(1-z)^{\beta_i^h}(1+\gamma^{h_i^h}(1-z)^{\delta^{h_i^h}_i} \tag{1.28}$$

For pions, all favored FFs are considered equal and two different FFs are used in the unfavored case. For kaons, all favored and unfavored FFs are not equal and a total of 5 FFs is fitted (in addition the heavy quark FFs) (table 1.6).

	Pions
Favored	$D_u^{\pi^+} = D_{\bar{d}}^{\pi^+} = D_{\bar{d}}^{\pi^-} = D_{\bar{u}}^{\pi^-}$
Unfavored	$D_{\bar{u}}^{\pi^+} = D_{\bar{d}}^{\pi^+} = D_{\bar{d}}^{\pi^-} = D_u^{\pi^-}, D_s^{\pi^+} = D_{\bar{s}}^{\pi^+} = D_s^{\pi^-} = D_{\bar{s}}^{\pi^+}$
Gluons	$D_g^{\pi^+} = D_g^{\pi^-}$
	Kaons
Favored	$D_u^{K^+} = D_{\bar{u}}^{K^-}$
Unfavored	$D_{\bar{u}}^{K^+} = D_u^{K^-}, \quad D_s^{K^+} = D_{\bar{s}}^{K^-}, \quad D_d^{K^+} = D_{\bar{d}}^{K^+} = D_d^{K^-} = D_{\bar{d}}^{K^-}$
Strange	$D_{\bar{s}}^{K^+} = D_s^{K^-}$
Gluons	$D_g^{K^+} = D_g^{K^-}$

Table 1.6: *Assumptions in the AKK parameterization.*

DSS parameterization . This was the first collaboration that determined the individual FFs for quark and antiquark. The functional form is more complex than the HKNS one and is more flexible:

$$D_i^h(z,Q^0) = \frac{N_i^h z^{\alpha_i^h}(1-z)^{\beta_i^h}\left[1+\gamma_i(1-z)^{\delta_i}\right]}{B[2+\alpha_i,\beta_i+1]+\gamma_i B[2+\alpha_i,\beta_i+\delta_i+1]} \tag{1.29}$$

For pions, the two favored (unfavored) FFs are related via a proportionality factor N. In addition, the isospin symmetry is considered only for the unfavored case but a slightly different normalization is assumed in the $q+\bar{q}$ sum ($D_{d+\bar{d}}^{\pi^+} = N_{\pi^-} D_{u+\bar{u}}^{\pi^+}, \quad D_{d+\bar{d}}^{\pi^-} = N_{\pi^-} D_{u+\bar{u}}^{\pi^-}$). In total four, FFs are fitted, two favored and two unfavored. For kaons, two independent favored FFs are considered. For the unfavored fragmentation the same functional form is assumed for all flavors (table 1.7). The FFs of heavy quarks is also considered.

The favored and unfavored FFs extracted by HKNS (red lines) and DSS (black lines) at LO for $Q^2 = 3(\mathrm{GeV}/c)^2$ are plotted as a function of z for π^+ and K^+ in figs. 1.10a and 1.10b, respectively. $D_u^{\pi^+}$ differ in the z region (0.1-0.75). The difference is smaller for $D_{\bar{d}}^{\pi^+}$. The DSS favored FFs (two top figures in fig. 1.10a) are slightly lower than the ones obtained by HKNS. The unfavored FF $D_{\bar{u}}^{\pi^+}$ (bottom left plot in fig. 1.10a) differ significantly at large z. A different behavior is observed in the gluon FF into pions in both parameterizations. In the DSS case the gluon contribution is better constrained with respect to the HKNS

	Pions
Favored	$D_u^{\pi^+} = N_{\pi^+} D_{\bar{d}}^{\pi^+}, \quad D_d^{\pi^-} = N_{\pi^-} D_{\bar{u}}^{\pi^-}$
Unfavored	$D_{\bar{u}}^{\pi^+} = D_d^{\pi^+}, \quad D_{\bar{s}}^{\pi^+} = D_s^{\pi^+} = N'_{\pi^+} D_{\bar{u}}^{\pi^+}$ $D_{\bar{d}}^{\pi^-} = D_u^{\pi^-}, \quad D_{\bar{s}}^{\pi^-} = D_s^{\pi^-} = N'_{\pi^-} D_u^{\pi^-}$
Gluons	$D_g^{\pi^+}, \quad D_g^{\pi^-}$
	Kaons
Favored	$D_u^{K^+}, \quad D_{\bar{u}}^{K^-}$
Unfavored	$D_{\bar{u}}^{K^+} = D_d^{K^+} = D_{\bar{d}}^{K^+} = D_s^{K^+}$ $D_{\bar{s}}^{K^-} = D_u^{K^-} = D_d^{K^-} = D_{\bar{d}}^{K^-}$
Strange	$D_{\bar{s}}^{K^+}, , \quad D_s^{K^-}$
Gluons	$D_g^{K^+}, \quad D_g^{K^-}$

Table 1.7: *Assumptions in the DSS parameterization.*

due to the use of pp collision data. In the kaon case, HKNS and DSS FFs differ in a more important way than in the pion case, especially for the $\bar{s}$ fragmentation into K^+ where the shapes are very different. The disagreement between the different parameterization sets can be understood by the different assumptions used. In the case of kaons the lack of experimental data is an important factor.

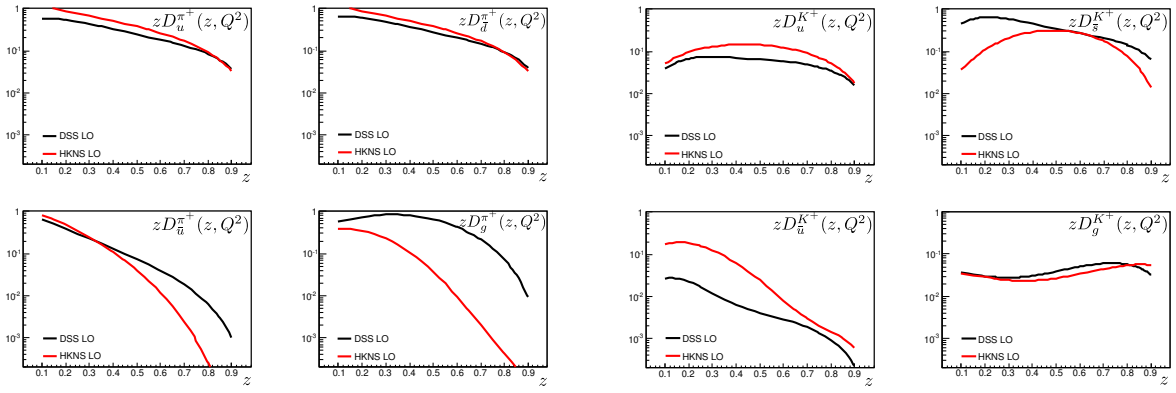


Figure 1.10: *Comparison between D_u^h, $D_{\bar{s}}^h$, $D_{\bar{u}}^h$ and $D_{\bar{g}}^h$ FFs extracted by HKNS (red lines) and DSS (black lines) parameterizations at LO for $Q^2 = 3(GeV/c)^2$. Left: π^+. Right:K^+. [49].*

1.8 Summary

The DIS was presented as a powerful tool to study the structure of the nucleon. The unpolarized PDFs are well constrained in a wide kinematic domain for the u, $\bar{u}$, d and $\bar{d}$ quarks. However, in the case of the s and $\bar{s}$ quarks large uncertainties still subsist. The quark FFs are defined as universal and process independent objects which parametrize the quark hadronization. They can be accessed via the analysis of data where hadron production is involved such as electron-positron annihilation, hadron-hadron collision or SIDIS.

Several parameterizations of quark FFs already exists based on LO and NLO analyzes of various data sets. They differ significantly for the strange quark sector. The COMPASS data presented in this thesis will provide a significant amount of new data which will help solving the discrepancies.

Chapter 2

The COMPASS Experiment at CERN

In this chapter a brief description of the COMPASS[1] experiment is provided. The generalities of the experimental apparatus are briefly described in section 2.1. The beam and target are presented in sections 2.2 and 2.3, respectively. The description of detectors used in particle tracking is in section 2.4 while the detectors used for the particle identification are described in section 2.5. The trigger system is described in section 2.6. The two final sections are dedicated to the data acquisition system and data reconstruction. The information provided in this chapter is inspired from [50].

2.1 General Overview

COMPASS is a high energy, high rate and fixed target experiment at the Super Proton Synchrotron (SPS) at CERN. The purpose of this experiment is the study of hadron structure and hadron spectroscopy with high intensity muon and hadron beams.

The COMPASS spectrometer illustrated in fig. 2.1 covers a wide kinematic domain: values of Q^2 up to 100 $(\mathrm{GeV}/c)^2$ and values of x down to 10^{-5} for incident energies (muon beam) of 160 GeV/c.

The apparatus can be divided in three parts: the first part contains the detectors in charge of the incoming beam detection and is located upstream the target location. The second and third parts are located downstream of the target and extend over a total length of 50 m. The second part is built around the analyzing magnet SM1 and is known as the *Large Angle Spectrometer* (LAS); it has been designed to ensure 180 mrad polar acceptance. The third section is the *Small Angle Spectrometer* (SAS). It is built around the analyzing magnet SM2 and measures the particles emitted at small angles ($\pm$30mrad).

In 2006, the data taking was performed with a 160 GeV/c muon beam impinging on a deuteron target (^{6}LiD).

[1] COmmon Muon and Proton Apparatus for Structure and Spectroscopy

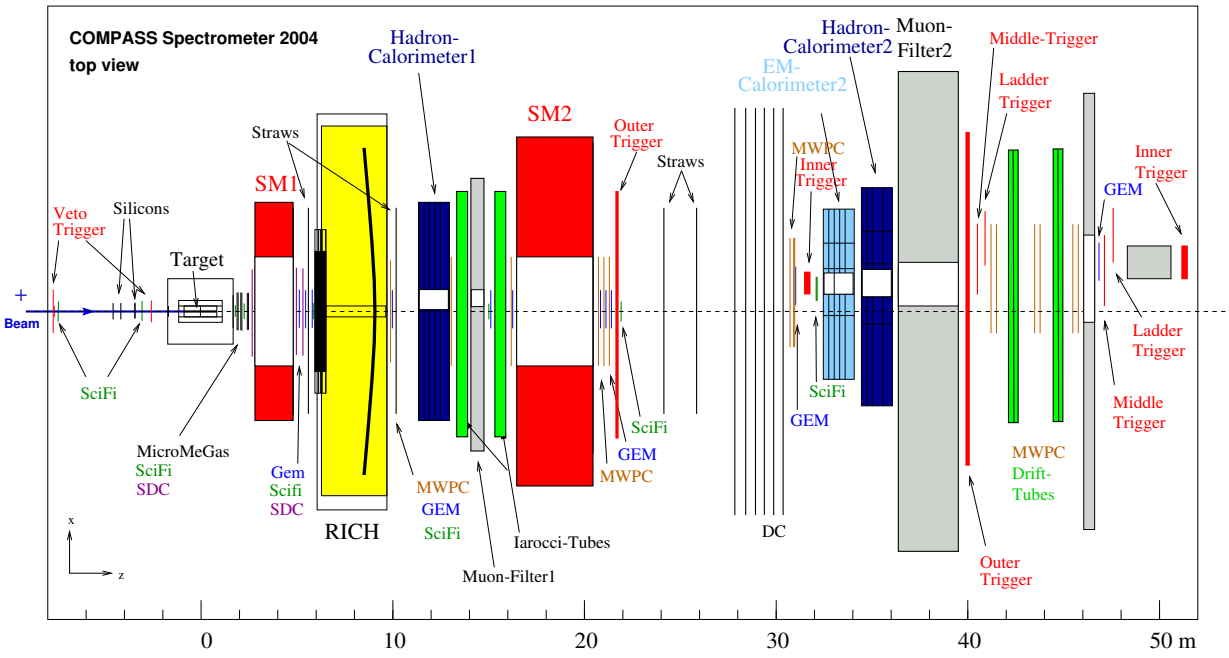


Figure 2.1: *COMPASS 2004 muon setup top view. Figure taken from [50].*

2.2 Muon Beam

The muon beam used by COMPASS is obtained from a primary proton beam, accelerated in the SPS up to 400 GeV/c. The proton beam interacts with a 50 cm thick beryllium target (known as the T6 target), producing pions and kaons. The proton beam is delivered onto the T6 target within a time period window of 4.8 s (*spill* time). The produced pions are transported along a 600 m long channel of the M2 beam line [51] leading to around 5% of the pions to decay into muons and neutrinos. Both pions and a large fraction of the muons produced in the decays are transported until the muons are focused and the hadrons are stopped in a hadron absorber. After the absorber, a system of magnets is used to select and focus the muons of 160 GeV/c. The muons coming from the $\pi \to \mu\nu$ weak decay channel are polarized, however this feature is not exploited in the present analysis.

The muon beam which arrives at the experimental area has transverse dimensions of $\sigma_x \times \sigma_y \approx 8 \times 8$ mm^2 and an angular divergence of $\sigma_{\theta x} \times \sigma_{\theta y} \approx 0.5 \times 1$ mrad^2. At each spill $2 \cdot 10^8$ muons reach the experimental area. The beam is accompanied by a large muon *halo* that extends transversely up to several meters of distance with respect to the beam line, its intensity decreases with the distance. The high intensity halo near the beam line is measured by a dedicated veto counter (size 30×30 cm^2), with a 4 cm diameter central hole) and it represents $\sim 16\%$ of the muon beam. The far halo, whose intensity is lower, is measured by a large veto counter, with a central hole of 30×30 cm^2, where it amounts to 7% of the nominal muon beam.

The *Beam Momentum Station* (BMS) is used for the determination of the incident muon momentum. It consists of six scintillator hodoscopes (BM01-BM06) located symmetrically

upstream and downstream a bending magnet (B6: three consecutive dipole magnets) which is surrounded by four quadrupole systems (Q29-Q32). The BMS is illustrated in fig. 2.2.

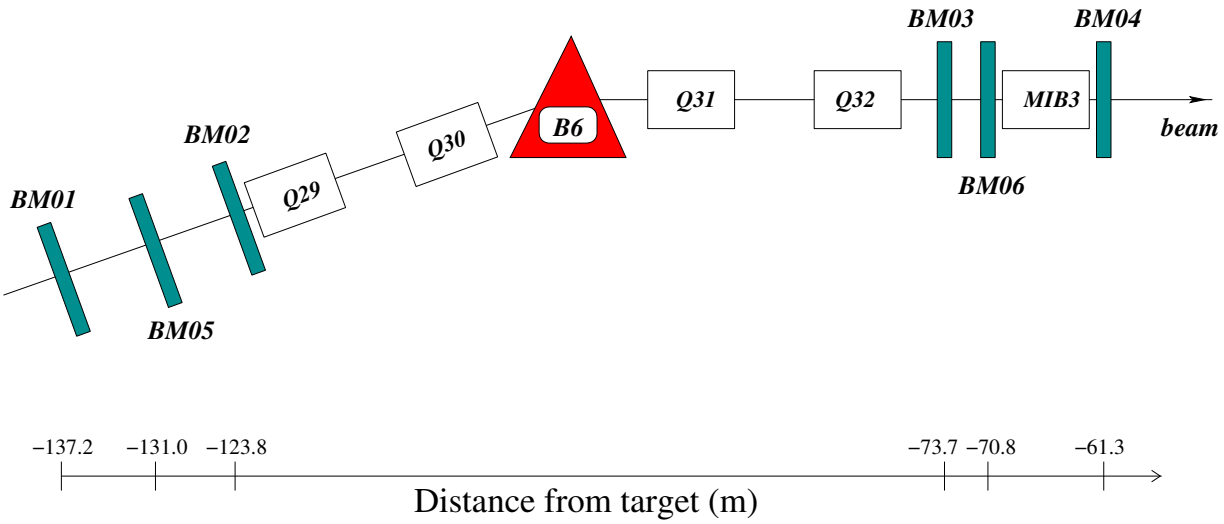


Figure 2.2: *Layout of the Beam Momentum Station for the COMPASS muon beam. Figure taken from [50].*

The BMS system was designed to measure the momentum of more than 10^8 individual particles per spill with a relative precision of 0.5%. In order to eliminate ambiguities in the reconstruction of particle trajectories, their time of transit is measured with a resolution of 50 ps.

2.3 Target

The target is composed of three cylindrical target cells. Each cell contains small pieces of ^{6}LiD and is filled with a ^{3}He/^{4}He cooling liquid mixture. The central cell is 60 cm long and the up/downstream cells are 30 cm long; the gaps between cells are 5 cm long (fig. 2.3).

The COMPASS target is designed to be a polarized target however this feature is not relevant for the present analysis. The main components of the target system are: the refrigerator system which is used to conserve the target temperature between 55 and 95 mK, a superconducting solenoid which produces a homogeneous magnetic field of 2.5 T along the beam direction with an angular aperture of $\sim$ 180 mrad and a dipole coil magnet with a 0.42 T magnetic field.

2.4 Tracking Detectors

The *Large Angle Spectrometer* (LAS) and *Small Angle Spectrometer* (SAS) are equipped with different types of tracking detectors. Depending on the detection cover area of each

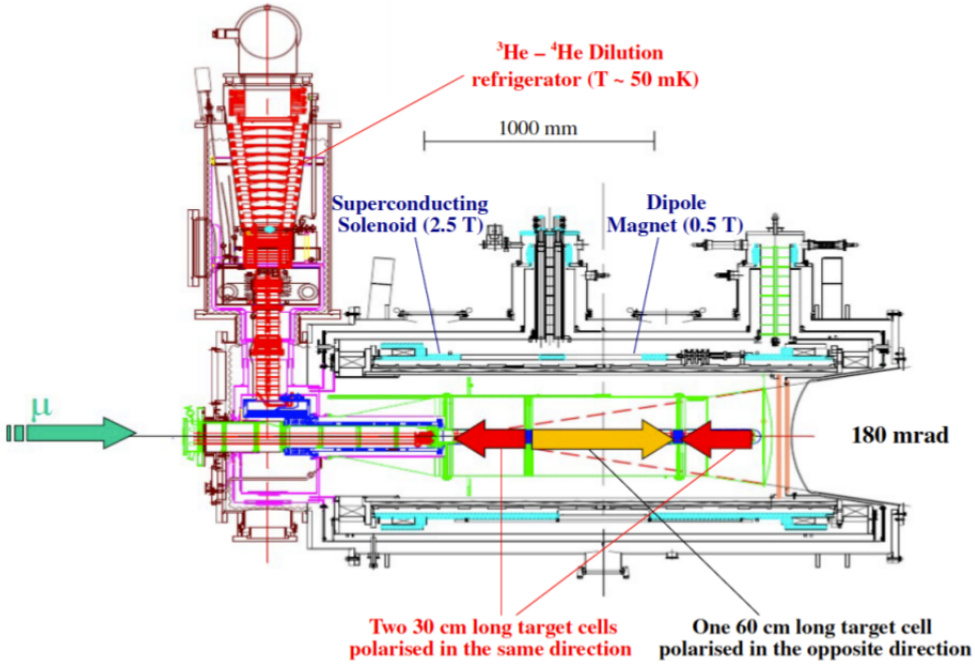


Figure 2.3: *2006 COMPASS target. Figure taken from [52].*

detector they are classified as: *very small area trackers*, *small area trackers* and *large area trackers*.

2.4.1 Very Small Area Trackers

The very small area trackers cover the transverse beam size up to ~ 3 cm. In this region the particle rate is very high (10^5/s/mm^2 in the center of the muon beam) thus the tracking detectors must have an excellent time and (or) position resolutions.

The scintillating fibre detectors are the smallest ones in the spectrometer and their size vary between ~ 16 cm^2 and ~ 144 cm^2. The best characteristic of this kind of detectors is their time resolution ~ 500 ps. All along the apparatus there are 9 stations composed by two or three scintillating fibre detectors. In each station the orientation of the detectors is orthogonal. Where there is a third detector, this last one is orientated 45° with respect to the others.

The silicon detector size is 30 cm^2 with a space and time resolution of ~ 10 μm and < 2.5 ns. The measurement of the incoming muon beam track is improved with these detectors in addition to the BMS. There are 3 silicon detectors stations located upstream the target; each one of these stations is composed by two silicon detectors where one of them is rotated by 5° with respect to the other.

2.4.2 Small Area Trackers

The radial region between 2.5 cm and 20 cm is covered by the small area trackers. Two types of gaseous detectors which combine a high rate capability ($\sim 10^4$/s/mm^2) and good spatial resolution (< 100 μm) with minimal material budget are used: Micromegas (MICRO Mesh Gaseous Structure) and GEM (Gas Electron Multiplier) detectors.

A GEM is a thin polyimide foil (50 μm)with Cu cladding on both sides, into which numerous microholes ($\sim 10^4$/cm^2) with a diameter of 70 μm has been chemically etched using lithographic techniques. The COMPASS GEM detection principle is shown in fig. 2.4; it is composed by three GEM amplification stages which are separated by thin grids of 2mm height. Across each foil a high voltage (several 100 V) is applied to generate the avalanche multiplication of the electron through the holes. The fast signal is induced by the electron cloud emerging from the last GEM foil on an anode segmented into two sets of 768 orthogonal strips (pitch 400 μm). A GEM station consists in two detectors (oriented 45° relative to each other) and there are in total 11 stations (located at SAS). The active area for the GEM detectors is 31×31cm^2. The central area (5 cm diameter) can be activated to align the detector with low intensity beams. The detection efficiency is $\sim 97\%$ with a spatial and time resolution of ~ 70 μm and 12 ns respectively.

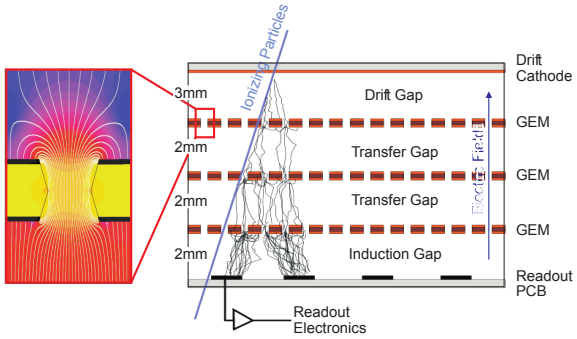


Figure 2.4: *COMPASS GEM detection principle. Figure taken from [50].*

There are three Micromegas stations (located at LAS), each one is composed by four detectors with different orientations: horizontal, vertical and two rotated by $\pm 45°$ with respect to the vertical. Each plane has an active area of 40×40 cm^2 with a deactivated central zone (5 cm diameter). The basic principle of the detector is illustrated in fig. 2.5: the particle ionizes the gas in the conversion gap, the resulting primary electrons drift in a moderate field $\approx$ 1.5kV/cm (to prevent important secondary ionizations) towards the amplification gap. The field in the amplification zone is large enough to accelerate the electrons to produce an avalanche. The conversion and amplification gaps are separated by a *micromesh*, which collects the positive ions produced during the avalanche in a short period of time (< 100 ns). This feature is possible because of the small width of the gap (~ 100 μm). All Micromegas detectors operate with a detection efficiency of 98% and with a spatial resolution < 100 μm.

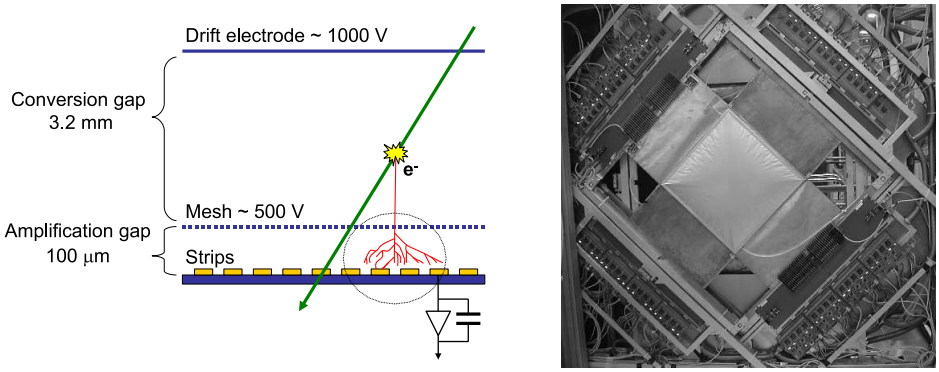


Figure 2.5: *Micromegas detection principle (left). Micromegas detector (right). Figure taken from [50].*

2.4.3 Large Area Trackers

The large area trackers cover all the spectrometer acceptance with a good spatial resolution. As the particle rate in the region covered by the large area trackers is small in comparison to the central region (10^2/s/mm^2), the use of detectors such as drift chambers (DC) [53], straw drift tubes [54] and multiwire proportional chambers (MWPC) is possible. These detectors have large active size area ($\sim$ m^2) with a central dead zone of few cm^2.

Each DC consists of eight layers of wires with four different inclinations: vertical, horizontal and rotated $\pm 20°$ with respect to the vertical direction. Two consecutive planes with the same inclination are staggered by 3.5 mm to solve the left-right ambiguities, and the ordering of the plane with different orientations is such to minimize the fake track combinations.Two of the DC are located before SM1 and have an active area of 180×127 cm^2; the third one is located downstream SM1 and has a larger active area 204×204 cm^2. All of the three DC have a central dead zone of 30 cm. The central dead zone can be activated for alignment purposes with a low intensity beam. The average spatial resolution of a single DC layer is 270 μm and the efficiency is above 95%.

A straw detector station consists in three straw detectors with different orientations (vertical, horizontal and rotated by 10° with respect to the vertical); there are 5 stations in total located between SM1 and SM2. Each detector is composed by two layers of straw tubes with the same orientation. The straw tubes consist in two layers of thin plastic film, one is coated with carbon loaded Kapton, the other one is an aluminised Kapton foil. The active area for the straw detectors is 320×280 cm^2 and have a central dead zone of about 20×20 cm and their average resolution is 190μm.

There are three types of MWPC installed in the spectrometer, each one of them have a different quantity of layers: A-type, consists of three layers vertical and two tilted by $\pm 10°$ with respect to the vertical axis; A*-type has four layers, 3 identical to the A-type and an additional one orientated horizontally. These detectors have an active area of 178×120 cm^2. The B-type is composed of two double layers detectors with opposite orientation fixed together, but only three layers are read out (the vertical layer and the

two tilted ones). The active area is 178×180 cm^2. All layers have a wire length of about 1 m, a wire diameter of 20μm and a pitch of 2 mm and are enclosed on both sides by graphite-coated Mylar foils. The central dead zone of each detector increases with respect to the detector position (the diameter varies from 16 to 22 cm). The average spatial resolution of the chambers is 1.6 mm.

2.5 Particle Identification

Several detectors are used to provide particle identification: the RICH detector allows to distinguish between pions, kaons and protons in a high intensity environment. To measure the energy of the hadrons, photons and electrons two types of calorimeters [50] are used: two hadron calorimeters (HCAL1 and HCAL2)[2] and two electromagnetic calorimeters[3] (ECAL1 and ECAL2). To perform muon identification two muon wall detectors (MW1 and MW2) are used, both consisting of tracking detectors combined with a hadron absorber. The identification detectors are briefly described in the following subsections because there were not used on the different analysis presented in this thesis. For the RICH detector, a dedicated chapter with the general characteristics is provided (chapter 3).

2.5.1 Hadron Calorimeters

A hadron calorimeter can distinguish between hadron and muon tracks depending on the energy deposited: a hadron deposits almost all its energy in a form of a hadronic shower while a muon only deposits a small energy fraction. The signal measured is used to determine the hadron energy by a linear relation [50]. Hadron calorimeters are also used by the trigger system described in section 2.6. HCAL1 and HCAL2 are simple calorimeters with a modular structure with iron and scintillator plates and are located before the muon filters.Their efficiency depends on the energy: for HCAL1 for hadrons with momenta above 5 GeV/c it is almost constant and close to 100%; for HCAL2 the same efficiency is reached for hadrons with momenta above 10 GeV/c.

2.5.2 Electromagnetic Calorimeters

The electromagnetic calorimeters are used to identify electrons and to detect photons. ECAL1 (located in the LAS) and ECAL2 (located in the SAS) consist of blocks of lead glass connected to photomultipliers with light guides. The incoming photon or electron interacts with the calorimeter and initiates an electromagnetic shower which produces Cherenkov radiation inside the lead glass. The detected light intensity is proportional to the deposited energy.

[2]The numbers 1 and 2 refers to the detector location: LAS and SAS respectively.

[3]ECAL1 was not installed in the 2004 configuration, that is why it is not shown in fig. 2.1.

2.5.3 Muon Wall Detector

The muon wall detectors MW1 and MW2 consist of a set of tracking stations and a hadron absorber: thick iron wall (60 cm) for MW1 and concrete wall (2.4 m) in MW2 case. Muons passing through the MW detector are detected before and after the absorber while other tracks (hadrons, electrons, etc.) are stopped.

2.6 The Triggering System

The triggering system [55] has the task to select physical event candidates in a high rate environment; it is composed by scintillating hodoscope stations, scintillator veto stations (to exclude halo muons) and calorimeters (to select events with hadron production). The distribution of all these components is illustrated in fig. 2.6a. The different triggers are formed combining these different components and they are summarized in table 2.1.

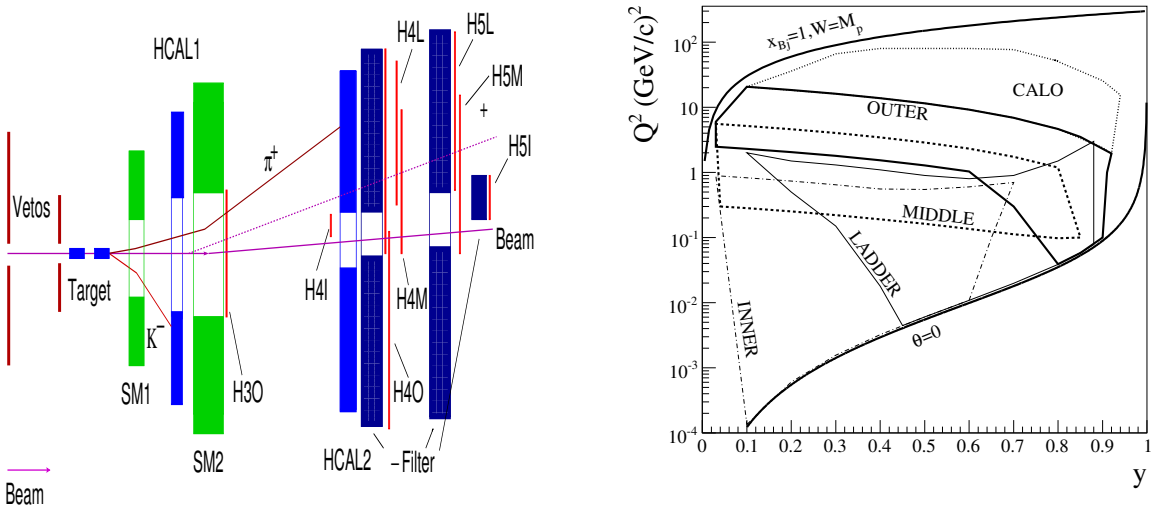
Table 2.1: *COMPASS triggers with the muon beam in 2006 run.*

Trigger name	Components	
Inner trigger (IT)	H4I, H5I, HCAL1, HCAL2	Inclusive triggers
Middle trigger (MT)	H4M, H5M, veto, HCAL1, HCAL2	
Ladder trigger (LT)	H4L, H5L, veto, HCAL1, HCAL2	
Inclusive MT (iMT)	H4M, H5M, veto	Semi-Inclusive triggers
Outer trigger (OT)	H30, H40, veto	
Pure calo trigger (CT)	HCAL1, HCAL2, veto	Calorimeter trigger

Depending on the event kinematics two different algorithms are used to determine the scattered muon kinematics. When the angle of the scattering muon is large enough ($Q^2 > 1$ (GeV/c)^2) to be measured by using only the hodoscope stations, the trigger signals are built using the vertical target pointing algorithm. The angle $(\theta_x^\mu, \theta_y^\mu)$ of the scattering muon is determined by using two hodoscopes with horizontal strips located at different positions along the beam direction and the vertical component θ_y is determined. If θ_y is compatible with the target position the event is triggered. The $y-z$ plane is selected since the particle track is not bent by the dipole magnet in the y-direction. In the case where the scattering muon angle is too low to be measured ($Q^2 < 1$ (GeV/c)^2), the bending angle of the magnet is used to determine θ_x^μ and to perform a *energy loss* criteria to fire the trigger. This is possible since the muon energy loss ν is translated into a deflection by the magnetic dipole field. In the low Q^2 region the information coming from

the hodoscopes is not enough to tag the event since there are several background processes in the scattered muon signal (elastic scattering off target electrons, elastic and quasi-elastic radiative scattering off target nuclei). To improve the trigger signal, an energy deposit in the hadron calorimeters (larger than 5.4 GeV) is required additionally. In the particular case where any scattered muon is detected, it is still possible to determine if an inelastic interaction has occurred; if a large deposit of energy (> 16.2 GeV) is registered in HCAL1 or (and) HCAL2 the event will fire the trigger.

The kinematic range covered by the trigger system is shown in fig. 2.6b as a function of y and Q^2. The trigger system is optimized to select *quasi-real photon* and DIS events; the quasi-real photon events correspond to the low Q^2 region (~ 0) and they will be detected by the *inner* (IT) and *ladder* (LT) triggers. A DIS event is characterized to have higher Q^2 values; the *middle* trigger (MT) is in charge of the DIS events with small to moderate Q^2. The *inclusive middle* trigger (iMT) covers the same region as MT but it not requires an energy deposit in the calorimeters. The *outer* trigger (OT) is designed to detect DIS events with high Q^2. The events with the largest Q^2 are covered by the *pure calorimeter* trigger (CT).



(a) **(b)**

Figure 2.6: *(a) Main elements of the trigger system. (b)Trigger system kinematical domain. Figure taken from [55].*

2.7 Data Acquisition

The data acquisition system (DAQ) [50] is in charge of the information management coming from more than 250,000 spectrometer electronic channels to be sent to permanent storage. At COMPASS the typical event size is 45 KB at a trigger rate of about 10 KHz. The pipeline used in the DAQ is illustrated in fig. 2.7. First the analogue signals coming

from the detectors are preamplified, then they are digitized directly at the front-end by ADCs (Analog to Digital Converters) or TDCs (Time to Digital Converters) according to the type of detectors the front-ends are coupled to. The data are then transferred to the readout driver modules CATCH (COMPASS Accumulate, Transfer and Control Hardware) or GeSiCA (GEM and Silicon Control and Acquisition) upon the arrival of a trigger signal (provided by the Trigger Control System: TCS). CATCH and GeSiCA combine the data from up to 16 cards (ADC or TDC) and transmit them via an optical S-Link to the computers named *Read-out Buffer* (ROBs, maximum through output 160 MB/s), where they are stored in 512 MB spill buffer cards. During the 4.8 s of beam time the data are written into the memory, during the rest of the full SPS cycle (12 s) they are read through a PCI interface. In this way the required bandwidth is reduced by a factor of three. The events are written to multiple 1 GB large files (chunks) labeled by the run number and their consecutive chunk number. Finally, the data are transferred to the CERN central data recording facility (CASTOR).

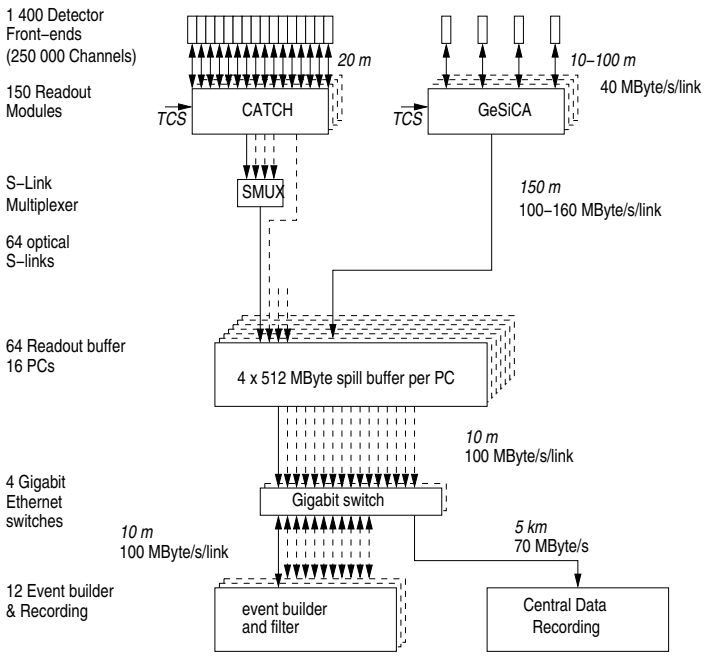


Figure 2.7: *General architecture of the DAQ system. Digitized data from the detector front-ends are combined on the readout modules named CATCH and GeSiCA. The storage of the data during the spill and the event building is performed locally. The data are recorded at the CERN computer center. Figure taken from [50].*

2.8 Event Reconstruction

The off line analysis of the events stored at CASTOR is performed by the COMPASS software CORAL[4] [50]. CORAL can also analyze the events generated by a Monte Carlo (MC) simulation (see chapter 6). CORAL has a modular architecture and is written in C++. The scheme of the steps followed by the reconstruction program are shown in fig. 2.8. First, the information on the fired detectors channels is extracted. This is known as decoding, in the case of MC data this step is known as digitization. In general, there are more than one detector channels fired by the same particle. In that case a clustering process is performed, *i.e.* the neighboring detector channels that were fired are grouped together, and the coordinate of the cluster in the experiment reference system is computed. At this stage the detector calibration and position are used to extract the information. The CORAL output is stored in a ROOT [56] structure called mDST (mini Data Summary Tape).

The physics information is extracted from the recorded mDST using the software package PHAST [57] (PHysics Analysis Software Tools). PHAST gives access to the reconstructed event information, and it provides a set of algorithms to compute the relevant physical variables of each event. The PHAST outputs are stored as a root file. These files are significantly smaller that the mDST and are used in the final physics analysis.

In COMPASS the experimental data are organized into several levels. The first level is known as *period*, it contains all the data taking information of a one week period. A period is divided in several *runs*, which are subsequently divided in *spills*. Each spill is composed of the *events* which are analyzed.

[4] COMPASS Reconstruction Algorithm Library

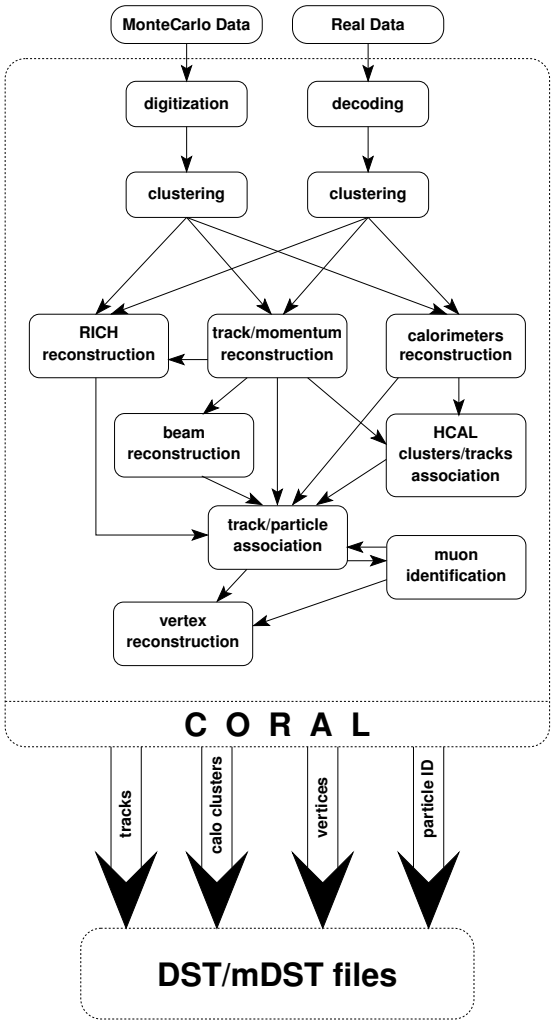


Figure 2.8: *Schematic representation of the COMPASS reconstruction software. Figure taken from [50].*

Chapter 3

RICH Detector

The particle identification is an important step in the hadron multiplicities extraction. In the COMPASS spectrometer, the hadron identification is performed by a large-size Ring Imaging Cherenkov detector (RICH) capable of separating pions, kaons and protons in a wide momentum range ($\sim$ 2GeV/$c-$ $\sim$ 60GeV/c) and in an angular acceptance of 0.01-0.4 radians.

In this chapter the RICH detection principle is presented as well as the description of the main components: the gas and mirror system (section 3.2.1 and section 3.2.2), the photon detectors (section 3.2.3), the readout electronics (section 3.2.4) and the data reconstruction (section 3.2.5).

3.1 Cherenkov Effect

When a charged particle is moving through a transparent medium with a speed v greater than the speed of light ($v_{light} = c/n$. n is the medium refractive index) a radiation known as *Cherenkov radiation* is produced by the medium.

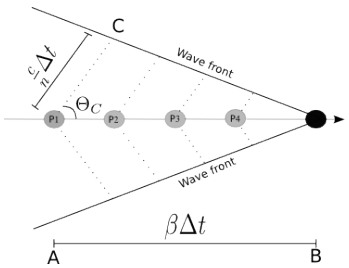


Figure 3.1: *Cherenkov radiation geometry.*

The Cherenkov radiation produced by a particle with a certain mass M_h and momentum p_h is emitted only at a particular angle (Θ_C) with respect to the particle track (fig. 3.1).

The coherence between waves (emitted between A and B) is achieved when the particle traverses $\overline{AB}$ at the same time as the radiation travels from A to C. The opening angle Θ_C is defined geometrically in eq. (3.1) where β is the particle velocity.

$$\cos\Theta_C = \frac{c/n\ \Delta t}{\beta c\ \Delta t} = \frac{1}{n\beta} \tag{3.1}$$

From eq. (3.1) some characteristics from the Cherenkov radiation can be inferred:

Threshold limit: If $\beta \leq 1/n$ no Cherenkov radiation will be emitted.

Maximum emission angle: $\cos\Theta = \frac{1}{n}$. It is reached for ultra-relativistic particles ($\beta = 1$).

In order to perform a particle identification with a RICH detector two variables are needed: Θ_C and the particle momentum p_h (which is measured independently in the spectrometer). Θ_C can be measured if the radiation (photons) emitted by the particle is detected. There are different techniques to recollect and transport the produced photons where light detectors are placed [58]. The resulting image in the detection plane is a ring with a radius $r \propto \Theta_C$. The particle identification is performed by a mass assignment given in eq. (3.2).

$$M_h = p_h\sqrt{n^2\cos^2\Theta_C - 1} \tag{3.2}$$

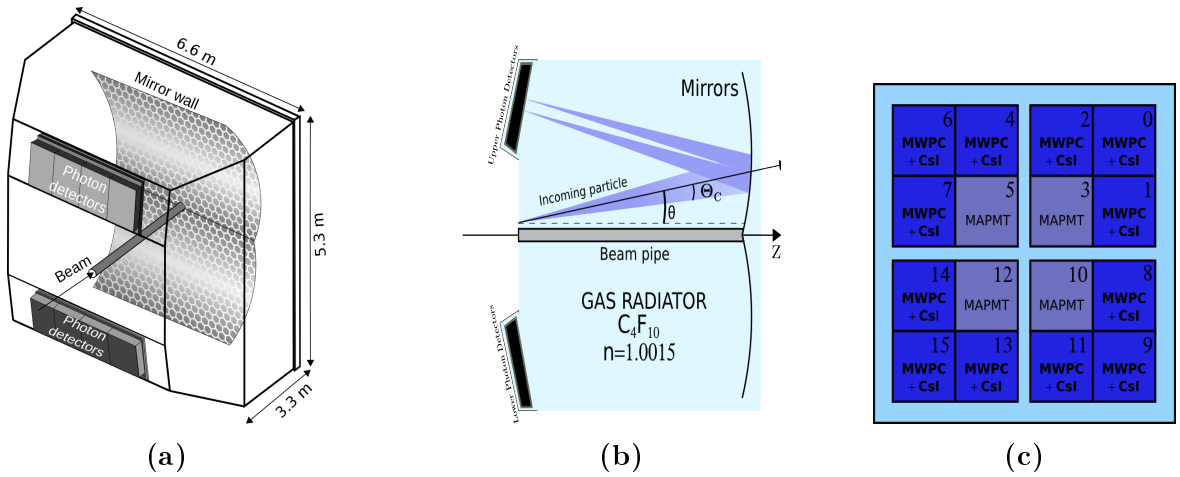
3.2 The COMPASS RICH Detector

The COMPASS RICH detector (fig. 3.2) is designed to distinguish between pions, kaons and protons in a high-intensity environment. The momentum range covered goes from the pion Cherenkov threshold (≈ 2.67 GeV/c) to ~ 60 GeV/c.

RICH is a large size detector ($\sim 3 \times 5 \times 6$ m^3), which is filled with a gaseous radiator. Two spherical mirror systems reflect the photons into an array of photon detectors (multiwire proportional chambers and multianode photomultipliers tubes), sensitive to a large wavelength range, from visible to far UV, placed outside the spectrometer acceptance, one above and the other below the beam line. The whole structure of the detector vessel is built mainly in thin aluminum in order to minimize the material budget.

3.2.1 Gas System

One of the principal elements of a RICH detector is the radiator. At COMPASS the radiator is gaseous, C_4F_{10}. This particular gas has a refractive index $n \approx 1.0015$ and



(a) **(b)** **(c)**

Figure 3.2: *Artistic view of the COMPASS RICH detector (a) and the Cherenkov effect (b). (c): photon detectors (not to scale).*

a low chromaticity[1] ($dn/dE \sim 5 \cdot 10^{-5}$ eV^{-1}). These characteristics allow the particle identification (PID) to be performed in a wide momentum range (~ 2.67 to ~ 60 GeV/c).

The propagation of the Cherenkov photons in the vessel can be affected by the presence of water vapor and oxygen (high UV light absorption cross section). In order to remove these impurities, the gas is constantly circulating and filtered at a constant pressure (1 mbar higher than the atmospheric pressure) in a dedicated gas system [60]. The overpressure of the vessel is needed to prevent the air contamination and to avoid mechanical stress to the detector, given its large size. Other circulation system (known as *fast circulation* system) allows a reshuffling of the gas inside the vessel, to avoid stratification, that may cause a gradient in the value of the refractive index from top to bottom.

In order to absorb the photons emitted by the muon beam, a 10 cm diameter pipe filled with helium is positioned inside the vessel on the beam path.

3.2.2 Mirror System

The RICH optical system covers a total area of ~ 21 m^2. It consists of two spherical surfaces, each one contains 58 spherical mirrors of different shapes (34 hexagons and 24 pentagons); the final arrangement of the mirrors is shown in fig. 3.3. All the mirrors have a reflectance above 80% in the UV region.

The mirror system has a radius of curvature of 6.6 m. The photon image is focused outside the spectrometer acceptance where the photon detectors are located (fig. 3.2). The radius of curvature is not the same for each mirror ($R = 6600 \pm 1\%$ mm), in consequence the reflected image maybe be slightly blurred. This effect is more pronounced for particles incident at larger angles than for particles at small angles [61]; this aberration contribute

[1] Dependence of the refractive index of a dielectric medium on the photon wavelength [59].

to the dispersion of the photon angle with respect to the angle of emission which affect the detection resolution [62].

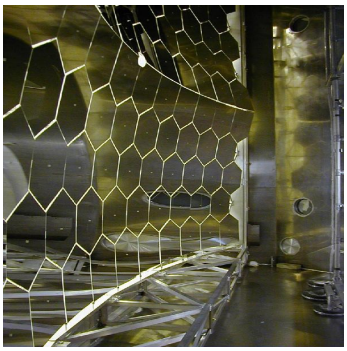


Figure 3.3: *Photograph of the COMPASS RICH detector optical system.*

3.2.3 Photon Detectors

The photon detector array consists of two symmetric parts with respect to the beam direction, each one is composed of 8 modules as is shown in fig. 3.2c. The modules located in the external region consist on MultiWire Proportional Chambers (MWPC) equipped with solid state CsI photocathodes [63]. The central area (quadrants 3, 5, 10,12 fig. 3.2c) is composed by MultiAnode Photomultipliers Tubes (MAPMT) [64] coupled to individual telescopes of fused silica lenses. The use of two different photon detector types employing different photon converters results in the detection of photons in two different wavelength regions, < 200nm for the MWPCs and $\sim 200 - 650$ nm for MAPMT. The low momentum particles are mainly detected in the outter part of the detector and the high momentum particles in the central region [62].

The spherical mirrors will focalize all the photons emitted parallel in the same point, as a result of this, the image reflection in the photon detectors will be a ring. An example of a RICH event is shown in fig. 3.4. In this particular event there are 6 visible rings and a few active pads due to background.

3.2.4 Readout Electronics

The readout electronics for the central area (MAPMT) consists of three elements: a chip card MAD4, a bus board (Roof) and a DREISAM card. The MAD chip amplifies and discriminates the signal. The discriminated signal is transported to the chip card DREISAM. This chip functions as a time to digital converter. The signal acquired in a window of 100 ns centered on the time of the trigger is transformed into a temporal information with respect to the absolute time of the experiment with a resolution better than 130 ps. The Roof board acts as a bridge to carry the information from the MAD to DREISAM and, in

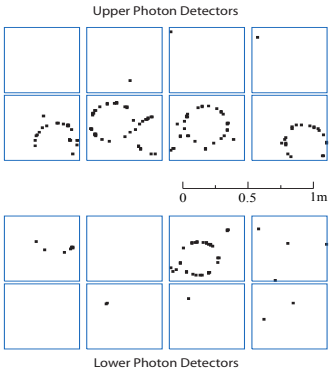


Figure 3.4: *Illustration of a typical RICH event. Figure taken form [50].*

the opposite way, to deliver to MAD information on the thresholds. The readout system is free from cable connections to minimize the electrical noise.

For the peripheral area, characterized by a lower population of events and lower background level, the electronic of the readout system is based on a APV25 chip (128-channel preamplifier/shaper ASIC with analogue pipeline) [65, 66]. This chip has an integration time of the signal < 400 ns.

The large number of RICH electronic channels (82,944) corresponds to $\sim 40\%$ of the total number of COMPASS electronic channels. To reduce the data flow, empty channels are suppressed at the front end stage and only the amplitude signals above threshold are read out in local FIFO arrays[2]. Data are then transmitted with optical fibers to the general acquisition system at a rate of 40 MB/s.

3.2.5 RICH Event Reconstruction

RICHONE is a package contained in CORAL software (see section 2.8). It is in charge of RICH event reconstruction, *i.e* reconstructs the physical variables from the RICH active pads for each event. The reconstruction is divided in several stages. The first one regards the data decoding[3] and clustering[4]. Then the reconstruction of the Cherenkov angle for each individual photon is done. It is also possible to perform the ring reconstruction which is used for studies on the apparatus. The particle identification (PID), in which a mass hypothesis is assigned to each particle, is based on a maximum likelihood calculation. The PID will be explained in more detail in chapter 4.

[2]It means that elements are transmitted from the array in the same order they arrive - first in, first out (FIFO).

[3]The information on the fired detector channel is extracted.

[4]The detector channels that are fired by the same particle are grouped together.

Decoding and clustering

There are two different types of photon detectors and they have different decoding systems and clustering algorithms. Regarding the MWPC the analog signal comes from channels that have a signal (this signal is sampled three times) [67]. Since more than one channel can produce a signal, a clustering procedure is used. First, the pad with the maximum pulse height has to be found. All the adjacent pads with a pulse height smaller than the maximum are included in the cluster [62]. During the clustering process the mean position of each active pad is evaluated in the cluster, weighting the signals with their maximum pulse height, to determine the center of gravity of the cluster. For the MAPMT signal decoding is enough to read the time information coming from the PMT that was hit. The signal coming from the MAPMT does not require clustering, since the probability to have correlated hits in adjacent pads is negligible [65].

The cluster or hit position will be used to determine the trajectory of the photon. In addition, the time information coming from the MAPMT is used to reject out-of-time photons while the amplitude information from the MWPC serves to reduce the background both from out-of-time photons and from electronic noise [62].

Cherenkov angle and ring reconstruction

To reconstruct Θ_C, the trajectory of each Cherenkov photon is calculated with respect to the plane containing the particle track and its virtual reflection in the mirror [68]. All the photons emitted by the particle are expected to have the same angle Θ_C and to be uniformly distributed in φ[5], whereas the photons emitted by others particles or background have a flat Θ_C distribution. The emitted photons with the same (Θ_C, φ) pair are reflected on the same point on the focal surface (neglecting the spherical aberration), resulting in a ring image on the photon detector (fig. 3.4). Since the emission point of the photon along the particle trajectory is not known, the middle point between the detector and the mirror is taken. A good determination of the track trajectory parameters and the momentum of the particle are mandatory in order to extract Θ_C with good precision.

To characterize the RICH, for example to determine the angular resolution, the ring reconstruction of the emitted photons is needed. The ring reconstruction is based on the search of a peak in the Θ_C distribution: small intervals of $\pm 3\sigma$ (where σ is the single photon resolution: $\sigma_{MAPMT} = 2.0$ mrad and $\sigma_{MWPC} = 2.5$ mrad)[6] on an overall range of $0 - 70$ mrad are considered. The interval with the maximum number of entries defines the ring. This procedure associates a ring to each track and in order to reject tracks with only background photons a minimal amount of 4 photons per ring is required [62]. The resolution of the Cherenkov angle measurement provided by each single photon as a function of the particle momentum is illustrated in fig. 3.5. In the high momentum region where the Cherenkov angle saturates the resulting resolution is ~ 1.2mrad.

[5] φ is the angle between the photon trajectory and X on the XY plane

[6] The MWPC resolution is bigger that the MAPTMT because of the MWPC smaller pad size. Also the

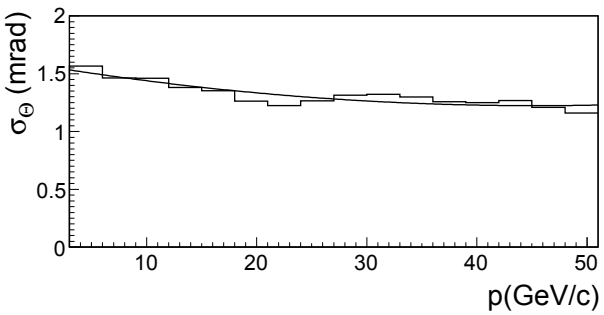


Figure 3.5: *Resolution of the Cherenkov angle for the reconstructed ring images, provided by each single photon, versus the particle momentum for a sample of identified pions. Figure taken from [50].*

Mass separation

For the physical analysis of RICH events the PHAST software is used. The available information at PHAST is summarized in table 3.1.

The measured values of Θ_C as a function of p_h for the RICH detector are shown in fig. 3.6; in this figure the number of kaons and protons are scaled in order to be visible. In the low momentum region, the RICH detector is only sensitive to electrons, muons and pions. The bands corresponding to kaons and protons start to be visible at $p_h \approx 9.45$ GeV/c and $p_h \approx 17.95$ GeV/c respectively. For high momentum values, above 40 GeV/c, a saturation of the Cherenkov angle is observed, principally for pions and kaons.

Table 3.1: *RICH information available in PHAST.*

Background, pion, kaon, proton, electron and muon likelihood
(LH(bg), $LH(\pi)$, $LH(K)$, $LH(p)$, $LH(e)$, $LH(\mu)$)
Pion, kaon, proton , electron and muon likelihood derived
Maximum likelihood angle
Reconstructed ring angle, Θ_{Cj}
Number of photons used in Θ_C reconstruction, N
Fitted ring angle, Θ_C
Ring, pion, kaon, proton, electron and muon χ^2
Mean ring time from photons in the PMT part
Number of photons per ring in the PMT part

The final particle identification is performed using a likelihood method. This likelihood is constructed for five different hypotheses (e, μ, $_pi$, K and p) and the corresponding explanation is given in section 4.2.

resolution in the peripheral region is diminish because of the spherical aberration and the poorer particle momentum resolution

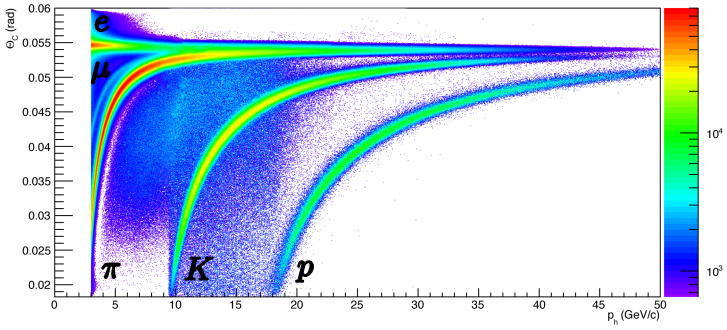


Figure 3.6: *Measured Cherenkov angle Θ_C as a function of p_h. π threshold ≈ 2.67 GeV/c, K threshold ≈ 9.45 GeV/c and p threshold ≈ 17.95 GeV/c.*

Chapter 4

RICH Performance Study and Particle Identification

The final goal of this thesis is the extraction of pion and kaon multiplicities. The hadron identification is done using the information from the RICH detector; the detection efficiency is not perfect, in consequence the misidentification probability is not zero, so it is necessary to determine the RICH response for the different hadrons. The RICH performance study consists in measuring the identification and misidentification efficiencies (probabilities) for pions, kaons and protons. The process followed to determine the RICH performance is described in section 4.1; as a first step *pure* samples of pions, kaons or protons have be selected without the use of the RICH detector (section 4.1.1); the RICH detector response is determined from these samples (section 4.1.2) *i.e.* the hadrons are identified using the RICH information and the identification and misidentification probabilities are extracted in the hadron phase space (p_h, θ) (where p_h is the hadron momentum and θ is the polar angle) for each type of particle ($\pi^{\pm}$, $K^{\pm}$, p and $\overline{p}$). The particle identification is performed via a likelihood estimator for different hypotheses (section 4.2); as a first approach, the mass assignment corresponds to the largest likelihood among them; this identification is improved by requiring a better separation between all the hypotheses, this stage is presented in section 4.2.1 for pions, 4.2.3 for kaons and 4.2.2 for protons.

The results obtained for the RICH performance are given in section 4.3. Finally, the method to identify pions, kaons and protons using the full RICH information (likelihood cuts and RICH performance matrix) is described in section 4.4.

4.1 Determination of RICH Detector Performance

The identification (misidentification) probability is given by the ratio of the number of particles correctly (wrongly) identified out of a *pure* sample of known hadrons over the total number of particles composing the pure sample.

$$P(t \to i) = \frac{N(t \to i)}{N(t)} \tag{4.1}$$

In other words, $P(t \to i)$ is the probability that a particle of type t is identified as a particle of type i. In eq. (4.1), $N(t)$ is the number of particles t and $N(t \to i)$ is the number of particles t identified as i. The identification and misidentification probabilities ($P(t \to t)$ and $P(t \to i)$) are a pure detector properties and can be presented in a matrix form (eq. (4.2)), where the diagonal elements are the identification probabilities and the off-diagonal ones the misidentification probabilities.

$$M_R = \begin{pmatrix} P(\pi \to \pi) & P(K \to \pi) & P(p \to \pi) \\ P(\pi \to K) & P(K \to K) & P(p \to K) \\ P(\pi \to p) & P(K \to p) & P(p \to p) \\ P(\pi \to X) & P(K \to X) & P(p \to X) \end{pmatrix} \tag{4.2}$$

In order to determine these probabilities a calibration sample is needed; three different calibration samples where used to study the RICH response.

4.1.1 Selection of ϕ_{1020}, K^0_S and Λ Samples

The RICH performance analysis is based on the study of the pion, kaon and proton ($\overline{p}$) samples coming from ϕ_{1020}, K^0_S and Λ ($\overline{\Lambda}$) decay. In table 4.1 the branching ratios for these mesons and baryons are summarized.

Table 4.1: *Decay modes for ϕ_{1020}, K^0 and Λ [69].*

ϕ_{1020}		K^0				Λ	
		K^0_L		K^0_S			
K^+K^-	48.9%	$\pi^\pm e^\mp \nu_e$	40.6%	$\pi^+\pi^-$	69.2%	$p\pi^-$	63.9%
$K^0_L K^0_S$	34.2%	$\pi^\pm \mu^\mp \nu_\mu$	27.0%	$\pi^0\pi^0$	30.7%	$n\pi^0$	35.8%
$\rho\pi + \pi^+\pi^-\pi^0$	15.3%	3 π^0	19.5%				
		$\pi^+\pi^-\pi^0$	12.5%				

For the RICH calibration sample analysis, the same 6 periods used in the multiplicities analysis[1] of the 2006 data taking were used. To ensure a good sample selection some requirements had to be fulfilled by the event in order to be kept. In the following the selection criteria are explained.

Spill selection. During data taking there can be some instabilities in the beam behavior or detector performances. Bad or unstable data from these periods are rejected on a spill by spill basis. To determine if a spill should be removed many variables are analyzed: the number of reconstructed beam tracks per vertex, the number of outgoing tracks per vertex, the number of primary vertexes per event, detectors efficiencies, etc. If the value of one of

[1] W40, W41, W42, W43, W45 and W46

these variables deviates from the majority of the neighboring spills the spill is considered as a *bad spill* and is rejected. In fig. 4.1 the number of outgoing tracks per vertex is shown as a function of the spill number; the red points correspond to the *good spills*. In black the spills that are rejected. In table 4.2 the percentage of spills rejected for each of the periods used in this analysis is summarized.

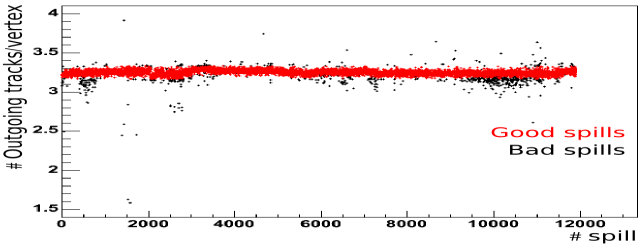


Figure 4.1: *Number of outgoing tracks per vertex as a function of spill number. In red, the points corresponding to good spills, in black the rejected spills.*

Best primary vertex. The first event rejection criterion is related to the existence of an interaction point known as *vertex* in the event. A *primary vertex* is defined as the intersection of two tracks, one incoming track (associated to the muon beam μ) and one outgoing track (related to the scattered muon μ'). If more than one primary vertex is found in the event, the vertex with the maximum number of outgoing tracks is chosen. In case of equal number of tracks, the vertex with the smallest χ^2 is selected.

Table 4.2: *Number of total spills and fraction of* bad *spills for the different periods used in this analysis.*

Period	# spills	% bad spills
W40	24900	6.5%
W41	17412	14.6%
W42	15538	5.2%
W43	18841	7.4%
W45	15543	7.2%
W46	10559	11.8%

Track selection. Due to the different decay nature of the mesons used in this analysis leads to different requirements for the track selection:

- Strong decay. The ϕ_{1020} mean lifetime is $\sim 10^{-23}$ s so it will decay before it travels $\sim 10^{-13}$ cm in consequence it is not possible to differentiate between the primary vertex and the decaying vertex (secondary vertex) that is why the primary vertex is request to have only 2 outgoing tracks of opposite charge in addition to the scattered muon μ' to be considered as a ϕ_{1020} candidate.

The inability of differentiate between the primary and secondary vertex has as consequence the contamination of the ϕ_{1020} sample coming from hadrons generated in the muon-nucleon interaction and not from the meson decay. If the hadron pair is required to carry all (or almost all) the virtual photon energy this contamination is reduced; the ϕ_{1020} sample composed by this particular kind of hadrons is known as exclusive sample ($\phi^{\rm excl}_{1020}$). In order to select the $\phi^{\rm excl}_{1020}$ sample, a cut in the reaction inelasticity (I) is required. The inelasticity is defined in eq. (4.3), where $M_{\rm p}$ is the proton mass and $M_{missed} = M_{\rm p} + M_\gamma - M_\phi$ where M_ϕ is the invariant mass of the K^+K^- system. The $\phi^{\rm excl}_{1020}$ yield is coming from the region where $I < 2.5$ GeV/c^2

$$I = \frac{M^2_{missed} - M^2_{\rm p}}{2M_{\rm p}} \tag{4.3}$$

The $\phi^{\rm excl}_{1020}$ sample is a very pure kaon sample, the hadron contamination coming from the other decay channels is almost negligible. However, the $\phi^{\rm excl}_{1020}$ do not reflect the contamination that is observed in data used in the multiplicity analysis where the hadron sample is mostly composed by pions. In order to perform the RICH response analysis in a similar kaon sample, the inclusive ϕ_{1020} sample ($\phi^{\rm incl}_{1020}$) is defined by events with $I \geq 2.5$ GeV/c^2.

- Weak decay. The K^0_S and $\Lambda(\overline{\Lambda})$ are long-living particles ($\sim 10^{-10}$ s), they decay *far* away from the primary interaction point. The events with K^0_S and $\Lambda(\overline{\Lambda})$ candidates are required to have at least one secondary vertex (V_0) reconstructed with 2 outgoing tracks of opposite charge.

  For the final selection of V_0 additional requirements are asked, these are summarized in table 4.3. The first one involves the distance between the primary vertex (PV) and V_0 ($PV - V_0$). If a V_0 decay does not have a large enough distance from its primary interaction, the decay particles cannot be properly reconstructed. Only the V_0's with a decay length ($PV - V_0$) larger than 2 times the error of the decay length (σ_{PV-V_0}) are kept. In addition, a collinearity requirement is demanded. The *collinearity cut* rejects events for which the cosine of the angle between the reconstructed V_0 momentum and the vector defined by PV and V_0 positions is larger than 0.9996 for K^0_S and 0.9999 for Λ ($\overline{\Lambda}$). The cut in the invariant mass value is used to remove the overlapping zone of K^0_S and $\Lambda(\overline{\Lambda})$ samples (fig. 4.2).

Table 4.3: *V_0 selection cuts.*

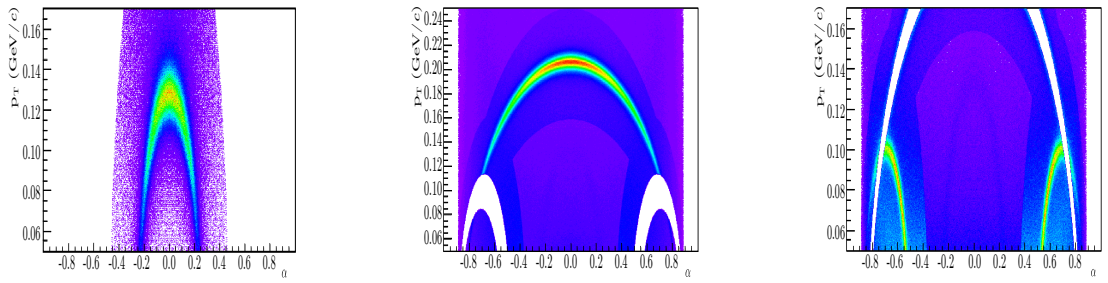
	K^0_S	Λ or $\overline{\Lambda}$
Distance PV $- V_0$	$> 2\sigma_{PV-SD}$	$> 2\sigma_{PV-SD}$
Collinearity	$cos\theta_{col} > 0.9996$	$cos\theta_{col} > 0.9999$
Invariant Mass	$\lvert M_\Lambda \rvert < 10$ MeV/c^2	$\lvert M_{K^0_S} \rvert < 10$ MeV/c^2

The track momentum is limited to the region: $[2.8-50\ \mathrm{GeV}/c]$. The lower cut is defined by the Cherenkov threshold for pions. The upper limit is chosen to avoid the region where the Cherenkov effect saturates (pion and kaon bands overlap) (fig. 3.6). In addition, the track incident angle is limited to be in $[0 < \theta < 0.4\ \mathrm{rads}]$ region, to eliminate the particles outside RICH acceptance.

The presence of the RICH pipe degrades the RICH performance [67] (see also Appendix C.2). The particles close to RICH pipe are removed from the analysis by applying a radial cut at the RICH entrance defined in eq. (4.4).

$$x^2 + y^2 > r^2_{PIPE}, \qquad r_{PIPE} = 5cm \tag{4.4}$$

The final selection of K^0_S ($\sim 4.4 \times 10^7$ events), ϕ_{1020}($\sim 1.0 \times 10^6$ events) and Λ ($\overline{\Lambda}$) ($\sim 1.5 \times 10^7$ events) are shown in the Armenteros diagram in fig. 4.2 [70]. The white regions in figs. 4.2b and 4.2c correspond to the cut in the invariant mass distribution defined in table 4.3.



(a) ϕ_{1020}. **(b)** K^0_S. **(c)** Λ *and* $\overline{\Lambda}$.

Figure 4.2: *Two dimensional plot of the transverse momentum of decay products versus the longitudinal momentum asymmetry* ($\alpha = \frac{p_{1L}-p_{2L}}{p_{1L}+p_{2L}}$) *for* ϕ_{1020}, K^0_S, Λ *and* $\overline{\Lambda}$ *samples.*

4.1.2 $P(t \to i)$ Efficiencies and Sample Purity

In this section the procedure to determine $N(t \to i)$ and $N(t)$ from eq. (4.1) is explained in detail only for pions (the process is analogue for kaons and protons). The probabilities are evaluated for positive and negative hadrons because of the different RICH response for different charged particles [67].

The positive pion sample is built with the requirement that the negative decay particle (*spectator*) is properly identified by the RICH (for the negative sample the spectator particle is required to be positive) [71]. This procedure is applied in order to reduce the background under the peak in the invariant mass distribution. This mostly concerns the kaon sample, since the ϕ_{1020} sample is highly contaminated with pions (see ϕ_{1020} decay in table 4.1).

The identification and misidentification probabilities for π^+ are defined in eqs. (4.5) to (4.8). All equations have the same denominator $N_{K^0_S}(\pi^-)$ (the particle in the parenthesis is the *spectator*). In the numerator, the first particle in the parenthesis is the *spectator* and the second one is the particle for which the probability is being calculated.

$$P(\pi^+ \to \pi^+) = \frac{N(\pi^+ \to \pi^+)}{N(\pi^+)} = \frac{N_{K^0_S}(\pi^-\pi^+)}{N_{K^0_S}(\pi^-)} \tag{4.5}$$

$$P(\pi^+ \to K^+) = \frac{N_{K^0_S}(\pi^-K^+)}{N_{K^0_S}(\pi^-)} \tag{4.6}$$

$$P(\pi^+ \to p) = \frac{N_{K^0_S}(\pi^-p)}{N_{K^0_S}(\pi^-)} \tag{4.7}$$

$$P(\pi^+ \to X) = \frac{N_{K^0_S}(\pi^-X)}{N_{K^0_S}(\pi^-)} \tag{4.8}$$

In this analysis 5 subsamples are extracted from the K^0_S calibration sample. In the first one only the spectator particle π^- is identified with the RICH, its corresponding invariant mass distribution is shown in fig. 4.3a. In the other three subsamples, in addition to the spectator, the second particle is identified as a π^+, K^+ or p. Their invariant mass distribution is plotted in fig. 4.3b, fig. 4.3c and fig. 4.3d. The last subsample contains all the second particles that were not identified by the RICH (fig. 4.3e). In the following only the name of the particle which is being studied will be called to made reference to the invariant mass distribution.

The functions used to fit the invariant mass distribution of ϕ_{1020}, K^0_S, Λ and $\overline{\Lambda}$ subsamples are summarized in table 4.4. The fit is performed simultaneously for all subsamples. A constraint is added to the fit in order to always ensure that $\sum_i N(t \to i) < N(t)$ when the fit of invariant mass distribution is performed for $t = \pi^+$ and $i = \pi^+$, K^+, p; if the fit includes in addition the invariant mass distribution coming from the non-identified particles the constraint becomes $\sum_i N(t \to i) = N(t)$.

An example of the fits used in the identification and misidentification probabilities determination for positive hadrons is shown in figs. 4.3 to 4.6. The number of events in the invariant mass peak is represented by the green dotted line, the background contribution by the red one and the blue solid line corresponds to the sum of both contributions. The corresponding fit examples for negative hadrons can be found in Appendix A.

The overall results for positive and negative hadrons ($p_h \in [10, 40]$(GeV/c) and $\theta \in [0.01, 0.4]$(rad)) are presented in eqs. (4.9) and (4.10) where M_R is truncated.

$$M_R^+ = \begin{pmatrix} 0.89 & 0.01 & 0.04 \\ 0.02 & 0.91 & 0.04 \\ 0.01 & 0.06 & 0.90 \end{pmatrix} \tag{4.9}$$

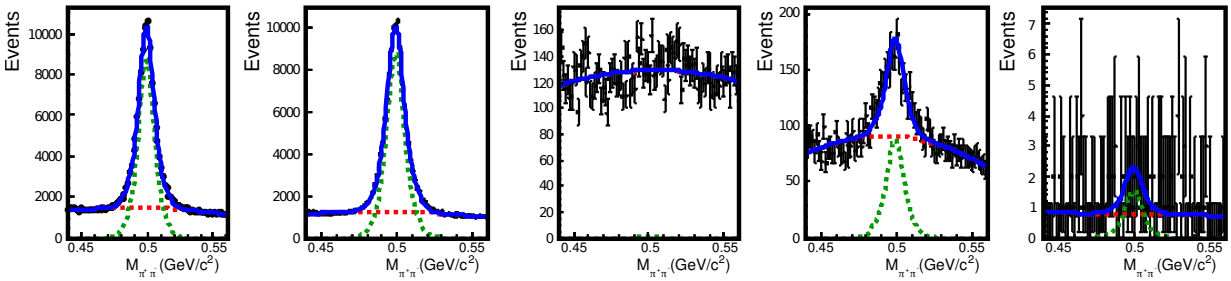


Figure 4.3: *K^0_S invariant mass distribution. In green the signal component, in red the background and in blue the sum of both components.*

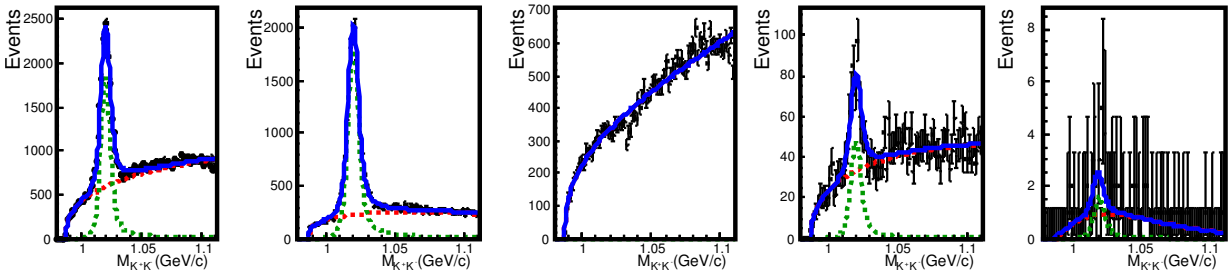


Figure 4.4: *Same as fig. 4.3 for ϕ^{incl}_{1020}.*

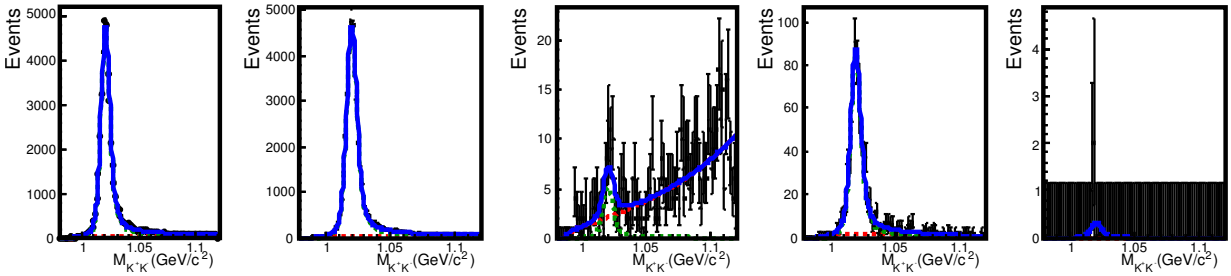


Figure 4.5: *Same as fig. 4.3 for ϕ^{excl}_{1020}.*

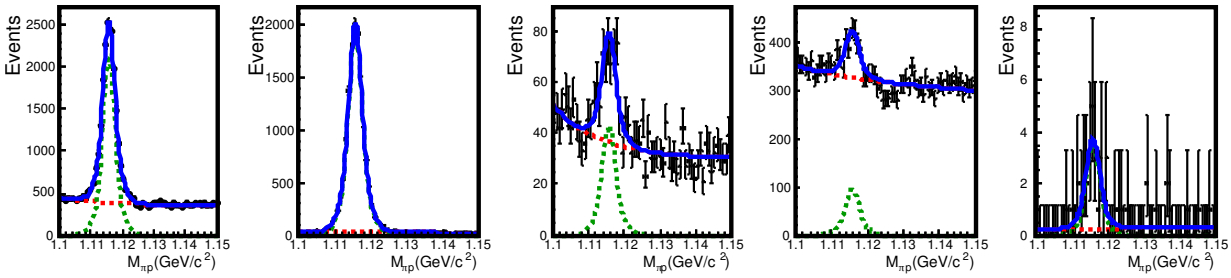


Figure 4.6: *Same as fig. 4.3 for Λ.*

Table 4.4: *Functions used to fit the ϕ_{1020}, K^0_S and $\Lambda(\overline{\Lambda})$ invariant mass distributions.*

	Background	Signal
ϕ_{1020}	$(x-2m_K)^n exp(-a1(x-2m_K))$	Gaussian $\otimes$ Relativisitc Breit-Wigner [72]
K^0_S	2nd degree Chebyshev polynomial	Two Gaussians
Λ $(\overline{\Lambda})$	$(x-(m_p+m_\pi))^n exp(-a1(x-(m_p+m_\pi)))$	Two Gaussians

$$M_R^- = \begin{pmatrix} 0.89 & 0.01 & 0.04 \\ 0.02 & 0.87 & 0.02 \\ 0.01 & 0.07 & 0.93 \end{pmatrix} \tag{4.10}$$

The last row where the $P(X \to h)$ probabilities are located is not shown, this presentation is adopted since M_R relate the number of identified hadrons $(\overrightarrow{I_h})$ to the true number of hadrons $(\overrightarrow{T_h})$ via the eq. (4.11); this system of equations is overdetermined (there are only 3 unknowns and 4 equations), however it can be solved since the last row can be expressed in terms of the others rows because of the $\sum P(\pi \to i) = 1$ condition.

$$\begin{pmatrix} I_\pi \\ I_K \\ I_p \\ X \end{pmatrix} = \begin{pmatrix} P(\pi \to \pi) & P(K \to \pi) & P(p \to \pi) \\ P(\pi \to K) & P(K \to K) & P(p \to K) \\ P(\pi \to p) & P(K \to p) & P(p \to p) \\ P(\pi \to X) & P(K \to X) & P(p \to X) \end{pmatrix} \begin{pmatrix} T_\pi \\ T_K \\ T_p \end{pmatrix} \tag{4.11}$$

Writing explicitly the expression for the identified kaons (eq. (4.12)) shows the contribution to the kaon sample from the pion and proton yields. From this relation is clear that the misidentification probabilities must be kept as low as possible.

$$I_K = P(\pi \to K)T_\pi + P(K \to K)T_K + P(p \to K)T_p \tag{4.12}$$

In order to measure the contributions from the different hadron species t to the identified sample i the *purity* of the sample $Q(t \to i)$ is defined (probability that a particle which is identified as a type i is truly of type t) (eq. (4.13)).

$$Q(t \to i) = \frac{T_t}{I_i} P(t \to i) \qquad i,j = \{\pi, K, p\} \tag{4.13}$$

The purity depend not only on the identification (misidentification) probabilities but also on the hadron rates. The hadron sample for the multiplicity analysis is mostly composed of pions ($\sim 74\%$), so the pion contribution will be the dominant one. In average $\frac{I_\pi}{I_K} \sim 7$ and $\frac{I_\pi}{I_p} \sim 20$ this means that even a small value of $P(\pi \to K)$ or $P(\pi \to p)$ causes a large pion contamination in the kaon and proton sample. The ratio $\frac{I_\pi}{I_K}$, $\frac{I_\pi}{I_p}$ and $\frac{I_K}{I_p}$ in the multiplicities binning are shown in Appendix B.

The results give in eqs. (4.9) and (4.10) are only an overview of the probabilities. In section 4.2 is shown the particle identification depends on the hadron phase space and because the probabilities depend on the particle identification, the M_R matrix also depends on the hadrons phase space. In order to take into account this dependence the RICH response is characterized by a set of M_R matrices binned in the hadrons phase space.

4.2 Likelihood Tuning

The particle identification is performed via a likelihood estimator [73] for different mass hypotheses (e, μ, π, K and p) and also for the background hypothesis (bg). The likelihood function is:

$$LH(M) = \exp[-(S_M + B)] \prod_{j=1}^{N} (s_M(\Theta_{Cj}, \varphi_j, \Theta_M) + b) \tag{4.14}$$

Where N is the number of photons used to reconstruct Θ_C (section 3.2.5), Θ_{Cj} is the measured Cherenkov angle for j photon and Θ_M is extracted from eq. (3.2) ($\Theta_M = \Theta_C$) for all mass hypothesis. The likelihood function depends on the probability for a photon to belong to the signal (s_M) or to the background (b) and their integral over the fiducial region [2]. The s_M probability takes into account the resolution of the photon detectors and lose of photons by absorption (C_4F_{10} impurities, beam pipe), holes in the mirror wall and photon detector dead zones. The background probability b is extracted directly from real data [62]. In order to extract the likelihood estimator for the background, s_M is set to zero in eq. (4.14).

As a consequence of the two different kinds of photon detectors installed in the RICH (section 3.2.3), the angular resolution is not constant in all RICH regions. The likelihood determination is directly affected by this angular resolution because of Θ_{Cj} and φ_j determination. The track parameters such as the particle momentum p_h and the polar angle θ at the RICH entrance determine the phase space where Θ_{Cj} and φ_j are measured. In consequence the particle identification is p_h and θ dependent.

As a first identification the largest value of $LH(\pi)$, $LH(K)$, $LH(p)$, $LH(e)$ and $LH(\text{bg})$ is taken as the good hypothesis. However, there are events where one or more likelihood values are very similar. In these particular cases is not enough to compare the likelihood

[2] This fiducial region is defined to keep only photons with $\theta_{Cj} < 70$ mrad

because the identification could be ambiguous. To improve the identification is possible to require $LH(h)/LH(\text{2nd})$ ($LH(\text{2nd})$: second largest likelihood) and $LH(h)/LH(\text{bg})$ distributions to be larger than some cut. With this requirements a good separation between likelihoods is ensured. The value of these cuts is tuned in order to privilege some physic channels. Since pions are the dominant hadrons in the multiplicity sample, the selection of the cuts is focused in diminishing the pion misidentification probabilities and to improve the purity of the kaon sample.

In the next subsections the procedure followed to determine the cuts in $LH(h)/LH(\text{2nd})$ and $LH(h)/LH(\text{bg})$ (likelihood tuning) distributions is explained. The different momentum p_h range where the pion, kaon or proton identification is possible with the likelihood information leads to a different approach for the likelihood tuning depending on the hadron type that is going to be identified.

4.2.1 Pion and Kaon Likelihood Tuning

A very detailed analysis of a tuning to enhance the purity of the kaon sample, while keeping its statistics as high as possible were performed for the 2007 COMPASS data in [74]. The results of this analysis were the starting point to tune the 2006 likelihood, a summary of the analysis is given in this section.

The first step consists of the extraction of the identification and misidentification probabilities with different sets of cuts on the $LH(h)/LH(\text{2nd})$ distribution. The identification and misidentification probabilities as a function of p_h for pions and kaons are shown in fig. 4.7. The identification probabilities for $p_h < 30$ GeV/c do not change their values with the different likelihood cuts. In the higher momentum range (where the Cherenkov angle saturates) the impact of the different cuts starts to be visible. An increase in the cut value is reflected in a decrease in the misidentification probabilities but also in the identification ones.

The pion and kaon purity of the sample was also extracted and they are plotted in figs. 4.8a and 4.8b respectively as a function of x. The pion purity values are high around 0.99 and the impact of the different likelihood set cuts is negligible. For kaons the purity value varies between 0.88 and 0.95 and its value is sensitive to the different likelihood cuts specially at low x.

To perform the likelihood tuning two Figures of Merit (FoM) were defined in terms of the efficiencies, purities and hadron rates:

1. $FoM1 = purity * efficiency$: depends only on p_h value (figs. 4.9a and 4.9b).
2. $FoM2 = purity * \frac{N^{ID}}{N_{all}}$: depends on the likelihood cut and p_h value(figs. 4.9c and 4.9d).

To find the optimal cut for the $LH(h)/LH(\text{2nd})$ distributions, a good compromise between the identification probability and purity of the sample were searched. The final values for the 2007 data tuning are summarized in table 4.5.

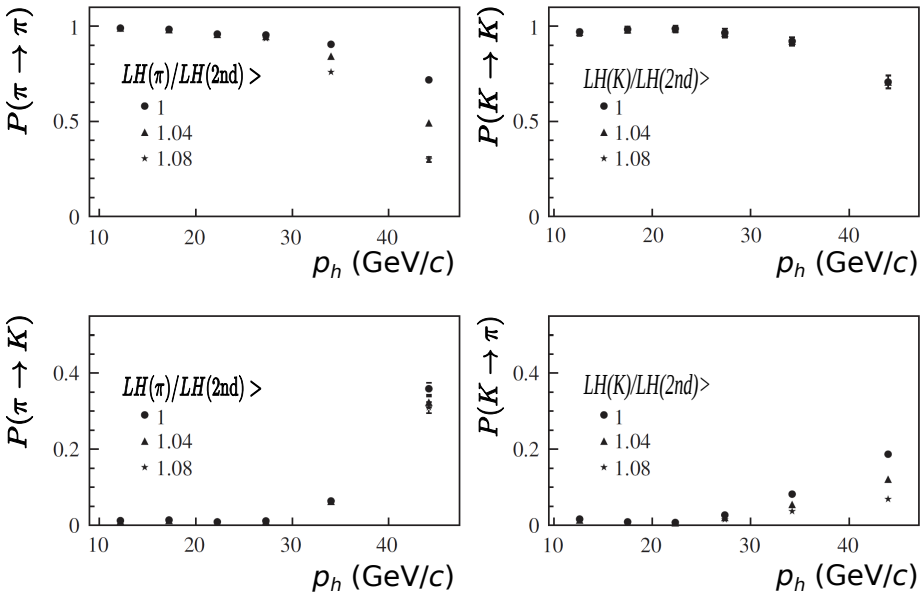


Figure 4.7: *Top raw: PID efficiency of pion (left) and kaons (right), as a function of the particle momentum p_h. Bottom raw: misidentification probability of kaons (left) and pions (right), as a function of the particle momentum p_h. Different markers correspond to different cuts on the $LH(h)/LH(2nd)$. Figure extracted from [74].*

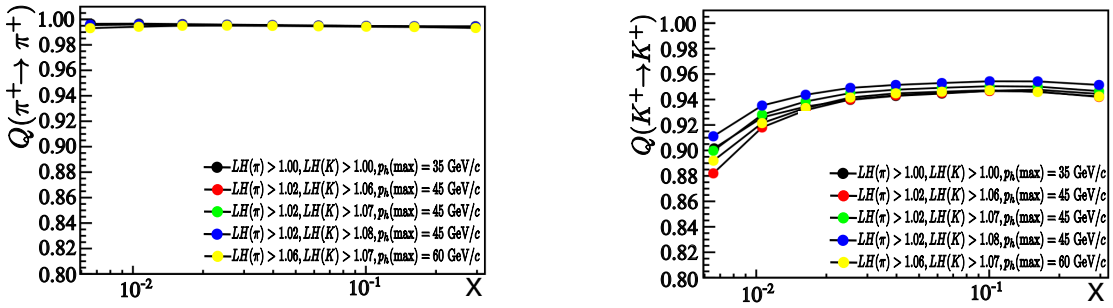
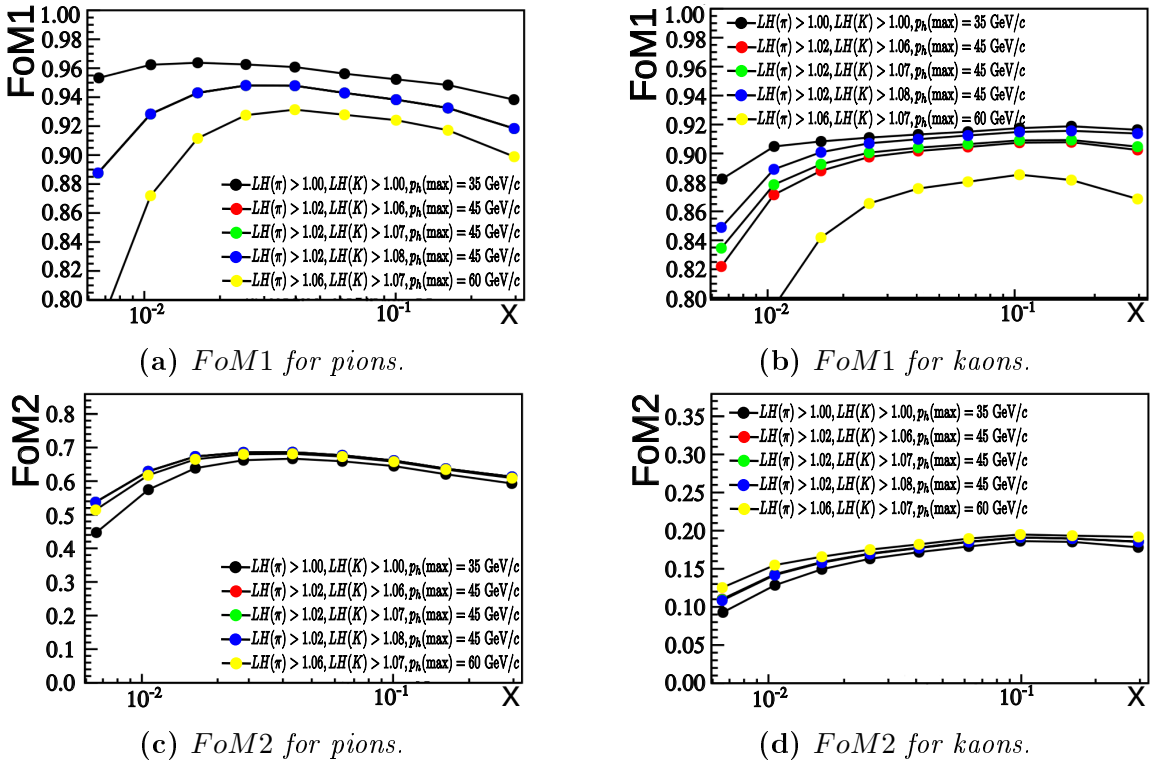


Figure 4.8: *Pion (a) and kaon (b) purity samples as a function of x. Figure extracted from [74].*

Table 4.5: *Summary of the RICH selection cuts for 2007 data.*

Pions	$LH(\pi)/LH(\text{2nd})1.02$
Kaons	$LH(K)/LH(\text{2nd}) > 1.08$

With RICH it is also possible to identify electrons. From fig. 3.6 is clear that is possible to separate electrons from pions only at low momentum values (< 8 GeV/c). A fine tuning



(a) *FoM1 for pions.* **(b)** *FoM1 for kaons.*

(c) *FoM2 for pions.* **(d)** *FoM2 for kaons.*

Figure 4.9: *Example of the two FoM for positive pions(left) and kaons (right) as a function of x. Each color corresponds to a different $LH(h)/LH(2nd)$ threshold and p_h. Figure extracted from [74].*

of this identification has been also done in [74]: it was found that is possible to distinguish electrons from pions if $LH(e) > 1.8 * LH(\pi)$. In the opposite case, the distinction is not possible and the electron hypothesis is not taken into account in the $LH(h)/LH(2\text{nd})$ calculation.

An analysis of the M_R matrices using the set of likelihood cuts defined in table 4.5 using the 2006 data were performed. The values were compared to the 2007 results and a good agreement was found.

4.2.2 Proton Likelihood Tuning

In eq. (4.12) it is shown that the protons also contaminate the kaon sample ($\sim$ 8% of the hadron sample used in the hadrons multiplicity are protons) so that, a tuning in the proton likelihood is also necessary to reduce the proton contamination. The method used to perform the likelihood tuning for protons (antiprotons) is different from the one described in the previous section since the proton threshold value is relatively high $\sim$ 17.95 GeV/c ($thr(p)$) in comparison to the kaon or pion threshold.

The identification of protons is performed directly from the proton likelihood information for particles with momentum $p_h > 17.95$ GeV/c. The protons with $p_h < thr(p)$ do

not emit Cherenkov photons and therefore the proton likelihood value will be zero. Nevertheless, is still possible to perform a particle identification. Two different ways to identify protons in the region below threshold are possible: if the proton goes through the RICH but no photons are detected in the fiducial region all the likelihoods values will be null, *i.e.* $LH(e) = LH(\mu) = LH(\pi) = LH(K) = LH(p) = LH(\text{bg}) = 0$. In the opposite case, if some photons are detected in the fiducial region, they will be considered to perform the likelihood estimation; if in addition the background likelihood value is the largest, it is possible to identify protons using the background information.

For the proton identification below threshold the likelihood tuning is performed on the $LH(\pi)/LH(\text{bg})$ and $LH(K)/LH(\text{bg})$ distributions. The FoM used to tune is defined in eq. (4.15), where S and B are the number of events in the signal and background, respectively in the Λ ($\overline{\Lambda}$) invariant mass distribution. The FoM and the signal events for Λ and $\overline{\Lambda}$ are shown in fig. 4.10 and fig. 4.11 respectively. The values that maximize the FoM are shown in each plot and they are summarized in table 4.6. In addition, particles other than protons are rejected (*i.e.* $\pi^{\pm}$, $K^{\pm}$).

$$FoM = \frac{S}{\sqrt{S+B}} \tag{4.15}$$

Table 4.6: *Summary of the likelihood cuts to identify protons for $p_h < 17.95$ GeV/c ($thr(p)$).*

p	$\overline{\text{p}}$
Not to be a π or K	
$\frac{LH(\pi)}{LH(Bg)} < 2.2$ $\frac{LH(K}{LH(Bg)} < 2.9$	$\frac{LH(\pi)}{LH(Bg)} < 2.1$ $\frac{LH(K}{LH(Bg)} < 2.8$
or ALL LH=0	

Different identification methods are applied in the different p_h momentum regions. Three p_h momentum regions are considered:

- $p_h \in [thr(K),\ thr(p) - 5)$ (GeV/c). Only the cuts provided in table 4.6 are used.
- $p_h \in [thr(p) - 5,\ thr(p) + 5]$ (GeV/c) This region is called *gray zone*. In this region, the proton candidate must satisfy the cuts from table 4.6 or $LH(p)/LH(\text{2nd}) > 1$.
- $p_h \in (thr(p) + 5, 50]$ (GeV/c). Only $LH(p)/LH(\text{2nd}) > 1$ is required. No further tuning in this distribution was studied since the Cherenkov angle saturation is significantly less important than for pions and kaons.

The importance of the *gray zone* lies in minimizing the loss of protons caused by the change in the identification algorithm. The proton (antiproton) identification probability

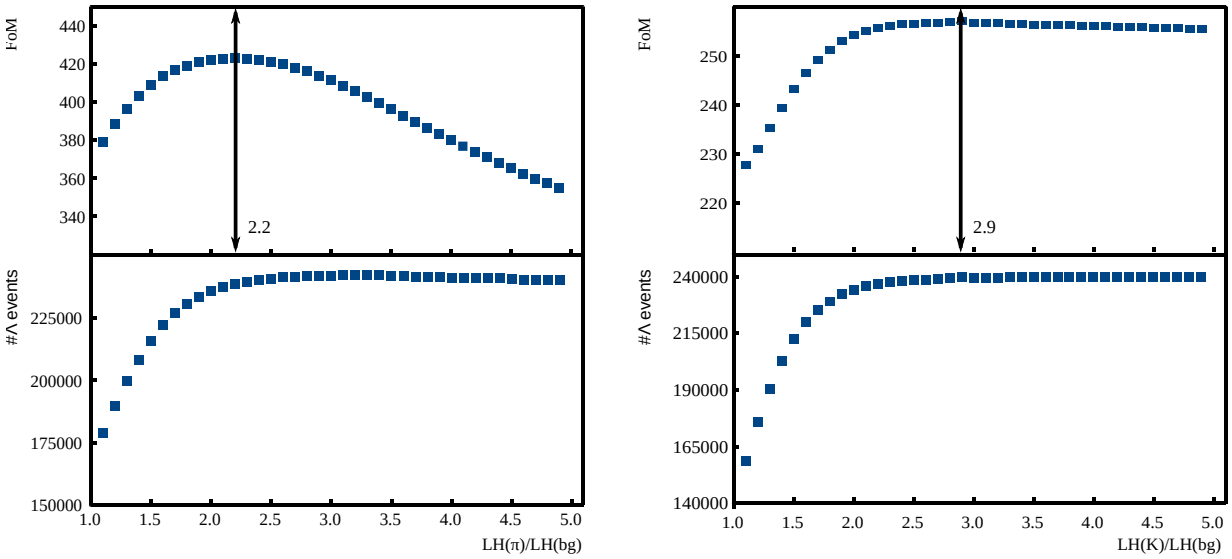


Figure 4.10: *FoM (top row) and number of Λ (bottom row) as a function of $LH(\pi)/LH(bg)$ (left) and $LH(K)/LH(bg)$ (right) in the momentum range [Thr(K),Thr(p)]. The maximum value of the FoM are shown in the plot.*

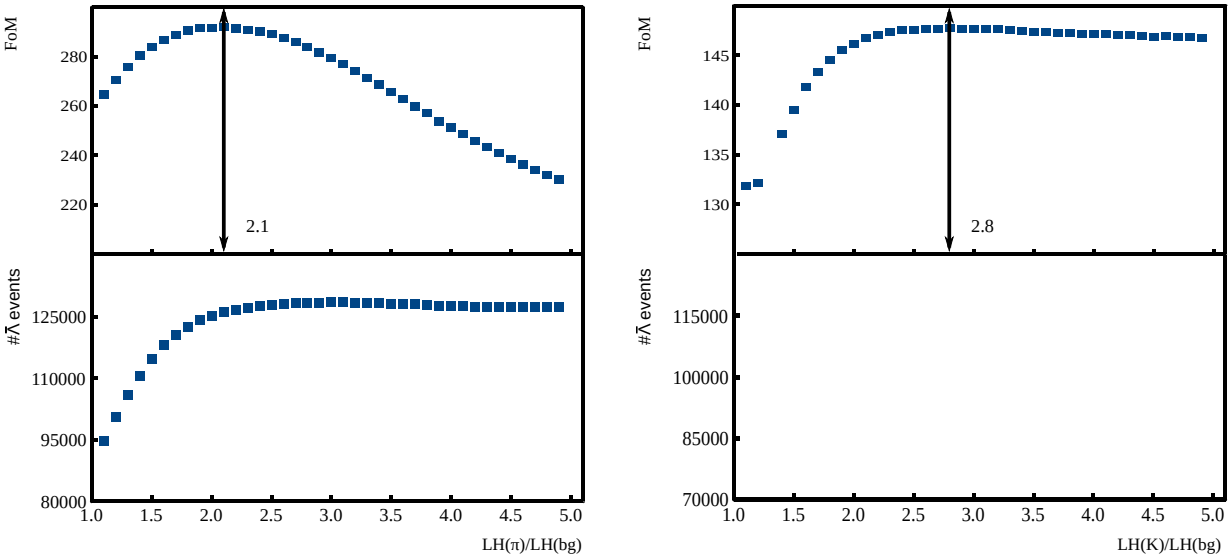
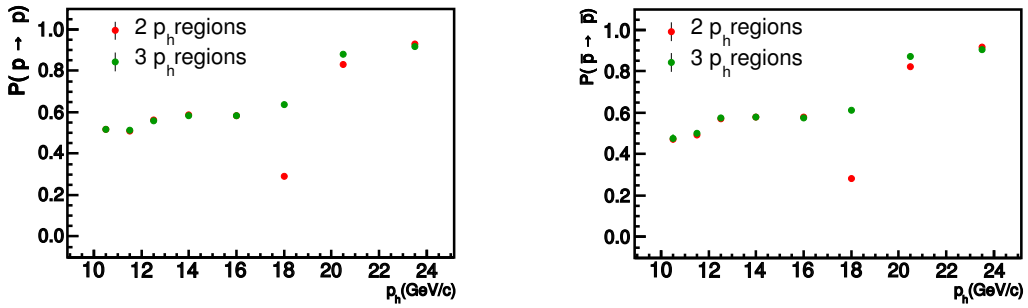


Figure 4.11: *Same as Fig. 4.10 for $\overline{\Lambda}$.*

is plotted as a function of p_h in fig. 4.12. In red, the points corresponding to $P(p \to p)$ (right) and $P(\bar{p} \to \bar{p})$ (left) when only two regions in the particle momentum range are considered ($p_h \leq thr$ and $p_h > thr(p)$); the green ones correspond to the identification probability when the p_h range is divided in three zones. The difference is visible in the identification probability value for the p_h bin which contains $thr(p)$. When the gray zone is not considered the $P(p \to p) \sim 0.30$ while if it is considered the identification probability is ~ 0.60.

The overall result for the proton PID efficiencies is shown as a function of p_h in six



(a) $P(p \to p)$ **(b)** $P(\overline{p} \to \overline{p})$

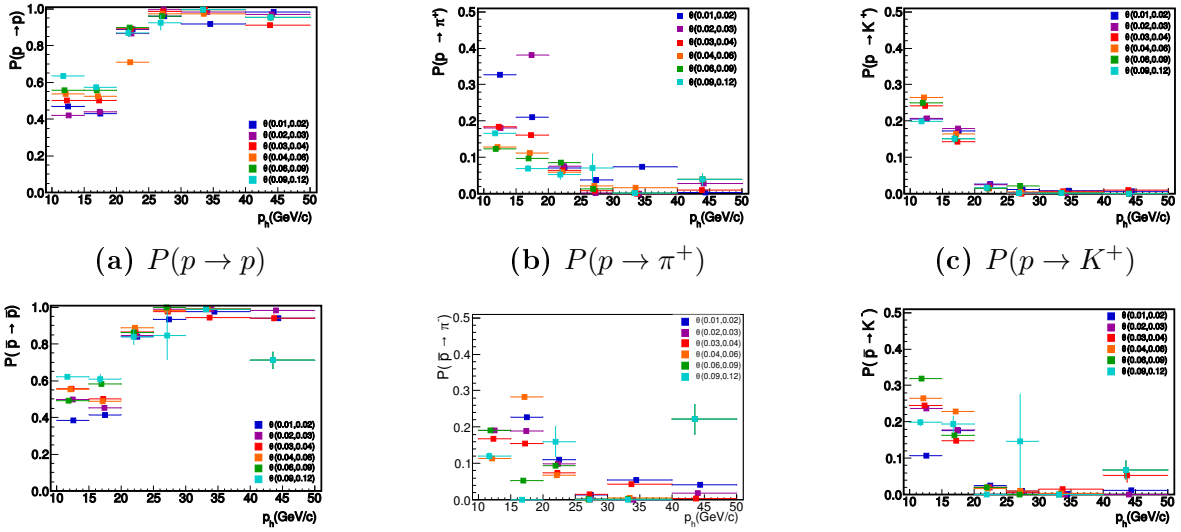
Figure 4.12: *Identification probabilities for p ($\overline{p}$) as a function of p_h. The green markers correspond to identification efficiencies when the gray zone is included in the p_h division, the red ones when p_h range is divided only in two zones.*

different θ bins in fig. 4.13. For $p_h > 25$ GeV/c the identification probability is > 0.90 for p and $\overline{p}$. The proton and antiproton misidentification as kaons or pions in this same region is < 0.05 for all θ bins.

For the smallest values of p_h (below the proton threshold) the identification probability $P(p \to p)$ and $P(\overline{p} \to \overline{p})$ are very unstable for the different θ bins: varies between $0.38 - 0.60$; the smallest identification probability is found in the largest θ bin (0.09-0.12 rad). The misidentification probabilities also present a high θ dependence. For the proton (antiproton) misidentification as pion $P(p \to \pi)$ the largest discrepancy between values is found in the second p_h bin ($15 - 20$ GeV/c), for p varies between 0.06-0.40 and 0.00-0.30 for $\overline{p}$. In the kaon misidentification case an important difference between p and $\overline{p}$ values is observed in the first p_h bin ($10 - 20$ GeV/c); $P(p \to K)$ varies from 0.2 to 0.28 while $P(\overline{p} \to K^-)$ varies from $0.11 - 0.32$. These misidentification probability values are not optimal for the kaon and pion purity. It can be concluded that the $LH(\pi, K)/LH(\text{2nd})$ tuning is not enough to separate $LH(\pi, K)$ form $LH(\text{bg})$ if we want to identify protons below threshold. The procedure followed to optimize the proton misidentification efficiencies below the proton threshold is described in the following.

The number of photons used in the RICH ring reconstruction (section 3.2.5) has a direct impact on the likelihood estimation and in consequence in the $LH(\pi, K)/LH(\text{bg})$ distribution[3] (see eq. (4.14)). The number of photons used in the ring reconstruction for pions and kaons with $10 < p_h < 20$ GeV/c in the 0.01-0.12(rad) θ range in Λ and $\overline{\Lambda}$ sample are plotted in figs. 4.14c and 4.14d respectively. The blue histogram corresponds to the events belonging to the peak (signal) distribution and the pink one to the background events. It is clear that most of the signal events are coming from likelihood estimated with a low number of photons, < 40 for pions and < 30 for kaons. The second blue peak in fig. 4.14c is the consequence of the intrinsic pion contamination in the Λ and $\overline{\Lambda}$ samples (table 4.1).

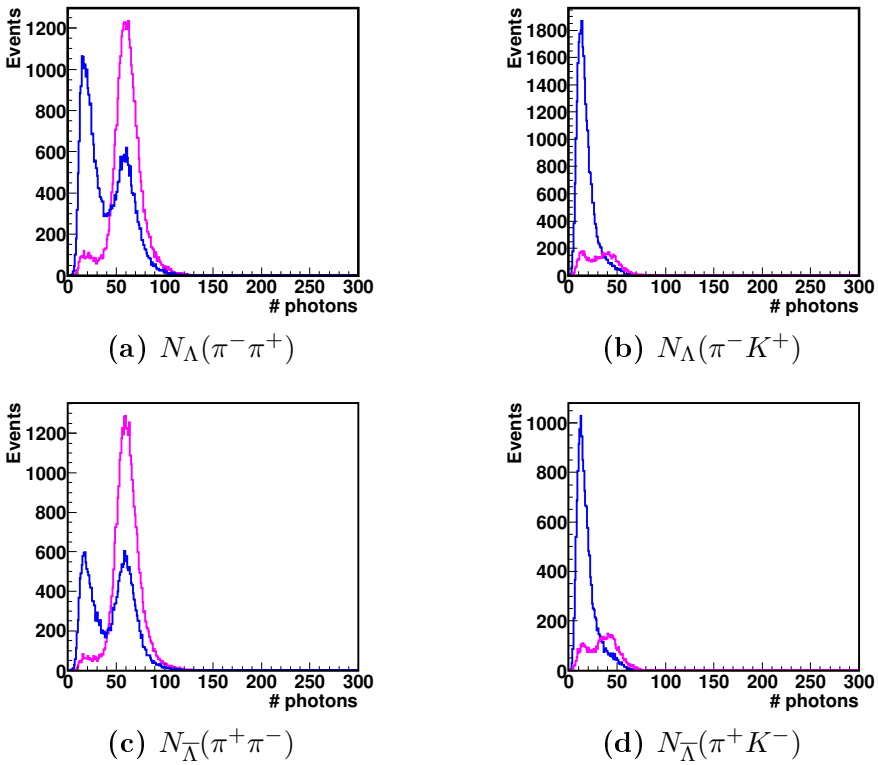
[3]The likelihood estimator is related to the number of photons used in the ring reconstruction.



(a) $P(p \to p)$ (b) $P(p \to \pi^+)$ (c) $P(p \to K^+)$

(d) $P(\overline{p} \to \overline{p})$ (e) $P(\overline{p} \to \pi^-)$ (f) $P(\overline{p} \to K^-)$

Figure 4.13: *Identification and misidentification probabilities for p (top) and $\overline{p}$ (bottom) as a function of p_h. Each different color corresponds to a different θ bin*

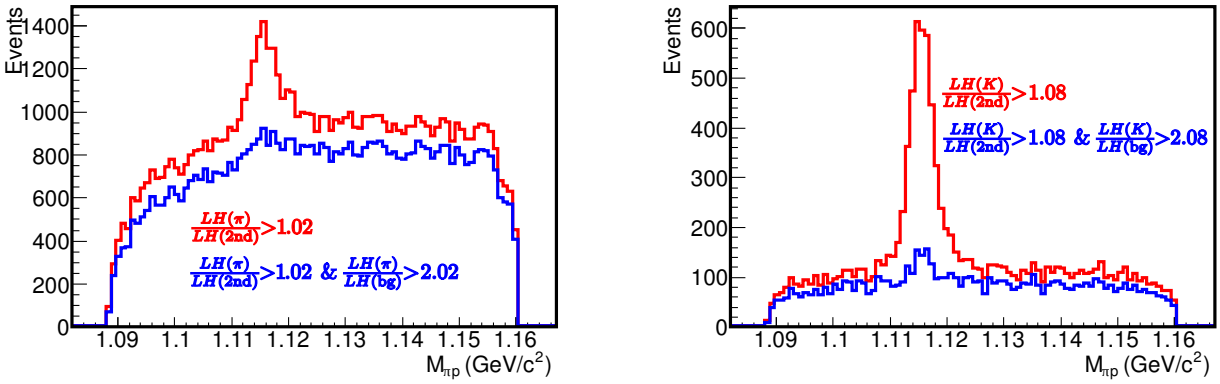


(a) $N_\Lambda(\pi^-\pi^+)$ (b) $N_\Lambda(\pi^- K^+)$

(c) $N_{\overline{\Lambda}}(\pi^+\pi^-)$ (d) $N_{\overline{\Lambda}}(\pi^+ K^-)$

Figure 4.14: *Number of photons used to reconstruct the RICH ring in the p_h region $10 - 20$ (GeV/c). Blue histogram: events from signal. Pink histogram: events from background.*

Different cut values in $LH(\pi, K)/LH(\text{bg})$ distributions have been tested and the chosen ones are $LH(\pi)/LH(\text{bg}) > 2.02$ and $LH(K)/LH(\text{bg}) > 2.08$ since the background in both distributions remains almost unchanged while the signal distribution is reduced drastically . The change in the invariant mass distribution in the sample used to extract the proton misidentification probabilities before and after this tuning are shown in figs. 4.15a and 4.15b. The red histogram corresponds to the invariant mass distribution when only the two ratios, $LH(\pi)/LH(\text{2nd})$ and $LH(K)/LH(\text{2nd})$, are considered to identify pions and kaons. The same distribution is shown in blue when $LH(\pi, K)/LH(\text{bg})$ are considered additionally. The proton (antiproton) misidentification probabilities with this new requirement are presented in section 4.3.



(a) $N_{\overline{\Lambda}}(\pi^+\pi^-)$ **(b)** $N_{\overline{\Lambda}}(\pi^+K^-)$

Figure 4.15: $\overline{\Lambda}$ *invariant mass distribution in the sample used to extract the misidentification probabilities. In red only the cut in* $LH(\pi, K)/LH(2nd)$ *is required to identify* π *and* K*. In blue the same distribution with* $LH(\pi)/LH(bg) > 2.02$ *(right) and* $LH(K)/LH(bg) > 2.08$ *(left).*

It is also important to study the impact of the new requirements in the $LH(\pi)/LH(\text{bg})$ and $LH(K)/LH(\text{bg})$ distributions on the pion and kaon identification probabilities. This study is shown in fig. 4.16 for pions and kaons. The identification probabilities are extracted in the θ range $0.01-0.12$ (rad) bins in the momentum p_h range $10-40$ GeV/c. The impact of this additional cut in the identification probabilities was found to be small: < 0.01 for pions and < 0.03 for kaons.

4.2.3 Kaon Identification and Misidentification Efficiencies

The likelihood tuning presented in the two previous sections was performed to minimize the pion and proton contamination in the kaon sample. In addition, a specific study of the RICH response for kaons was performed.

4.2.3.1 Kaon Threshold Influence

The number of photons available to perform the likelihood evaluation is related to the distance between p_h and the Cherenkov threshold limit. In the case of kaons, this limit

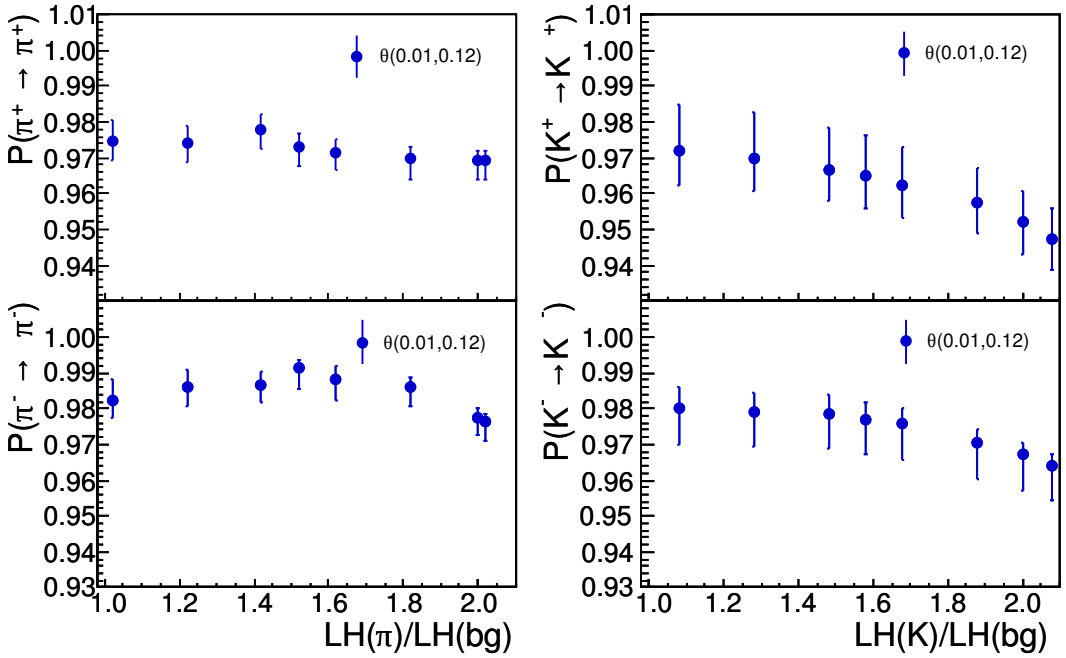


Figure 4.16: *First (second) row : Positive (negative) pion and kaon identification efficiency as a function of $LH(h)/LH(bg)$ cut value. The four plots are extracted globally in the 10-50 GeV/c p_h range in all θ range* $0.01 - 0.12$ *(rad).*

is ≈ 9.45 GeV/c. The closer the particle momentum value is to the threshold value few Cherenkov photons are emitted. A study was performed to determine if the $P(K \to i)$ probabilities values are affected by their proximity to the kaon threshold. Two types of binning in p_h have been defined in the $10 - 20$ GeV/c momentum region with a different number of bins near to the kaon threshold. The first one has only 4 bins $\{10, 12, 15, 17, 20\}$ and the second one has 10 bins $\{10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20\}$.

The kaon identification and misidentification probabilities $P(K^{\pm} \to i^{\pm})$ are plotted as a function of p_h in the θ range $(0.01, 0.12)$ in fig. 4.17, the different colors correspond to the different p_h binning. The vertical axis is located at the kaon threshold value (9.45 GeV/c). Fort the positive (left) and negative (right) case the probability to identify kaons increases as p_h value moves away from the kaon threshold (the photon production increases as the momentum p_h value gets larger). For momentum values $p_h > 13$ GeV/c the identification probability values are stables.

In the particular case where the p_h binning is finer is possible to see in more detail the influence of the kaon threshold in $P(K^{\pm} \to K^{\pm})$. For the first three bins $(10 - 13$ GeV/$c)$, the identification probability values varies from 0.70 to 0.95 for the positive case and from 0.75 to 0.94 for the negative one. Comparing the identification probability value between the first p_h bin and the fourth one the difference reach ~ 0.29 and ~ 0.21 for positive and negative kaons respectively. The misidentification probabilities of kaons as pions or protons are not affected by the kaon threshold. It can be concluded that the kaon identification is sensitive to the kaon threshold for momentum values $p_h < 12$ GeV/c.

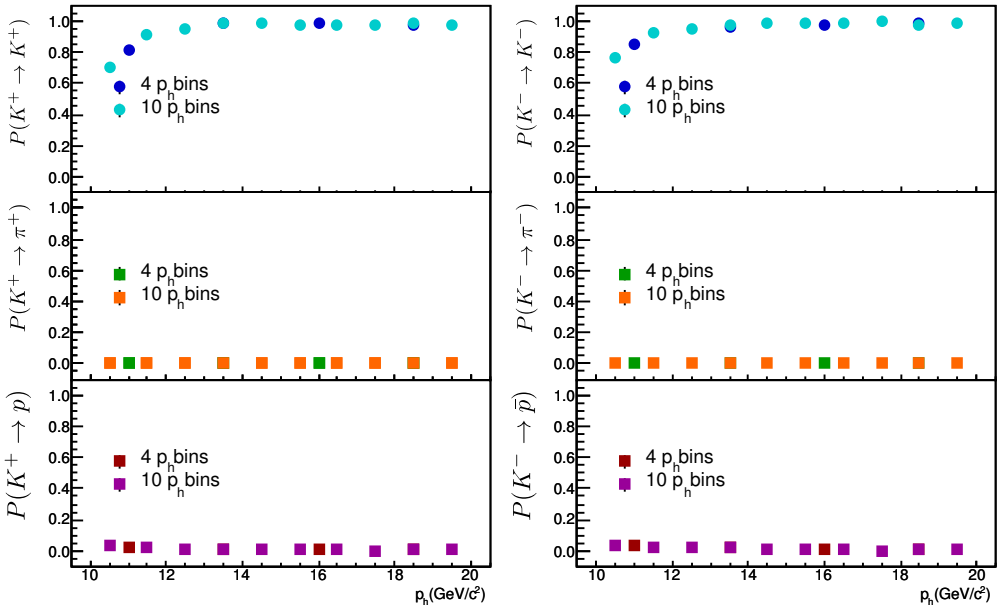


Figure 4.17: *Kaon identification efficiency as a function of p_h between $10-20$ GeV/c, in the θ range $(0.01, 0.12)$. The horizontal line corresponds to the kaon PID threshold (~ 9.45 GeV/c). The blue markers correspond to the 4 p_h binning the red ones to the 10 p_h binning.*

4.2.3.2 Event Multiplicity

The event multiplicity is defined as the total number of outgoing tracks within the event. The two subsamples extracted from the ϕ_{1020} sample show a very different multiplicity, this difference is illustrated in fig. 4.18. The $\phi^{\text{excl}}_{1020}$ multiplicity is shown in red and the $\phi^{\text{incl}}_{1020}$ in blue. For $\phi^{\text{excl}}_{1020}$ most events have a multiplicity $= 3$ (outgoing μ and the hadron pair) while in $\phi^{\text{incl}}_{1020}$ the event multiplicity varies between $3-10$. The event multiplicity is reflected in the amount of background in the invariant mass distribution. The background in the exclusive sample $\phi^{\text{excl}}_{1020}$ (fig. 4.5) is significantly lower than the background in the inclusive sample $\phi^{\text{incl}}_{1020}$ (fig. 4.4).

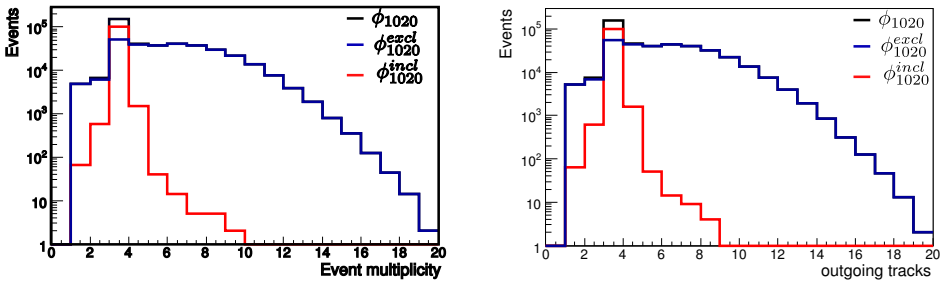


Figure 4.18: *Event multiplicity from ϕ_{1020} decay. Left: K^+. Right: K^-*

The identification and misidentification probabilities were extracted in the $\phi^{\text{excl}}_{1020}$ and $\phi^{\text{incl}}_{1020}$ samples. The difference between the probabilities values $\Delta P = P(\phi^{\text{incl}}_{1020}) - P(\phi^{\text{excl}}_{1020})$

is plotted in fig. 4.19 as a function of the particle momentum p_h. The different colors in each plot correspond to the probabilities measured in different θ bins.

The $P(K^{\pm} \to K^{\pm})$ and $P(K^{+(-)} \to p(\bar{p}))$ probabilities are affected by the event multiplicity in the p_h bins near to the kaon threshold. For both probabilities the effect is slightly larger on the first θ bin. The difference on the identification probabilities due to the event multiplicity reaches 0.25 for positive and negative kaons. For $P(K^{+(-)} \to p(\bar{p}))$ this difference is twice bigger. In both cases the difference is consequence of the amount of background in each sample. The kaon misidentification as pions behaves similar regardless of the event multiplicity.

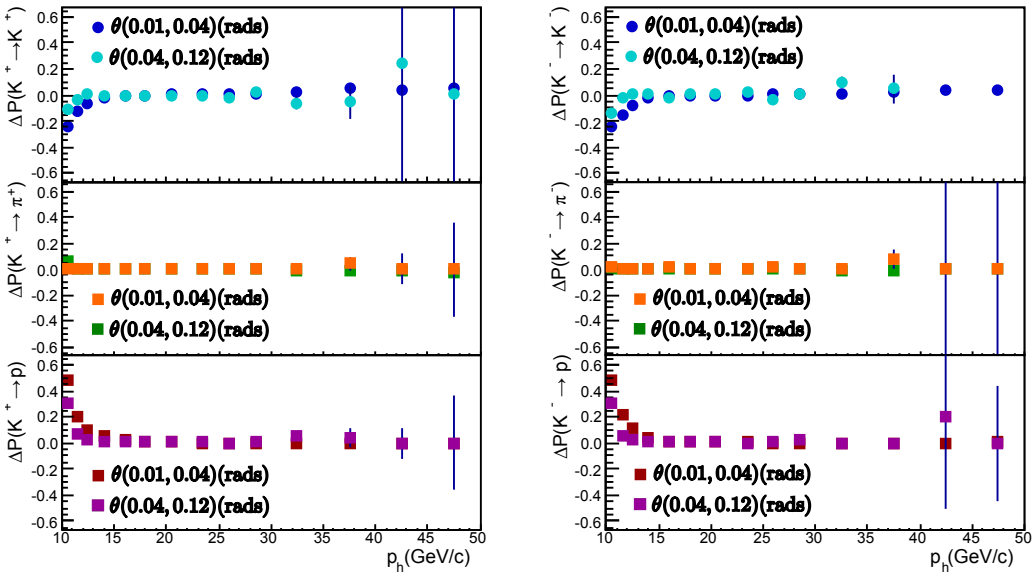


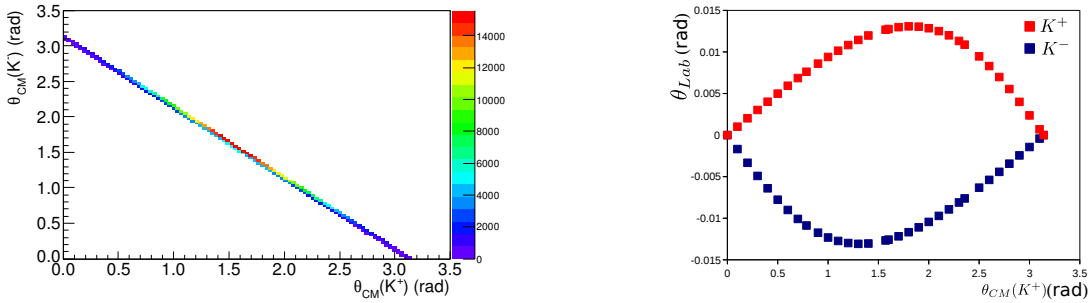
Figure 4.19: *Difference between ϕ^{incl}_{1020} and ϕ^{excl}_{1020} for kaon RICH identification (1^{st} row) and misidentification efficiencies (2^{nd} and 3^{rd} row). Left: K^+. Right: K^-.*

Even if this study was performed only on the kaon sample it can imply that RICH performance depends on the event multiplicity. An improvement to the work presented in this thesis will be an extraction of the identification and misidentification probabilities also as a function of the event multiplicity.

4.2.3.3 Further Studies

A study of the RICH response with respect to the $\phi^{\text{excl}}_{1020}$ decay angle was performed. This study is interesting because there are some regions where the information of the two decay particles is not available; for example if the two kaons travel near to the pipe region, but one of them is rejected by the RICH pipe cut the identification probability will be affected. The decay angle is measured at the center of mass for K^+ ($\theta_{CM}(K^+)$) and K^- ($\theta_{CM}(K^-)$). The correlation between the two decay angles is shown in fig. 4.20a; in the horizontal axis the decay angle values for K^+ and on the vertical one those corresponding to K^-. Most of

the statistics lies near to the $\theta_{CM} = \pi/2$ region. This in the laboratory frame correspond to the maximal decay angle. The decay angle of the two particles in the laboratory frame as a function of the decay angles in the ϕ_{1020} center of mass frame is shown in fig. 4.20b. To produce this plot, ϕ_{1020} is assumed to have a 20GeV/c momentum.



(a) **(b)**

Figure 4.20: *a) Decay angle correlation between K^+ ($\theta_{CM}(K^+)$) and K^- ($\theta_{CM}(K^-)$) coming from ϕ_{1020} decay. b) Decay angle in the frame where ϕ_{1020} is moving with momentum 20 GeV/c. x-axis: K^+ decay angle in the center of mass frame. y-axis: decay angle of each kaon relative to phi_{1020} direction.*

The identification (misidentification) probabilities are extracted in three different θ_{CM} bins: [0, 1.3, 1.7, 3.14]. The results for K^+ and K^- probabilities are shown in fig. 4.21 as a function of p_h for the (0.01,0.04) θ bin. The global probabilities are plotted in blue. The orange, green and red markers correspond to the different θ_{CM} bins. In general, the misidentification probabilities are not affected by the different decay angles. For the identification probabilities only two momentum bins are affected by the decay angle: the first one and the last one. In both cases the difference is consequence of the lack of statistics. We can conclude the RICH response is independent of the decay angle.

Another study was performed to verify if the RICH response is affected by the magnetic field generated by SM1 along the particle track. The study is very similar to the one previously explained. In this case the identification and misidentification probabilities as a function of the azimuthal angle (φ_{Az}) were studied. The binning chose for this analysis is $[-3.14, -1.2, 1.2, 3.14]$. The correlation between the azimuthal angle φ_{Az} for K^+ and K^- is shown in fig. 4.22.

The identification and misidentification probabilities are plotted in the same way as the θ_{CM} analysis in fig. 4.23. Non visible effect is observed for the misidentification probabilities. For the identification probabilities, in the two last momentum bin, a slight difference between the lower and the largest φ bins is observed.

4.3 RICH Performance Results

The RICH performance depends on the hadron phase defined by: the particle polar angle at RICH entrance (θ) and the hadron momentum (p_h). In this section the RICH performance

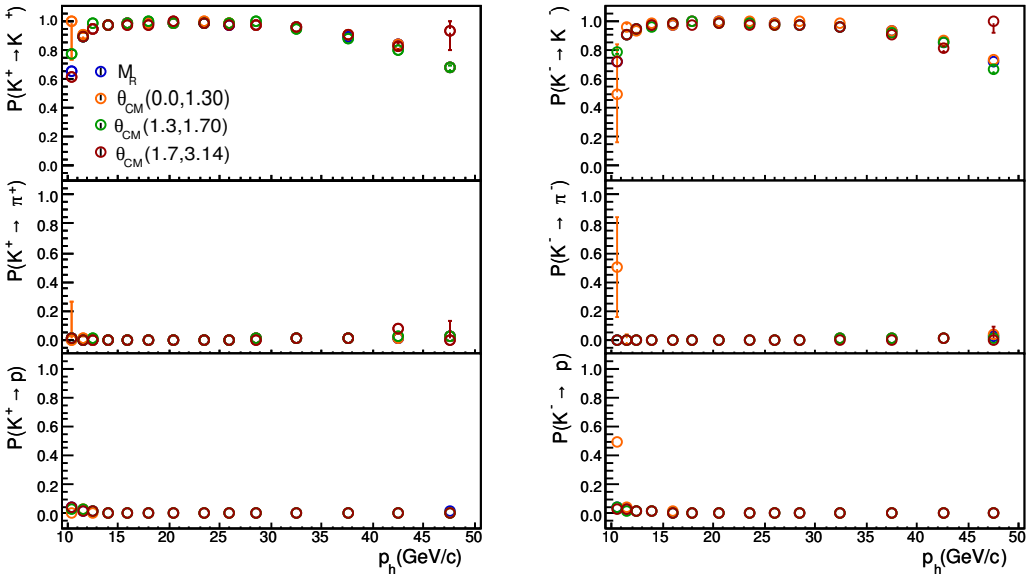


Figure 4.21: *$P(K^+ \to K^+)$ as a function of p_h. The green markers correspond to the final values. The blue, orange and yellow markers corresponds to the different θ_{CM} bins.*

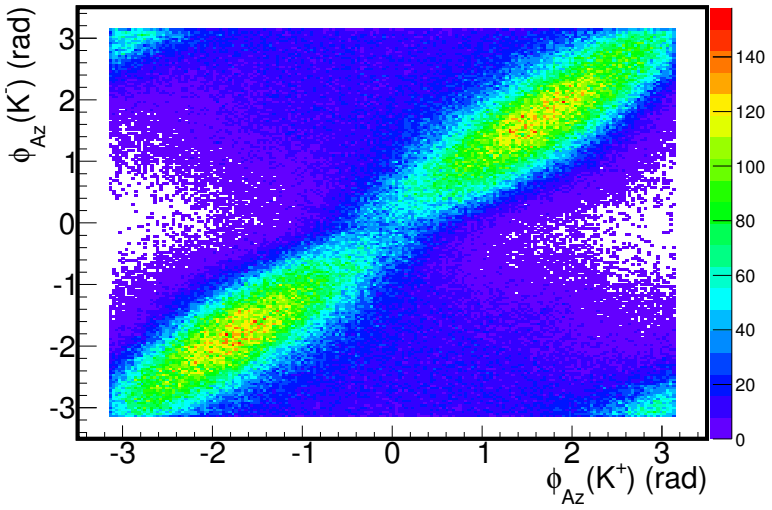


Figure 4.22: *Same as fig. 4.20 for φ_{Az}.*

results for the binning shown in table 4.7 are discussed. This binning was specific defined for the hadron multiplicity analysis, in order to reach this binning different sets of binning were studied and their corresponding results can be found in Appendix C.2. In these binnings it was observed that the RICH performances depend mostly on p_h; that is why the binning for the multiplicities analysis has 14 bins in p_h and only 2 bins in θ. The RICH identification and misidentification efficiencies for positive and negative hadrons are plotted in figs. 4.24 and 4.25, respectively.

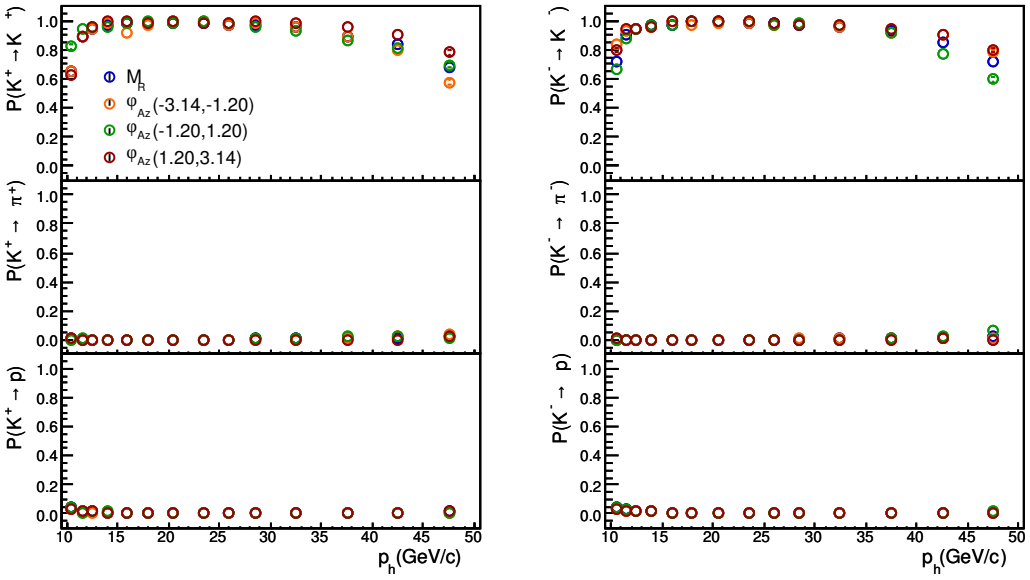


Figure 4.23: *Same as fig. 4.21 for φ_{Az}.*

Table 4.7: *Binning in p_h and θ used to extract the RICH response.*

p_h (GeV/c)	$10, 11, 12, 13, 15, 17, 19, 21, 25, 27, 30, 35, 40, 45, 50$
θ (rad)	$0.01, 0.04, 0.12$

For pions the identification efficiency is $\geq 97\%$ up to 30GeV/c, afterwards the efficiency starts to decrease and reach values below 50% in the last p_h bin. The misidentification of pions as kaons and protons remains below 3% for both of them in the $p_h < 30$GeV/c region. For higher p_h values the pion misidentification as kaons starts to increase its value as a result of the Cherenkov angle saturation explained in section 3.2.5. This behavior is not observed in the proton misidentification whose values remain constants.

In the $10 - 13$ GeV/c region, the identification efficiency for kaons varies between $40\% - 98\%$. This can be explained by the kaon threshold proximity (~ 9.45 GeV/c). In the same region, the kaon misidentification as protons reach its largest values $\sim 55\%$. In general the population of kaons is bigger than the protons one, in consequence the proton sample will be largely contaminated by kaons in this region.

The proton identification probability varies between $80\% - 99\%$ for all p_h domain. The bump in the middle p_h region is related to the different PID algorithms. The proton misidentification as pions is always bigger that the kaons ones for $p_h < 40$ GeV/c, this can be explained because the Λ sample is not a pure proton sample.

For the three particles the lack of statistics for high p_h and θ values has as a consequence the increment in the error bars.

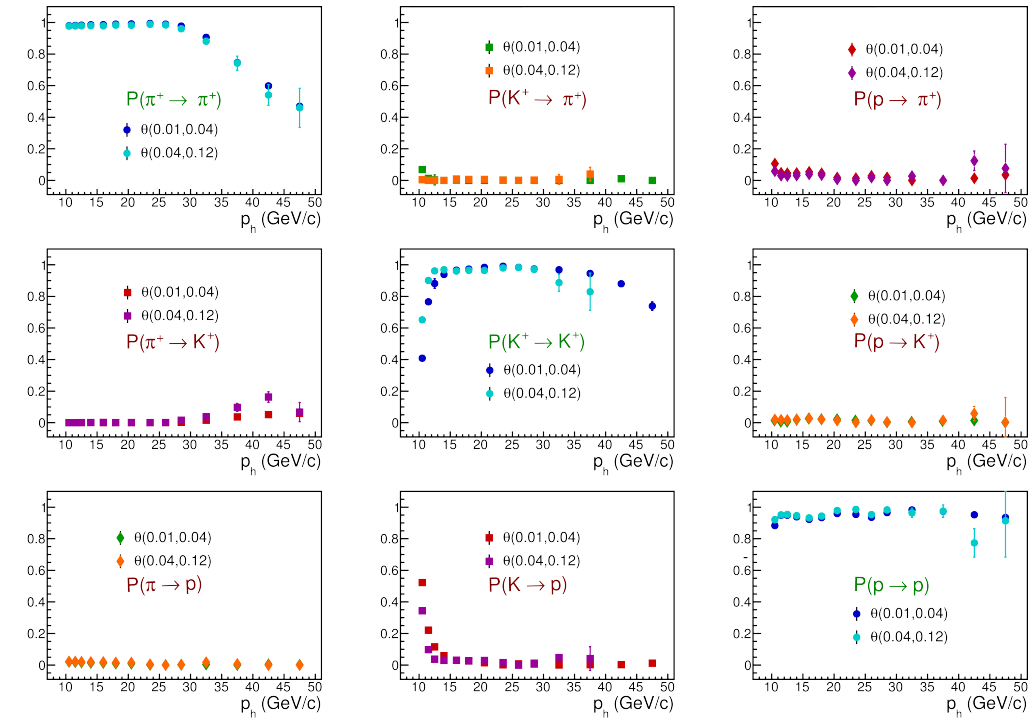


Figure 4.24: *Positive identification and misidentification probabilities for positive hadrons. The probabilities are plotted as a p_h function, the different color in each plot corresponds to a different θ bin.*

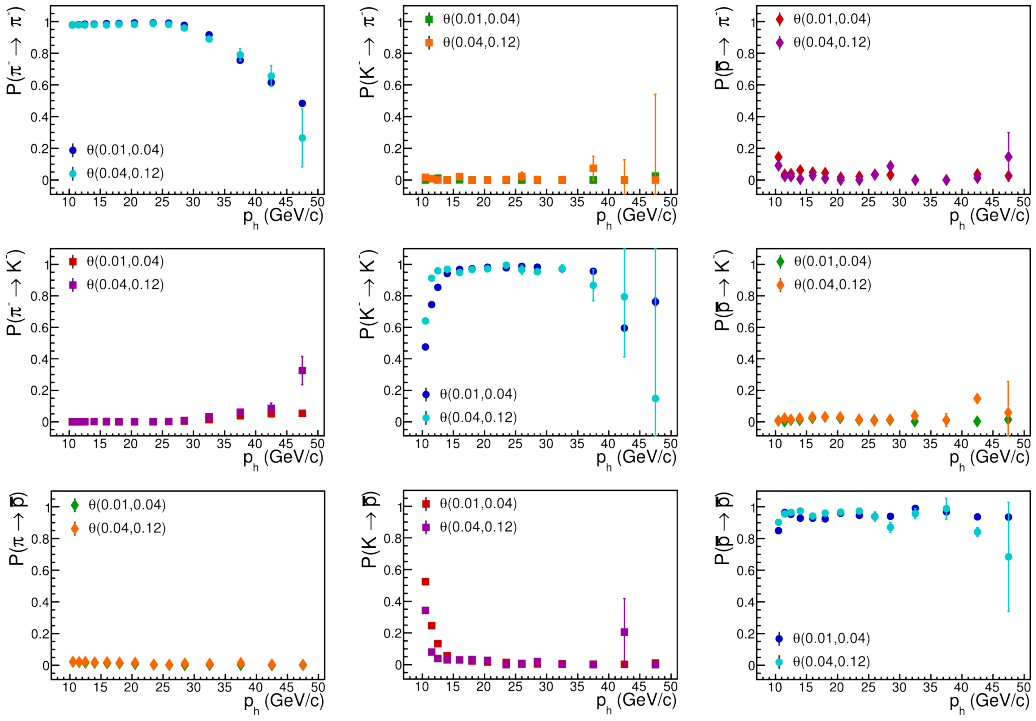


Figure 4.25: *Same as fig. 4.24 for negative hadrons.*

4.4 Particle Identification with RICH Detector

The particle identification is divided in two stages: in the first one pions, kaons and protons are identified with the requirements summarized in table 4.8. In the second stage the number of identified hadrons is corrected by the RICH efficiencies. This method is known as *unfolding*.

Table 4.8: *Summary of the RICH cuts to identify pions, kaons and protons in* $p_h \in [10, 40] GeV/c$. *Indentically to the ones used in the analysis.*

π	$\frac{LH(\pi)}{LH(\text{2nd})} > 1.02$ & $\frac{LH(\pi)}{LH(\text{bg})} > 2.02$		
K	$\frac{LH(K)}{LH(\text{2nd})} > 1.08$ & $\frac{LH(K)}{LH(\text{bg})} > 2.08$		
	$p_h < thr(p) - 5$	$p_h \in [thr(p) - 5, thr(p) + 5]$	$p_h > thr(p) + 5$
$p(\overline{p})$	Not to be a π or K $\frac{LH(\pi)}{LH(Bg)} < 2.2$ (2.1) $\frac{LH(K)}{LH(Bg)} < 2.9$ (2.8) or All $LH = 0$	Not to be a π or K $\frac{LH(\pi)}{LH(Bg)} < 2.2$ (2.1) $\frac{LH(K)}{LH(Bg)} < 2.9$ (2.8) or All $LH = 0$ or $\frac{LH(p)}{LH(\text{2nd})} > 1$	$\frac{LH(p)}{LH(\text{2nd})} > 1$

RICH unfolding

The unfolding method is used to correct the yields of identified hadrons for RICH detection performance. The true number of hadrons $\vec{T_h}$ can be solved from eq. (4.11) leading to the relation:

$$\overrightarrow{T_h} = M_R^{-1} \overrightarrow{I_h}, \tag{4.16}$$

where M_R^{-1} is the inverse of the M_R matrix truncated already presented in section 4.1.2. The unfolding correction is only possible if the M_R determinant is different from zero.

The T_h error propagation is taken from [75] and is given in eq. (4.17).

$$cov(T_i, T_j) = I_\alpha I_\beta \; cov(M_{i\alpha}^{-1}, M_{j\beta}^{-1}) + M_{ik}^{-1} M_{jk}^{-1} \; cov(I_k, I_l) \tag{4.17}$$

with

$$cov(M_{\alpha\beta}^{-1}, M_{ab}^{-1}) = \; cov(M_{ij}, M_{kl}) M_{\alpha i}^{-1} M_{j\beta}^{-1} M_{ak}^{-1} M_{lb}^{-1} \tag{4.18}$$

and

$$cov(I_k, I_l) = \sigma_k^2 \delta_{kl} \tag{4.19}$$

4.5 Summary

The RICH detector performances are determined from real data, using samples of π, K or p coming from the decay into two charged particles of K_S, ϕ or Λ. To take into account the dependence of the RICH performance on the hadron phase space, the performances are extracted in bins of the particle momentum (p_h) and the track polar angle (θ) at the RICH entrance.

A likelihood tuning for minimizing the pion contamination in the kaon sample was presented. In addition, a likelihood tuning to perform the proton identification below and above threshold was described. The optimal cuts for the hadron identification are summarized in table 4.8.

High identification probabilities are reached for hadron momentum p_h below 30 GeV/c: for pions > 0.97, for protons > 0.80 and for kaons > 0.90 (excluding the zone near to the kaon threshold ~ 9.7GeV/c). The identification probability values drop for larger p_h values in the pions and kaons case due to the Cherenkov angle saturation ($\beta \to 1$). As a consequence the misidentification probabilities become important in this.

Chapter 5

Extraction of Hadron Raw Multiplicities

The goal of the present work is the extraction of the charged hadron multiplicities from muon Deep Inelastic Scattering events (DIS) on a deuterium isoscalar[1] target (^6LiD). The multiplicities are defined as the mean number of charged hadrons produced per DIS event, and the goal is to measure them with high accuracy in a very fine binning.

The multiplicity analysis is divided into stages. In the first step the DIS and hadron yields (N^{DIS} and $N^h,\ h = h^\pm, \pi^\pm, K^\pm$ respectively) are extracted from experimental data in order to determine the "raw "multiplicities ($M^h_{raw} = N^h/N^{DIS}$) in an adequate binning of the relevant kinematical variables. The steps followed to extract M^h_{raw} are developed in detail in this chapter: DIS sample selection, hadron sample selection and particle identification. Also the binning selected to extract the multiplicities is discussed.

The value of M^h_{raw} is affected by the geometrical limitations of the spectrometer, the detector performance and by the data reconstruction efficiency. In this chapter only the corrections associated to the particle identification (performed by the RICH detector) and QED radiative effects are applied. The corrections which need a Monte Carlo simulation, such as the correction for the limited geometric and kinematic acceptance of the spectrometer are discussed in chapter 6. The last stage of the analysis consists in correcting M^h_{raw} to obtain the final multiplicity results which are discussed in chapter 7.

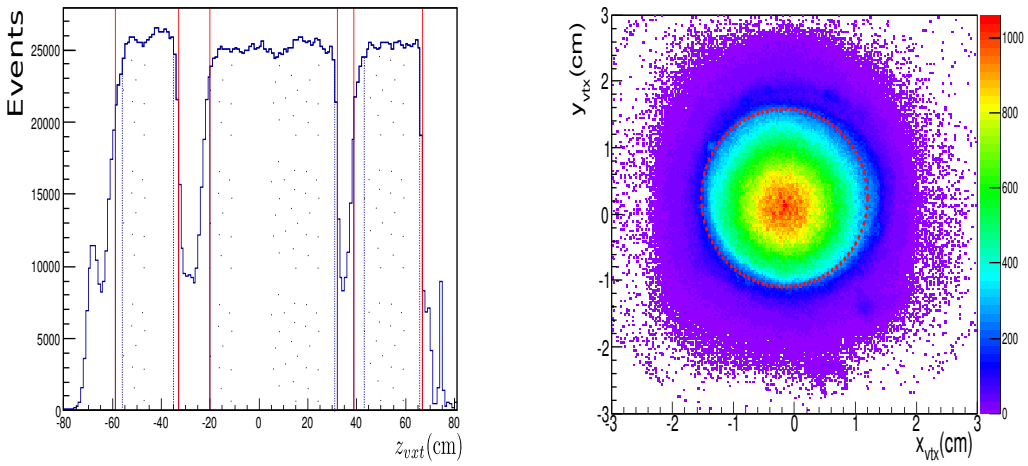
5.1 DIS Event Sample Selection

5.1.1 Preselected Sample

The hadron multiplicity analysis was performed using 6 weeks of 2006 data[2]. To improve the analysis speed, a preselected data sample including the following cuts was used. The

[1] Same number of protons and neutrons.
[2] W40, W41, W42, W43, W45 and W46

first cut consists in rejecting the bad spills defined in table 4.2. The remaining events must have at least one primary vertex reconstructed. The primary vertex is asked to have two tracks; one of the tracks corresponds to the incoming muon μ and the other one to the scattered muon μ'. If the μ' tracks is not reconstructed, the primary vertex is rejected; this cut removes half of the statistics. The μ energy is restricted to the $140 - 180$ GeV range. The primary vertex longitudinal position z_{vtx} is required to be within one of the three target cells *i.e.* $z_{vtx} \in [-59, -33] \cup [-20, 32] \cup [39, 67](cm)$. In fig. 5.1a the z_{vtx} distribution is shown (however in the final DIS selection); the three cell target location is represented by the vertical red lines. An additional radial cut in the vertex position is applied, defined as $x_{vxt}^2 + y_{vtx}^2 < 1.4^2 cm^2$. As it can be seen in fig. 5.1b the target is not centered at zero. The shifted position of the target is taken into account before the cut is performed. The red circle (dotted line) delimits the events which fulfill the radial cut. Some additional preselection cuts are applied on the event kinematics, $Q^2 > 0.8(\mathrm{GeV}/c)^2$ and $0.05 < y < 0.95$. The size of the effect of all the preselection cuts is summarized in table 5.1.



(a) (b)

Figure 5.1: *Primary vertex distribution. (a) z_{vtx} distribution. The red lines correspond to the target location. The dotted area represents the events of the DIS final sample (section 5.1.2). (b) y_{vtx} as a function of x_{vxt}. Events inside the red circle correspond to the DIS final sample (section 5.1.2).*

5.1.2 DIS Event Statistics

For the final data sample more requirements are applied. A stricter cut on the longitudinal position of the vertes z_{vtx} is applied to remove the bad filling at cell target extremities: $z_{vtx} \in [-56, -35] \cup [-20, 31] \cup [43, 66](cm)$. These new limits are represented in fig. 5.1a by the blue dotted vertical lines.

In order to ensure the same muon beam flux along the whole target, the beam track

Table 5.1: *Number of events and fraction of remaining events after each cut in the preselected event sample.*

	# events	Fraction (%)
Total events	1487206272	100.00
Primary vertex (PV)	1385088896	93.13
μ' reconstructed	776427136	52.21
$E_\mu \in [140 - 180]$ GeV	769273152	51.73
PV inside target region	407598080	27.41
$Q^2 > 0.8(\mathrm{GeV}/c)^2$	49767444	3.35
$y \in [0.05, 0.95]$	41106093	2.76

and its extrapolation are required to cross all target cells inside the radial target region. This is requested only to be consistent with previous spin asymmetry analysis [1, 76].

To select the DIS kinematic domain, the following cuts are applied on the kinematic variables Q^2, y and W (their impact in the sample size are summarized in table 5.2):

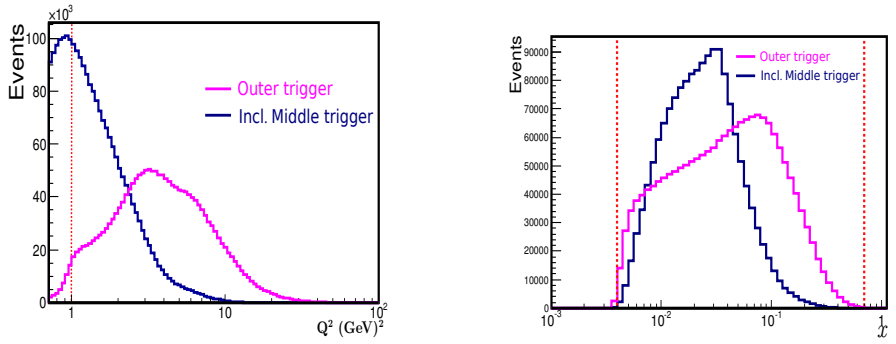
- $Q^2 > 1(\mathrm{GeV}/c)^2$ and 5 GeV $< W <$ 17 GeV . To select the deep inelastic scattering regime. The lower W limit is to avoid the region of nucleon resonance. The upper W is to avoid low statistic region.
- $0.1 < y < 0.9$. The lower limit ($y > 0.1$) removes events with poorly reconstructed μ' (momentum resolution degradation at low angle) and the misidentification of beam halo muons as μ'. The upper limit ($y < 0.9$) eliminates events where large radiative corrections are needed ($> 20\%$). .

It is also important to reject the events for which the incoming momentum is badly reconstructed[3].

The final step to select the DIS events is performed using the information provided by the trigger system described in section 2.6. For this analysis only the inclusive middle trigger (iMT) and the outer trigger (OT) are used. These two triggers were chosen because of their inclusive nature, *i.e.* only a scattered μ' is required to fire the triggers. The covered region for each of the triggers is shown as a function of Q^2 in fig. 5.2a and x in fig. 5.2b. The inclusive middle trigger (blue lines) covers the low Q^2 region while the outer trigger (magenta lines) covers the high Q^2. The very low and very high x regions are covered by the outer trigger, while the middle x region is dominated by the inclusive middle trigger.

In 2006, a significant region of lower efficiency was found for the iMT. This inefficiency region could not be reliably described by the Monte Carlo simulation, it was decided to remove the problematic region by using a geometrical cut [77]. The events where the μ' track passes through inefficient slabs are rejected.

[3]For this the reconstruction program PHAST function PaParticle::Chi2CutFlag() is used. If PaParticle::Chi2CutFlag() is true, the event is rejected.



(a) (b)

Figure 5.2: *Event distribution for the inclusive middle (iMT in blue) and outer (OT in magenta) triggers as a function of Q^2 (left) and x (right). The red dotted lines delimit the kinematical region covered in this analysis.*

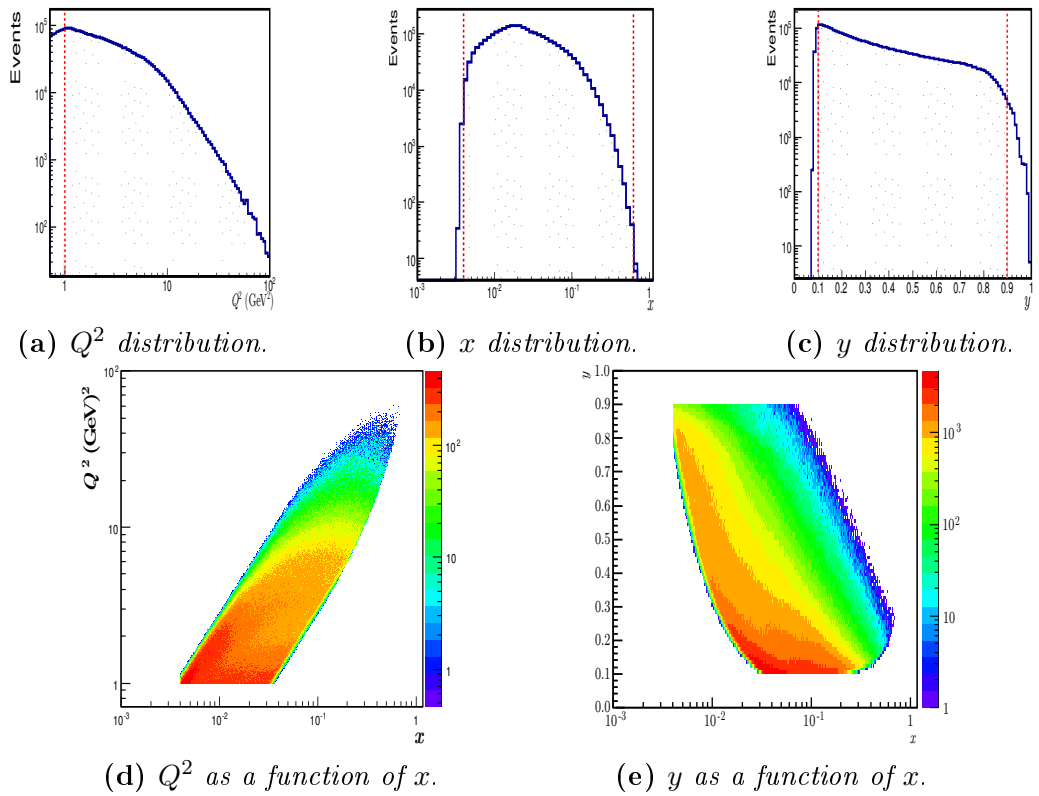
The Q^2, x and y distributions are shown in fig. 5.3, with the final DIS sample delimited by the red dotted lines. The $Q^2 - x$ correlation is illustrated in fig. 5.3d where the strong correlation between these two variables is visible. Most of the statistics is located at low $Q^2 - x$ values. For the present analysis the domain is restricted from 0.004 to 0.7. At high x values, Q^2 values up to 80 $(\mathrm{GeV}/c)^2$ are reached.

Table 5.2: *DIS event statistics after each cut and fraction of remaining events.*

	# events	Fraction (%)
Total events (preselection)	41106093	100.00
z_{vtx} extra cut	37286389	90.71
μ cross all target	36709511	89.3
$Q^2 > 1(\mathrm{GeV}/c)^2$	33766580	74.57
$y \in [0.1, 0.9]$	27348839	60.38
$W \in [5, 17](\mathrm{GeV})$	24694434	60.07
μ momentum reconstruction	21984586	48.53
Trigger	13044824	31.73
$x \in [0.004, 0.7]$	13033506	31.71

5.2 Hadron Selection: $h^{\pm}$, $\pi^{\pm}$ and $K^{\pm}$

The criteria to the hadron selection is presented in this section. We count the total number of charged hadrons h^+ and h^- (also called *unidentified* hadrons) and the total number of identified π^+, π^-, K^+ and K^- (subsamples of h^+ and h^-).



(a) Q^2 *distribution.* **(b)** x *distribution.* **(c)** y *distribution.*

(d) Q^2 *as a function of* x. **(e)** y *as a function of* x.

Figure 5.3: *Top:* Q^2*,* x *and* y *distribution. The red dotted lines delimit the final DIS sample. Bottom:* $Q^2 - x$ *and* $y - x$ *correlations for the final sample only.*

All the outgoing tracks in the primary vertex are studied to find hadron candidates. The tracks already tagged as μ' are rejected. The amount of material traversed by the track has to be smaller than 15 radiation lengths ($XX0$); the purpose of this cut is to reduce the muon contamination. To ensure a precise momentum reconstruction, the first measured point of the track is requested to be before SM1 ($Z_{first} < 350\ cm$) and the last one (Z_{last}>350) are requested to be after SM1. The cut in Z_{last} helps also to reject tracks reconstructed only in the fringe field of SM1, keeping only the tracks with good momentum resolution.

The fractional energy of the hadron z is required to be: $z \in [0.2, 0.85]$ (fig. 5.4). The lower cut avoids the contamination from target remnant fragmentation (section 1.6.2), while the upper cut excludes muons wrongly identified as hadrons and rejects regions with high diffractive contributions. The diffractive contributions will be discussed further on in section 6.2.

As the multiplicities for charged identified hadrons (π and K) will be extracted, additional cuts linked to the RICH detector are applied. In particular, the hadrons which travel near the RICH pipe are excluded since the RICH efficiency in this zone is not well determined (see section 4.1) with eq. (4.4). To finalize the hadron selection, cuts on the momentum p_h and on the θ angle (track polar angle at the RICH entrance with respect to the beam direction) are applied. The momentum cut restricts the p_h range common

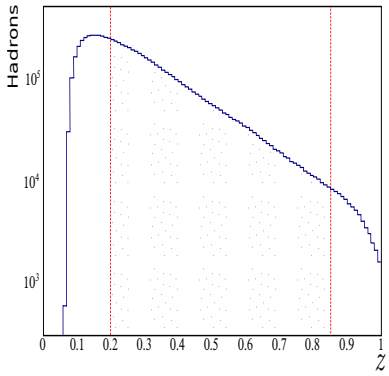


Figure 5.4: *Fractional energy of the hadron z distribution.*

to pion, kaon and proton optimal identification . The angular cut is applied to take into account the RICH acceptance. These cuts reject about half of the hadron sample, leaving $\sim 4 \times 10^6$ of h^+ and h^-. More detail about these cuts is given in the following. The impact of all these cuts is summarized in table 5.3. In total 3.5×10^6 of yet *unidentified* hadrons h^+ and h^- are kept for the 1.3×10^7 DIS events.

Table 5.3: *Impact of cuts on the hadron sample statistics after all cuts.*

	# events	Fraction (%)
Outgoing tracks in preselection	156334512	100.00
no μ' (preselected sample)	111578922	73.92
Hadron candidates (within DIS sample section 5.1.2)	39835022	100.00
XX0<15	39718776	99.71
$z \in [0.2, 0.85]$	7836554	19.68
RICH pipe cut	7466469	18.74
$p_h \in [12, 40](\mathrm{GeV}/c), \theta \in [0.01, 0.12]$	3580470	8.98

Pion, kaon and proton identification

In order to work where the RICH identification efficiencies are high and well determined and where the sample contamination due to misidentification efficiencies is low, the hadron phase-space (p_h, θ) is restricted. The first constrain in the p_h hadron momentum range is determined by the RICH detection threshold: pions ~ 2.67 GeV/c, kaons ~ 9.45 GeV/c and protons ~ 17.95 GeV/c. Since the principal goal of this analysis is the kaon multiplicity extraction, the p_h momentum range is restricted to $12 - 40$ GeV/c (p_h range common to pion, kaon and proton identification). The upper limit in p_h range is set to 40 GeV/c, in order to avoid the region where the identification efficiencies are not well

determined due to the Cherenkov effect saturation (fig. 3.6). The angular range is restricted to $\theta \in [0.01, 0.12]$ for pions, kaons and protons. The lower cut $\theta > 0.01$ rejects the zone where the RICH response is degraded by the photons produced by the muons of the beam far halo (Appendix C.2). The upper cut $\theta < 0.12$ was set to be consistent with the previous multiplicity analysis [76] and with the range where RICH performance matrices were extracted (section 4.3).

The particle identification (PID) is done using the RICH detector information. The identification procedure is described in chapter 4 and the PID selection criteria for π, K and p are summarized in table 4.8.

The unidentified hadron, pion and kaon phase space distributions (not final) are shown in figs. 5.5 and 5.6. The final pion sample constitutes $\sim 74\%$ of the total *unidentified* hadron sample, the kaon sample $\sim 14\%$ and the proton one $\sim 8\%$. In average, this corresponds to ~ 0.3 (0.24π, $0.05K$ and 0.03) hadron per DIS event. The final statistics used in the analysis is summarized in table 5.4. These values still need to be corrected by the RICH detection performance. For this the unfolding method introduced in section 4.4 is applied.

Table 5.4: *Number of* unidentified *hadrons, pions, kaons and protons for the 6 periods analyzed.*

DIS	h^+	h^-	π^+	π^-	K^+	K^-	p	$\overline{p}$
13033506	1980932	1599538	1393318	1241431	367710	236466	178157	87986

5.3 Binning

The raw multiplicities (eq. (5.1)) are extracted in a x, y and z binning.

$$M^h_{raw}(x, y, z) = \frac{N^h(x, y, z)}{N^{DIS}(x, y)}. \tag{5.1}$$

The choice of x and z variables is natural because the PDFs depend upon x and the fragmentation functions upon z. The y variable was chosen instead of Q^2 because of the large correlation between x and Q^2. (fig. 5.3d). In addition, y is more adapted to the evaluation of the acceptance via Monte Carlo data (section 6.1.4).

5.4 QED Radiative Effects

The DIS kinematic variables are based on the single photon exchange process (only first order QED process). However, higher order QED processes are also involved, such as the

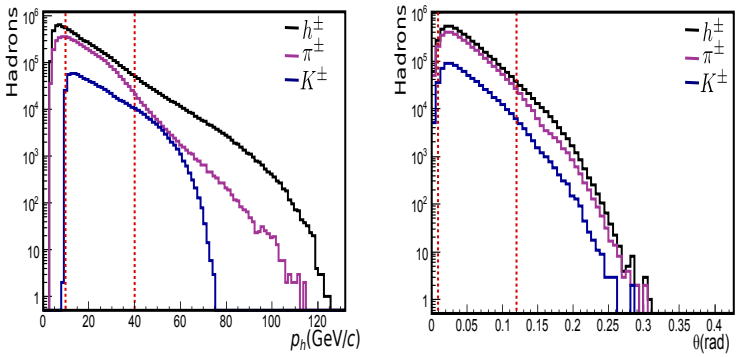


Figure 5.5: *Left: p_h distribution for unidentified hadrons, pions and kaons. Right: θ distribution. The red dotted lines indicate: $\theta \in [0.01, 0.12]$, $p_h \in [12 - 40]$. (Not final selection)*

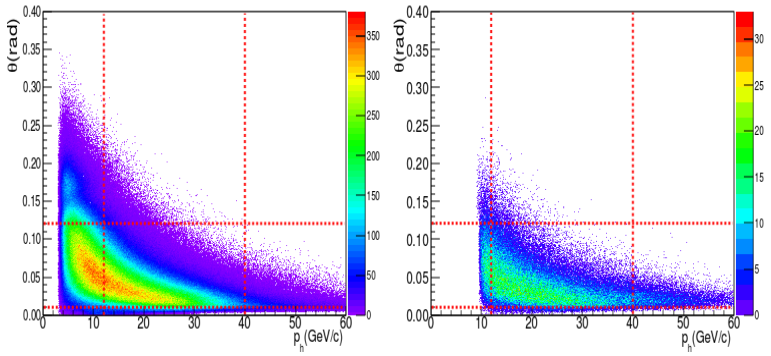


Figure 5.6: *Correlation between the polar angle θ and p_h for the pion (left) and kaon(right). The red dotted lines delimit the phase-space for this analysis.*

Table 5.5: *Multidimensional binning for the multiplicity extraction.*

Variable	Binning
x	0.004, 0.01, 0.02, 0.03, 0.04, 0.06, 0.1, 0.15, 0.7
y	0.10, 0.15, 0.20, 0.30, 0.50, 0.70, 0.90
z	0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.85

initial and final state radiations (fig. 5.7), vertex corrections and vacuum polarization [78]. These corrections are taken into account via the radiative correction factor defined as:

$$\eta(x,y) = \frac{\sigma_{1\gamma}(x,y)}{\sigma_{measured}(x,y)} \tag{5.2}$$

where $\sigma_{1\gamma}(x,y)$ is the cross section defined for the single virtual photon exchange and $\sigma_{measured}(x,y)$ is the observed cross section. It can be written as:

$$\sigma_{measured} = \sigma_{1\gamma} + \sigma_{initial\ state\ radiation} + \sigma_{final\ state\ radiation} + ... \tag{5.3}$$

The radiative correction factor $\eta(x,y)$ is taken from the precalculated tables given in [78] for a 160 GeV muon beam and the COMPASS target. The correction values are different for $N^h(x,y,z)$ and $N^{DIS}(x,y)$. If only μ and μ' are detected ($N^{DIS}(x,y)$) and the scattering is accompanied by an energetic bremsstrahlung photon (not measured in the experiment), an scattering event can be mistaken for a DIS event. This does not happen for a SIDIS event.

For DIS events ($N^{DIS}(x,y)$) the region more affected by the radiative corrections is comprised between $y > 0.7$ and $y < 0.9$ and $x < 0.02$ where the correction value is above 20%. At high x, for y between 0.10 and 0.15, the correction factor is negligible. For the remaining domain, the correction varies between 3% and 18% . The correction factor for $N^h(x,y,z)$ varies between 2% and 15%. Values are plotted in Appendix E individually for unidentified hadrons, pions and kaons and given in table E.1 for DIS events.

(a) *Initial state radiation.* **(b)** *Final state radiation.*

Figure 5.7: *Higher order QED contributions involved in a DIS process.*

5.5 Results for Raw Multiplicities ($h^{\pm}$, $\pi^{\pm}$ and $K^{\pm}$)

The raw multiplicity results shown in this section are corrected for the QED radiative effects. For the pion and kaon multiplicities the RICH unfolding correction is applied. The final statistics after the radiative corrections and the RICH unfolding is summarized in table 5.6.

Table 5.6: *DIS and* unidentified *hadrons statistics after RC corrections (section 5.4). Number of pions, kaons and protons after RC and RICH unfolding correction (section 4.4).*

DIS							
11805536.30							
h^+	h^-	π^+	π^-	K^+	K^-	p	$\overline{p}$
1810061.8	1457011.0	1315491.5	1170296.3	339646.5	215745.2	148567.4	65305.8

The unidentified hadron, the pion and the kaon multiplicities are shown as a function of z in different x and y bins ($12z \times 8x \times 7y$) in figs. 5.8 to 5.10. Each column corresponds to a different x bin and each row to a y bin. The gray shaded area correspond to the z bins which are excluded due to the momentum extrapolation section 6.1.6.

The statistical error of eq. (5.1) is extracted via the error propagation equation [79]:

$$\sigma^2_{M^h_{raw}} = \left(\frac{\partial M^h_{raw}}{\partial N^h}\sigma_{N^h}\right)^2 + \left(\frac{\partial M^h_{raw}}{\partial N^{DIS}}\sigma_{N^{DIS}}\right)^2 + 2\sigma_{N^h N^{DIS}}\frac{\partial M^h_{raw}}{\partial N^h}\frac{\partial M^h_{raw}}{\partial N^{DIS}} \tag{5.4}$$

where the correlation between N^h and N^{DIS} is taken as $\sigma_{N^h N^{DIS}} = \sigma^2_{N^h}\sigma^2_{N^{DIS}}$.

The charged hadron multiplicities strongly depend upon z as expected while a slow increase in the values is observed as a function of y. The multiplicities shown a weak dependence on x. In the low y region (4 first rows where $y < 0.3$) the raw multiplicity shape decreases for small z values. This change can be understood because of the p_h domain restriction ($12 - 40$ GeV/c).

The h^+,h^-, π^+ and π^- multiplicities are very similar. However, a difference between the positive and negative charged hadron multiplicity values is observed for high x values. For $M^{h^\pm}$ and $M^{\pi^\pm}$ the difference can be explained by the dominance of u valence quarks in the high x region. In the kaon case, $M^{K^+}_{raw} > M^{K^-}_{raw}$ this behavior reflects the fact that $K^-(\bar{u}s)$ cannot be produced from the nucleon valence quarks (uud or udd).

In total, there are 367 (367) M^{h^+} (M^{h^-}) points, 365 (364) M^{π^+} (M^{π^-}) points and 359 (352) M^{K^+} (M^{K^-}) points.

5.6 Systematic Studies

The precise estimation of the systematic uncertainties on the multiplicities is crucial because of the small size of the statistical uncertainties. For the raw multiplicities, two sources of systematic uncertainties were studied: the first one regards the compatibility in time of the experimental data. The second one is related to the particle identification and its corresponding unfolding to correct for RICH inefficiencies.

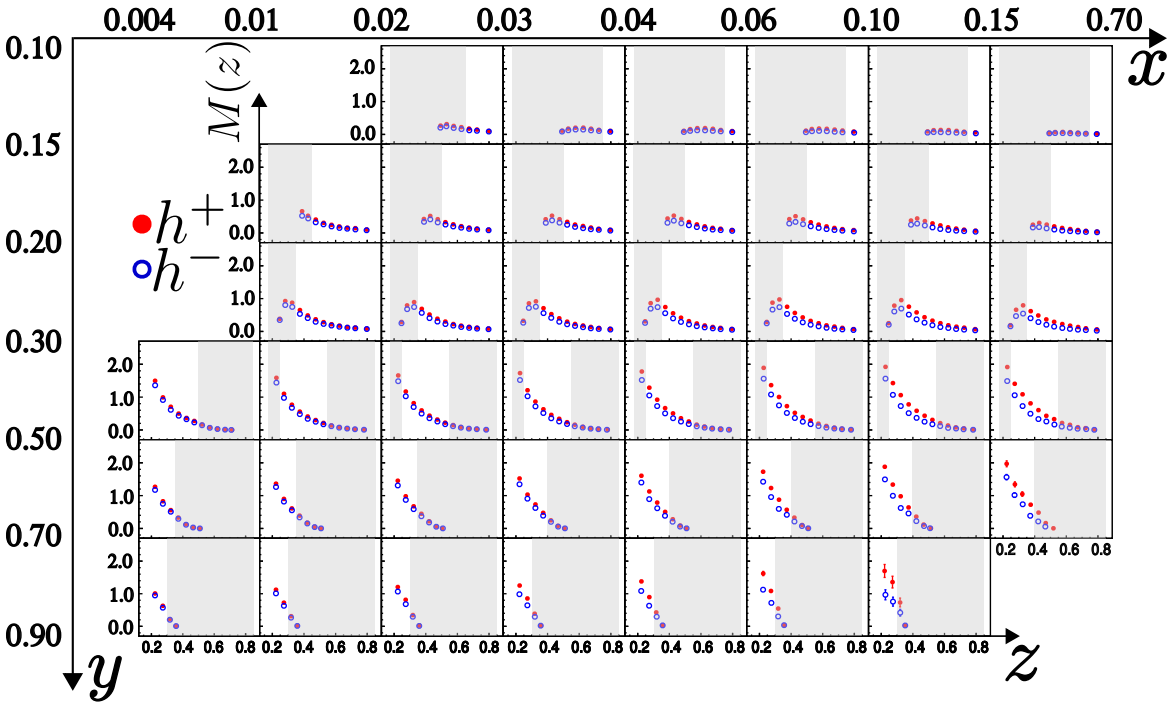


Figure 5.8: *Unidentified hadron raw multiplicities as a function of z in different xy bins. h^+ (red solid circles) and h^- (blue opne circles). Each column corresponds to a different x bin (8 in total) and each row to a y bin (6 in total). The gray shaded area indicates bins which are excluded from final analysis section 6.1.6. Statistical error is shown but is smaller than the points in most of the cases.*

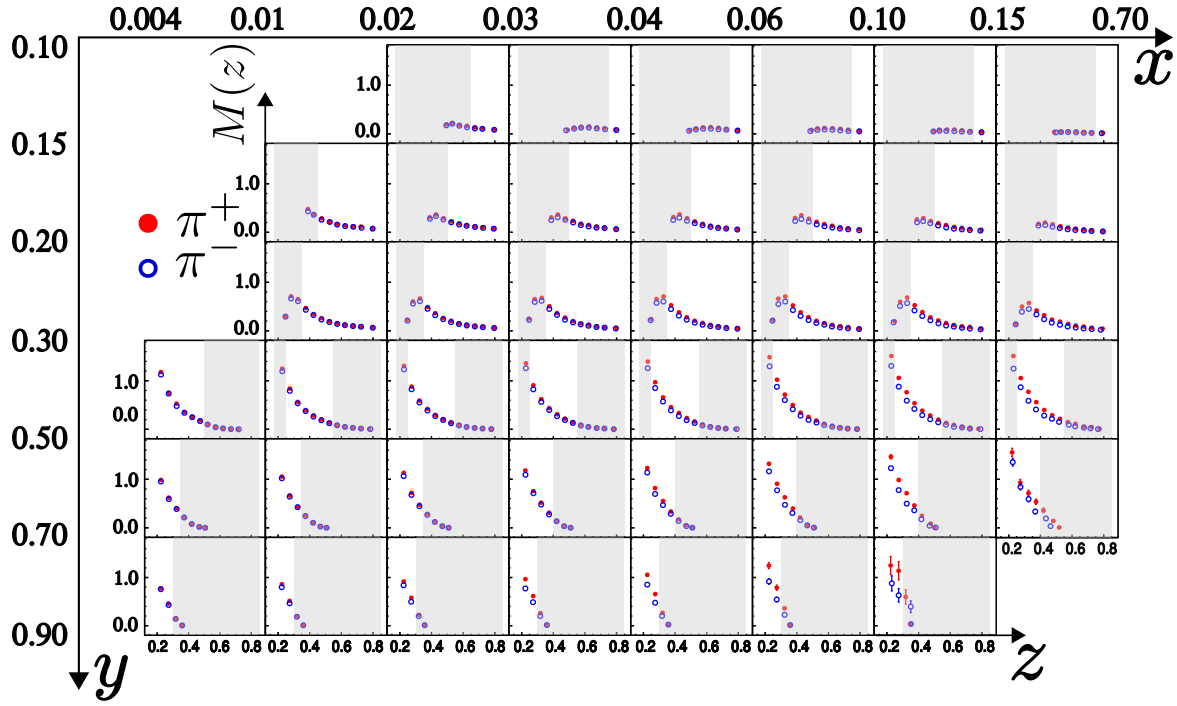


Figure 5.9: *Same as fig. 5.8 for π^+ and π^-.*

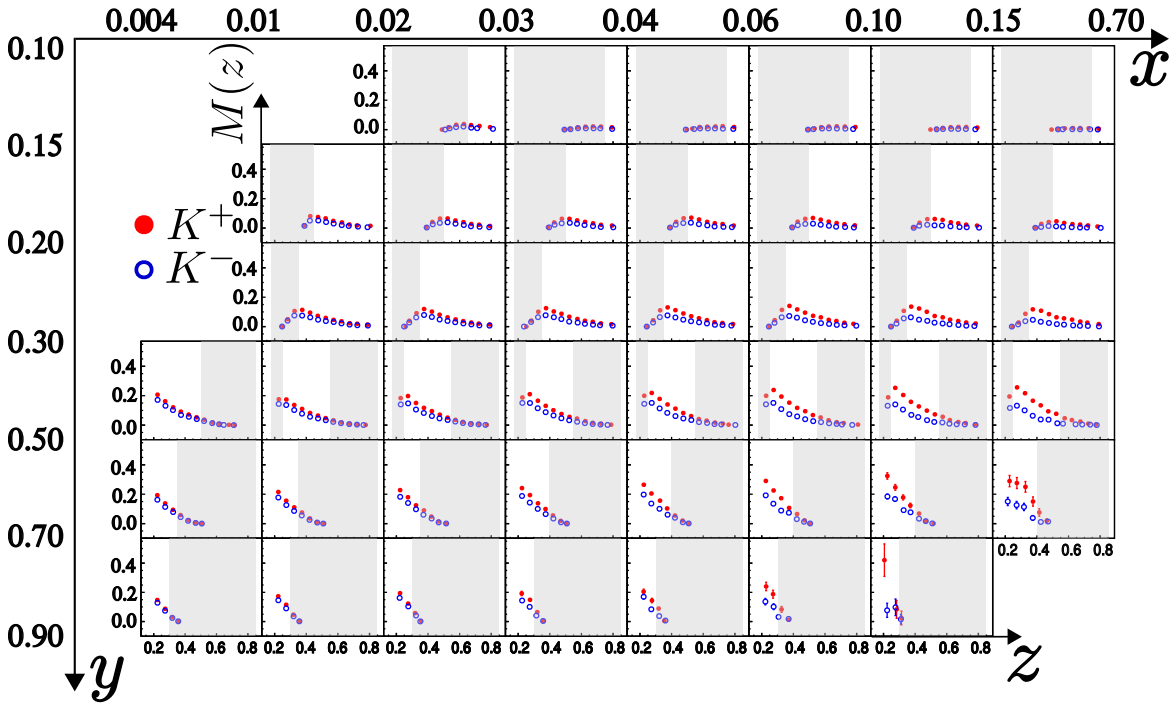


Figure 5.10: *Same as fig. 5.8 for K^+ and K^-.*

5.6.1 Stability Over Time

The SIDIS sample used in the present analysis consists of data taken during six different periods of about one week each. Since the experimental conditions may have slightly varied with time a study of the compatibility of the results from the various periods was performed.

The ratio between the extracted raw multiplicities in the different periods and their average is fitted with a constant in the same kinematics bins (x, y, z) defined in section 5.3. An example of this fit is shown for unidentified hadrons, pions and kaons in figs. 5.11 to 5.13 respectively for the period *W41*. Similar results were found for the other periods. The points within the gray region are not taken into account for the fit because they are later on rejected (see section 6.1.6).

The compatibility between the results of six periods was found to be compatibly within statistical fluctuations for unidentified hadrons, pions and kaons. Consequently, no systematic error is assigned for the data time compatibility.

5.6.2 Systematics Related to the Use of the RICH

The first stage of pion and kaon identification is based on the selection of the cuts for the likelihood ratios: $LH(h)/LH(\text{2nd})$ and $LH(h)/LH(\text{bg})$. These cuts were optimized to minimize the pion and proton misidentification as kaons. The systematic error associated

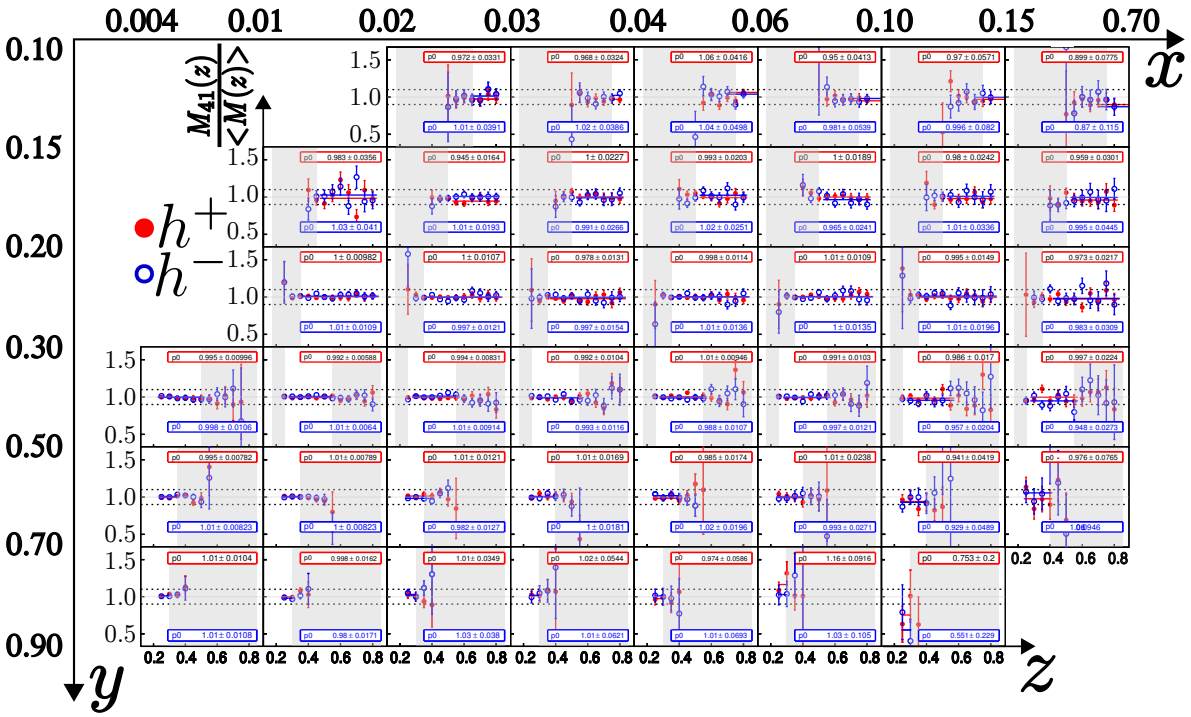


Figure 5.11: *Ratio of W41 raw unidentified hadron multiplicities over the six period average* $< M_{raw} >$*. The red markers correspond to positive hadrons and the blue ones to negative hadrons. Each column corresponds to a different x bin (8 in total) and each row to a y bin (6 in total). Gray zone: same as fig. 5.8.*

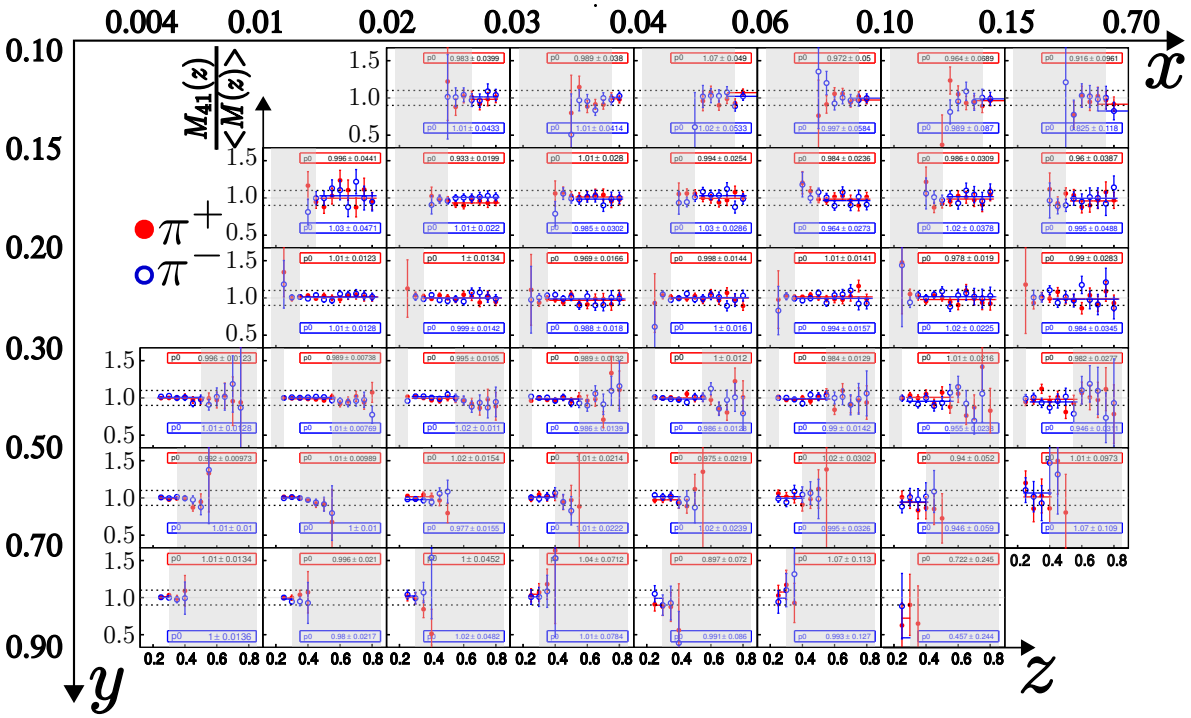


Figure 5.12: *Same as fig. 5.11 for* π^+ *and* π^- *multiplicities.*

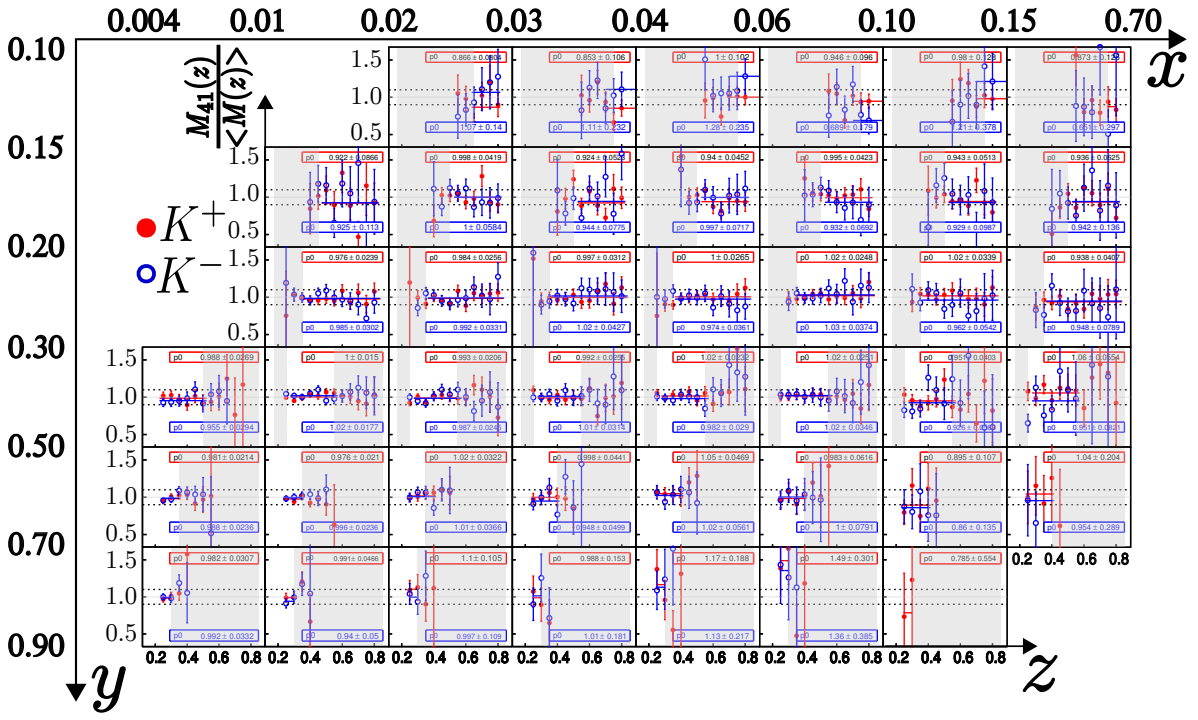


Figure 5.13: *Same as fig. 5.11 for K^+ and K^- multiplicities.*

to the selection of these cuts is performed varying the cuts around the optimized values. Two sets of cuts *loose* and *severe* were used (table 5.7). The difference between the corresponding RICH identification (or misidentification) probabilities and the optimal ones are plotted in fig. 5.14. For pions and protons the largest differences ($< 5\%$) are in the higher momentum p_h region. For kaons the difference reaches 4% at low p_h; small differences ($< 1\%$) are observed at the highest p_h value.

The pion and kaon identification is performed using the loose and severe sets of likelihood cuts and the corresponding multiplicities are extracted ($M_{raw}^{h^\pm\ loose}$ and $M_{raw}^{h^\pm\ severe}$ respectively). The largest difference between $M_{raw}^{h^\pm\ loose}$ and $M_{raw}^{h^\pm\ severe}$ with the nominal multiplicity $M_{raw}^{h^\pm}$ is taken as an estimate of the systematic error (eq. (5.5)).

$$\sigma_{syst}^{RICH\_LH} = MAX(|M_{raw}^{h^\pm\ loose} - M_{raw}^{h}|; |M_{raw}^{h^\pm\ severe} - M_{raw}^{h}|) \tag{5.5}$$

A further study is performed to estimate the impact of the statistical uncertainty. The values of the matrix probabilities are varied by their statistical uncertainties. When the identification probabilities goes up the misidentification values goes down as shown in eqs. (5.6a) to (5.6c). This procedure takes into account the constrain $\sum_{j=\pi,K,p} P(i \to j) < 1$ to avoid unphysical values.

$$P(\pi \to \pi) \pm \sigma_{P(\pi\to\pi)} \Rightarrow \quad P(\pi \to K) \mp \sigma_{P(\pi\to K)}, P(\pi \to p) \mp \sigma_{P(\pi\to p)} \tag{5.6a}$$

$$P(K \to K) \pm \sigma_{P(K\to K)} \Rightarrow \quad P(K \to \pi) \mp \sigma_{P(K\to\pi)}, P(K \to p) \mp \sigma_{P(K\to p)} \tag{5.6b}$$

$$P(p \to p) \pm \sigma_{P(p\to p)} \Rightarrow \quad P(p \to \pi) \mp \sigma_{P(p\to\pi)}, P(p \to K) \mp \sigma_{P(p\to K)} \tag{5.6c}$$

Table 5.7: *Set of loose and severe cuts to evaluate the RICH systematic errors.*

	Loose			Severe		
	π	K	$p(\overline{p})$	π	K	$p(\overline{p})$
$\frac{LH(\pi)}{LH(\mathrm{2nd})}$	>1.00	–	–	>1.06	–	–
$\frac{LH(\pi)}{LH(\mathrm{bg})}$	>2.00	–	<2.3 (2.2)	>2.04	–	<2.0 (1.9)
$\frac{LH(K)}{LH(\mathrm{2nd})}$	–	>1.06	–	–	>1.10	–
$\frac{LH(K)}{LH(\mathrm{bg})}$	–	>2.00	<3.0 (2.9)	–	>2.16	<2.7 (2.6)
$\frac{LH(p)}{LH(\mathrm{2nd})}$	–	–	>0.98	–	–	>1.06

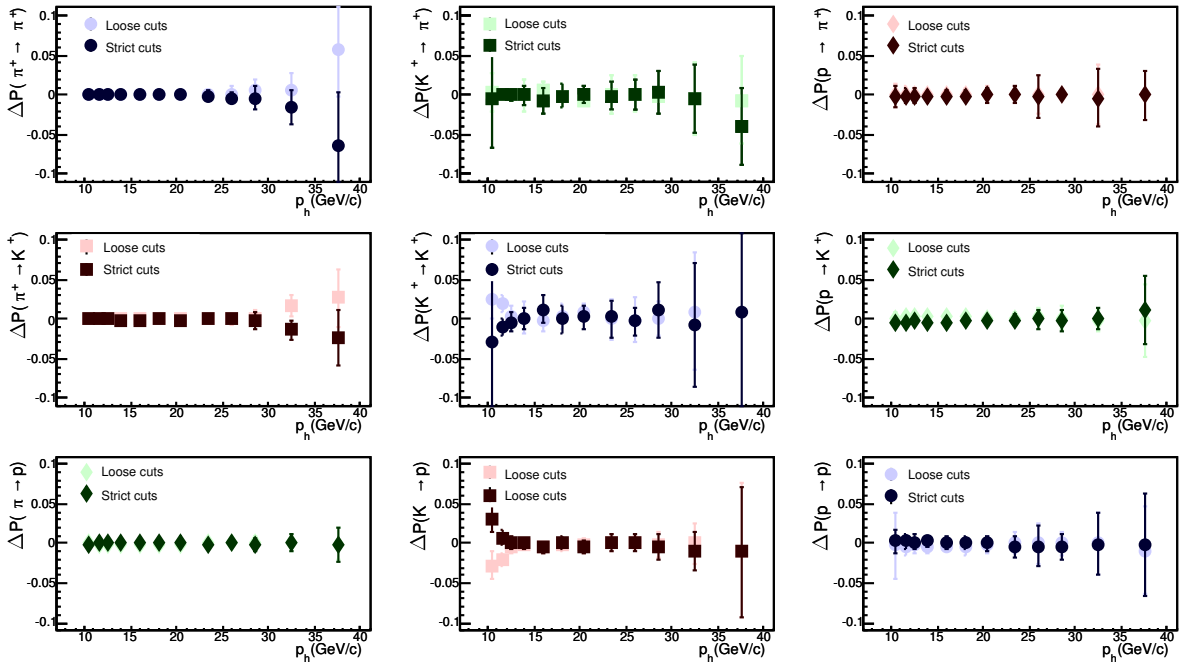


Figure 5.14: *Difference between the identification and misidentification probabilities of* loose *and* severe *cuts with the optimal ones.*

Two sets of RICH probability matrices M_R^+ and M_R^- are generated and the observed number of pions and kaons is corrected using eq. (4.16) with these matrices. The raw multiplicities $M_{raw}^{h^\pm}{}^+$ and $M_{raw}^{h^\pm}{}^-$ are recalculated. The largest difference between $M_{raw}^{h^\pm}{}^+$ and $M_{raw}^{h^\pm}{}^-$ with $M_{raw}^{h^\pm}$ is taken as the systematic error (eq. (5.7)).

$$\sigma_{syst}^{RICH\_stat} = MAX(|M_{raw}^{h^\pm}{}^+ - M_{raw}^{h}|; |M_{raw}^{h^\pm}{}^- - M_{raw}^{h}|) \tag{5.7}$$

The final systematic uncertainty associated to the particle identification and unfolding correction (σ_{syst}^{RICH}) is the largest value between $\sigma_{syst}^{RICH\_stat}$ and $\sigma_{syst}^{RICH\_LH}$. The relative error σ_{syst}^{RICH} is shown in figs. 5.15 and 5.16 for pions and kaons respectively in the same (x,y,z) binning as before. The red color zones indicate the systematic error for the positive hadrons and the blue zones for the negative ones. The gray framed region corresponds to the z bins rejected to avoid a model dependence in the multiplicity final results.

For pions in the region $y < 0.3$ (three upper rows) the systematic error is smaller than $< 0.01\%$ for all x and z bins. For higher y values, the error reaches the maximum ($\sim 2\%$) and only for the high z bins, where the multiplicities are small.

In the kaon case, the largest errors ($\sim 10\%$) are reached for high x (8th column) for large z values where the kaon multiplicities values are small. Elsewhere ther are of the order of below 5%.

5.7 Summary

The raw multiplicities for unidentified hadrons, pions and kaons coming from muon deep inelastic scattering on a deuterium isoscalar target (^{6}LiD) have been extracted in a (x,y,z) binning from the 2006 COMPASS data. The data cover a wide kinematic domain defined by: $Q^2 > 1$ (GeV/c)^2, $y \in [0.1, 0.9]$, $x \in [0.004, 0.7]$ and $W \in [5, 17]$ GeV. The hadron energy fraction z is limited to the [0.2,0.85] region. The hadron momentum range lies between 12 GeV/c and 40 GeV/c.

The raw multiplicities $M_{raw}^{h^\pm}$, $M_{raw}^{\pi^\pm}$ and $M_{raw}^{K^\pm}$ are corrected for radiative effects. The pion and kaon raw multiplicities are also corrected by the RICH identification and misidentification efficiencies. The acceptance correction and other MC corrections are not yet applied on the results showed in this chapter. The statistical error is the dominant one ($\sigma/M_{raw}^{h} < 1\%$).

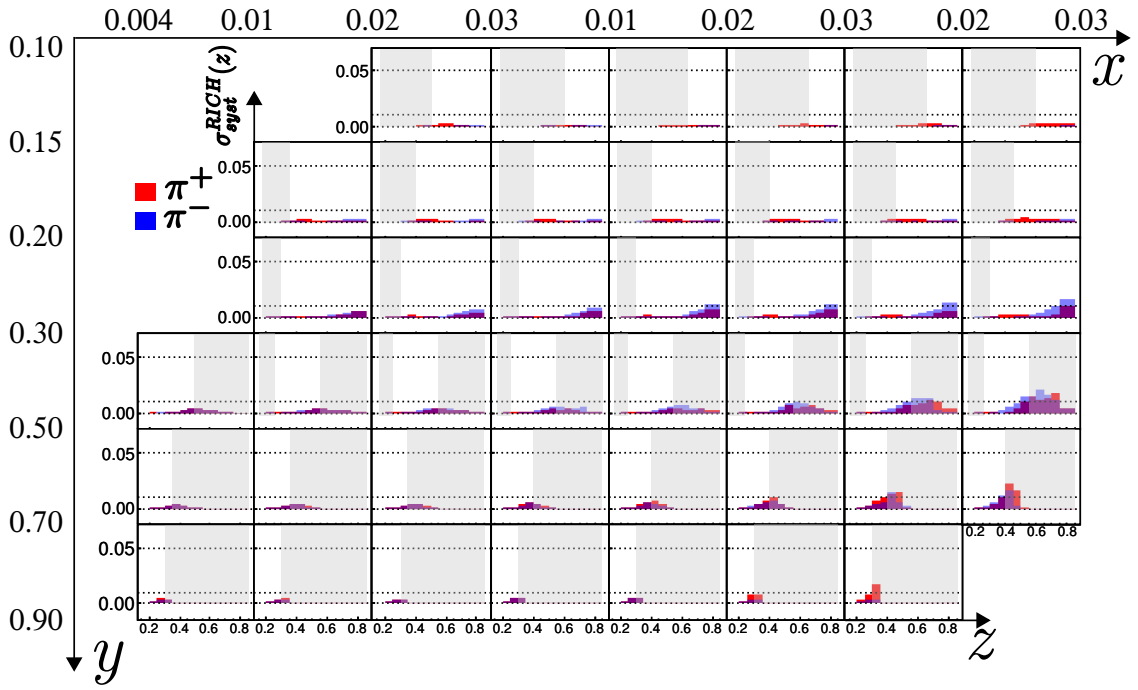


Figure 5.15: *Band associated to the systematic error (relative) due to the RICH unfolding correction in the pion sample. The red band corresponds to π^+ and the blue one to π^-. Gray zone as in fig. 5.8. The horizontal doted lines are placed in 0.01 and 0.05.*

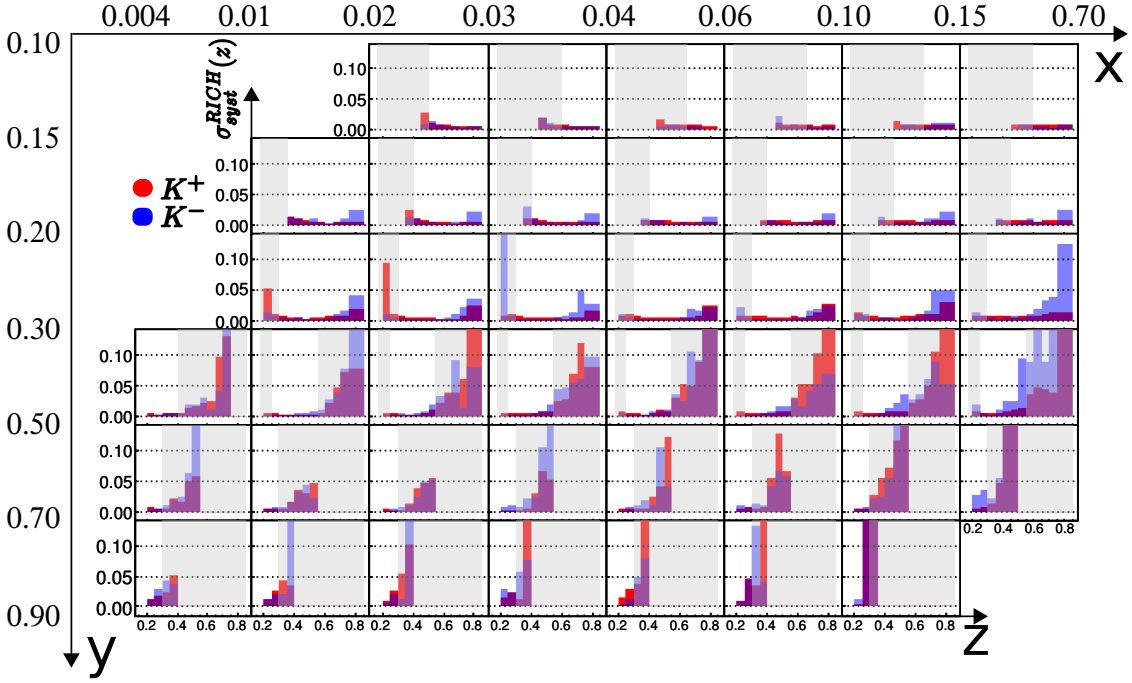


Figure 5.16: *Same as fig. 5.15 for K. The horizontal doted lines are placed in 0.05 and 0.10.*

Chapter 6

Acceptance and Other Monte Carlo Corrections

The *raw* charged hadron multiplicities presented in the previous chapter need to be *corrected* for acceptance and other effects before being used in a QCD global fit . In this chapter it is shown how the corrections were determined within the COMPASS collaboration [80, 81].

We describe the *acceptance* correction *i.e* the correction due to the geometrical limitations of the spectrometer and the data reconstruction efficiency. A brief description of the event generation, the spectrometer simulation and the MC data production is provided. The correction related to the electron contamination is evaluated. The raw multiplicities ($M^{h^\pm}$, $M^{\pi^\pm}$ and $M^{K^\pm}$) extracted in the previous chapter are restricted to the momentum domain 12 to 40 GeV/c because of the RICH particle identification limitations. The extrapolation of the results to the unmeasured range is explained . The data are corrected for the contribution of diffractive meson production.

6.1 Determination of the Spectrometer Acceptance

The acceptance is evaluated from a MC simulation. The first step consists in simulating events from the physics process of interest (SIDIS in this case) with an event generator such as LEPTO [82]. A simulation of the experimental setup has to be done (COMGEANT). The events coming from LEPTO are propagated through the experimental setup simulation. The output of this chain is referred to as *generated* event sample. These events are use as an input of the COMPASS code for event reconstruction and are analyzed in the same way as the real data. This new sample is referred to as *reconstructed* event sample. The final acceptance is defined as the ratio between the *reconstructed* events and the *generated* ones.

6.1.1 SIDIS Event Generation with LEPTO

The LEPTO generator simulates complete events in deep inelastic lepton-nucleon scattering. The lepton chosen for the event generation is μ^+. The muon momentum phase space is recorded during data taking and the generated muon momentum is chosen randomly from this information as well as transverse coordinates of the interaction. The interaction in the ^{6}LiD target is simulated by choosing the interacting nucleon as a proton or a neutron (with equal probability) and with momentum equal to zero.

The parton flavor depends on the parton distribution function (PDF) used to generate the events. For the present analysis the MSTWLO 2008 [83, 84] parametrization is used since its validity kinematic domain ($x \in [10^{-6}, 1]$ and $Q^2 \in [1, 10^9](GeV/c)^2$) comprises the COMPASS one.

The hard scattering processes (muon-parton interaction) are simulated according to the electromagnetic cross section. The fragmentation of the partons in the final state is simulated using the Lund string model [22] (section 1.6.1). The JETSET default values describe the *Large Electron-Positron Collider* (LFP) data; in [77] these values were optimized for the COMPASS data. The simulation also includes the decay of unstable particles based on measured branching ratios.

6.1.2 Monte Carlo Simulation of the Spectrometer

The COMPASS apparatus description *COMGEANT* is based on the simulation tool Geant 3.21 [85]. The different apparatus components such as the target (position and material distribution), the detector geometries (position, dimension and material distribution), the magnet fields intensities as well as the relevant processes of the particles interacting with the materials (energy loss, changes on the trajectory, generation of new particles, etc.) are described.

The COMGEANT output contains the energy deposits of the particle in the detector as well as the hit coordinates, the information on the particle identification and the particle kinematics. The COMGEANT output is given as a input to CORAL the COMPASS code for the event reconstruction, where the energy deposits and hit coordinates are used to simulate the detector response. The CORAL software records the information from the generated and reconstructed events.

6.1.3 Monte Carlo Sample from LEPTO

The MC sample for the determination of the acceptance was produced in an enlarged kinematic domain: $Q^2 > 0.8$ GeV^2, $x \in [10^{-4}, 0.99]$, $y \in [0.05, 0.95]$ and $\nu \in [0, 250]$ GeV. This in order to be able to take into account smearing effects. The total number of events generated is $\sim 9.3 \times 10^7$.For the event reconstruction, the same selection are applied as for

the real data. The total number of reconstructed DIS events is $\sim 1.4 \times 10^7$ (same order as experimental data). The inclusive kinematic variables Q^2, x and y are shown in fig. 6.1 for real (black markers) and reconstructed MC data (red markers). The ratio between real and reconstructed MC data is plotted in blue. A good agreement, $\sim 90\%$ is observed.

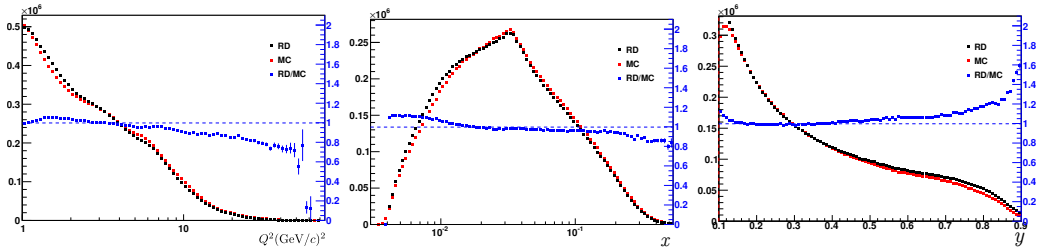


Figure 6.1: *Q^2, x and y distributions of real (black) and reconstructed MC data (red) with the ratio RD/MC in blue markers. Figure taken from [81].*

6.1.4 Acceptance Calculation

The acceptance is calculated from the ratio of the number of *reconstructed* (N_r^h) and the number of *generated* hadrons (N_g^h). The generated hadrons are taken with the kinematical cuts: $z \in [0.2, 0.85]$ and $p_h \in [12, 40]$(GeV/c) while the reconstructed ones must satisfy in addition the requirements defined in section 5.2. For generated and reconstructed hadrons, the hadron identification is taken from the information provided by the generator. The acceptance is defined as:

$$A^h(x_r, y_r, z_r) = \frac{N_r^h(x_r, y_r, z_r)}{N_g^h(x_g, y_g, z_g)} \tag{6.1}$$

For the inclusive variables x and y, the reconstructed values are used to fill the corresponding bin (i.e $x_g = x_r$ and $y_g = y_r$). For the semi-inclusive variable z, the reconstructed value is used for the nominator and the generated value for the denominator. In other words, a generated event is considered only if the corresponding reconstructed DIS event exists. The raw multiplicities are corrected for every (x,y,z) bin by the correction A^h as:

$$M^h_{cor\_acc}(x, y, z) = \frac{M^h_{raw}(x, y, z)}{A^h(x, y, z)}. \tag{6.2}$$

The acceptance correction factors $A^h(x, y, z)$ for unidentified hadrons ($h^\pm$), pions ($\pi^\pm$) and kaons ($K^\pm$) are shown in figs. 6.2 to 6.4. The acceptance is mostly flat in z and shows a dependence on y. For unidentified hadrons, pions and kaons, the acceptance varies between $50\% - 78\%$.

Lately, with an imporoved MC sample, a detailed study on the acceptance were done to check the results obtained versus the longitudinally position of the vertex (z) along the target. It was found that the multiplicities obtained separately for the three targets cells agree well, with better than $\pm 3\%$ [80].

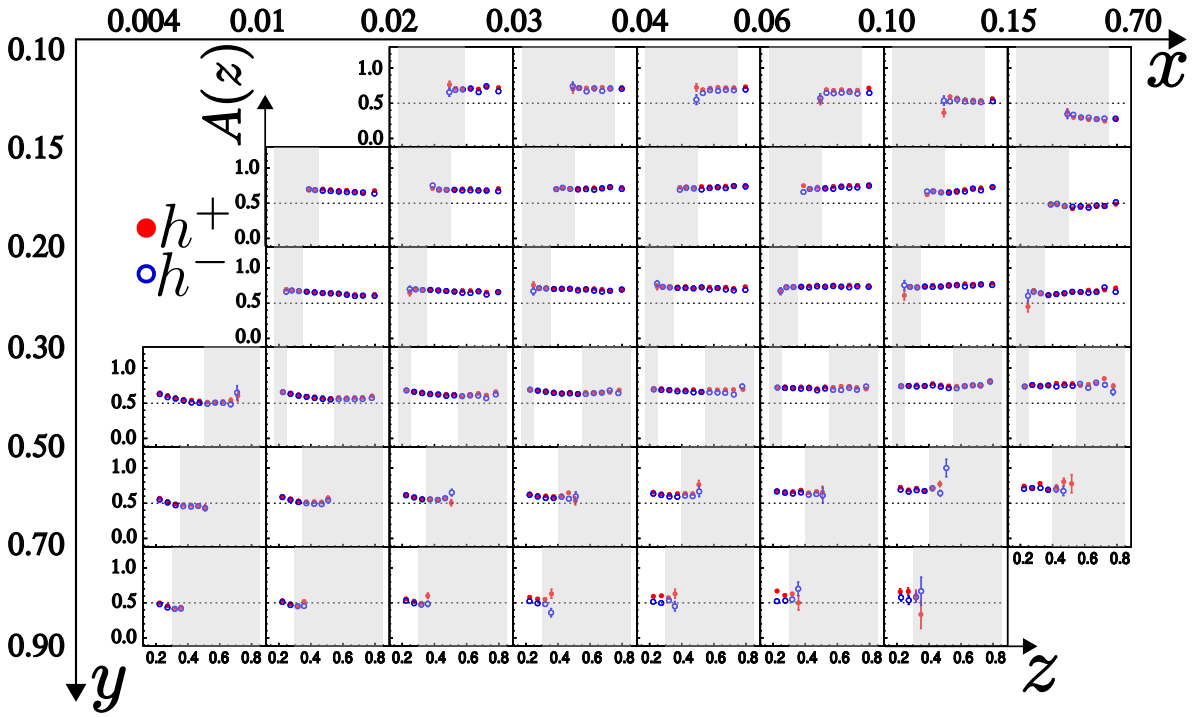


Figure 6.2: *Acceptance correction factor for unidentified hadrons versus z in different xy bins. A^{h^+} in red and A^{h^-} in blue. Each column corresponds to a different x bin (8 in total) and each row to a y bin (6 in total). The gray shaded area indicates bins which are excluded from final analysis (section 6.1.6).*

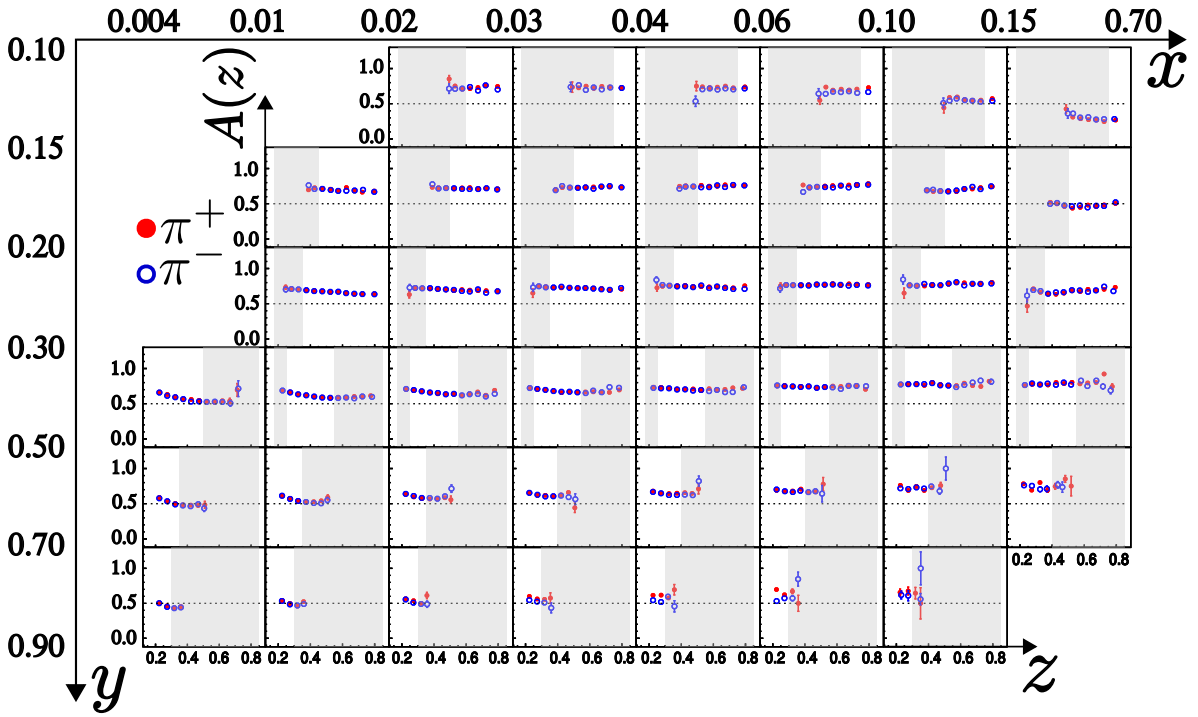


Figure 6.3: *Same as fig. 6.2 for pions.*

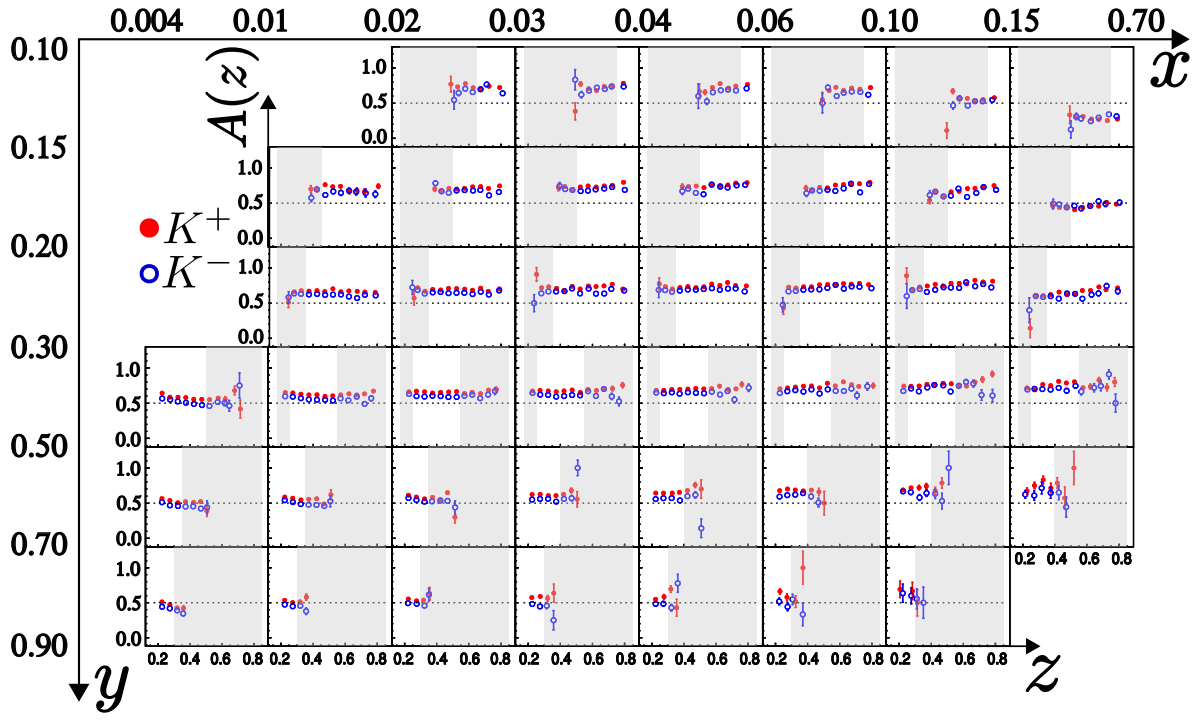


Figure 6.4: *Same as fig. 6.2 for kaons.*

6.1.5 Electron Contamination

A dedicated study was performed to estimate the electron contamination in the hadron sample [80]. Two different momentum regions were studied:

- $3-8\text{GeV}/c$. In this region the RICH detector can discriminate electrons from pions, so that a comparison between real data and MC data is possible.
- Momentum region $12-40\text{GeV}/c$. The RICH detector can no longer discriminate electrons from pions, kaons or protons. The electron contamination is based only on MC.

The MC sample used in this analysis is the one described in section 6.1.3. Only reconstructed events are used. In order to perform the electron identification two sets of likelihood cuts were used: $\frac{LH(e)}{LH(bg)} > 2$ and $\frac{LH(e)}{LH(bg)} > 3$. The fraction of electrons(positrons) to the total sample is about 9.85(8.2)% while in MC it is 8.6(7.3)%. The difference between real data and MC is only $\sim 13\%$. Given the overall small size of this correction it is reasonable to use the MC sample at momenta the electron identification cannot be provided by the RICH detector. The fraction of electrons in the hadron and pion sample in the momentum range between 12 and 40 GeV/c obtained from MC alone is shown in figs. 6.5 and 6.6 respectively (this contamination does not affect the kaon sample) . This contamination decreases from 8% at low z to below 1% at high z (excluding bins in the gray regions). This correction is incorporated into the acceptance correction of section 6.1.4.

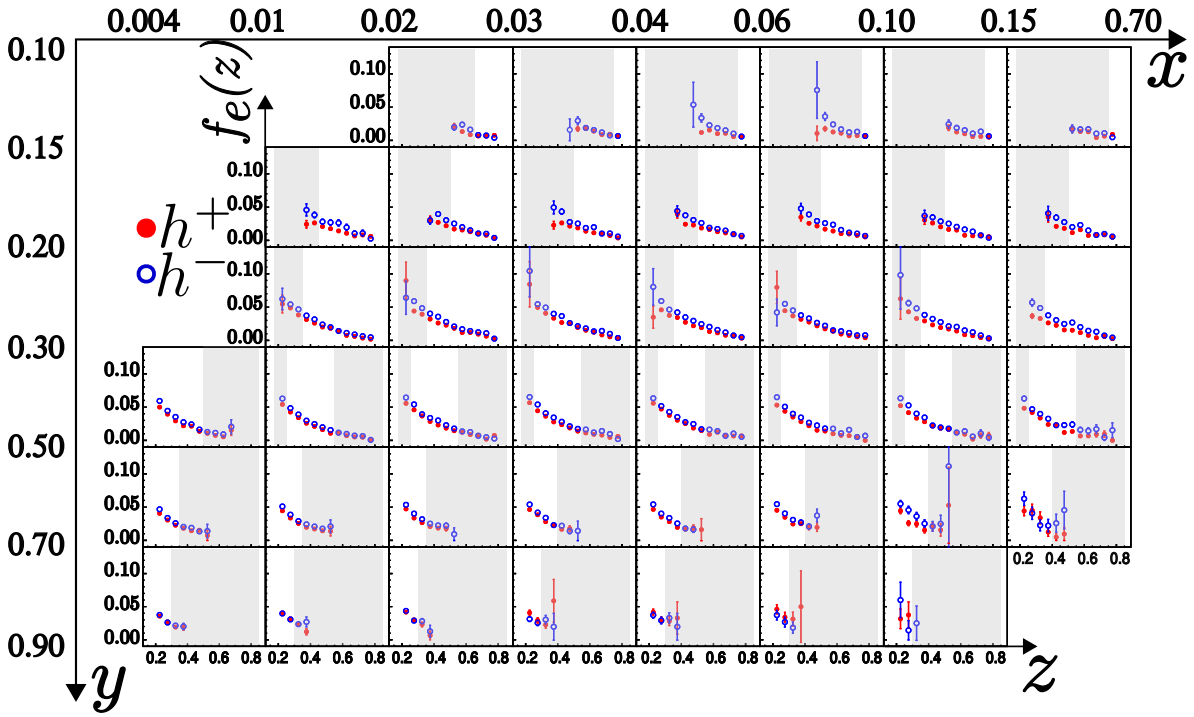


Figure 6.5: *Fraction of electron contamination for unidentified hadron as a function of z in different xy bins. Red markers:h^+. Blue markers:h^-. Each column corresponds to a different x bin (8 in total) and each row to a y bin (6 in total). Gray zone: same as fig. 6.2*

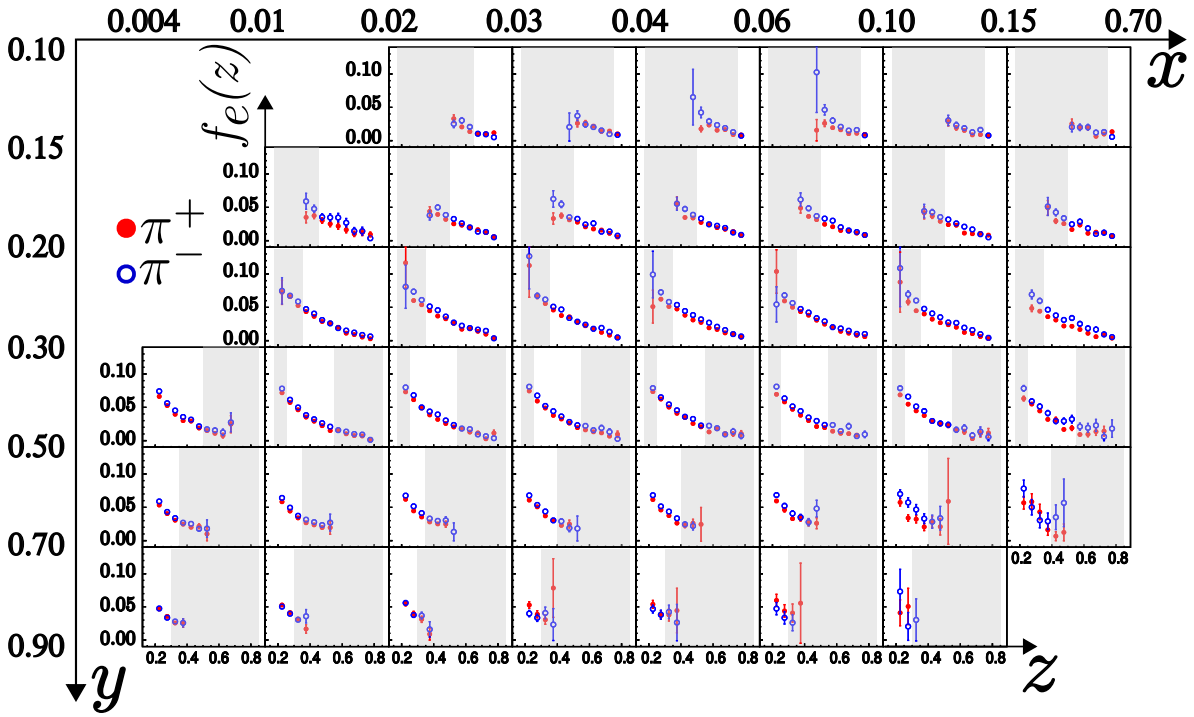


Figure 6.6: *Same as fig. 6.2 for pions.*

6.1.6 Extrapolation to Unmeasured Momentum Range

The multiplicities were extracted in a truncated momentum range ($12-40\mathrm{GeV}/c$) defined by the cut for the particle identification with the RICH detector. In order to be able to use our multiplicities to extract fragmentation functions in a global QCD analysis we must evaluate the impact of this rejection. In a first step, the non-measured range is completed by using the generated sample. Two multiplicities are calculated from this sample, one of them restricted to the same momentum range as the raw multiplicities ($M^h_{LEPTO}(12-40)$) and the other one without restriction (M^h_{LEPTO}). To obtain the extrapolation factor the following difference is used:

$$M^h_{extra}(x_g, y_g, z_g) = M^h_{LEPTO}(x_g, y_g, z_g) - M^h_{LEPTO}(12-40)(x_g, y_g, z_g) \tag{6.3}$$

The multiplicity in the non-covered momentum range M^h_{extra} is added to the measured multiplicity (when all corrections have been applied) in each (x, y, z) bin.

$$f(x, y, z) = \frac{M^h(x, y, z)}{M^h(x, y, z) + M^h_{extra}(x, y, z)} \tag{6.4}$$

The fraction $f(x, y, z)$ is plotted for $h^\pm$, $\pi^\pm$ and $K^\pm$ in figs. 6.7 to 6.9 respectively. The regions where f is below 0.9 are shown as gray shaded regions. They are further discarded from the analysis to limit the model dependence. The same shaded area is used in all other plots.

6.2 Diffractive ρ^0 and ϕ Production

The scattering of a lepton off a nucleon can produce DIS but also diffractive vector mesons as shown in fig. 6.10. These are unstable particles which decay into pions and kaons that cannot be distinguished from the ones resulting from the fragmentation of a quark belonging to the nucleon. However, in a MC simulation it is possible to differentiate the π and K origin and thus estimate the correction factor . The full description of this study can be found in [81].

The dominant channels at COMPASS energies are [87]:

$$\gamma^* N \to \rho^0 N \to N\pi^+\pi^- \tag{6.5}$$

$$\gamma^* N \to \phi^0 N \to N K^+ K^- \tag{6.6}$$

6.2.1 MC simulated with HEPGEN

The HEPGEN[1] generator [88] simulates the ρ^0 and ϕ diffractive vector meson production based on the ν, Q^2 and p_T^2 distributions measured in NMC [89]. The cross section is

[1] Hard Exclusive Production GENerator

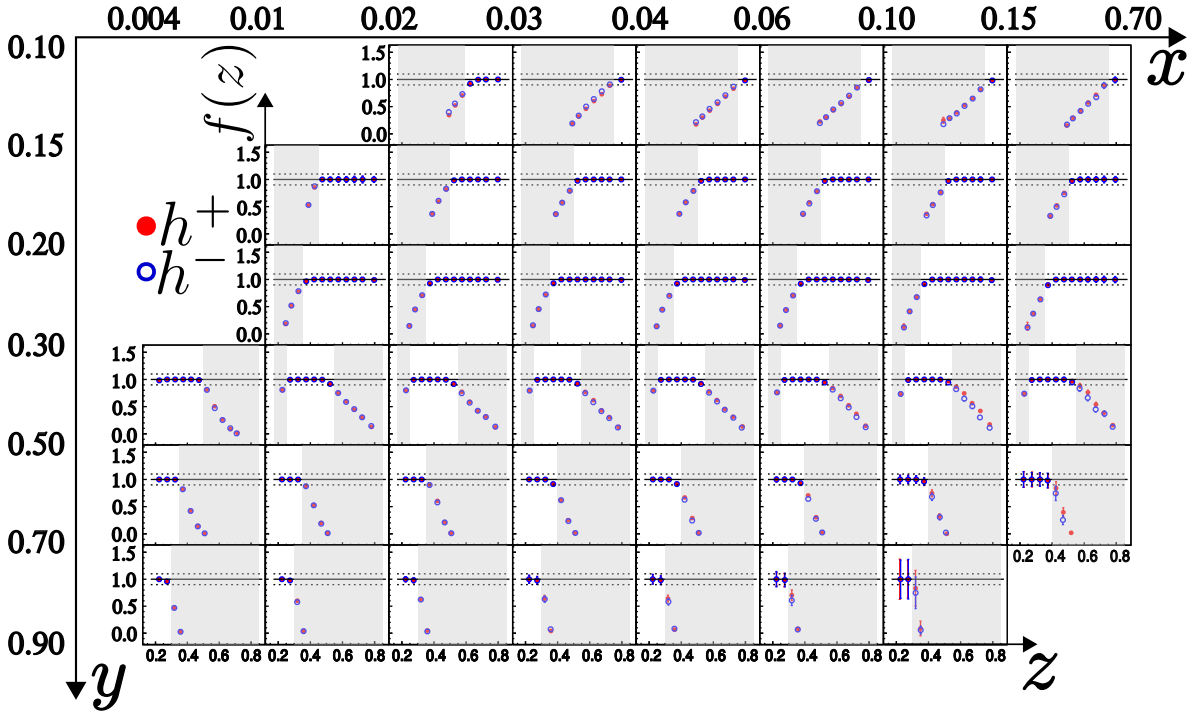


Figure 6.7: *Ratio between the measured multiplicities (after all corrections are applied) with and without the extrapolation in the non-covered momentum. The points inside the gray zone have a LEPTO contribution* $> 10\%$ *and they are excluded from final analysis.*

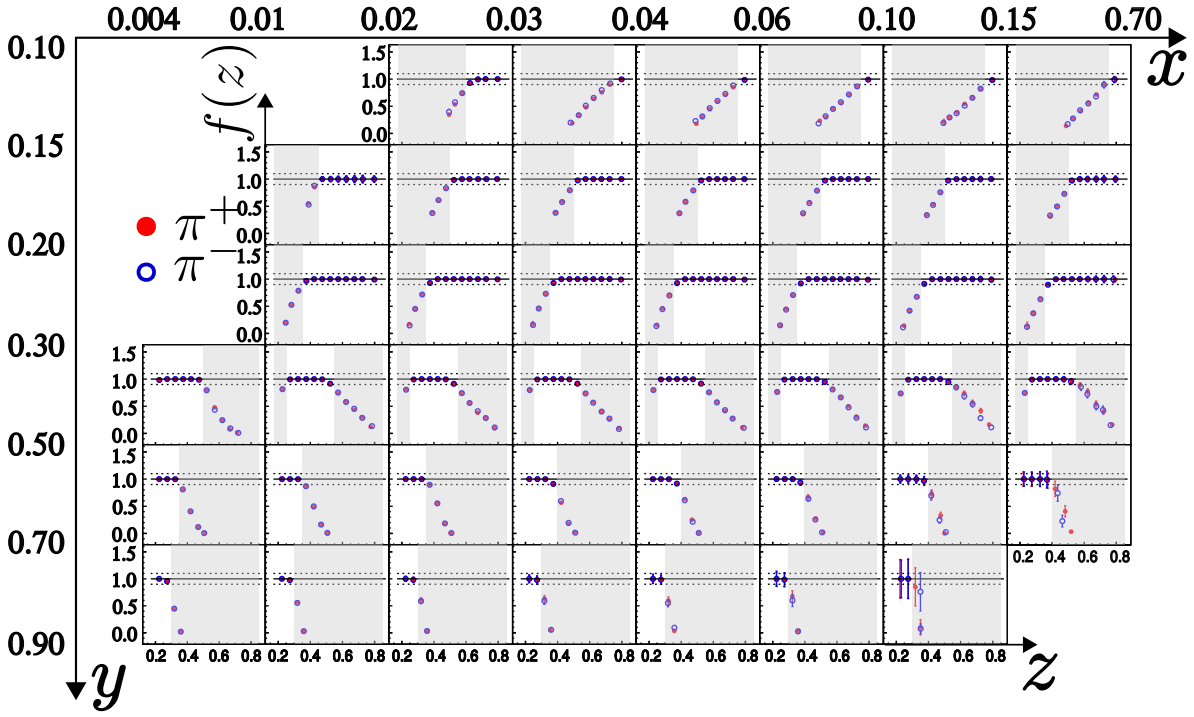


Figure 6.8: *Same as fig. 6.7 for pions.*

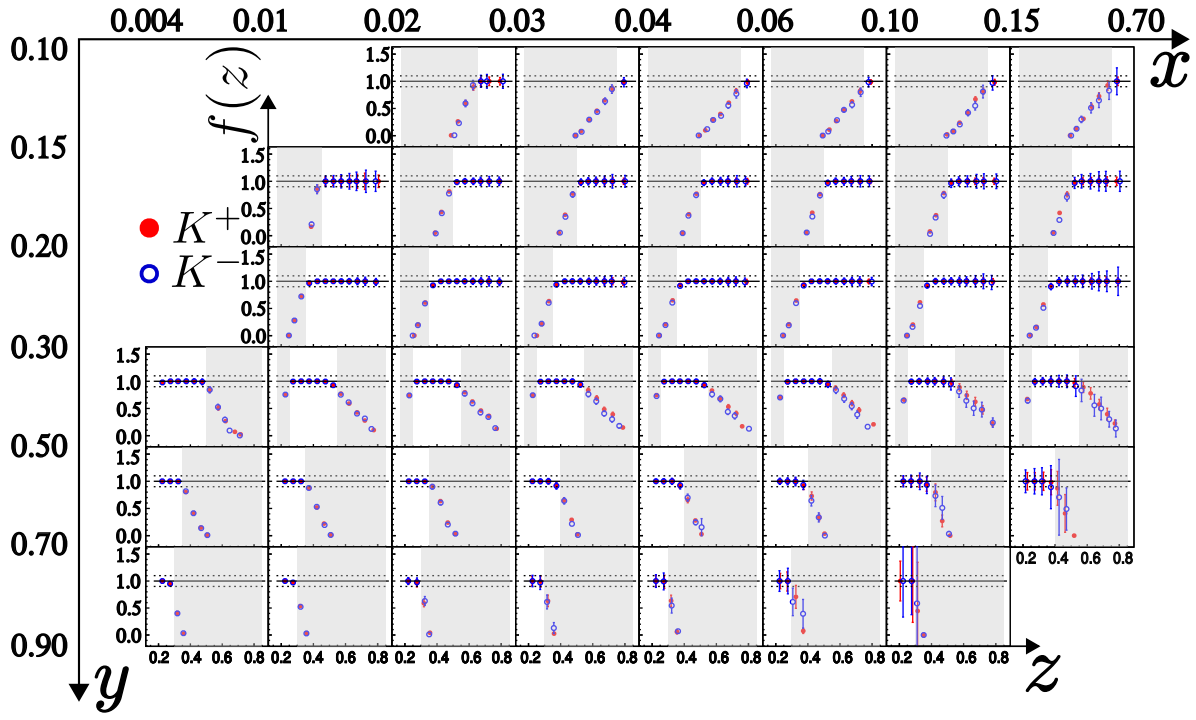


Figure 6.9: *Same as fig. 6.7 for kaons.*

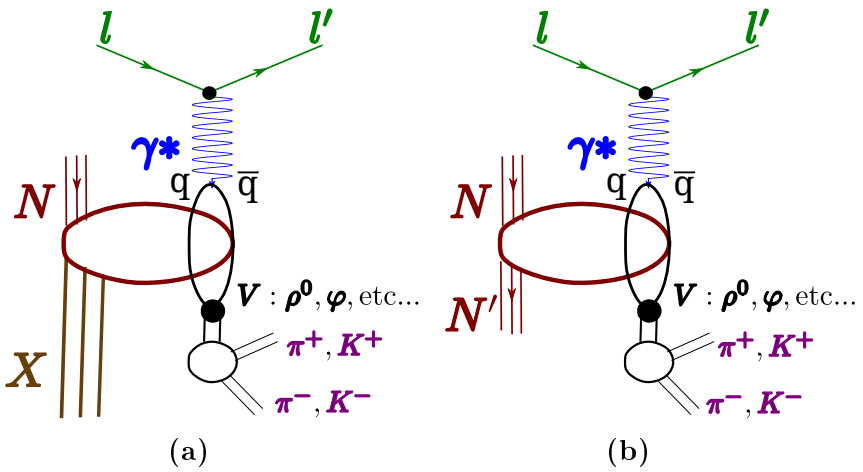


Figure 6.10: *Vector-meson (V) diffractive production. In the VMD model [86], the $\gamma*$ virtual photon fluctuates in a $q\bar{q}$ pair with compatible quantum numbers. (a) Vector-meson exclusive production (the same nucleon N is found in the final state). (b) Vector-meson productions with nucleon diffractive dissociation.*

normalized by the Kroll and Goloskokov GPD model [90]. The incident muon momentum and coordinates are selected in the same way as for LEPTO and the target is equally simulated. For the calculation of the cross sections the ν and Q^2 variables are chosen randomly in the phase-space. A weight proportional to the cross section is associated to the event. All hadrons coming from the ρ^0 and ϕ decay and all particles produced (mostly

pions) by diffractive dissociation of the target nucleon ($\sim 10\%$ of the cases) are kept .

As previously, the MC chain HEPGEN + COMGEANT + CORAL is used to produce the HEPGEN *generated* and *reconstructed* events.

The MC sample was produced in the kinematic domain: $Q^2 \in [0.5, 80](\mathrm{GeV}/c)^2$ and $\nu \in [5, 155](\mathrm{GeV})$. The total number of events generated is $\sim 6.1\times10^6$ and $\sim 2.2\times10^6$ for ρ^0 and ϕ respectively.The total number of weighted events in the ρ^0 and ϕ DIS reconstructed sample are $\sim 1.7 \times 10^5$ and $\sim 5.3 \times 10^4$.

6.2.2 Diffractive Vector Meson Correction

The fraction of pions (kaons) resulting from a diffractive ρ^0 (ϕ) is calculated in the same binning as the raw multiplicities as:

$$f^{\pi}_{\rho^0}(x,y,z) = \frac{N^{\pi}_{HEPGEN}(x,y,z)}{N^{\pi}_{LEPTO}(x,y,z) + N^{\pi}_{HEPGEN}(x,y,z)} \tag{6.7}$$

$$f^{K}_{\phi^0}(x,y,z) = \frac{N^{K}_{HEPGEN}(x,y,z)}{N^{K}_{LEPTO}(x,y,z) + N^{K}_{HEPGEN}(x,y,z)} \tag{6.8}$$

where N^{π}_{HEPGEN}, N^{π}_{LEPTO}, N^{K}_{HEPGEN} and N^{K}_{LEPTO} are the number of pions and kaons reconstructed from the HEPGEN and LEPTO MC samples normalized by the corresponding MC luminosity (L_{MC}). The luminosity defined depends on the event weighting and the process cross-section σ_{int} (DIS for LEPTO events and diffractive vector meson production for HEPGEN events) eq. (6.9). The final weigthed number of unidentified hadrons, pions and kaons are summarized in table 6.1.

$$\sum_{events} w_i = L_{MC} \cdot \sigma_{int} \tag{6.9}$$

Table 6.1: *Weigthed number of unidentified hadrons, pions and kaons used for the diffractive vector meson correction.*

Generator	N^{h^+}	N^{h^-}	N^{π^+}	N^{π^-}	N^{K^+}	N^{K^-}
LEPTO	65 168	54 922.1	4 184	3 670.3	906.9	646.3
HEPGEN	710.9	682.2	157.5	153.3	27.2	26.1

The diffractive ρ^0 and ϕ can also lead to a contamination in DIS events (eqs. (6.10) and (6.11)), with the fraction of the contamination expressed in eqs. (6.10) and (6.11). In contrast to eqs. (6.7) and (6.8) the denominator only includes the DIS events from the LEPTO generator because the cross section used to generate the LEPTO sample already takes into account the diffractive contribution.

$$f^{\rho^0}_{DIS}(x,y) = \frac{N^{HEPGEN}_{\rho^0}(x,y)}{N^{DIS}_{LEPTO}(x,y)} \tag{6.10}$$

$$f^{\phi}_{DIS}(x,y) = \frac{N^{HEPGEN}_{\phi}(x,y)}{N^{DIS}_{LEPTO}(x,y)} \tag{6.11}$$

The total contribution from the diffractive vector-meson contribution (f^{VM}_{DIS}) to the DIS sample is the sum of $f^{\rho^0}_{DIS}$ and f^{ϕ}_{DIS}. The final correction reads as follows:

$$\begin{aligned} B^h(x,y,z) = &\frac{N^{\pi}(x,y,z)}{N^h(x,y,z)} B^{\pi}(x,y,z) + \frac{N^{K}(x,y,z)}{N^h(x,y,z)} B^{K}(x,y,z) \\ &+ \frac{N^{p}(x,y,z)}{N^h(x,y,z)} \frac{1}{1 - f^{VM}_{DIS}(x,y)} \end{aligned} \tag{6.12}$$

$$B^{\pi}(x,y,z) = \frac{1 - f^{\pi}_{\rho^0}(x,y,z)}{1 - f^{VM}_{DIS}(x,y)} \tag{6.13}$$

$$B^{K}(x,y,z) = \frac{1 - f^{K}_{\phi^0}(x,y,z)}{1 - f^{VM}_{DIS}(x,y)} \tag{6.14}$$

Results for B^h are shown in bins of (x, y, z) in figs. 6.11 to 6.13. For pions, the correction is in general small ~ 1, but can reach high values $\sim 40\%$ in some high z bins where the multiplicity values are small.

6.3 Systematic Uncertainties Associated to the Monte Carlo Simulations

The various corrections described previously (acceptance, vector-meson production, etc.) constitute a source of systematic uncertainties.

6.3.1 Monte Carlo Sample: LEPTO Dependence

To determine the acceptance dependence upon the physical model chosen, different PDF parameterizations were used to generate different MC sample. In addition, different parameters were used in JETSET [76, 77]. For each sample, the acceptance is recalculated and the hadron multiplicities are corrected by this new acceptance A'^h. The systematic uncertainty on the acceptance is estimated for each kinematic bin (x,y,z) as:

$$\sigma^{A^h}_{syst} \leq \frac{\frac{A'^h}{A^h} - 1}{2} M^h_{cor\_acc} \tag{6.15}$$

The corresponding value of $\sigma^{A^h}_{syst} \sim 5\%$ found for the systematic error on the acceptance is the largest contribution to the total systematic error on the multiplicities.

The same MC samples are used to study the systematic uncertainty due to the multiplicity extrapolation to the unmeasured momentum range (section 6.1.6). The contribution of this extrapolation to the systematic uncertainties is $0.1 M^h_{extra}$, where M^h_{extra} is the multiplicity in the non-covered momentum range obtained from the generated MC sample.

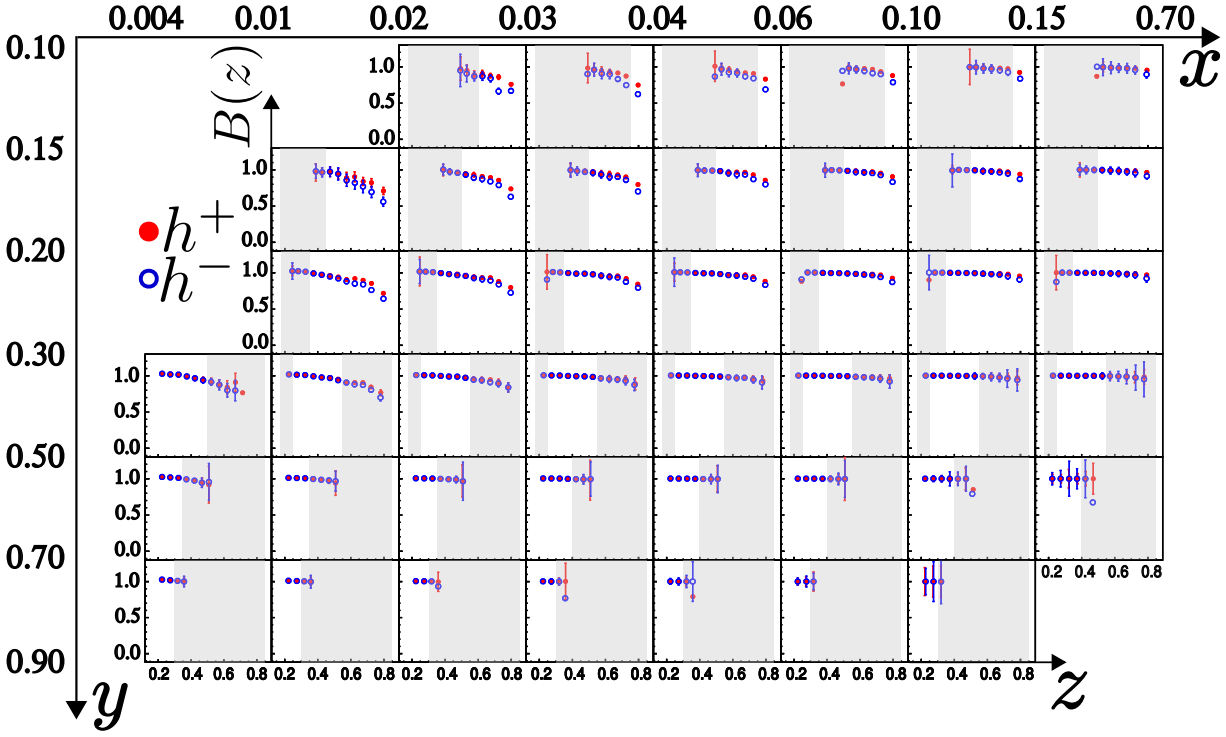


Figure 6.11: *Correction factor B^h for the ρ^0 and ϕ as a function of z for $(x-y)$ bins.*

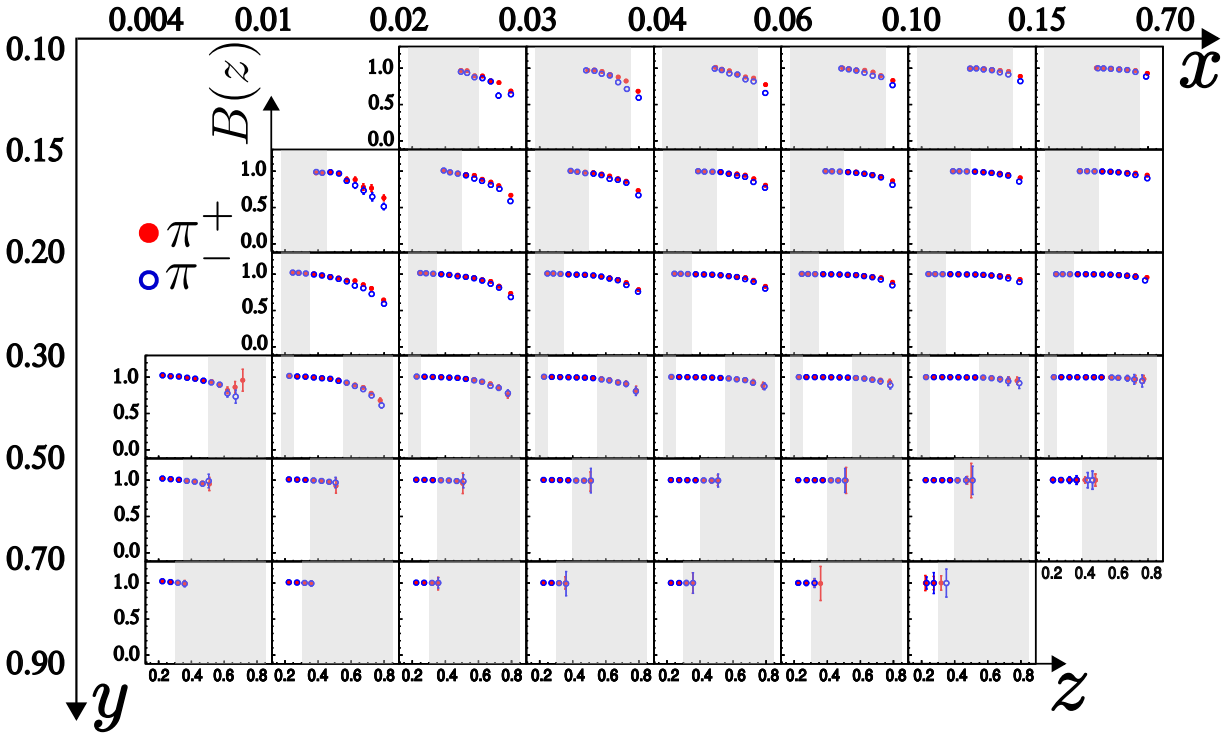


Figure 6.12: *Same as fig. 6.11 for pions.*

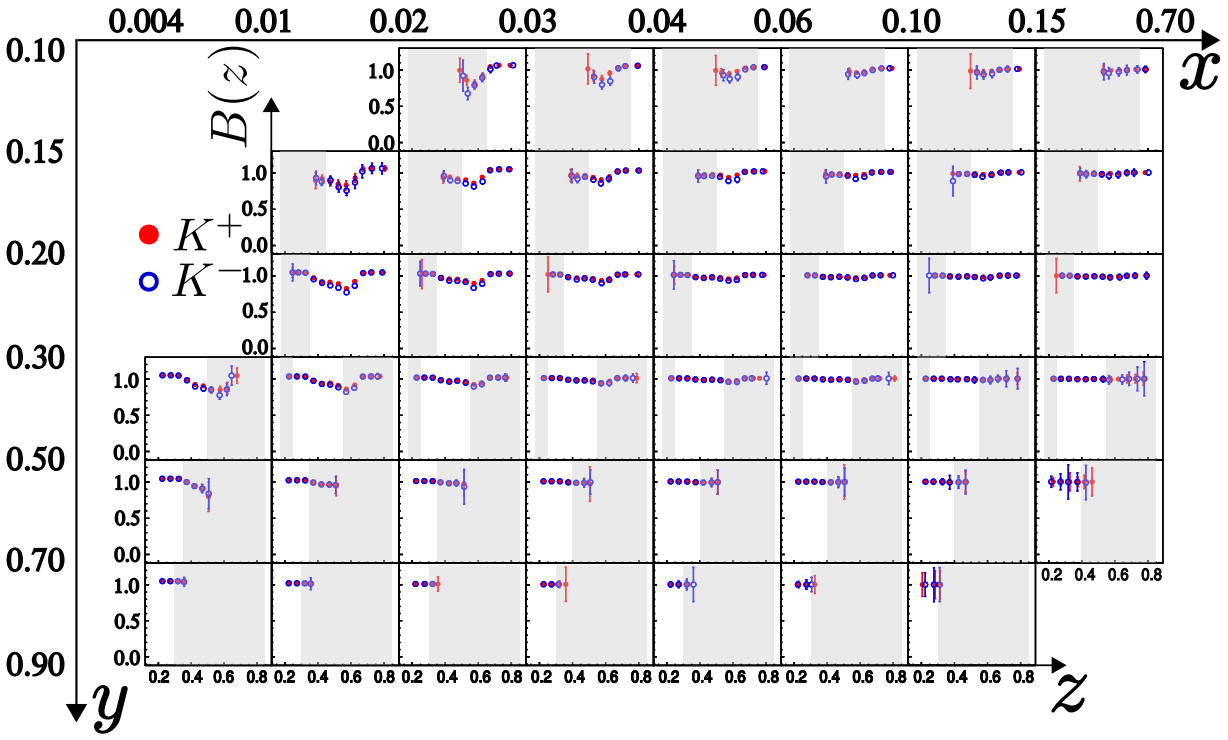


Figure 6.13: *Same as fig. 6.11 for kaons.*

6.3.2 Monte Carlo Sample: HEPGEN Dependence

The cross-section for exclusive ρ^0 production is normalized to the GPD model of Goloskokov and Kroll. The theoretical uncertainty on the predicted cross-section close to COMPASS kinematics is around 30% [90]. The uncertainty affects the values $N_{\rho^0}^{HEPGEN}$, N_{ϕ}^{HEPGEN}, N_{HEPGEN}^{π} and N_{HEPGEN}^{K}. Propagating this uncertainty in the correction factor leads to at most, a systematic uncertainty of 13% for the pions and below 10% for the kaons [81].

6.4 Summary

The most important correction is the acceptance $A(x, y, z)$, which takes into account the geometrical limitations of the spectrometer and the efficiency of the detectors and of the data reconstruction; this correction varies between 50% and 78% for $h^{\pm}$, $\pi^{\pm}$ and $K^{\pm}$.

The correction factor corresponding to the vector meson contamination ($B^h(x, y, z)$) varies from 0 to 40% for pions but it mainly affects the regions where the multiplicities are very low (high z). For the kaons, the maximum correction reaches 25% for medium z values. This correction is model dependent.

Chapter 7

Results for Charged Hadron Multiplicities and Extraction of Quark Fragmentation Functions

This chapter contains two parts: in the first one, the final multiplicities of unidentified charged hadrons, charged pions and kaons extracted from semi-inclusive deep inelastic scattering of 160 GeV muons off an isoscalar target (^{6}LiD) are presented. These results are calculated from the *raw multiplicities* extracted in chapter 5 and the correction factors determined in chapter 6.

In the second part, the quark FFs into π and K are extracted from the charged pion and kaon multiplicities. The results are compared to other FFs parameterizations.

7.1 Charged Hadron Multiplicities

The final results are given with and without the subtraction of the diffractive meson contribution (B^h, because of the uncertainty of the correction. See eq. (7.1)). The multiplicities presented in the following sections are all corrected for the diffractive meson contribution (section 6.2.2).

$$M^h_{VM}(x,y,z) = M^h_{raw}(x,y,z)\frac{1}{A^h(x,y,z)}B^h(x,y,z) + M^h_{extra}(x,y,z) \tag{7.1}$$

The x, y and z binning used is defined in table 5.5. A^h corresponds to the acceptance correction (section 6.1.4) and M^h_{extra} to the extrapolation accounting for the unmeasured momentum range (section 6.1.6). The error propagation of eq. (7.1) was performed assuming A^h and B^h are independent [79, 91]:

$$\left(\frac{\sigma_{M^h_{VM}}}{M^h_{VM}}\right)^2 = \left[\left(\frac{\sigma_{M^h_{raw}}}{M^h_{raw}}\right)^2 + \left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 + \left(\frac{\sigma_{M^h_{extra}}}{M^h_{VM}}\right)^2\right] \tag{7.2}$$

The corresponding systematic errors from the different sources are added in quadrature. The largest contribution in most bins comes from the uncertainty on the acceptance correction (see Appendix F).

7.1.1 Unidentified Hadron Multiplicities

The unidentified hadron multiplicities $M_{VM}^{h^\pm}$ are shown in fig. 7.1 as a function of z for different x, y bins. In the low y region (3 first rows) minor differences between h^+ (red markers) and h^- (blue markers) multiplicities are observed. These differences become larger in the high $y-x$ region (three last rows and 4 last columns). A strong z dependence is observed for all (x,y) bins while the x dependence is weak as expected. The Q^2 values range from 1 to 30 (GeV/c)^2. After the correction for the unmeasured momentum range a total of 204 and 205 points are left for M^{h^+} and M^{h^-} respectively in comparison to the 367 points presented in fig. 5.8. The statistical uncertainties are too small to be visible in almost all kinematic bins. The bands at the bottom are the systematic errors.

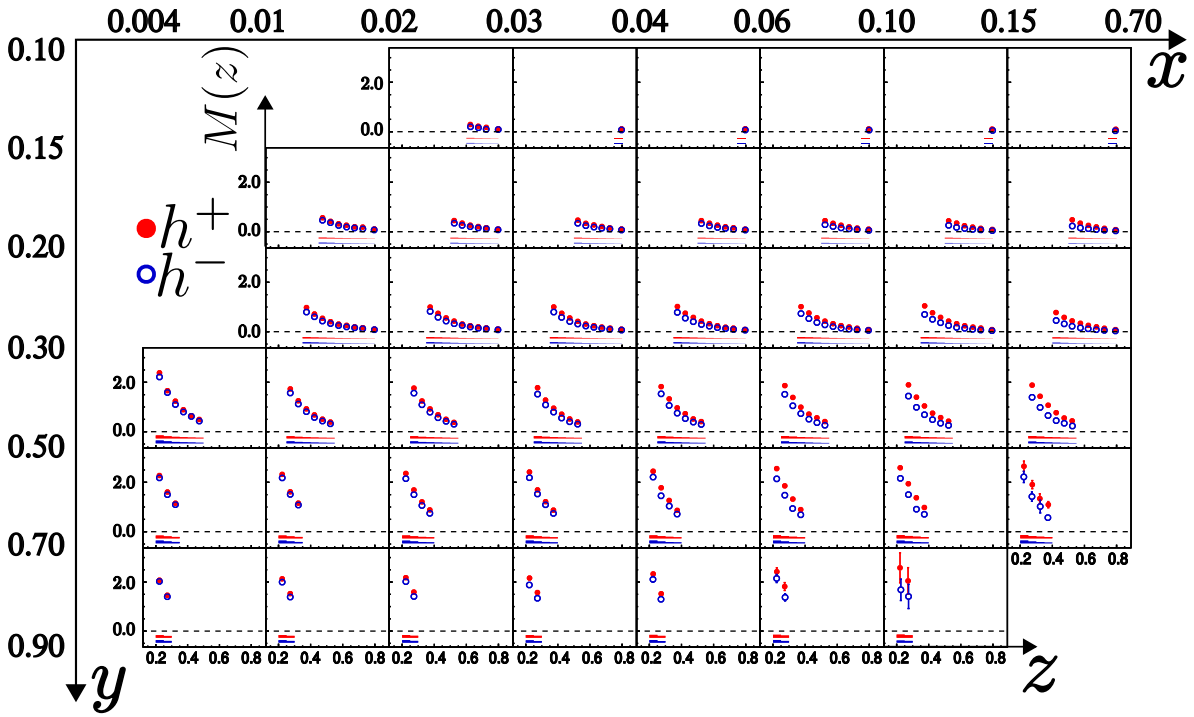


Figure 7.1: *Unidentified hadron multiplicities (with diffractive VM correction) as a function of z in different xy bins. Red markers: $M_{VM}^{h^+}$. Blue markers: $M_{VM}^{h^-}$.*

An alternative presentation of the charged hadron multiplicities M^{h^+} and M^{h^-} is given in fig. 7.2a, where the multiplicities are averaged over y using the square of the statistical errors as weight. To illustrate the impact of the diffractive vector-meson (VM) correction the corresponding values with (M_{VM}^h: open circles) and without (M^h: closed squares) correction are plotted together in figs. 7.2b and 7.2c for positive and negative unidentified

hadrons respectively in logarithmic scale. Each panel corresponds to a different x bin , The larger correction is found at high z where the multiplicities are small. The corrected multiplicities are more linear than the not corrected one in a log scale.

7.1.2 Charged Pion Multiplicities

The results for $M^{\pi^+}_{VM}$ and $M^{\pi^-}_{VM}$ are shown in fig. 7.3. The are 205 and 206 points for π^+ and π^- respectively. Positive and negative pion multiplicities are very similar. A strong dependence on z and a weak one in x is observed. The statistical error starts to be visible only at large x and y values.

The π^+ and π^- multiplicities are shown averaged over y in fig. 7.4a as a function of z in different x bins. In this representation a difference between the positive and negative pion multiplicities is visible for $x > 0.04$. This excess of π^+ compared to π^- can be understood by a dominant u quark distribution in the target: the fragmentation process u $\to \pi^+$ ($u\bar{d}$) is more probable than u $\to \pi^+(\bar{u}d)$. The comparison between the VM corrected and uncorrected multiplicities is shown in figs. 7.4b and 7.4c for π^+ and π^-, respectively. The correction has the strongest impact at high z ($>$0.5) and low x ($<$0.04). In this region the correction leads to a more linear distribution.

7.1.3 Charged Kaon Multiplicities

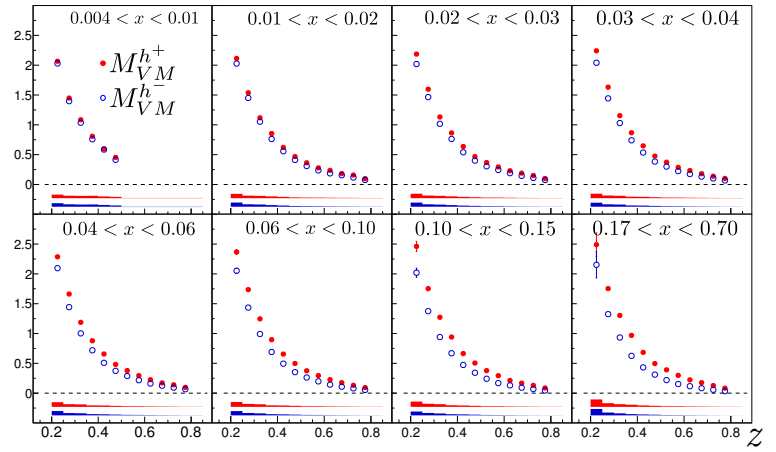
The results for positive and negative kaon multiplicities corrected by the vector-meson production (eq. (7.1)) are plotted in fig. 7.5. There are 207 points for K^+ and 206 for K^-. A systematic difference between the positive and negative multiplicities is observed over all the kinematic domain. This behavior reflects the fact that the fragmentation process to produce K^- is *unfavored* (u $\to K^-(\bar{u}s)$, needs quarks $\bar{u}$ and s from the quark sea) while the K^+ production is *favored* (u $\to K^-(u\bar{s})$ needs only $\bar{s}$ from the quark sea).

The VM correction for kaons is maximal for medium z values and for low x values ($<$0.04) where it can reach 20%. In the comparison between the kaon multiplicities with and without corrections shown in figs. 7.6b and 7.6c (K^+ and K^-, respectively) the difference is almost imperceptible in all x domain.

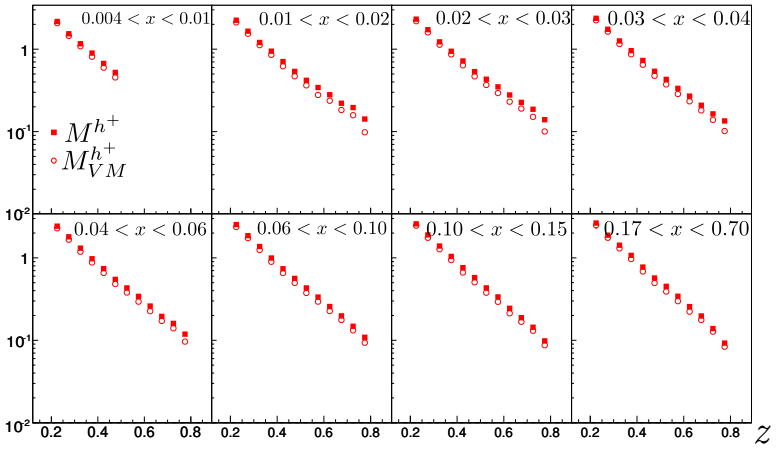
7.2 Sums of Charged Hadron Multiplicities

Before the global extraction of the FFs it is possible to study some specific quantities (at leading order) like the ratios and the sums of the various multiplicities. Following eq. (1.26), in an isoscalar target and neglecting the heavy quark contribution, the charged hadron multiplicities can be written as:

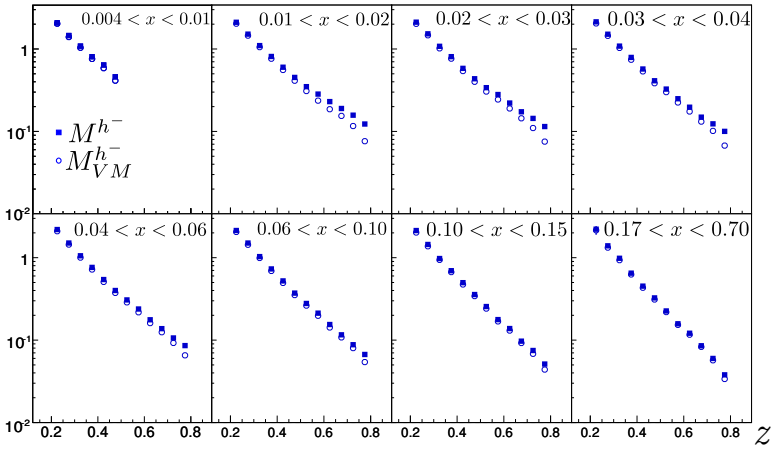
$$M^h = \frac{(q_u + q_d)(4D^h_{q_u} + D^h_{q_d}) + (q_{\bar{u}} + q_{\bar{d}})(4D^h_{q_{\bar{u}}} + D^h_{q_{\bar{d}}}) + 2q_s D^h_s + 2q_{\bar{s}} D^h_{\bar{s}}}{5(q_u + q_{\bar{u}} + q_d + q_{\bar{d}}) + 2(q_s + q_{\bar{s}})} \tag{7.3}$$



(a)



(b) h^+



(c) h^-

Figure 7.2: *Top: h^+ (red) and h^- (blue) multiplicities versus z in 8 x bins. Data are averaged in y. Middle: h^+ multiplicities without (full squares) and with (empty circles) VM correction. Bottom: Same as middle for h^-.*

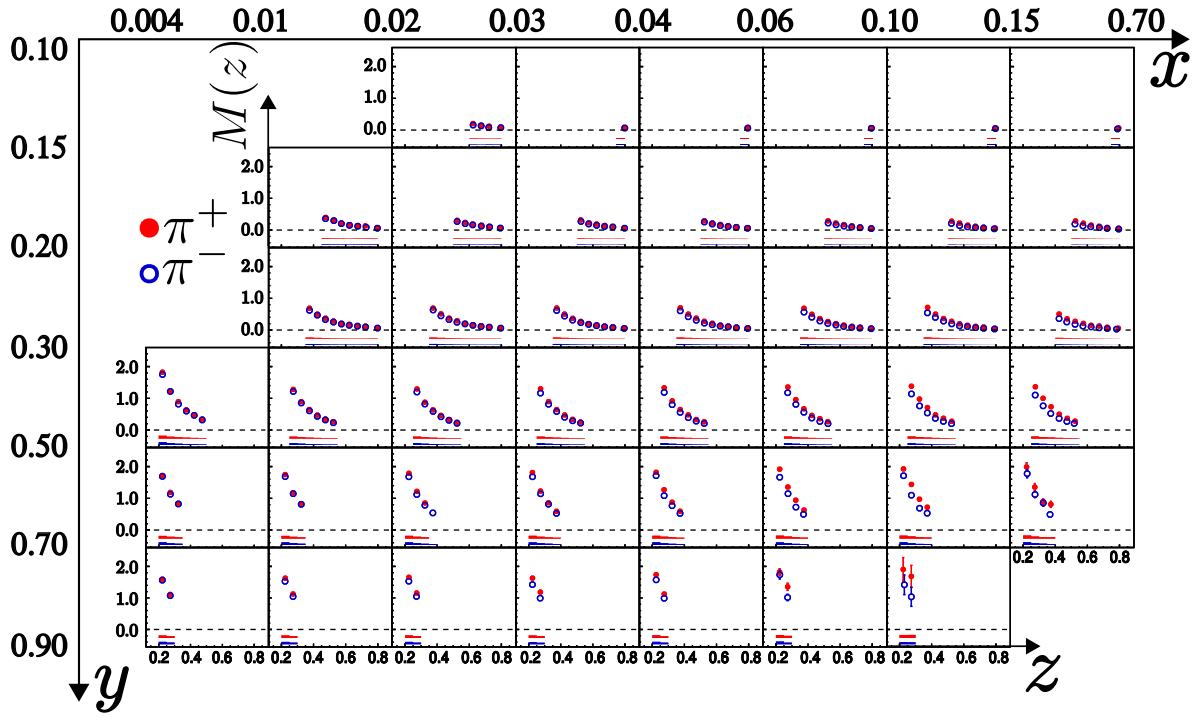


Figure 7.3: *Same as fig. 7.1 for pions:* $M^{\pi^+}_{VM}$ *and* $M^{\pi^-}_{VM}$.

In particular, the sum of charged kaon multiplicities versus x, gives a simple access to the strange quark distribution in the nucleon $q_s(x)$ and the strange quark fragmentation into kaons $D^K_s(z)$. This sum can be written as

$$\int_{z_{min}}^{z_{max}} M^{K^+}(z) + M^{K^-}(z)dz = \frac{Q(x)\int_{z_{min}}^{z_{max}} D^K_Q(z)dz + S(x)\int_{z_{min}}^{z_{max}} D^K_S(z)dz}{5Q(x) + 2S(x)} \tag{7.4}$$

where

$$Q(x) = q_u(x) + q_{\bar{u}}(x) + q_d(x) + q_{\bar{d}}(x), \quad D^K_Q(z) = 4D^K_{fav}(z) + 6D^K_{unf}(z) \tag{7.5a}$$

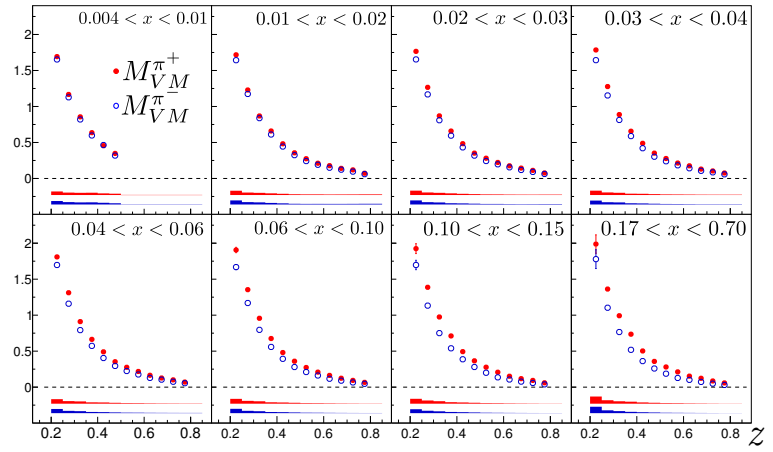
$$S(x) = q_s(x) + q_{\bar{s}}(x), \quad D^K_S(z) = 2D^K_s(z) \tag{7.5b}$$

$$D^K_{fav} = D^{K^+}_{q_u} = D^{K^-}_{q_{\bar{u}}} \tag{7.5c}$$

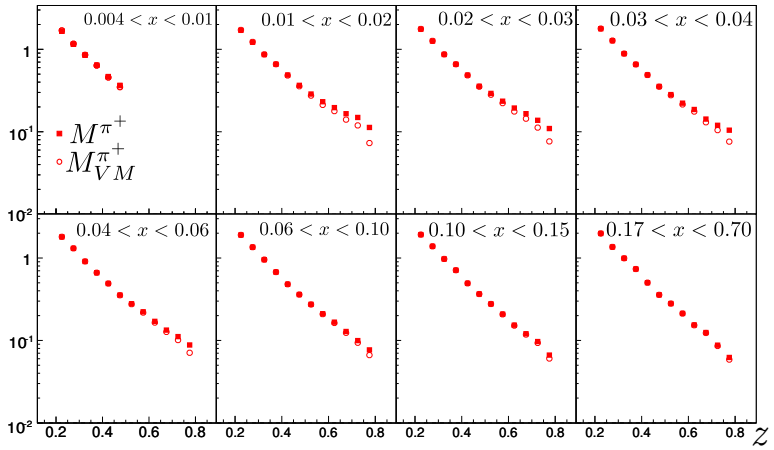
$$D^K_s = D^{K^+}_{q_s} = D^{K^+}_{q_{\bar{s}}} = D^{K^-}_{q_s} = D^{K^-}_{q_{\bar{s}}} \tag{7.5d}$$

$$D^K_{unf} = D^{K^+}_{\bar{u}} = D^{K^+}_{d} = D^{K^+}_{\bar{d}} = D^{K^-}_{u} = D^{K^-}_{d} = D^{K^-}_{\bar{d}} \tag{7.5e}$$

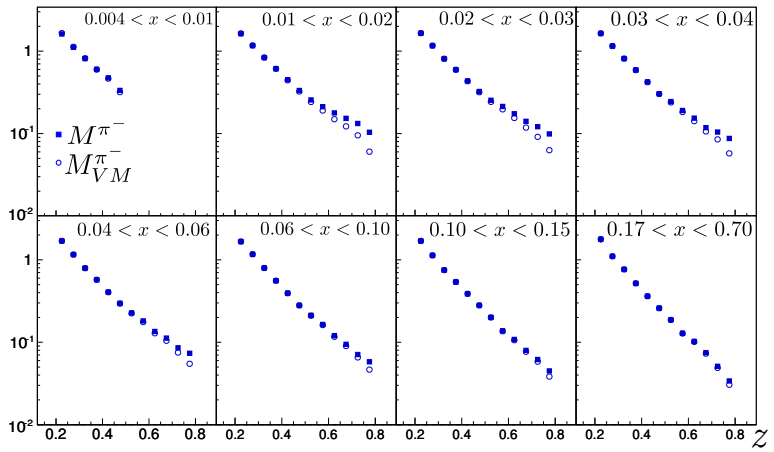
The kaon multiplicities integrated over z are shown in fig. 7.7 for the [0.01,0.70] x region. The first x bin [0.004,0.10] is not considered since the z range covered in this bin is smaller than [0.20,0.85]. The results are plotted with (red open circles) and without (red full squares) the diffractive vector-meson correction. However they cannot be distinguished. The full blue circles in fig. 7.7 represent the charged kaon multiplicity obtained from a MC sample at LO using LEPTO (Lund fragmentation) and MSTW PDFS. HERMES results [37] with and without the vector-meson correction are plotted using full black



(a)



(b) π^+



(c) π^-

Figure 7.4: *Same as fig. 7.2 for pions.*

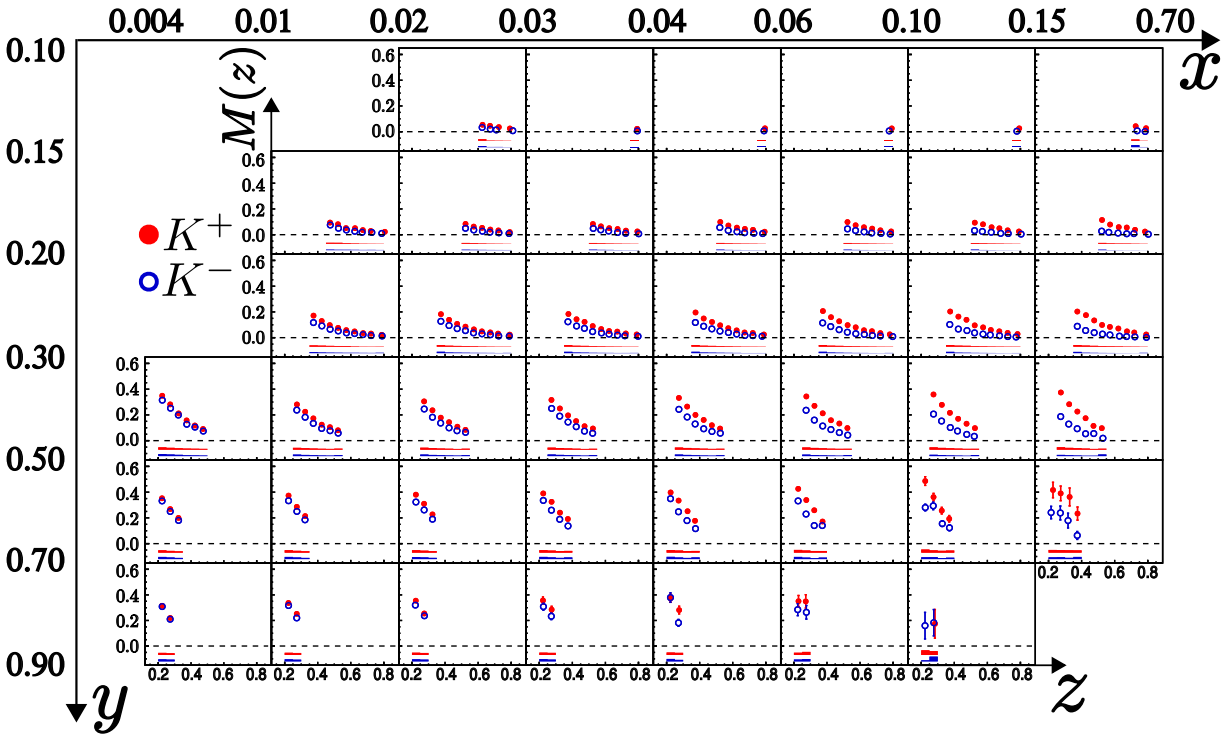


Figure 7.5: *Same as fig. 7.1 for kaons:* $M_{VM}^{K^+}$ *and* $M_{VM}^{K^-}$.

squares and open black circles respectively (again indistinguishable); the corresponding vertical error bars includes the statistical and systematic errors added in quadrature. The present results differ from those of HERMES [37]; as a consequence the conclusions on the strange quark FFs and PDFs differ. If we assume $2S(x) << 5Q(x)$, eq. (7.4) can be rewritten as:

$$\begin{aligned} M^{K^+ + K^-} &= \int_{0.2}^{0.85} M^{K^+}(z) + M^{K^-}(z) dz \\ &\approx \frac{Q(x)\int_{0.2}^{0.85} D_Q^K(z)dz + S(x)\int_{0.2}^{0.85} D_S^K(z)dz}{5Q(x)} \end{aligned} \tag{7.6}$$

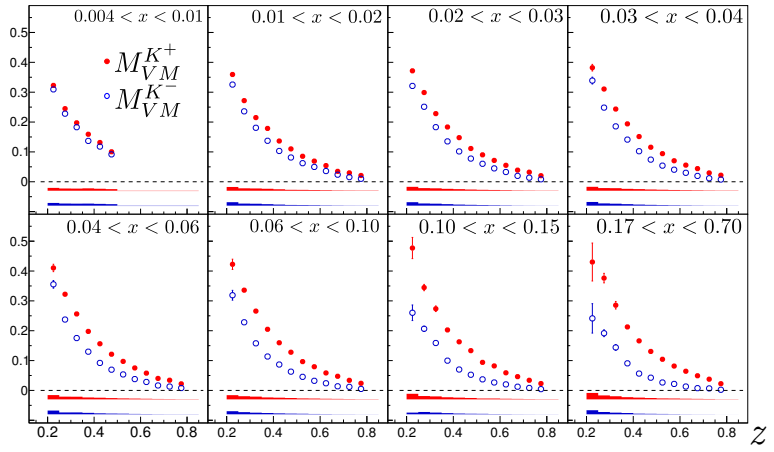
In the high x region, $S(x) \sim 0$ and therefore eq. (7.6) becomes only z dependent:

$$M_{\text{high } x}^{K^+ + K^-} \approx \frac{1}{5}\int_{0.2}^{0.85} D_Q^K(z)dz \tag{7.7}$$

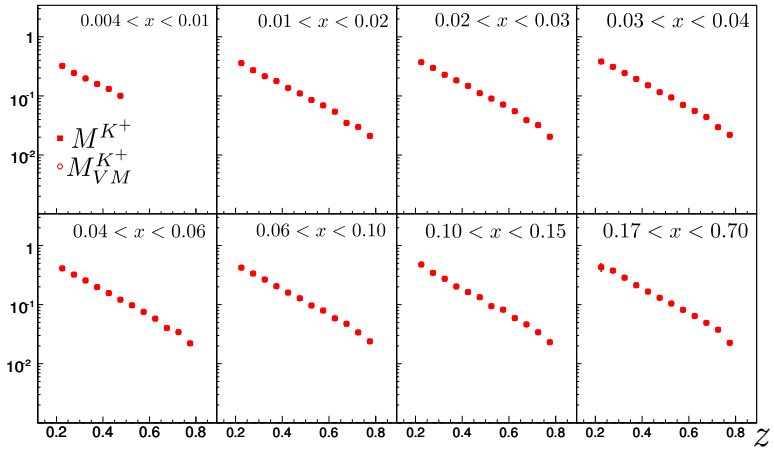
In the low x region $S(x) \neq 0$ and eq. (7.6) becomes

$$M_{\text{low } x}^{K^+ + K^-} \approx \frac{1}{5}\left(\int_{0.2}^{0.85} D_Q^K(z)dz + \frac{1}{Q(x)}S(x)\int_{0.2}^{0.85} D_S^K(z)dz\right) \tag{7.8}$$

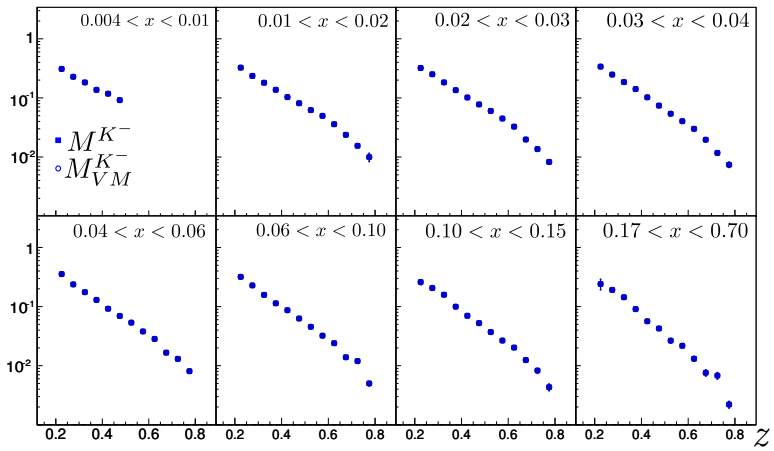
Using eqs. (7.7) and (7.8) the contribution of $S(x)\int_{0.2}^{0.85} D_S^K(z)dz$ can be inferred (for the z measured range 0.2-0.85). If $M_{\text{low } x}^{K^+ + K^-} > M_{\text{high } x}^{K^+ + K^-}$, then SD_S^K is likely positive, while



(a)



(b) K^+



(c) K^-

Figure 7.6: *Same as fig. 7.2 for kaons.*

if $M^{K^+ + K^-}_{\text{low } x} < M^{K^+ + K^-}_{\text{high } x}$ it is negative. If a small or null difference between $M^{K^+ + K^-}_{\text{low } x}$ and $M^{K^+ + K^-}_{\text{high } x}$ is found, one can expect that SD^K_S is close to zero. From COMPASS data (fig. 7.7a) one might conclude that SD^K_S is rather small since $M^{K^+} + M^{K^-}$ does not rise at low x, while HERMES results do.

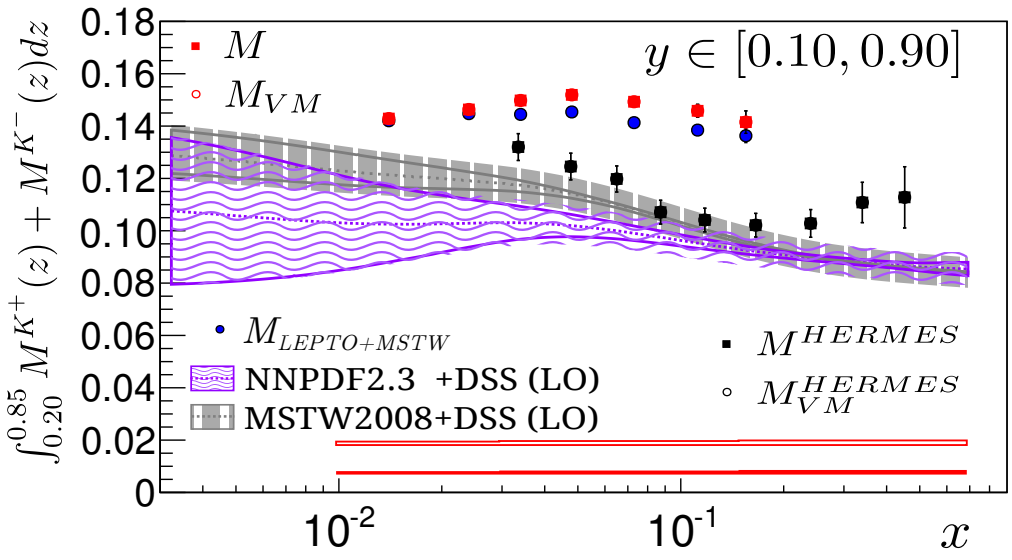
Another observation from this plot is the difference which cannot be explained by statistical or systematic fluctuations between the multiplicity extracted from the LEPTO generator data and the experimental multiplicities in the x region [0.02,0.15]. In order to study this discrepancy, the sum of M^{K^+} and M^{K^-} was extracted in a smaller y domain ($[0.10, 0.30]$) which corresponds to lower Q^2 values. This sum is shown in fig. 7.7b. The decrease in $M^{K^+} + M^{K^-}$ value in comparison to fig. 7.7a is due to the different z covered domain. In the lower y region only $z \in [0.35, 0.85]$ is covered. A very good agreement between the experimental multiplicities and the ones extracted from LEPTO generator is found in fig. 7.7b in contrast to fig. 7.7a. The agreement between the low y multiplicities and the LEPTO ones was not expected since the LEPTO generator is not adjusted to follow the y dependence. This can be done as an improvement of the present analysis.

In figs. 7.7a and 7.7b two LO calculations are plotted in addition. In magenta, the PDF parameterization from NNPDF [92] and in gray the one from MSTW [17]. For both calculations, the DSS [43] parameterization was used for the FFs. The x and Q^2 values used in the prediction correspond to the COMPASS mean values, summarized in table 7.1. The region delimited by the full lines in each case corresponds to the error associated to the PDF parameterization. The largest error band is obtained when an error of 7% is associated to the fragmentation function value (this is estimated from the difference between COMPASS and DSS values). The discrepancies between COMPASS and HERMES data with the results based on FFs from DSS and PDFs from MSTW and NNPDF are substantial in fig. 7.7a. However, in fig. 7.7b the LO calculation with MSTW and NNPDF parameterizations are in reasonable agreement with our data especially at low x, below 0.04.

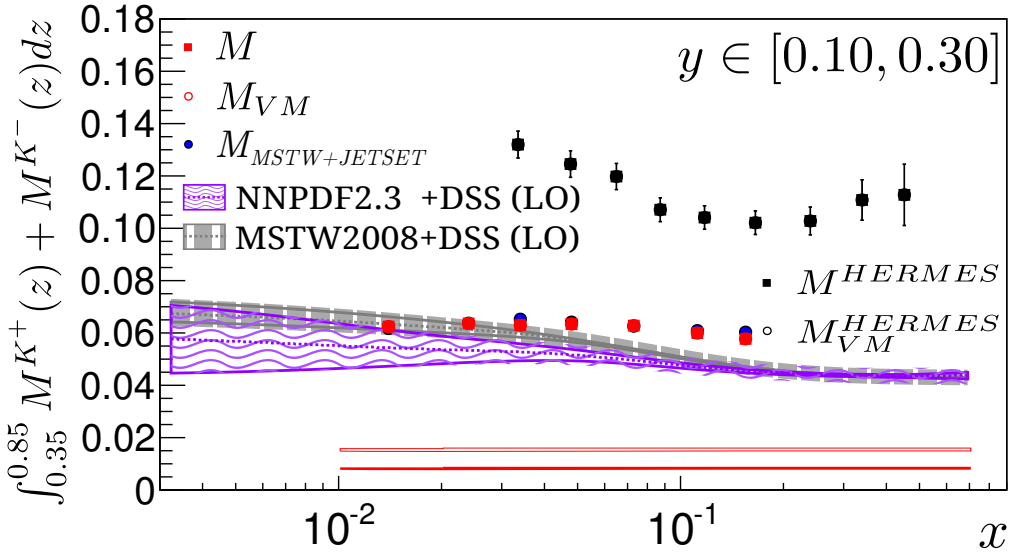
For completeness, the sum of π^+ and π^- multiplicities is showed in fig. 7.8. Here, a fair agreement between the LO calculations and the COMPASS pion multiplicities is observed in both y regions. Only the HERMES data disagree.

7.3 Ratio of Charged Pion Multiplicities

The ratio of the charged pion multiplicities M^{π^+}/M^{π^-} is shown in fig. 7.9 as a function of x between 0.01 to 0.7 and integrated in two z ranges [0.20,0.85] (fig. 7.9a) and [0.35,0.85] (fig. 7.9b). The ratio using the multiplicities with and without the exclusive correction are represented by the open red circles and red full squares respectively. In fig. 7.9a, the full y domain is considered [0.10,0.90] while in fig. 7.9b it is restricted to [0.10,0.30]. In addition, the results from the JLab E00-108 [36] (second entry) (blue full squares) and HERMES (black full squares and open circles) experiments are shown in both plots. A disagreement



(a)



(b)

Figure 7.7: *Sum of kaon multiplicities $M^{K^+} + M^{K^+}$ in two different y regions integrated over z. The present COMPASS data (in red) are compared to a MC simulation at LO (blue points, LEPTO+MSTW) and two parameterizations at LO (magenta and gray bands, see text). HERMES data [37] are shown only in the upper plot (in black). Data with (open points) and without (full points) VM corrections are not distinguishable.*

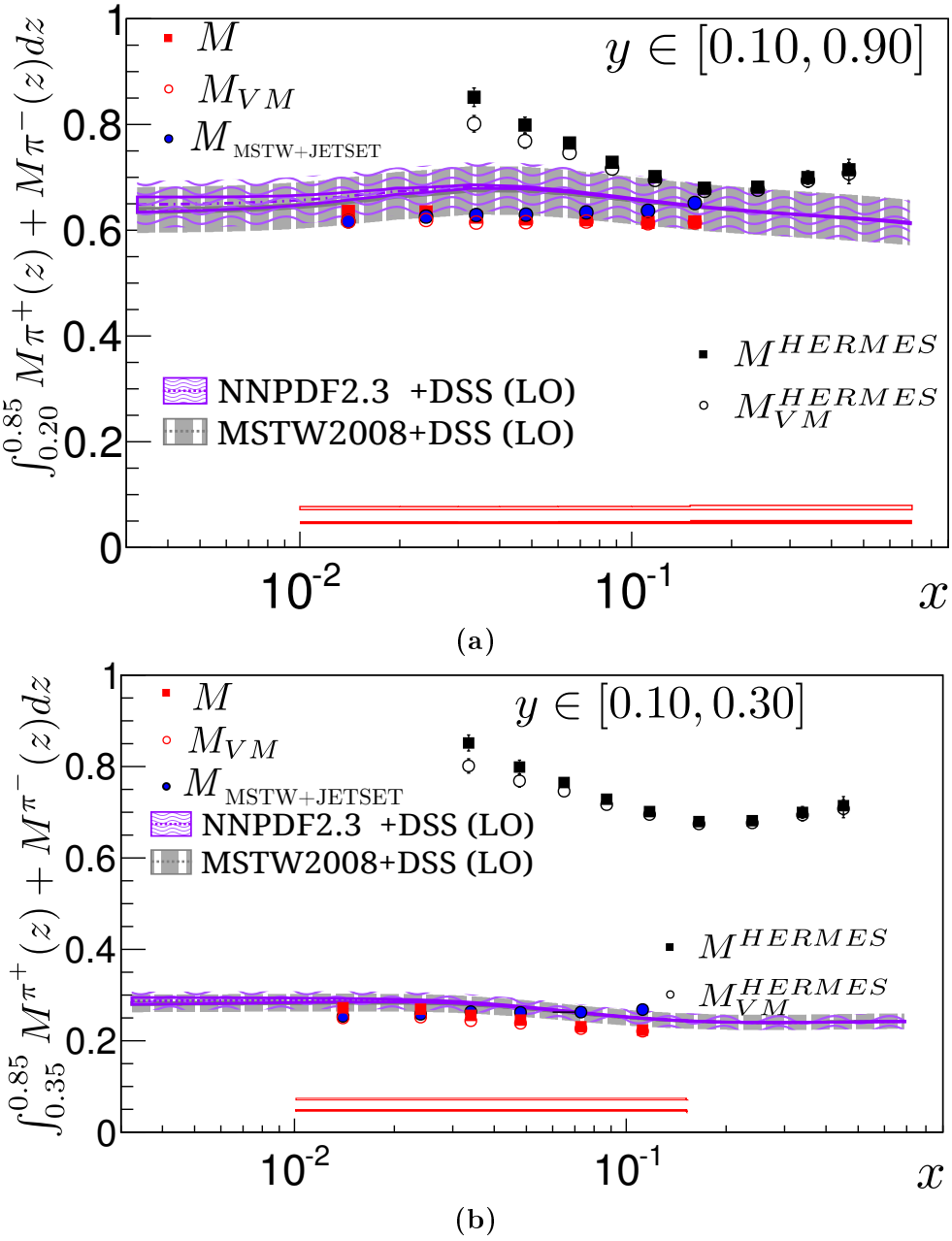


Figure 7.8: *Same as fig. 7.7 for pions.*

Table 7.1: *x and Q^2 mean values for COMPASS and HERMES [37].*

$< Q^2 >$ (GeV/c)^2				
COMPASS			HERMES	
$< x >$	$y \in [0.10, 0.90]$	$y \in [0.10, 0.30]$	$< x >$	
0.014	1.543	1.047		
0.024	2.434	1.167		
0.034	3.488	1.701	0.033	1.193
0.048	4.884	2.405	0.048	1.382
0.073	7.545	3.779	0.065	1.555
0.112	11.676	5.85	0.087	1.728
0.155	19.027	9.166	0.118	2.173
			0.166	3.167
			0.240	4.878
			0.340	7.477
			0.452	10.238

between the E00-108 and the COMPASS is found. This is expected because the E00-108 selected kinematic domain is restricted to $Q^2 \in (2, 4)$ (GeV/c)^2 and $x \in (0.2, 0.6)$. The ratio obtained from HERMES data is in reasonable agreement with the results obtained in this work in fig. 7.9a. As in fig. 7.7, two LO calculations are plotted. In both figures, the theoretical prediction is systematically above COMPASS and HERMES data at low x. However, a fair agreement between the predictions and the ratios at high x values for COMPASS and JLAB is observed in fig. 7.9a.

7.4 Quark fragmentation functions from charged pion and kaon multiplicities

The pion and kaon multiplicities have been used as an input to LO fits in view of the extraction of quark FFs [39, 80, 93]. In this chapter a summary of the results obtained is given.

In order to extract the fragmentation functions from the measured charged π and K multiplicities (separately) the corresponding expression at LO (eq. (7.3)) is used. For this study the PDFs are considered to be known. The LO MSTW parametrization given in [83] provided by the LHAPDF data group [94] is used. Only three independent FFs are considered for pions (including the gluon FF) and four for kaons, assuming isospin and charge summetry and also $D_S = D_{unf}$ for pions. The corresponding expressions are

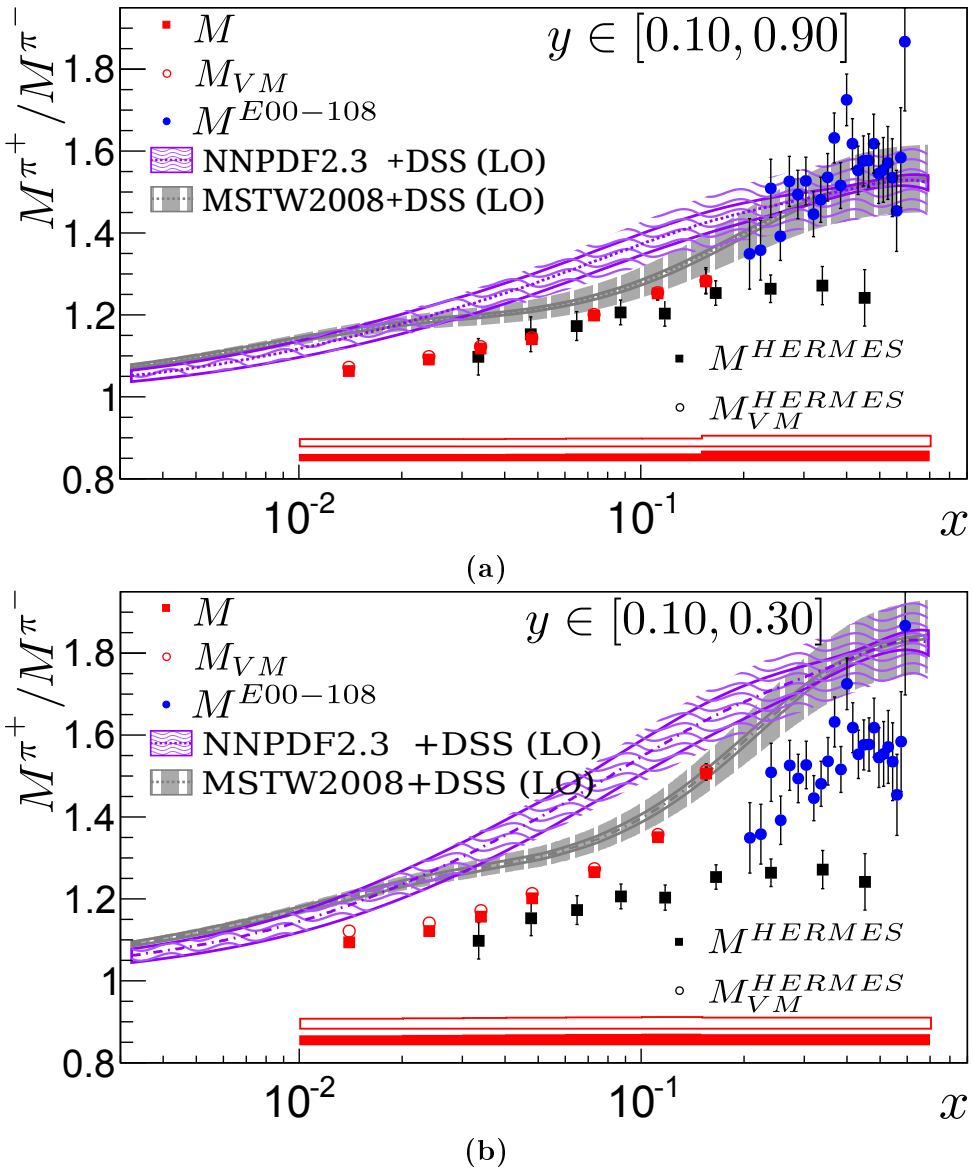


Figure 7.9: *Ratio of M^{π^+}/M^{π^-} for COMPASS (red), HERMES data [37](black) and JLab (blue) data compared to LO calculations as in fig. 7.7.*

defined for the reference value $Q_0^2 = 1$ (GeV/c)^2 in Table 7.2. The extraction of pion and kaon FFs is performed via a χ^2 minimization:

$$\chi^2 = \sum_{j=1}^{N} \left[\left(M_{\text{theo}}^{h^\pm}\left(x_j, Q_j^2\right) - M_{\text{exp}}^{h^\pm}\left(x_j, Q_j^2\right) \right) / \sigma_j \right]^2 \tag{7.9}$$

where N are the number of data points and $\sigma_j^2 = \sigma_{M_{\text{stat}}}^2 + \sigma_{M_{\text{sys}}}^2$. Finally $M_{\text{theo}}^{h^\pm}$ is extracted from eq. (7.3).

Table 7.2: *Quark FF parameterizations at $Q_0^2 = 1$ $(GeV/c)^2$. The normalization factor $|N|$ is calculated numerically at each iteration such as $|N| = \int_0^1 z D_q(z, Q_0^2) dz$*

FFs	Pions	Kaons
D_{fav}	$(N/\|N\|)z^\alpha(1-z)^\beta$	$(N/\|N\|)z^\alpha(1-z)^{\beta=\text{cte}}$
D_S		$(N/\|N\|)z^\alpha(1-z)^\beta$
D_{unf}	$(N/\|N\|)z^\alpha(1-z)^\beta$	$(N/\|N\|)z^\alpha(1-z)^\beta$

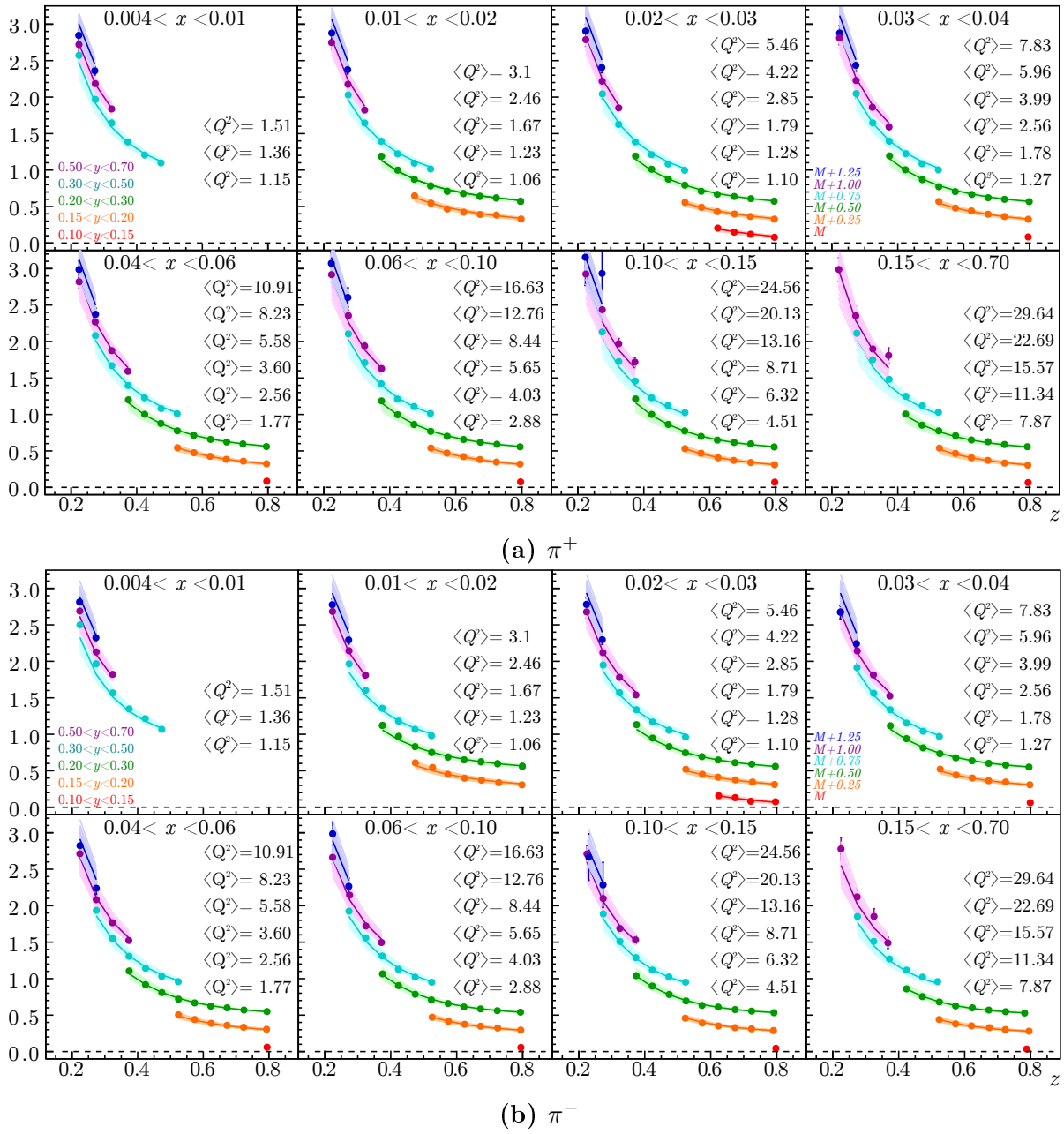
The actual calculation of FFs is performed by a customized version of the Q^2 evolution code developed in [41].

To determine the fit uncertainties the bootstrap method [95] is used. This method consists in generating replicas of the original multiplicity data set by repeatedly and randomly varying the original data points according to a Gaussian distribution centered on their mean values and with the width of the data point errors. Then for each replica the corresponding FFs are extracted using Eq. (7.9). The final error band due to the data uncertainties is determined as the mean value ± RMS of the extracted FFs from replicas. For the current analysis the number of 100 replicas is used.

7.4.1 Pions

The results for the z and y dependence of the π^+ and π^- multiplicities with vector-meson correction are presented in the eight bins of x in Figs. 7.10a and 7.10b. In each x bin the results for the different y are shifted by a 0.25 constant. The statistical and systematic uncertainties are shown as vertical bars (added in quadrature); however they are in general smaller than the size of the symbols. The curves correspond to the global LO fit of π^+ and π^- multiplicities described below to extract pion FFs. A good agreement between the fit and the data is found. The data represent a new precise SIDIS data set that will be used for flavor separation also in the NLO extractions of fragmentation functions.

The final parameters of the fit are given in Table 7.3. The extracted favored and unfavored FFs at $Q^2 = 3$(GeV/c)^2 are shown in Figs. 7.11a and 7.11b, respectively. The error



(a) π^+

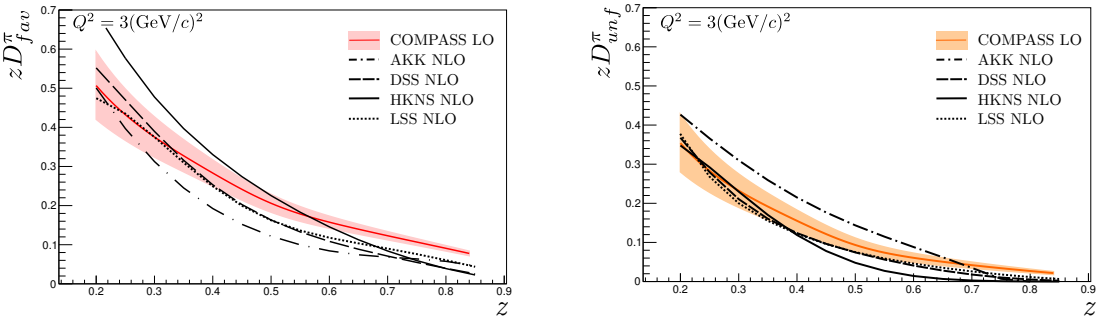
(b) π^-

Figure 7.10: *π^+ (a) and π^- (b) multiplicities versus z for 8 x bins and 5 y bins (corresponding to 5 Q^2 values and staggered by +0.25 for clarity). Combined statistical and systematic errors are plotted (vertical bars, but in general not visible). The curves correspond to the LO fit of $\pi^\pm$ multiplicities.*

bands from statistical and systematic uncertainties as well as the FFs extracted from the original multiplicity set (solid lines) are plotted. As expected, the favoured FFs are larger than the unfavoured. The FFs at NLO from DSS [43] (dashed line), HKNS [41](solid line), AKK [42](dash-dotted line) (second entry) and LSS [45](dotted line) parametrizations are also plotted. A good agreement between the COMPASS LO unfavored FF with DSS and LSS NLO parameterizations is found in almost all the z range ($< 25\%$ and $< 26\%$ respectively). Also, a fair agreement in the low z domain (z below 0.4) is observed for the favored FF (LSS $< 36\%$ and for DSS $< 30\%$). This was expected since this two parameterizations include the data from SIDIS. The agreement with LSS is sliglty better because they have included a preliminary version of the COMPASS data in their fit. The D^{π}_{fav} parametrization from HKNS differs from all the others parameterizations because it only includes electron-positron annihilation data. However, this discrepancy is less important for the unfavored FF. In the AKK parameterization (electron-positron annihilation and pp collision data), an agreement ($< 29\%$) with the COMPASS favored FF is observed at the low z ($z < 0.35$), while some discrepancy is observed at high z. For the unfavored FF, this parameterization is systematically above all the others parameterizations.

	|N|	α	β	χ^2/ndf	|N|	α	β	χ^2/ndf
	Without vector-meson correction				With vector-meson correction			
D^{π}_{fav}	0.162	-0.583	0.417	2.207	0.159	-0.777	0.391	1.265
D^{π}_{unf}	0.069	-0.757	0.706		0.064	-0.621	1.188	

Table 7.3: *Parameters of the favored and unfavored FFs for pions at* $Q^2 = 1(GeV/c)^2$.



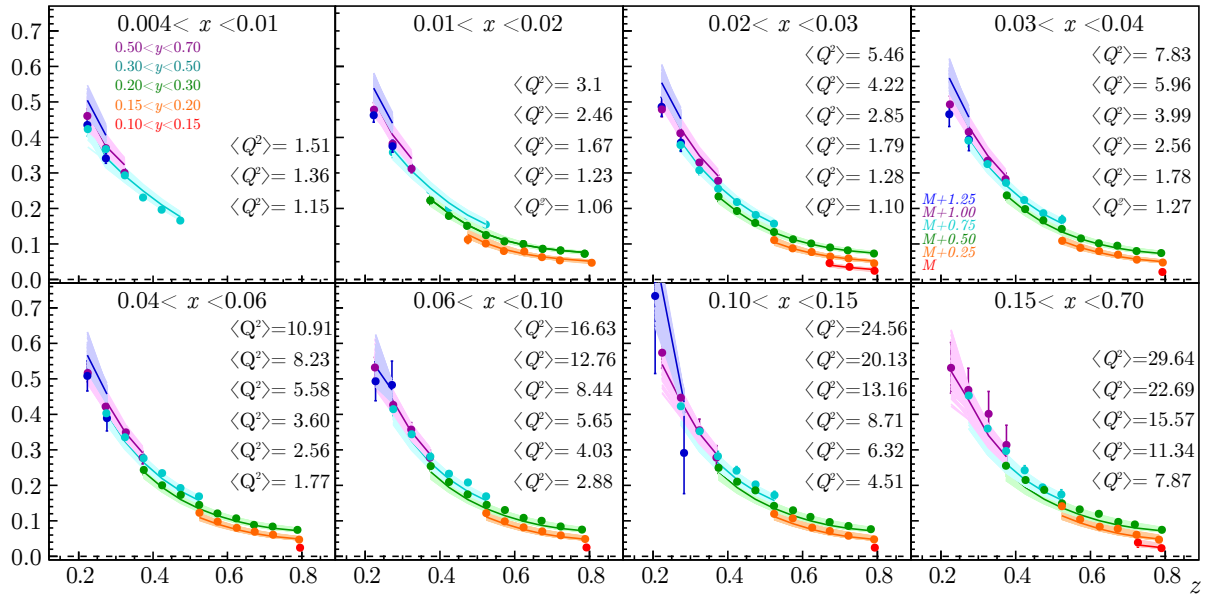
(a) **(b)**

Figure 7.11: *Favored (a) and unfavored (b) quark fragmentation functions into pions* zD^{π}_{fav} *and* zD^{π}_{unf} *versus z, extracted from a LO fit of pion multiplicities (corrected for VM contribution) at* $Q^2 = 3(GeV/c)^2$. *The error bands correspond to statistical and systematics uncertainties.*

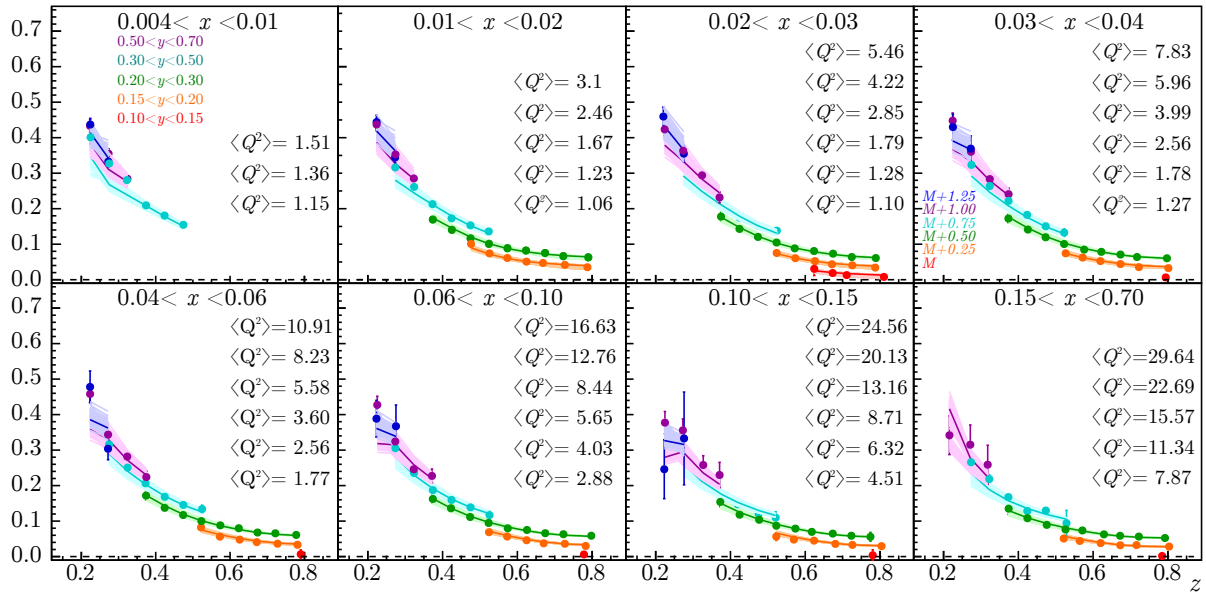
7.4.2 Kaons

The results for charged kaon multiplicities with vector-meson correction are presented in eight bins of x in Fig. 7.12. The data are well described by the fit at almost all

x values for y below 0.7. In the last y bin (0.7-0.9), some discrepancies between the experimental multiplicities and the ones extracted from the FFs start to be visible. For K^+, the experimental data are systematically below the LO calculations. In addition for x below 0.06, the K^+ multiplicities for the larger y bin are smaller than the precedent one ($y \in [0.5, 0.7]$). In the K^- case, a better agreement with the LO calculations is found, except in the x bins 0.04-0.06 and 0.10-0.15 for the higher y bin.



(a) K^+



(b) K^-

Figure 7.12: *Same as Fig. 7.10 for kaons*

The parameters obtained for D_{fav}^K, D_S^K and D_{unf}^K fits using the kaon multiplicities with

and without the vector-meson correction are given in Table 7.4. They are compared to the FFs at NLO from DSS [43] (dashed line), HKNS [41](solid line) and AKK [42](dash-dotted line)(second entry) in Fig. 7.13.

	|N|	α	β	χ^2	|N|	α	β	χ^2
	Without VM correction				With VM correction			
D^K_{fav}	0.059	-1.319	0.02(cte)		0.05	-0.952	0.02(cte)	
D^K_S	-0.074	7.991	33.378	7.369	0.191	6.868	12.076	5.591
D^K_{unf}	0.009	5.198	9.662		0.000	42.101	382.36	

Table 7.4: *Parameters of the favored, strange and unfavored FFs for Kaons at* $Q^2 = 1(GeV/c)^2$.

The results for the kaon favored FF shown in Fig. 7.13a is in a good agreement with the NLO calculations from the DSS parameterization (within 4 to 21%) in the z range 0.35-0.7. This is expected because the DSS fit uses K multiplicities from HERMES (SIDIS process). The COMPASS favored FF agrees also with the NLO HKNS parameterization within 3 to 23% without considering z above 0.80. A significant discrepancy between the AKK prediction is observed in the z range 0.25-0.85. For the unfavored FF (Fig. 7.13b) D^K_{unf}, the good agreement with DSS is conserved, but the compatibility with HKNS is lost at low z. A fair agreement with AKK is found ($< 50\%$). The D^K_S shape from the present study differs significantly from the three NLO parameterizations shown in Section 7.4.2.

7.5 Summary

The final unidentified hadron ($h^{\pm}$), pion ($\pi^{\pm}$) and kaon ($K^{\pm}$) multiplicities extracted from the 2006 COMPASS data (muon deep inelastic scattering on a ^{6}LiD target) were presented as a function of z for different x and y bins. In addition, a two dimensional projection (averaging over y) was also shown. The sum of charged pion and kaon multiplicities were compared to HERMES experiment results [37] and to LO calculation using the LO definition of hadron multiplicities within the quark-parton model improved with QCD defined in Section 1.7. Two sets of PDFs parametrization were used: MSTW [17] and NNPDF [92]. For the FFs only the DSS parameterization were considered. For the sum of charged pions, a good agreement between the LO calculations and COMPASS data were found. However, an important discrepancy with HERMES data was observed at low x. In the kaon case, only a good agreement with the LO calculations was found in the low Q^2 domain (low y). A discrepancy with the recent result from HERMES experiment is observed in all x range. For pions, the ratio of π^+ over π^- multiplicities was compared to HERMES [37] and E00-108 (JLab) [36] results. A good agreement with both data was found in the common x ranges. However, an agreement with the LO calculations was only found at high x.

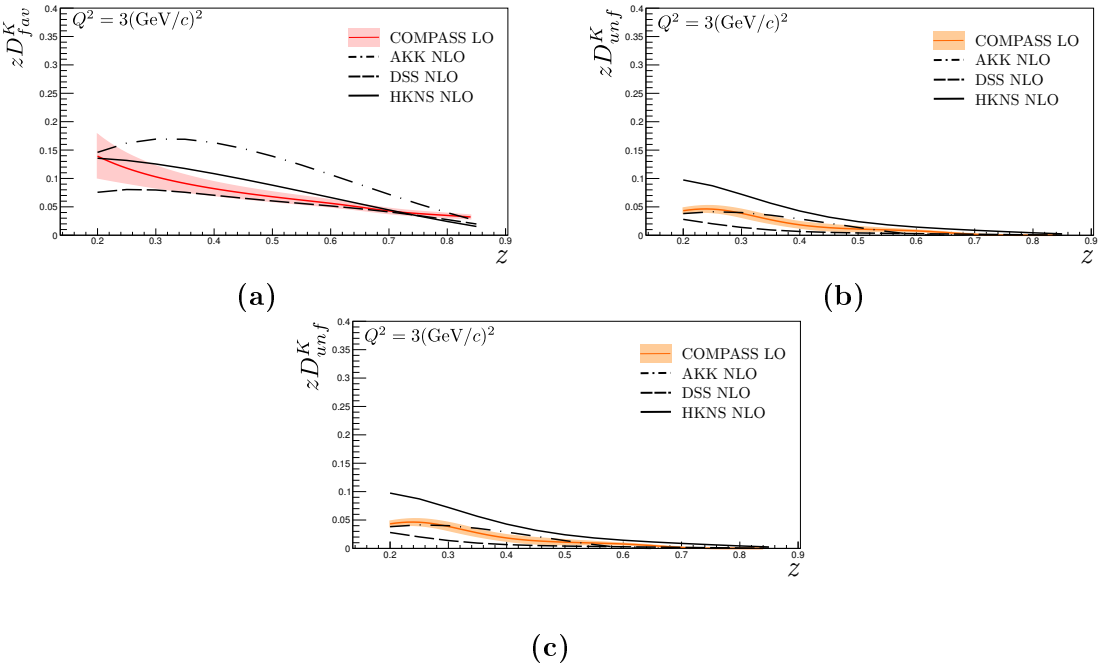


Figure 7.13: *The favored (a), strange (b) and unfavored (c) quark FFs into kaons zD^K_{fav}, zD^K_S and zD^K_{unf} versus z, extracted from a LO fit of kaon multiplicities (corrected for VM contribution) at $Q^2 = 3(GeV/c)^2$. The error bands correspond to statistical and systematics uncertainties.*

Pion and kaon fragmentation functions have been extracted from a fit at LO of the corresponding multiplicities. They have been extracted in the 0.20-0.85 z domain. They were compared to NLO predictions given by four parameterizations (which consider different sets of data to perform their fits): HKNS [41], AKK [42], DSS [43], and LSS [45]. In the case of pions two FFs were extracted: the favored (D^π_{fav}) and the unfavored (D^π_{unf}). The pion FFs obtained agree with the NLO DSS and LSS parameterization by 30% or less in certains z regions. As expected they differ from the ones of HKNS and AKK, which do not contain SIDIS data. For kaons, a total of three FFs were extracted: the favored (D^K_{fav}), the strange (D^K_S) and the unfavored (D^K_{unf}). The favored FF D^K_{fav}, agree with the NLO DSS parameterization within $< 17\%$ (for $z \in [0.40, 0.75]$). The HKNS prediction had a partial agreement for D^K_{fav} in the low and high z region while only agree in the high$z$ region for D^K_{unf}. The favored FF shows important discrepancies with AKK in all z domain while the AKK calculation of D^K_{unf} agrees within $< 36\%$ with COMPASS. In the particular case of the strange quark FF, D^K_S, the fit obtained from this work is not very stable and differs significantly from the existing calculations.

These data represent a high statistics data set which will be included in future global NLO QCD fits of world data. They will constrain the values of quark FFs into pions and kaons at moderate Q^2 values.

Chapter 8

Summary and Conclusion

In order to improve the knowledge on quark fragmentation functions into pions and kaons, the production of charged hadrons in semi-inclusive deep inelastic scattering on a isoscalar target (^{6}LiD) was studied.

The multiplicities of charged unidentified hadrons, pions and kaons have been determined in a fine binning of three kinematical variables: x, the fraction of momentum carried by the quark; y, fractional energy transfer to the virtual photon and z, the energy fraction of the virtual photon transferred to the produced hadron h. The measurement was performed using 6 weeks of 2006 COMPASS data ($\sim 1.1 \times 10^7$ DIS events and a hadron sample $\sim 3.2 \times 10^6$). The kinematic domain covered is the following: $Q^2 > 1$ (GeV/c)^2, $y \in [0.1, 0.9]$, $x \in [0.004, 0.7]$, $W \in [5, 17]$ GeV and $z \in [0.2, 0.85]$.

The hadron identification as pion, kaon and proton was achieved using the Ring Imaging CHerenkov (RICH) detector. The performance of the RICH detector was analyzed using data from the same 6 weeks. The good knowledge of the RICH response is essential to calculate the absolute value of the pion and kaon multiplicities. The criteria used to identify were optimized by tuning the RICH likelihood cuts. This was performed carefully to ensure the best compromise between the identification efficiency and the purity of the sample, specially the kaon purity. The hadron momentum range is restricted to $[12 - 40]$ GeV/c and the polar angle in the range 10 to 120 mrad. The lower cut in hadron momentum was selected to be well above the kaon identification threshold (~ 9.8GeV). The upper limit is to avoid the region where effects arising from saturation ($\beta \to 1$) are observed to ensure a good pion-kaon separation. The polar angle range is selected where the efficiencies are generally high and precisely measured. In this phase-space, high identification efficiencies are found in general for pions (> 0.97) and kaons (> 0.90) and low misidentification probabilities are found ($< 4\%$ for pions and $< 6\%$ for kaons).

The hadron multiplicities have been corrected for the geometrical limitations of the spectrometer and the data reconstruction efficiency. This global acceptance correction is estimated using a MC simulation. The contribution to the hadron yield from the vector-meson production was also estimated and the charged hadron multiplicities are provided with and without this contribution.

The sum of K^+ and K^- multiplicities is of special interest, since it is related, at least at LO in QCD, to a combination of the strange quark distribution $S(x, Q^2)$ in the nucleon and to the strange quark FFs into kaons $D_S^K(z, Q^2)$. The results for $M^{K^+} + M^{K^-}$ are in good agreement with LO calculations only at low Q^2. They will be of high interest to constrain the existing models in the full Q^2 range. To complete the sum of π^+ and π^- multiplicities is also calculated. Results of π^+ and π^- multiplicities ratio are also presented; a fair agreement with HERMES and JLab data is found if compared to a similar Q^2 range.

The π^+ and π^- multiplicities were used to extract the favored and unfavored quark FFs: D_{fav}^{π} and D_{unf}^{π}. this was done via a fit at LO, assuming the PDFs known. Similarly, using K^+ and K^- multiplicities, the favored D_{fav}^K, unfavored D_{unf}^K and strange D_S^K quark FFs were extracted at LO.

The present pion and kaon multiplicity data (about 800 points in total) constitute an important input for further global fit of world data to be done at NLO. In particular, the kaon data enlarge quite significantly the available data set.

Appendices

Appendix A

Invariant Mass Fits for Negative Hadrons

An example of the fits used in the identification and misidentification probabilities determination for negative hadrons. The p_h momentum bin is $17-19$ GeV/c and $0.01-0.04$ rads for the incident angle at RICH entrance θ. As in Section 4.1.2, the negative hadron sample is built with the requirement that the spectator particle (positive hadron) is properly identified by the RICH. The corresponding identification and misidentification probabilities (Eq. (4.2)) for the showed bin are:

$$M_R^- = \begin{pmatrix} 0.99 & 0.00\ (0.00) & 0.92 \\ 0.00 & 0.97\ (0.98) & 0.03 \\ 0.01 & 0.02\ (0.02) & 0.04 \\ 0.00 & 0.00\ (0.00) & 0.00 \end{pmatrix} \tag{A.1}$$

Two different values for the kaons identification and misidentification are shown on the second column in Eq. (A.1). The first values correspond to the probabilities extracted in the ϕ_{1020} inclusive sample and the ones inside the parenthesis come from the ϕ_{1020} exclusive sample.

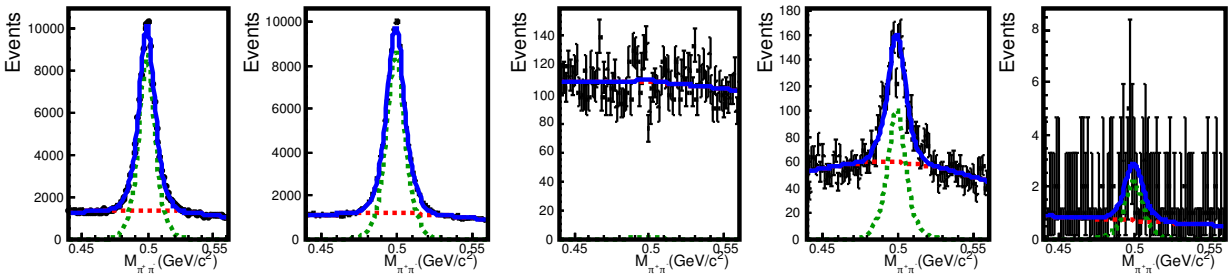


Figure A.1: *K^0_S invariant mass distribution. In green the signal component, in red the background and in blue the sum of both components.*

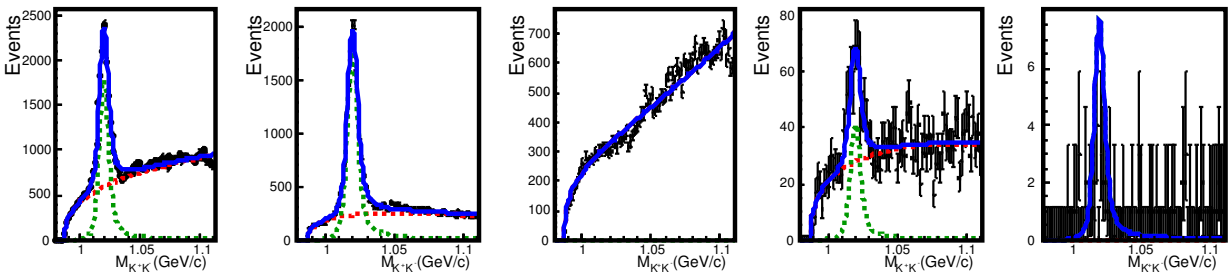


Figure A.2: *Same as fig. A.1 for ϕ^{incl}_{1020}.*

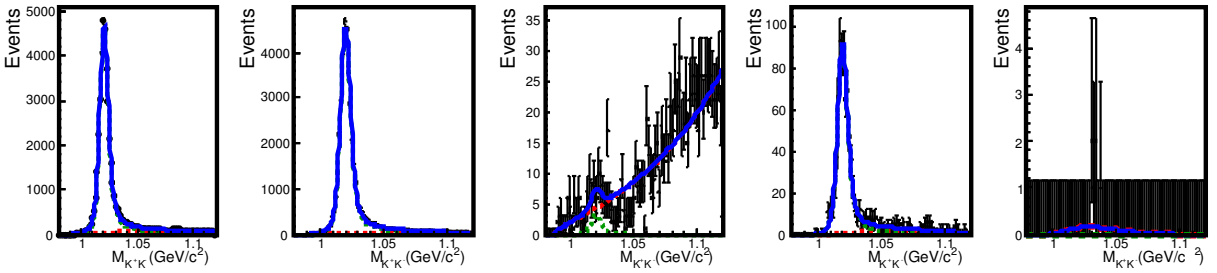


Figure A.3: *Same as fig. A.1 for ϕ^{excl}_{1020}.*

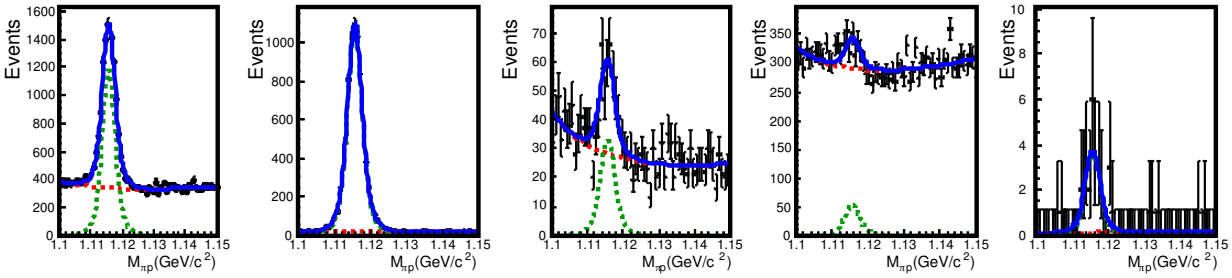


Figure A.4: *Same as fig. A.1 for Λ.*

Appendix B

Ratio of Identified Hadrons.

In order to optimize the kaon identification with the RICH detector, the purity of the different hadron samples (pion, kaon and protons) is needed. The sample purity (Eq. (4.13)) is defined in terms of the identification and misidentification probabilities and the ratios: $\frac{I_\pi}{I_K}$, $\frac{I_\pi}{I_p}$ and $\frac{I_K}{I_p}$, where I_h refers to number of observed hadrons. In this appendix, these ratios are shown in each (x, y) bin used in the hadron multiplicities analysis as a function of the particle momentum p_h and the incident angle at RICH entrance θ in Figs. B.1 to B.3 respectively.

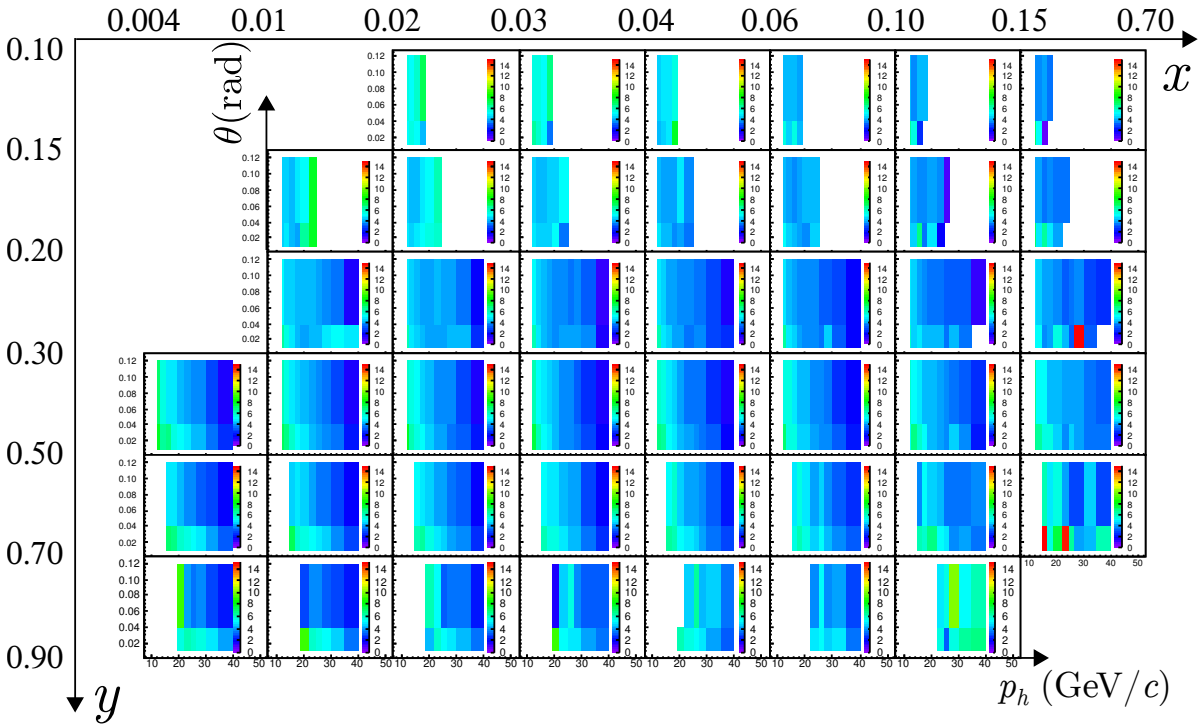


Figure B.1: *The ratio* $\frac{I_\pi}{I_K}$ *as a function of* p_h *and* θ *in the* x, y *binning used in the hadron multiplicities analysis.*

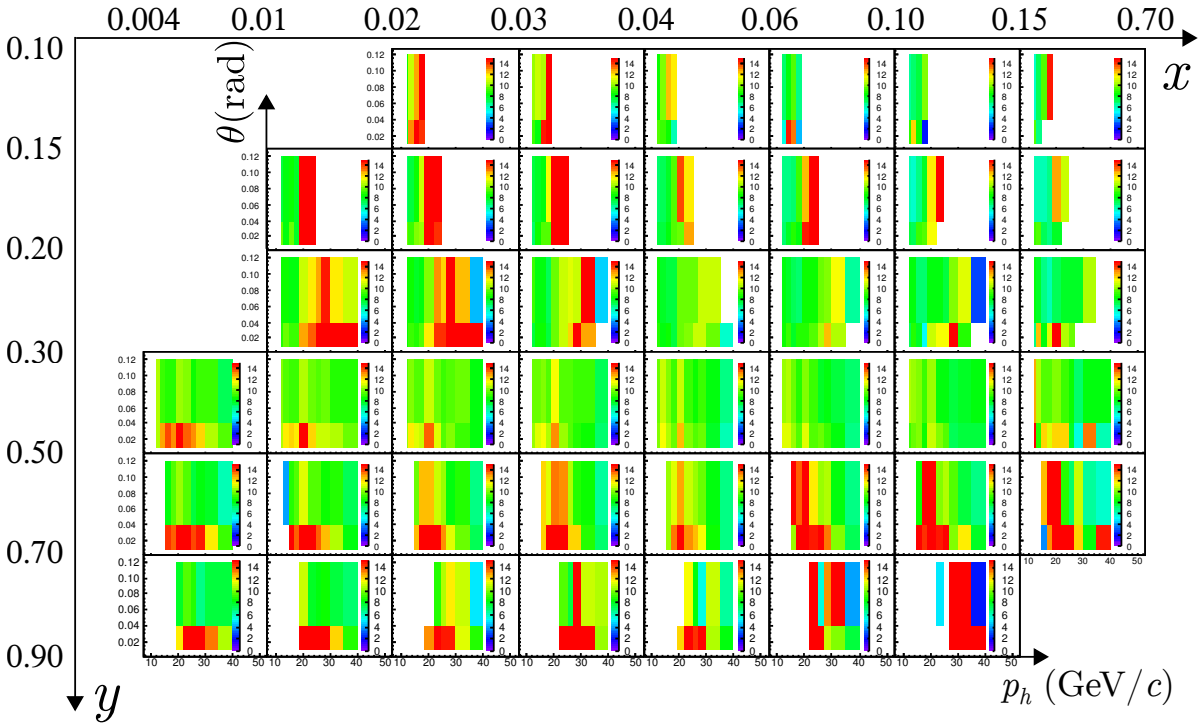


Figure B.2: *Same as Fig. B.1 for* $\frac{I_\pi}{I_p}$.

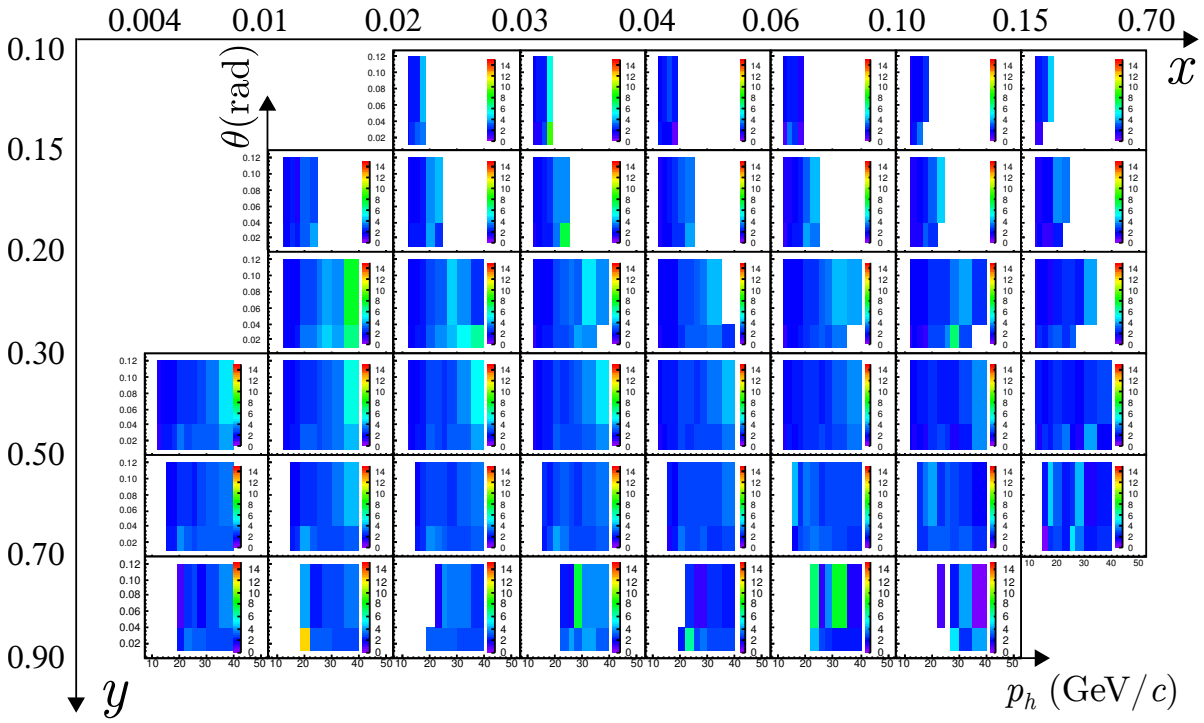


Figure B.3: *Same as Fig. B.1 for* $\frac{I_K}{I_p}$.

Appendix C

RICH Performances Matrix in Different Binning

C.1 Binning: $p_h(5) \times \theta(6)$

In this section the RICH performance results for the binning showed in Table C.1 are discussed.

Table C.1: *Binning in p_h and θ used to extract the RICH response.*

p_h (GeV/c)	$9.6, 13, 18, 25, 30, 50$
θ (rad)	$0.00, 0.01, 0.02, 0.03, 0.04, 0.06, 0.4$

In this binning an analysis of the influence of the cut of RICH pipe was performed. In figs. C.1 and C.2 the identification and misidentification efficiencies for positive and negative pions and kaons are shown as a function of p_h in different θ bins., the red markers correspond to $P(h \to h)$ when all the incoming particles at RICH are taking into account and the green ones when the particles near to the RICH pipe are eliminated. For π^+ and π^-, in the first θ bin (0.0,0.01) the identification efficiencies with and without the cut varies as a function of p_h between $\sim 0.20 - 0.45$. For the (0.01,0.02) bin this dependence decrease and only differs $\sim$0.10 for the first p_h bin (9.8,10) and is not visible any more for high p_h values (figs. C.1a and C.1b). The presence of the RICH pipe does not have any influence in $P(\pi \to K)$ as it can see in figs. C.1c and C.1d. For kaons a similar analysis was performed,but different distance from the RICH pipe where also studied. For $P(K \to K)$ (figs. C.2a and C.2b) for the (0.0,0.01) θ bin, for the first p_h bin the RICH pipe cut increase the identification value by ~ 0.7 for K^+ and ~ 0.15 for K^-, for the subsequent p_h bins the difference increase an can reach ~ 0.55. Contrary for pions the identification values are still affected by the RICH pipe presence in the second θ bin for all p_h bins. Regarding the

different distance for RICH pipe there is not an improvement in the values. The $P(K \to \pi)$ values are also sensitive to the pipe, but the difference are < 0.05. Only in the last p_h bin in $(0.00, 0.01)$ there is a difference between $0.10 - 0.25$ for the different distance, but all them lies within the error bars.

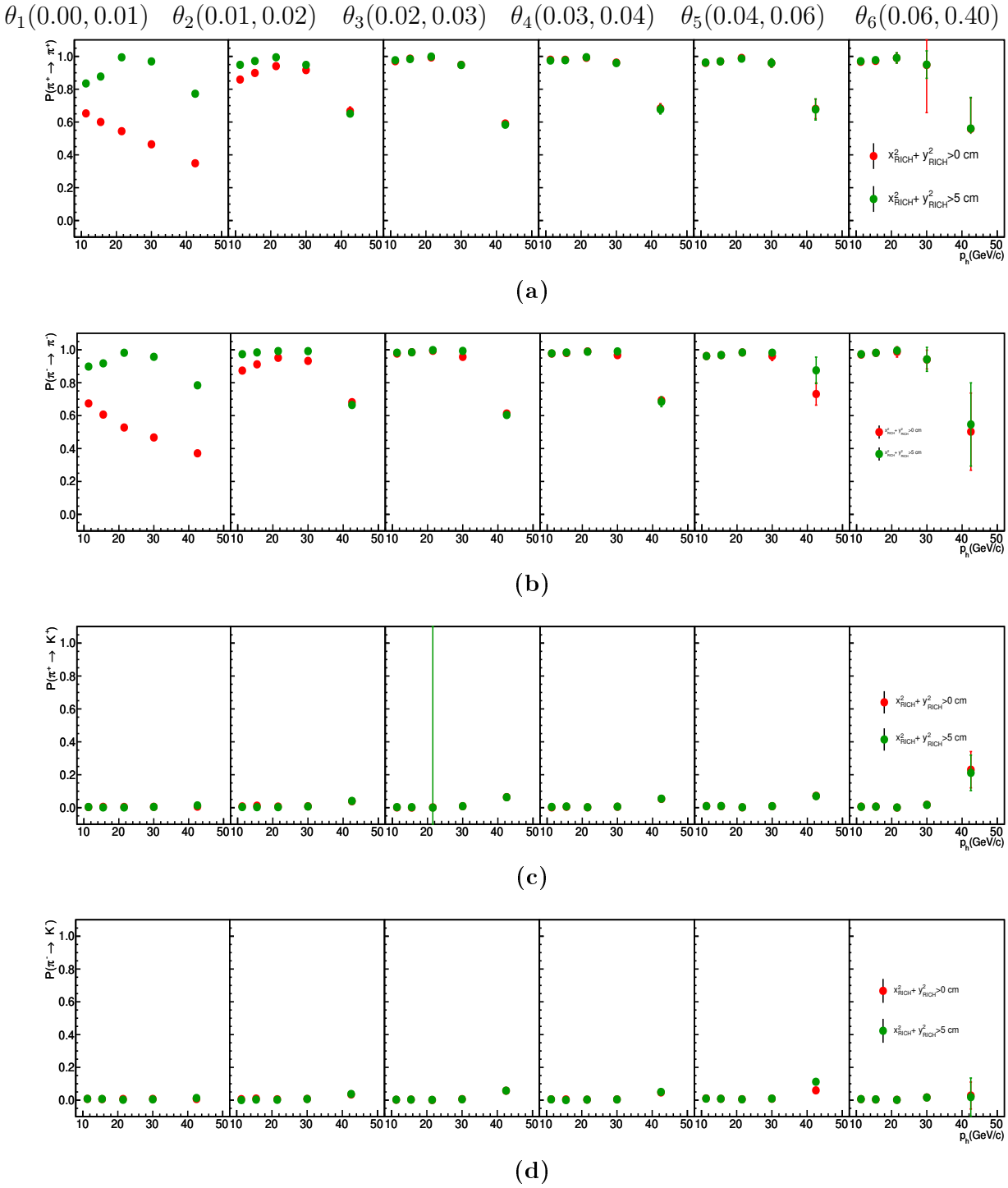
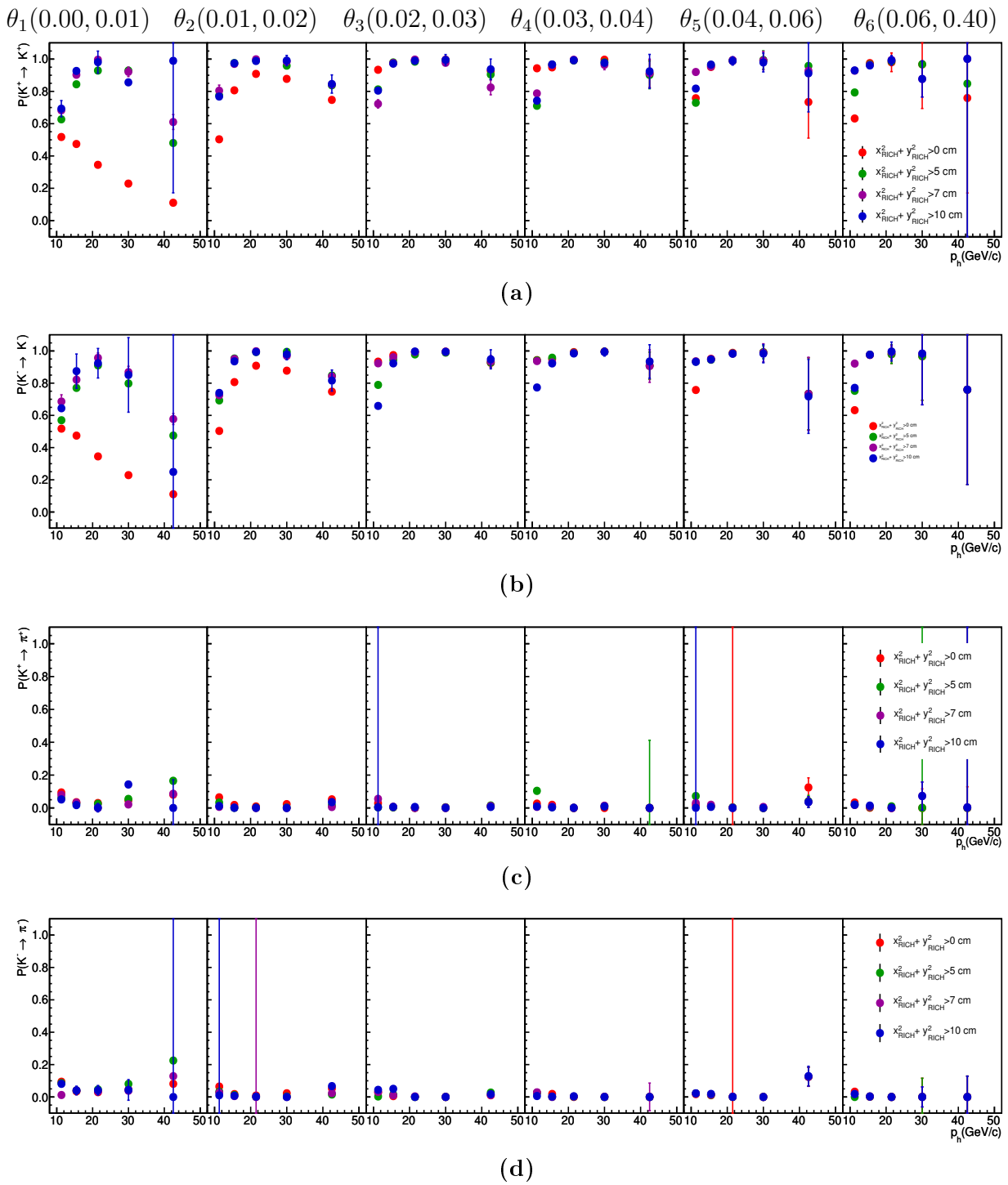


Figure C.1: *Pion RICH probabilities (a and b identification, c and d misidentification) as a function of p_h in six θ bins (each column corresponds to a different θ bin). The green markers correspond to the RICH efficiencies when the cut $x^2_{RICH} + y^2_{RICH} > 5^2 cm$ is applied and the red ones when any cut is applied.*



(a)

(b)

(c)

(d)

Figure C.2: *Same as fig. C.1 for kaons.*

C.2 Binning: $p_h(6) \times \theta(6)$

This binning was defined (Table C.2) in a previous multiplicity analysis [76]. The RICH efficiency matrices (M_R) were extracted in this binning in order to made a comparison between 2004 and 2006 multiplicity results.

Table C.2: *Binning in p_h and θ used to extract the RICH response.*

p_h (GeV/c)	$10, 15, 20, 25, 30, 40, 50$
θ (rad)	$0.01, 0.02, 0.03, 0.04, 0.06, 0.09, 0.12$

The first θ bin ($0 < \theta \leq 0.01$) is no showed anymore. In addition it was found that in a few bins, the statistics is too low to perform a correct analysis (summarized in table C.3). In those bins, the identification and misidentification values were taken equal to the neighboring bin (for example p_h(40,50) θ(0.09,0.12) = p_h(40,50) θ).

Table C.3: *List of bins where the statistics is too low to perform a good fit.*

	$p_h(6) \times \theta(7)$
π^+	$p_h(40, 50)$ $\theta(0.09, 0.12)$
K	$p_h(40, 50)$ $\theta(0.04, 0.06)$
	$p_h(30, 40)$ $\theta(0.06, 0.09)$
	$p_h(40, 50)$ $\theta(0.06, 0.09)$
	$p_h(20, 25)$ $\theta(0.09, 0.12)$
	$p_h(25, 30)$ $\theta(0.09, 0.12)$
	$p_h(30, 40)$ $\theta(0.09, 0.12)$
	$p_h(40, 50)$ $\theta(0.09, 0.12)$
p	$p_h(40, 50)$ $\theta(0.06, 0.09)$
	$p_h(30, 40)$ $\theta(0.09, 0.12)$
	$p_h(40, 50)$ $\theta(0.09, 0.12)$

Results for the M_R matrix elements are shown in figs. C.3 to C.5. The pion identification probabilities values are > 0.95 for p_h between 10-30 GeV/c in all θ bins. For $30 < p_h < 40$ GeV/c the values varying with θ between 0.80-0.93. In the last momentum bin the performance drops significantly to ~ 0.5 or lower values (figs. C.3a and C.3d). The behavior of the misidentification probability of pions into kaons ($P(\pi \to K)$) is highly correlated with the identification values (figs. C.3b and C.3e) if the identification values are high ($p_h < 30$) the pions misidentification as kaon will be low and viceversa. The misidentification of pions into protons ($P(\pi \to \mathrm{p})$) it is constantly below 0.01 for almost all bins. However

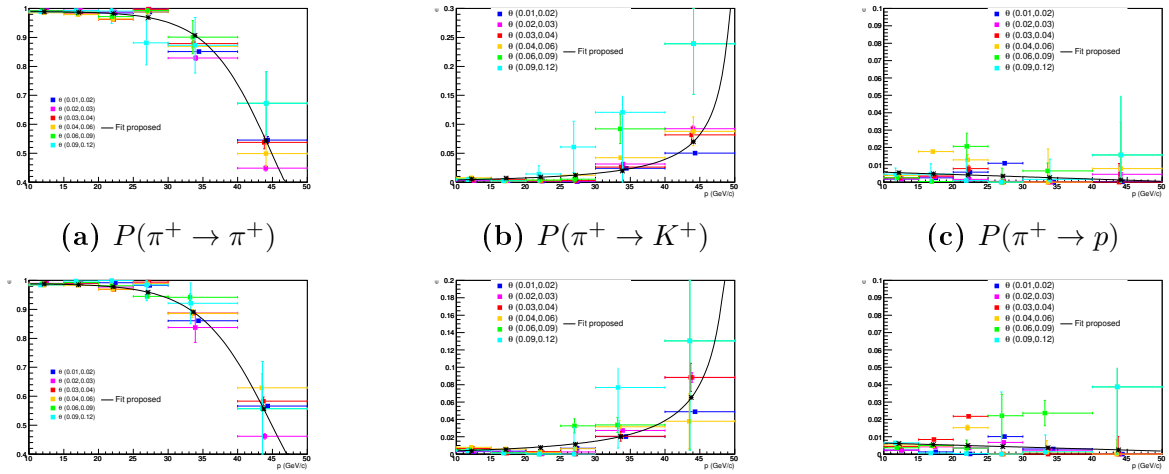
in the $20 < p_h < 25$ GeV/c the values vary an can reach values > 0.02, this can be explained because the K^0 sample used was contaminated by protons because it did not have $|M_\Lambda| < 10\text{MeV}/c^2$ cut.

For kaons in the first p_h bin (10-15 GeV/c), $P(K \to K)$ vary between 0.95-0.98 and is smaller in comparison with the subsequent bins values ($p_h < 30$ GeV/c)(figs. C.4a and C.4d). For higher p_h bins the efficiency drops to 0.75. The misidentification efficiency values are below 0.03 wihtin error bars for all bins. However in the last θ bin $P(K \to h)$ vary 0-0.09, the instability in this bin is related to the reduce number of K with high p_h and θ values (figs. C.4b and C.4f).

For protons, in the two first p_h bins $P(p \to p)$ vary between 0.38-0.60, $P(p \to \pi)$ between 0.0-0.50 and $P(p \to K)$ between 0.10-0.35.

Within error bars no deviation of the θ dependence is observed for any $P(h \to h)$ distribution, whence is possible to use a single parametrization in p_h averaged over all θ bins. The final parametrization is shown in black in Figs. C.3 to C.5 for pions, kaons and protons. In the pions case, the parametrization for $P(\pi \to \pi)$ and $P(\pi \to K)$ probabilities are performed using a Fermi function (figs. C.3a to C.3e). For the misidentification $P(\pi \to p)$ (figs. C.3c and C.3f) a first order polynomial was used. For the kaon identification probability (figs. C.4a and C.4d) the product of an exponential and a first order polynomial was used. The $1 - exp * pol(1)$ function was used for fit the misidentification probability to identify kaons as pions (figs. C.4b and C.4e). The misidentification of kaons as protons was fitted with a second order polynomial (figs. C.4c and C.4f). In the case of the proton identification and misidentification probabilities, two regions were considered: below and above proton threshold. For $P(p \to p)$, $P(\pi \to p)$ and $P(p \to K)$ probabilities in the region above threshold, a smoothed Heaviside step function was used for the fit. For the region below threshold a constant was used to fit the points on $P(p \to p)$ and $P(\pi \to p)$ while for $P(p \to K)$ a first order polynomial was used (figs. C.5a to C.5f).

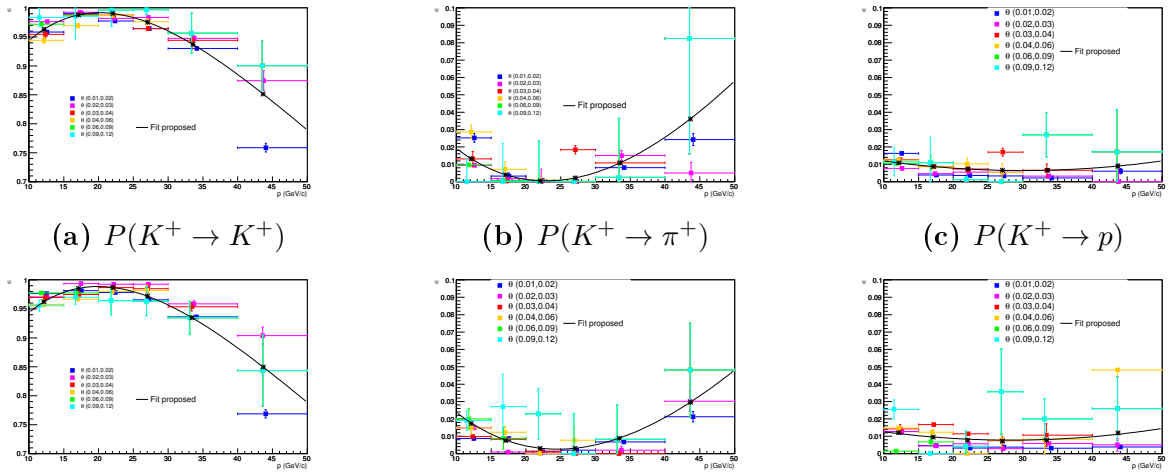
The results obtained from the parametrization at the central value of the p_h bins are given in tables Tables C.4 to C.6. The errors shown in the tables correspond to the RMS of the values obtained for the various p_h bins. These values were used in [76] to perform the pion and kaon identification.



(a) $P(\pi^+ \to \pi^+)$ **(b)** $P(\pi^+ \to K^+)$ **(c)** $P(\pi^+ \to p)$

(d) $P(\pi^- \to \pi^-)$ **(e)** $P(\pi^- \to \pi^-)$ **(f)** $P(\pi^- \to p)$

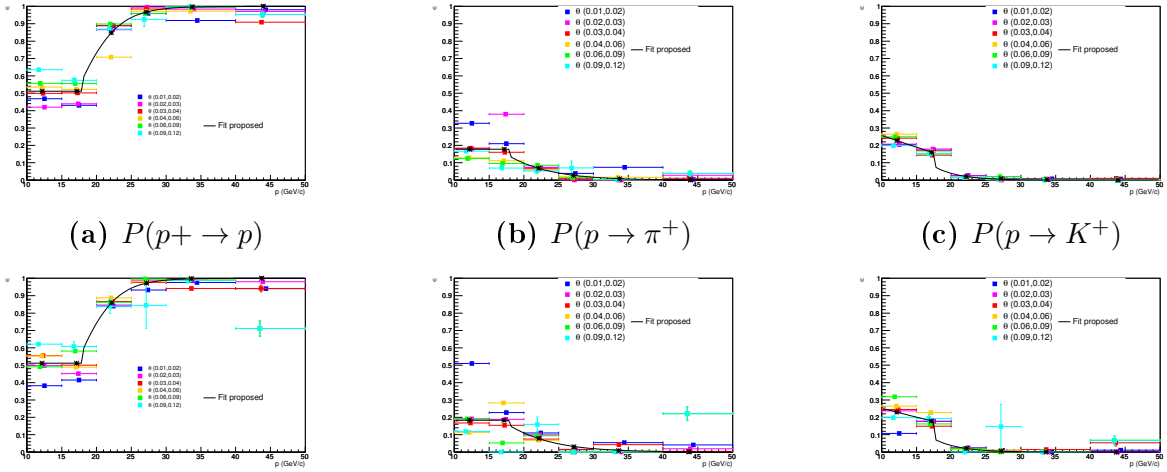
Figure C.3: *Values for RICH identification and misidentification probabilities* $P(\pi \to h)$*, as a function of* p_h *for various* θ *bins. Each color corresponds to a different* θ *bin. The solid line correspond to the fit function (see text) and the black points correspond to the fit value at the central value of the* p_h *bin*



(a) $P(K^+ \to K^+)$ **(b)** $P(K^+ \to \pi^+)$ **(c)** $P(K^+ \to p)$

(d) $P(K^- \to K^-)$ **(e)** $P(K^- \to \pi^-)$ **(f)** $P(K^- \to \bar{p})$

Figure C.4: *Same as fig. C.3 for* $P(K \to h)$



(a) $P(p+ \to p)$ (b) $P(p \to \pi^+)$ (c) $P(p \to K^+)$

(d) $P(\bar{p} \to \bar{p})$ (e) $P(\bar{p} \to \pi^-)$ (f) $P(\bar{p} \to K^-)$

Figure C.5: *Same as fig. C.3 for* $P(p \to h)$

Table C.4: *RICH response parametrization for pions.*

π^-			
Region	$FIT(\pi^- \to \pi^-)$	$FIT(\pi^- \to K^-)$	$FIT(\pi^- \to \bar{p})$
10-15 GeV/c	0.988 ± 0.0028	0.0037 ± 0.0024	0.0061 ± 0.0027
15-20 GeV/c	0.9852 ± 0.0077	0.0053 ± 0.0034	0.0056 ± 0.0033
20-25 GeV/c	0.9779 ± 0.0128	0.0077 ± 0.0073	0.005 ± 0.0094
25-30 GeV/c	0.9594 ± 0.03	0.0114 ± 0.0129	0.0044 ± 0.0091
30-40 GeV/c	0.8911 ± 0.0384	0.0200 ± 0.0267	0.0036 ± 0.0092
40-50 GeV/c	0.5562 ± 0.0549	0.0658 ± 0.0356	0.0024 ± 0.0164
π^+			
	$FIT(\pi^+ \to \pi^+)$	$FIT(\pi^+ \to K^+)$	$FIT(\pi^+ \to p)$
10-15 GeV/c	0.9887 ± 0.0028	0.0055 ± 0.0017	0.0054 ± 0.0027
15-20 GeV/c	0.9871 ± 0.0051	0.0069 ± 0.0042	0.0048 ± 0.0057
20-25 GeV/c	0.9824 ± 0.0127	0.009 ± 0.0061	0.0042 ± 0.0079
25-30 GeV/c	0.9688 ± 0.0420	0.0120 ± 0.0214	0.0035 ± 0.0043
30-40 GeV/c	0.9069 ± 0.0459	0.0195 ± 0.0520	0.0026 ± 0.0027
40-50 GeV/c	0.5433 ± 0.0745	0.0699 ± 0.0774	0.0013 ± 0.0072

Table C.5: *RICH response parametrization for kaons.*

K^-			
Region	$FIT(K^- \to K^-)$	$FIT(K^- \to \pi^-)$	$FIT(K^- \to \bar{p})$
10-15 GeV/c	0.9623 ± 0.0176	0.0173 ± 0.0087	0.0120 ± 0.0124
15-20 GeV/c	0.9855 ± 0.0206	0.0076 ± 0.015	0.0095 ± 0.0102
20-25 GeV/c	0.9871 ± 0.0184	0.0030 ± 0.0144	0.0078 ± 0.0075
25-30 GeV/c	0.9719 ± 0.0127	0.0030 ± 0.0032	0.0072 ± 0.0129
30-40 GeV/c	0.9349 ± 0.0151	0.0089 ± 0.0057	0.0077 ± 0.0029
40-50 GeV/c	0.8496 ± 0.0689	0.0296 ± 0.0144	0.0119 ± 0.0124
K^+			
	$FIT(K^+ \to K^+)$	$FIT(K^+ \to \pi^+)$	$FIT(K^+ \to p)$
10-15 GeV/c	0.9631 ± 0.0137	0.0132 ± 0.0099	0.0108 ± 0.0028
15-20 GeV/c	0.9881 ± 0.0079	0.0039 ± 0.0030	0.0087 ± 0.0029
20-25 GeV/c	0.9905 ± 0.0070	0.0005 ± 0.0004	0.0073 ± 0.0033
25-30 GeV/c	0.9755 ± 0.0124	0.0021 ± 0.0075	0.0065 ± 0.0057
30-40 GeV/c	0.937 ± 0.0122	0.0108 ± 0.0048	0.0066 ± 0.0106
40-50 GeV/c	0.8515 ± 0.0759	0.0361 ± 0.0403	0.009 ± 0.0089

Table C.6: *RICH response parametrization for protons.*

$\bar{p}$			
Region	$FIT(\bar{p} \to \bar{p})$	$FIT(\bar{p} \to \pi^-)$	$FIT(\bar{p} \to K^-)$
10-15 GeV/c	0.5119 ± 0.0872	0.23 ± 0.0917	0.1832 ± 0.0331
15-20 GeV/c	0.5119 ± 0.0706	0.1792 ± 0.0652	0.1832 ± 0.0151
20-25 GeV/c	0.8377 ± 0.0799	0.0207 ± 0.0508	0.0790 ± 0.0177
25-30 GeV/c	0.9484 ± 0.0387	0.0057 ± 0.0248	0.0301 ± 0.0109
30-40 GeV/c	0.9722 ± 0.015	0.0011 ± 0.0147	0.0084 ± 0.011
40-50 GeV/c	0.9749 ± 0.0282	0.0001 ± 0.0312	0.0011 ± 0.0047
p			
	$FIT(p \to p)$	$FIT(p \to \pi^+)$	$FIT(p \to K^+)$
10-15 GeV/c	0.5116 ± 0.0694	0.2271 ± 0.0645	0.1781 ± 0.0293
15-20 GeV/c	0.5116 ± 0.0645	0.1607 ± 0.0493	0.1781 ± 0.0147
20-25 GeV/c	0.8319 ± 0.0505	0.0199 ± 0.0272	0.0682 ± 0.0136
25-30 GeV/c	0.9489 ± 0.0593	0.0083 ± 0.0626	0.0273 ± 0.0252
30-40 GeV/c	0.9723 ± 0.0131	0.0026 ± 0.0143	0.0078 ± 0.0041
40-50 GeV/c	0.9805 ± 0.0251	0.0004 ± 0.0306	0.0011 ± 0.0089

Appendix D

Individual Periods Statistics

D.1 Data statistics

The number of DIS event and unidentified hadrons (with radiative correction) of each of the used periods is summarized in Table D.1. The total number for identified pions, kaons and protons is in Table D.2 (with radiative correction and after RICH unfolding).

Week	DIS	$h+$	$h-$
W40	2857330.79	440391.88	354876.46
W41	1938510.04	296539.50	239834.93
W42	1980746.22	303702.61	244895.8
W43	2122197.34	324914.26	261165.03
W45	1774610.90	271661.05	217152.38
W46	1132140.77	172852.49	139086.36

Table D.1: *Number of DIS events and unidentified hadrons for the 6 periods analyzed.*

D.2 Systematic study: Stability over time

To finalize the time systematic error analysis the ratio between the multiplicities extracted from the analysis of W40_W41_W45 and W42_W43_W45 is fitted with a constant (a) in all (x, y, z) bins. The results from the fit are summarized in Table D.3 for unidentified hadrons and pions and in Table D.4 for kaons. The compatibility between the two samples is at the per mile level within statistical fluctuations. Consequently, no systematic error will be assigned for the data compatibility.

Week	$\pi+$	$\pi-$	$K+$	$K-$	p	$\bar{p}$
W40	323004.38	286993.80	81053.01	52040.88	36486.67	15955.97
W41	215228.4	192037.09	55614.62	35637.89	24199.02	10822.50
W42	220418.03	196468.61	57030.74	36352.00	24891.15	10930.14
W43	235170.03	209401.41	61718.96	38735.70	26503.87	11617.13
W45	196350.17	173806.99	51581.39	32363.74	22310.72	9694.6
W46	125320.49	111588.42	32647.82	20614.96	14175.98	6285.45

Table D.2: *Number of identified hadrons for the 6 periods analyzed.*

x	0.004-0.01		0.01-0.02		0.02-0.03		0.03-0.04		0.04-0.06		0.06-0.1		0.1-0.15		0.15-0.7	
y	a	δa	a	δa	a	δa	a	δa	a	δa	a	δa	a	δa	a	δa
$h+$																
0.10-0.15		—		—	0.994	0.017	0.989	0.015	0.986	0.017	0.99	0.018	1.004	0.026	1.03	0.036
0.15-0.20		—	0.982	0.018	0.999	0.008	0.992	0.011	0.986	0.010	0.992	0.009	0.991	0.012	0.988	0.018
0.20-0.30		—	0.99	0.005	1.002	0.006	1.004	0.007	0.992	0.006	0.994	0.006	1.004	0.008	1.004	0.011
0.30-0.50	0.993	0.006	1.002	0.003	1.005	0.004	1.001	0.005	0.988	0.004	0.985	0.005	0.993	0.008	0.993	0.010
0.50-0.70	0.995	0.005	1.004	0.005	0.991	0.007	0.979	0.010	1.031	0.011	0.982	0.013	0.935	0.023	1.039	0.043
0.70-0.90	0.992	0.006	0.99	0.009	1.018	0.020	0.939	0.029	1.059	0.035	0.944	0.044	1.064	0.137	—	
$h-$																
0.10-0.15		—		—	1.000	0.020	0.974	0.018	0.975	0.018	1.003	0.025	0.960	0.036	0.982	0.053
0.15-0.20		—	0.958	0.02	0.995	0.010	0.997	0.013	0.995	0.012	0.994	0.012	0.963	0.019	0.971	0.026
0.20-0.30		—	1.001	0.006	0.989	0.007	0.982	0.008	1.002	0.007	0.987	0.007	0.985	0.010	1.004	0.014
0.30-0.50	0.995	0.006	0.99	0.003	0.999	0.004	0.999	0.005	0.997	0.005	1.008	0.006	1.025	0.010	1.043	0.014
0.50-0.70	1.002	0.005	1.001	0.005	1.003	0.008	0.989	0.011	0.976	0.012	1.015	0.017	1.026	0.031	1.054	0.056
0.70-0.90	1.009	0.007	1.026	0.011	1.007	0.022	0.997	0.037	1.049	0.042	1.110	0.067	1.124	0.213	—	
π^+																
0.10-0.15		—		—	0.970	0.020	0.984	0.018	0.983	0.018	0.991	0.023	1.026	0.033	1.027	0.046
0.15-0.20		—	1.000	0.023	1.000	0.011	0.973	0.014	0.989	0.013	1.002	0.012	0.981	0.018	0.969	0.023
0.20-0.30		—	0.988	0.007	1.003	0.008	0.999	0.009	0.992	0.008	0.981	0.008	0.995	0.011	1.008	0.014
0.30-0.50	0.992	0.007	1.001	0.004	1.002	0.005	1.004	0.006	0.987	0.006	0.979	0.006	0.999	0.010	0.983	0.013
0.50-0.70	0.999	0.006	1.001	0.006	0.982	0.010	0.976	0.013	1.033	0.014	0.949	0.018	0.928	0.029	1.002	0.056
0.70-0.90	0.980	0.008	0.982	0.013	1.023	0.028	0.911	0.039	1.071	0.049	0.975	0.063	1.068	0.193	—	

Table D.3: *Results for the constant fit for the ratio between the extracted multiplicities from periods* 40_41_45 *and* 42_W43_W45, *for the* $8x \times 6y$ *bins*

x	0.004-0.01		0.01-0.02		0.02-0.03		0.03-0.04		0.04-0.06		0.06-0.1		0.1-0.15		0.15-0.7	
y	a	δa	a	δa	a	δa	a	δa	a	δa	a	δa	a	δa	a	δa
π^-																
0.10-0.15		—		—	0.995	0.022	0.988	0.020	0.970	0.020	1.005	0.028	0.963	0.039	0.968	0.057
0.15-0.20		—	0.972	0.024	0.989	0.011	0.990	0.016	0.993	0.015	0.991	0.014	0.962	0.021	0.984	0.029
0.20-0.30		—	1.009	0.007	0.986	0.008	0.981	0.010	1.002	0.009	0.987	0.009	0.992	0.012	1.003	0.017
0.30-0.50	0.995	0.008	0.990	0.004	0.996	0.005	1.008	0.007	0.997	0.006	1.004	0.007	1.027	0.012	1.030	0.016
0.50-0.70	0.998	0.006	0.998	0.006	1.004	0.010	0.992	0.014	0.970	0.015	1.013	0.021	0.995	0.038	0.995	0.064
0.70-0.90	1.003	0.009	1.011	0.014	1.032	0.031	0.992	0.049	1.083	0.058	1.051	0.083	1.053	0.243		—
K^+																
0.10-0.15		—		—	1.074	0.084	1.032	0.091	0.902	0.070	1.030	0.077	0.992	0.028	0.950	0.128
0.15-0.20		—	0.935	0.065	0.977	0.031	1.081	0.045	0.972	0.035	0.985	0.032	1.003	0.014	0.990	0.049
0.20-0.30		—	1.004	0.018	1.012	0.020	1.022	0.024	1.003	0.020	1.004	0.019	0.971	0.010	0.974	0.031
0.30-0.50	0.998	0.021	1.007	0.011	0.981	0.015	1.038	0.020	0.987	0.017	1.016	0.019	1.002	0.012	1.077	0.044
0.50-0.70	0.996	0.016	1.019	0.016	1.038	0.027	0.980	0.033	1.015	0.035	1.071	0.050	1.023	0.034	0.991	0.151
0.70-0.90	0.997	0.029	1.010	0.036	0.895	0.068	0.966	0.113	0.842	0.115	0.821	0.155	1.064	0.239		—
K^-																
0.10-0.15		—		—	0.892	0.123	0.806	0.135	0.776	0.122	0.870	0.146	0.868	0.225	1.245	0.585
0.15-0.20		—	0.863	0.079	1.009	0.045	1.029	0.062	0.995	0.054	1.036	0.057	0.899	0.071	0.820	0.091
0.20-0.30		—	0.910	0.021	0.999	0.025	0.981	0.032	1.016	0.028	0.954	0.027	0.999	0.042	1.051	0.077
0.30-0.50	1.009	0.023	1.000	0.013	0.990	0.019	0.954	0.023	0.976	0.022	1.026	0.026	1.004	0.046	0.972	0.063
0.50-0.70	1.035	0.019	0.989	0.018	1.004	0.029	0.966	0.040	0.947	0.043	0.983	0.064	1.121	0.125	1.175	0.278
0.70-0.90	1.033	0.032	1.069	0.042	0.918	0.076	1.030	0.180	0.898	0.144	1.071	0.263	0.897	0.851		—

Table D.4: *Results for the constant fit for the ratio between the extracted multiplicities from periods* 40_41_45 *and* 42_43_45*, for the* $8x \times 6y$ *bins*

Appendix E

Radiative Corrections

The value of the radiative correction for the DIS events is summarized in table E.1 for the (x, y) binning used on the hadron multiplicities analysis. The radiative corrections for the unidentified hadrons, pions and kaons are represented in figs. E.1 to E.3 respectively.

y/x	0.004-0.01	0.01-0.02	0.02-0.03	0.03-0.04	0.04-0.06	0.06-0.1	0.1-0.15	0.15-0.7
0.10-0.15	—	—	0.937	0.944	0.950	0.96	0.972	0.99
0.15-0.20	—	0.925	0.931	0.937	0.943	0.952	0.963	0.983
0.20-0.30	—	0.916	0.922	0.927	0.933	0.942	0.953	0.973
0.30-0.50	0.886	0.898	0.905	0.911	0.916	0.925	0.937	0.956
0.50-0.70	0.850	0.864	0.875	0.881	0.888	0.898	0.912	0.93
0.70-0.90	0.767	0.798	0.812	0.823	0.832	0.85	0.878	0.896

Table E.1: *Radiative corrections values for the DIS events for* $8x \times 6y$ *bins.*

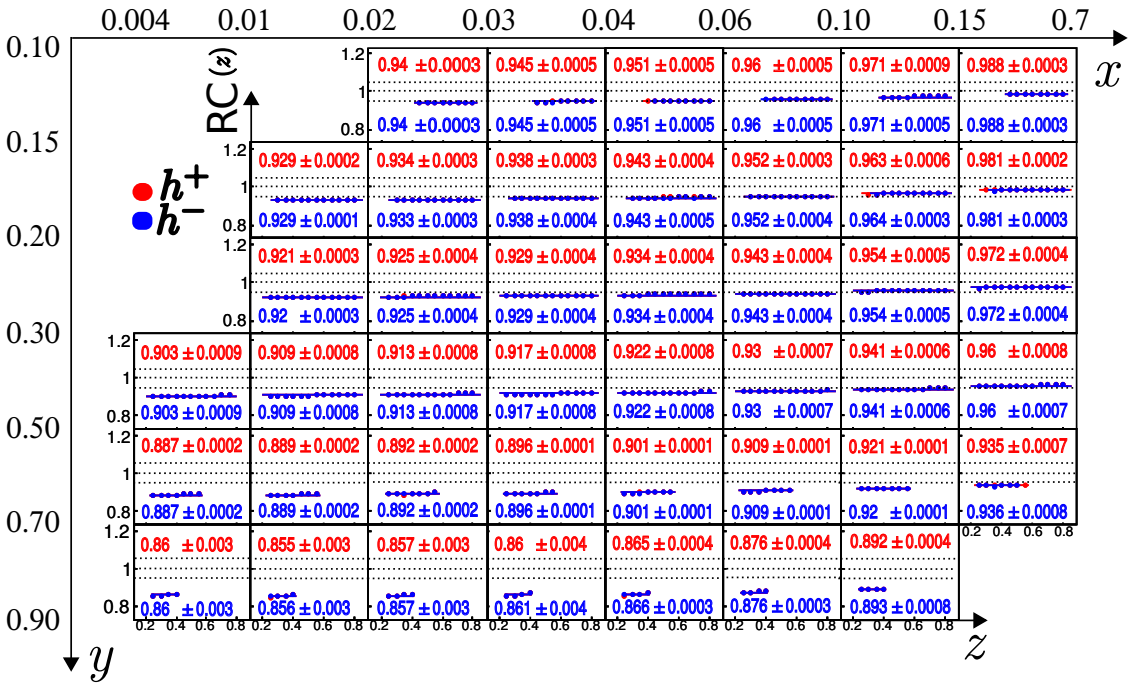


Figure E.1: *Hadron radiative corrections*

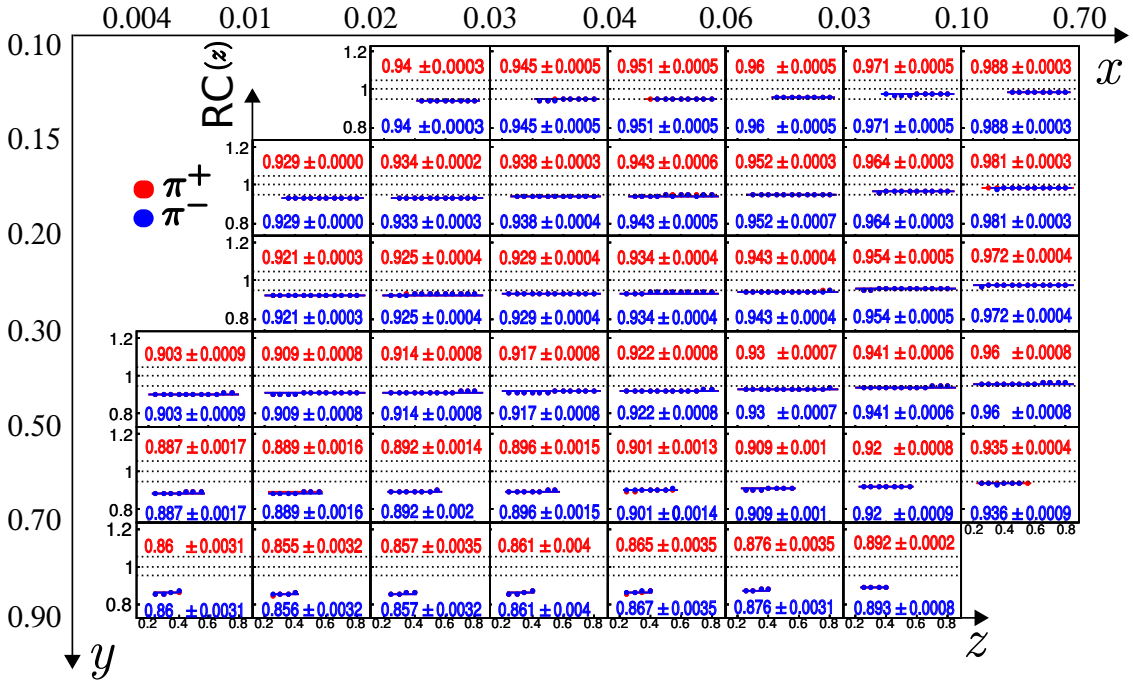


Figure E.2: *Same as fig. E.1 for* π.

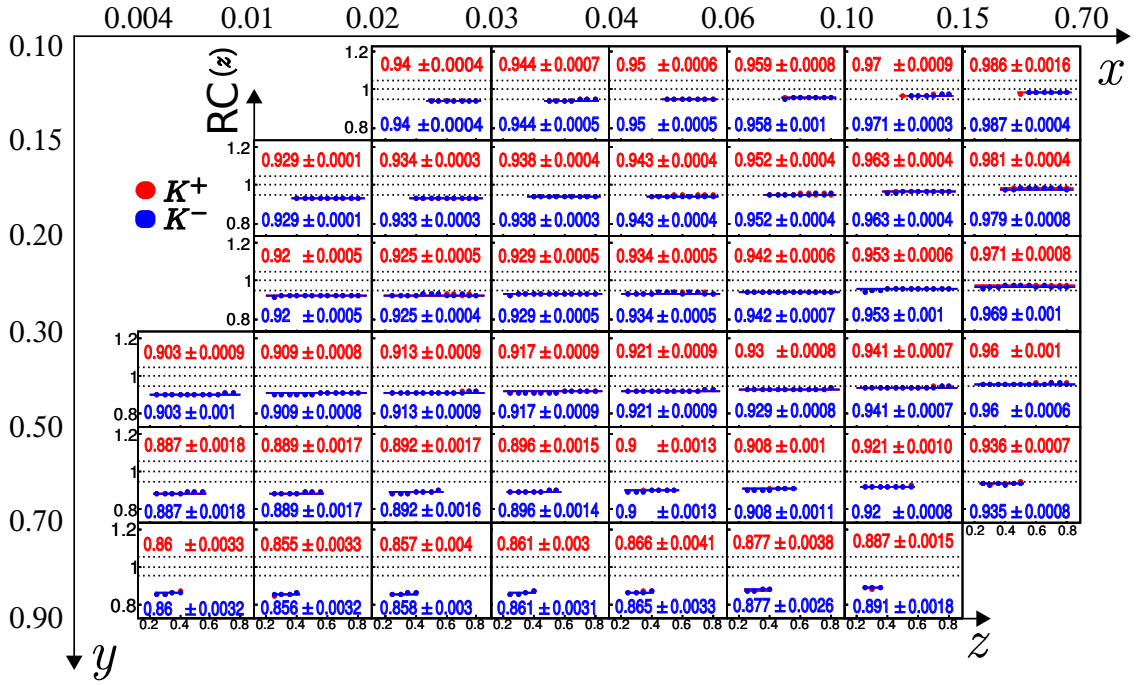


Figure E.3: *Same as fig. E.1 for K.*

Appendix F

Systmatic Errors

The ration between the systematic and the statistical errors is shown for unidentified hadrons, pions and kaon in figs. F.1 to F.3, respectively. In all cases the systematics are larger than the statistical errors.

The contribution to the systematics is shown for unidentified hadrons in fig. F.4. In this case three different contribution are considered: the acceptance, the contamination from the vector-meson decay and the extrapolation to the unmeasured momentum. The acceptance is the dominant error in almost all bins. The contribution from the vector-meson contribution becomes important for high z. For pions and kaons (figs. F.5 and F.6, respectively) the systematic from the RICH detector is also considered. However, the contibution from RICH identification is not as important as the acceptance and the vector-meson ones.

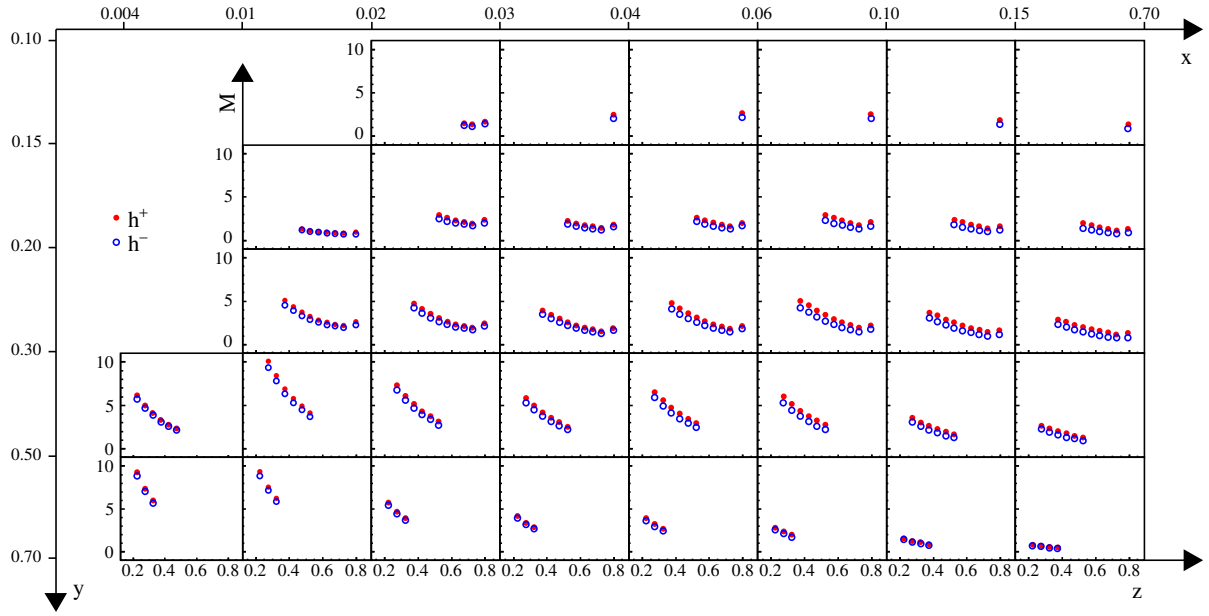


Figure F.1: *Ratio between the systematic and the statistical errors for unidentified hadrons as a function of z in different x and y bins. Red markers: h^+. Blue markers: h^-.*

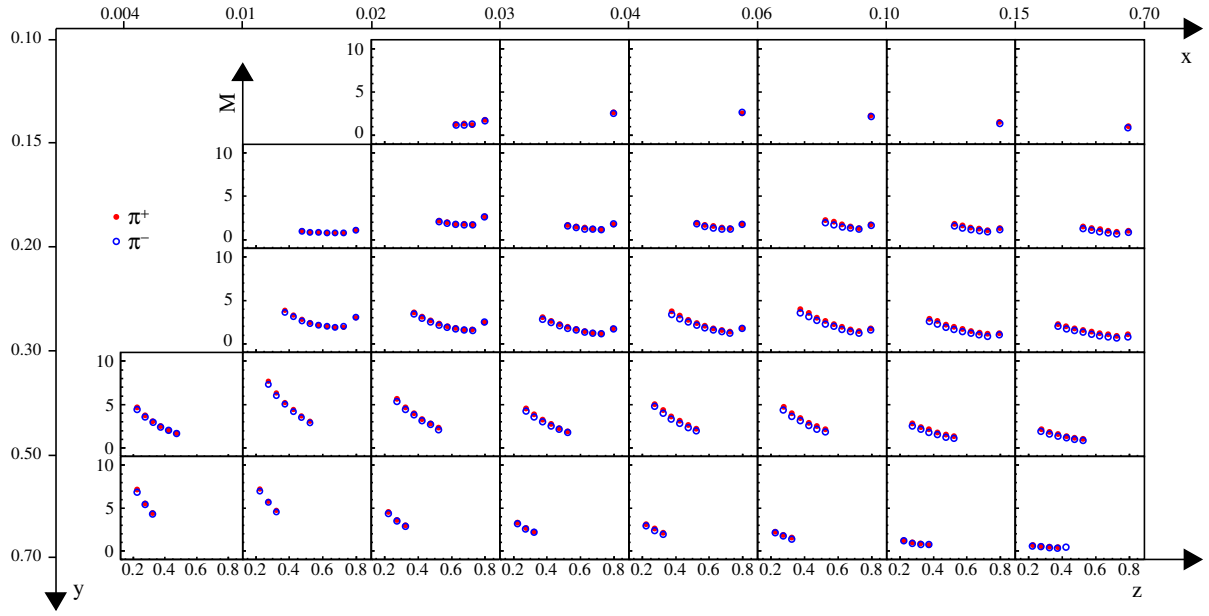


Figure F.2: *Same as fig. F.1 for pions.*

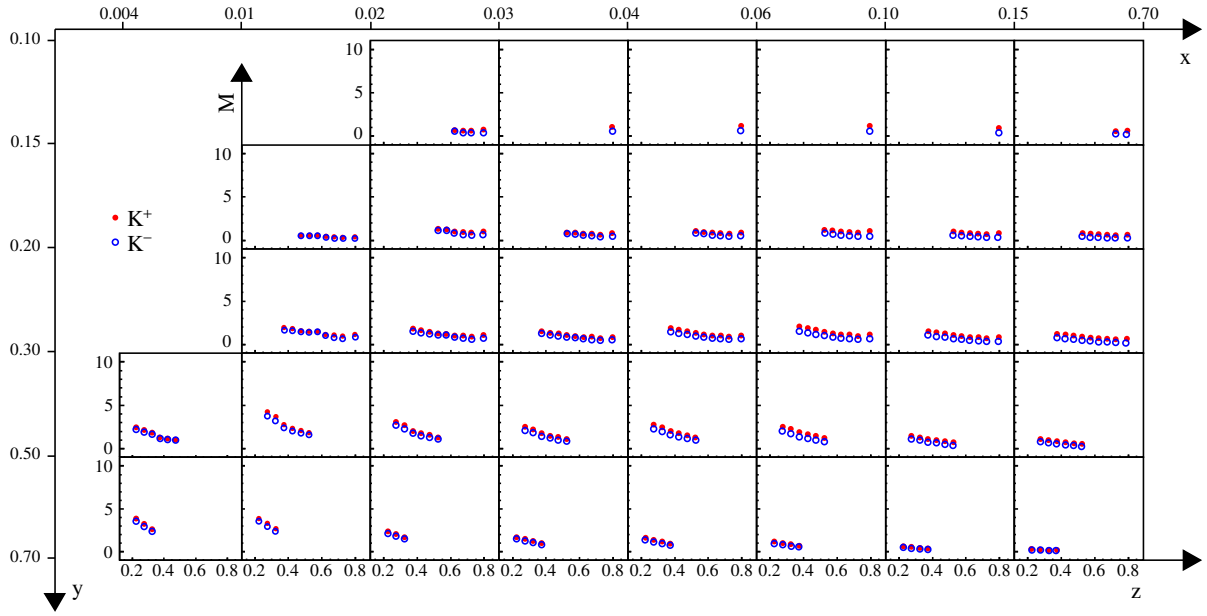


Figure F.3: *Same as fig. F.1 for kaons.*

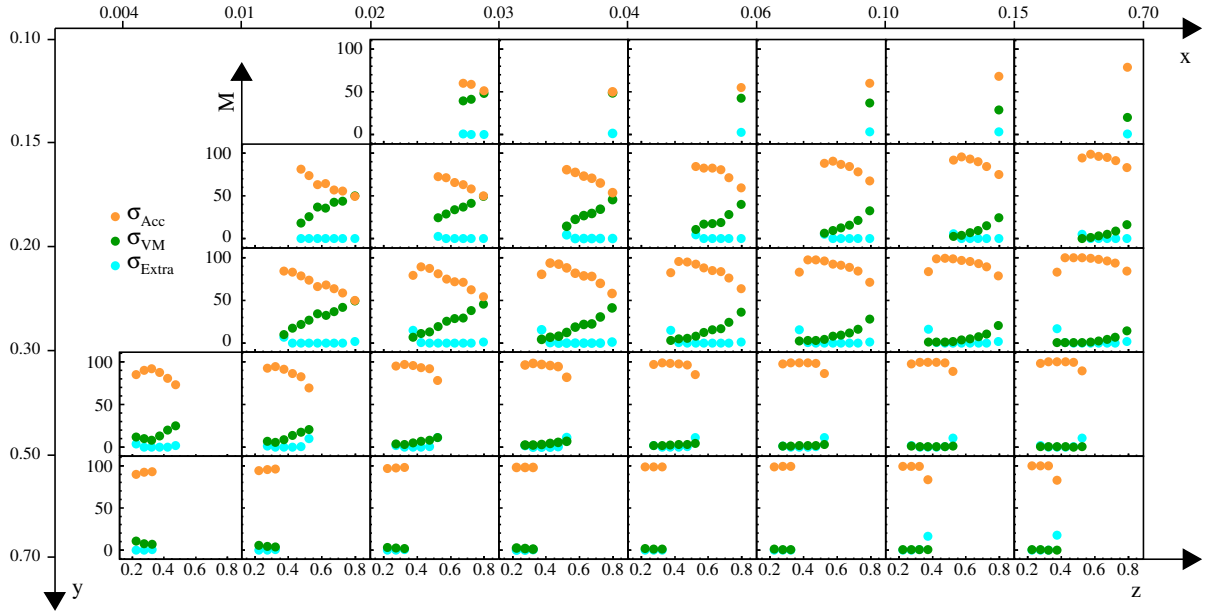


Figure F.4: *Fraction of the various contribution to the systematic as a function of z in different x and y bins. Yellow markers: acceptance. Green markers: Vector-meson production. Blue markers: Extrapolation to unmeasured momentum.*

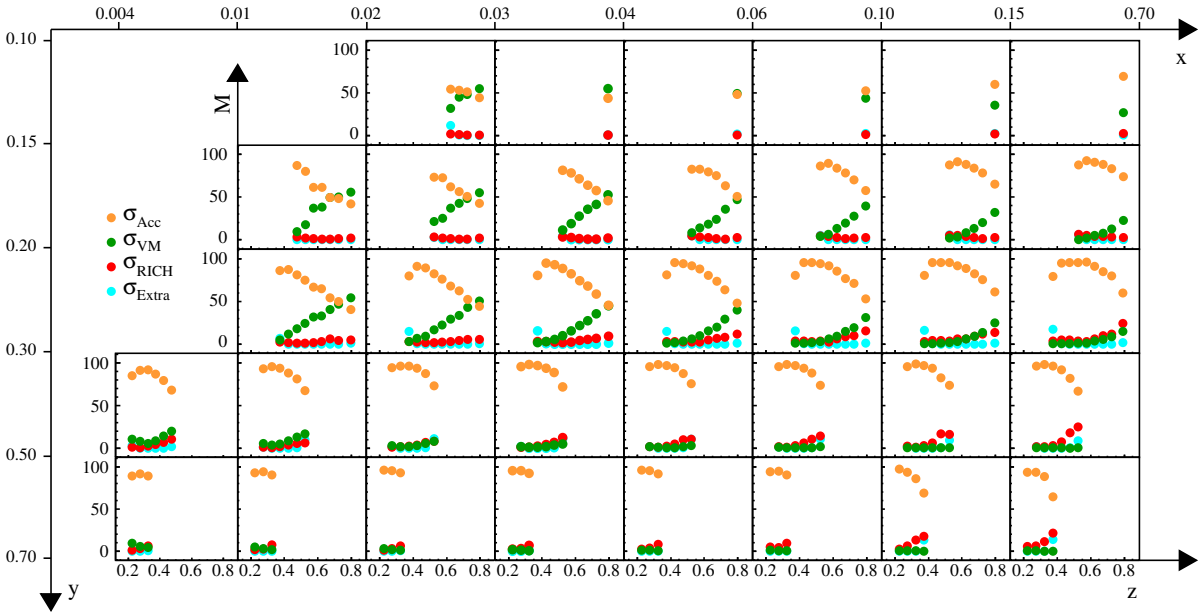


Figure F.5: *Same as fig. F.4 for pions. Red markers: particle identifiction with RICH.*

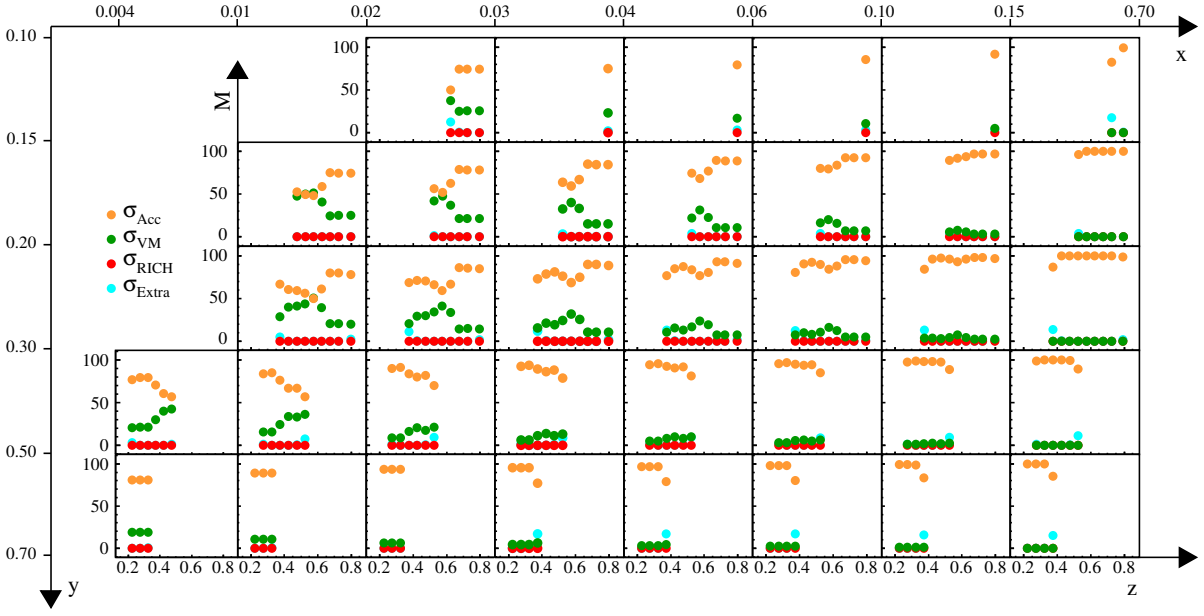


Figure F.6: *Same as fig. F.4 for kaons.*

Bibliography

[1] COMPASS collaboration: M.G. Alekseev *et. al.* Quark helicity distributions from longitudinal spin asymmetries in muon–proton and muon–deuteron scattering. *Physics Letters B*, 693(3):227 – 235, 2010.

[2] K.A. Olive et al. REVIEW OF PARTICLE PHYSICS. *Chinese Physics C*, 38(9):13 –18, 2014.

[3] Halzen F. and Martin A.D. *An Introductory Course in Modern Particle Physics.* John Wiley & Sons, Inc, 1984.

[4] J.D. Bjorken and E.A. Paschos. Inelastic electron –proton and γ –proton scattering and the structure of the nucleon. *Physical Review*, 185(54):1975 –1982, 1969.

[5] M. Anselmino, A. Efremov, and E. Leader. The theory and phenomenology of polarized deep inelastic scattering. *Physics Reports*, 261(1–2):1 –124, 1995.

[6] *Proceedings of the International School of Physics Enrico Fermi.* IOS Press, 2008.

[7] R.P. Feynman. *Photon-hadron Interactions.* Advanced Book Program Series. Addison-Wesley, 1998.

[8] C. G. Callan and David J. Gross. High-energy electroproduction and the constitution of the electric current. *Phys. Rev. Lett.*, 22:156–159, Jan 1969.

[9] Edmond L. Berger. SEMI–INCLUSIVE INELASTIC ELECTRON SCATTERING FROM NUCLEI. *ANL-HEP-CP-87-45*, 1987.

[10] A. Pich. Quantum chromodynamics. *arXiv:hep-ph/9505231*, May 1995.

[11] G. Sterman J.C. Collins, D.E. Soper. Factorization of hard processes in qcd. *arXiv:hep-ph/0409313*, Sep 2004.

[12] Sylvester Johannes Joosten. *Fragmentation and nucleon structure in semi–inclusive deep–inelastic scattering at the HERMES experiment.* PhD thesis, University of Illinois, 2013.

[13] Yu L. Dokshitser. Calculation of the structure functions for deep inelastic scattering and e^+ e^- annihilation by perturbation theory in quantum chromodynamics. *Journal of Experimental and Theoretical Physics*, 46(4):641 –652, 1977.

[14] V.N. Gribov and L.N. Lipatov. Deep inelastic *ep* scattering in perturbation theory. *Soviet Journal of Nuclear Physics*, 15:438, 1972.

[15] V.N. Gribov and L.N. Lipatov. e^+e^- pair annihilation and deep inelastic *ep* scattering in perturbation theory. *Soviet Journal of Nuclear Physics*, 15:675, 1972.

[16] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126(2):298 –318, 1977.

[17] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the lhc. *European Physical. Journal C*, 63:189–285, 2009.

[18] J. Levelt and P.J. Mulders. Quark correlation functions in deep–inelastic semi-inclusive processes. *Physical Review D*, 49(1):96 –113, 1994.

[19] R. Pankin. *Extraction of fragmentation functions from charged kaon and pion production at COMPASS*. PhD thesis, COMPASS, 2013.

[20] Tsuneo Uematsu. Q^2 dependence of quark and gluon fragmentation functions in a parton picture based on QCD. *Physics Letters B*, 79(1-2):97 – 102, 1978.

[21] B.R. Webber. Fragmentation and hadronization. *arXiv:hep-ph/9912292*, Dec 1999.

[22] B. Andersson *et al.* Parton fragmentation and string dynamics. *Physics Reports*, 97(2 –3):31 –145, 1983.

[23] Janusz Michael Niczyporuk. *Quark flavor tagging in polarized semi-inclusive deep inelastic scattering*. PhD thesis, Massachusetts Institute of Technology, 1997.

[24] B.A. Kniehl, G. Kramer, and B. Pötter. Testing the universality of fragmentation functions. *Nuclear Physics B*, 597(1-3):337 – 369, 2001.

[25] ALEPH Collaboration: D. Buskulic *et. al.* Measurement of α_s from scaling violations in fragmentation functions in $e + e-$ annihilation. *Physics Letters B*, 357(3):487 – 499, 1995.

ALEPH Collaboration: D. Buskulic *et. al.* Measurement of α_s from scaling violations in fragmentation functions in $e + e-$ annihilation [phys. lett. b 357 (1995) 487]. *Physics Letters B*, 364(4):247 – 248, 1995.

OPAL Collaboration: R. Akers *et. al.* Measurement of the longitudinal, transverse and asymmetry fragmentation functions at LEP. *Zeitschrift für Physik C Particles and Fields*, 68(2):203 – 213, 1995.

DELPHI Collaboration: P. Abreu *et al.* Measurement of the quark and gluon fragmentation functions in z^0 hadronic decays. *The European Physical Journal C - Particles and Fields*, 6(1):19–33, 1999.

[26] TPC/Two-Gamma Collaboration: H. Aihara *et. al.* Charged-hadron inclusive cross sections and fractions in e^+e^- annihilation at $\sqrt{s} = 29$ gev. *Physical Review Letters*, 61:1263 –1266, Sep 1988.

[27] BABAR Collaboration: J.P. Lees *et. al.* Production of charged pions, kaons, and protons in e^+e^- annihilations into hadrons at $\sqrt{s}$=10.54 GeV. *Phys. Rev. D*, 88:032011, Aug 2013.

[28] Belle Collaboration: M. Leitgab *et. al.* Precision measurement of charged pion and kaon differential cross sections in e^+e^- annihilation at $\sqrt{s}$=10.52 GeV. *Phys. Rev. Lett.*, 111:062002, Aug 2013.

[29] UA5 Collaboration: R.E. Ansorge *et. al.* Charged particle multiplicity distributions at 200 and 900 gev c.m. energy. *Zeitschrift für Physik C Particles and Fields*, 43(3):357–374, 1989.

[30] UA1 Collaboration: C. Albajar *et. al.* A study of the general characteristics of proton-antiproton collisions at âˆšs=0.2 to 0.9 tev. *Nuclear Physics B*, 335(2):261 – 287, 1990.

[31] ALICE Collaboration: K. Aamodt *et. al.* Charged-particle multiplicity measurement in protonâ€“proton collisions at $\sqrt{s}$ = 7Â tev with alice at lhc. *The European Physical Journal C*, 68(3-4):345–354, 2010.

[32] CMS Collaboration: V. Khachatryan *et. al.* Transverse-momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s}$ = 0.9 and 2.36 tev. *Journal of High Energy Physics*, 2010(2), 2010.

CMS Collaboration: V. Khachatryan *et. al.* Transverse-momentum and pseudorapidity distributions of charged hadrons in *pp* collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 105:022002, Jul 2010.

[33] PHENIX Collaboration: S.S. Adler *et. al.* Midrapidity neutral-pion production in proton-proton collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. Lett.*, 91:241803, Dec 2003.

STAR Collaboration: J. Adams *et. al.* Forward neutral pion production in $p+p$ and $d+$Au collisions at $\sqrt{s_{NN}} = 200$ GeV. *Phys. Rev. Lett.*, 97:152302, Oct 2006.

BRAHMS Collaboration: I. Arsene *et. al.* Production of mesons and baryons at high rapidity and high p_T in proton-proton collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. Lett.*, 98:252001, Jun 2007.

[34] ATLAS Collaboration: *et al.* Measurement of inclusive jet charged-particle fragmentation functions in $pb+pb$ collisions at $\sqrt{(}s_n n) = 2.76$ tev with the ATLAS detector. *arXiv:1406.2979*, June 2014.

[35] EMC Collaboration: J.J. Aubert *et. al.* A determination of the fragmentation functions of u quarks into charged pions. *Phys.Lett.*, B160:417 – 420, 1985.

[36] E00-108 Collaboration: T. Navasardyan *et. al.* Onset of quark-hadron duality in pion electroproduction. *Phys. Rev. Lett.*, 98:022001, Jan 2007.

E00-108 Collaboration: R. Asaturyan *et. al.* Semi-inclusive charged-pion electroproduction off protons and deuterons: Cross sections, ratios, and access to the quark-parton model at low energies. *Physical Review C*, 85:015202, Jan 2012.

[37] HERMES Collaboration: A. Airapetian *et. al.* Multiplicities of charged pions and kaons from semi-inclusive deep-inelastic scattering by the proton and the deuteron. *Physical Review D*, 87:074029, Apr 2013.

[38] N. Makke on behalf COMPASS Collaboration. Fragmentation functions measurement at COMPASS. *arXiv:1307.3407*, Juin 2013.

N. Makke on behalf COMPASS Collaboration. Study of the hadronisation process from single hadron and hadron-pair production in SIDIS at COMPASS. *Phys.Part.Nucl.*, 45:138–140, 2014.

[39] N. Du Fresne on behalf COMPASS Collaboration. Fragmentation functions measurement at COMPASS. *arXiv:1307.3407*, Juin 2014.

[40] S. Kretzer. Fragmentation functions from flavor inclusive and flavor tagged $e + e-$ annihilations. *Physical Review D*, 62:054001, 2000.

[41] M. Hirai, S. Kumano, T.-H. Nagai, and K. Sudoh. Determination of fragmentation functions and their uncertainties. *Physical Review D*, 75:094009, May 2007.

[42] S. Albino, B.A. Kniehl, and G. Kramer. Fragmentation functions for light charged hadrons with complete quark flavour separation. *Nuclear Physics B*, 725(1â€“2):181 – 206, 2005.

S. Albino, B.A. Kniehl, and G. Kramer. {AKK} update: Improvements from new theoretical input and experimental data. *Nuclear Physics B*, 803(1-2):42 – 104, 2008.

[43] Daniel de Florian, Rodolfo Sassot, and Marco Stratmann. Global analysis of fragmentation functions for pions and kaons and their uncertainties. *Phys. Rev. D*, 75:114010, Jun 2007.

[44] D. Florian, R. Sassot, M. Epele, R.J. Hernández-Pinto, and M. Stratmann. Parton-to-pion fragmentation reloaded. *arXiv:1410.6027v1*, October 2014.

[45] Elliot Leader, Alexander V. Sidorov, and Dimiter B. Stamenov. Possible resolution of the strange quark polarization puzzle? *Phys. Rev. D*, 84:014002, Jul 2011.

[46] A.V. Sidorov2 E. Leader and D.B. Stamenov. Importance of semi inclusive DIS processes in determining fragmentation functions. *arXiv:1312.5200v3*, June 2014.

[47] HERMES Collaboration: A. Airapetian *et. al.* Multiplicity of charged and neutral pions in deep-inelastic scattering of 27.5 gev positrons on hydrogen. *EUROPEAN PHYSICAL JOURNAL C*, 21(4):599–606, 2001.

[48] E. Christova. On fragmentation functions. 10th International Workshop on Hadrons Structure ans Spectroscopy, Venice, Italy, 2010.

[49] FF generator. `http://lapth.cnrs.fr/ffgenerator/`. Accessed: 2014-10.

[50] COMPASS collaboration: P. Abbon *et al.* The COMPASS experiment at CERN. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 577(3):455 – 518, 2007.

[51] N. Doble *et al.* The upgraded muon beam at the SPS. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 343(2 –3):351 –362, 1994.

[52] Celso Filipe Correia Franco. *Measurement of the Gluon Polarisation from Open –Charm Production at COMPASS*. PhD thesis, Universidade Técnica de Lisboa, 2011.

[53] H. Pereira Da Costa. *Développement de chambres à dérive pour les hauts flux dans expérience COMPASS*. PhD thesis, COMPASS, 2001.

[54] Alexander Zvyagin. *D –meson production by muons in the COMPASS experiment at CERN*. PhD thesis, COMPASS, 2010.

[55] COMPASS collaboration: C. Bernet *et al.* The COMPASS trigger system for muon scattering. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 550(1–2):217 – 240, 2005.

[56] Rene Brun and Fons Rademakers. ROOT –an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1–2):81 –86, 1997. New Computing Techniques in Physics Research V.

[57] S. Gerassimov. Physics analysis software tools. http://ges.web.cern.ch/ges/phast/.

[58] J.V. Jelley. *Cerenkov radiation and its applications*. Pergamon Press, 1958.

[59] T. Bellunato *et al.* Refractive index dispersion law of silica aerogel. *European Physical Journal C*, 52:759 –764, 2007.

[60] COMPASS collaboration: E. Albrecht *et al.* The radiator gas and the gas system of COMPASS RICH-1. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 502(1):266 – 269, 2003. Experimental Techniques of Cherenkov Light Imaging. Proceedings of the Fourth International Workshop on Ring Imaging Cherenkov Detectors.

[61] COMPASS collaboration: E. Albrecht *et al.* The mirror system of COMPASS RICH-1. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 502(1):236 – 240, 2003. Experimental Techniques of Cherenkov Light Imaging. Proceedings of the Fourth International Workshop on Ring Imaging Cherenkov Detectors.

[62] COMPASS collaboration: P. Abbon *et al.* Particle identification with COMPASS RICH-1. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 631(1):26 – 39, 2011.

[63] F. Piuz. Ring imaging CHerenkov systems based on gaseous photo-detectors: trends and limits around particle accelerators. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 502(1):76 – 90, 2003. Experimental Techniques of Cherenkov Light Imaging. Proceedings of the Fourth International Workshop on Ring Imaging Cherenkov Detectors.

[64] COMPASS collaboration: P. Abbon *et al.* Design and construction of the fast photon detection system for COMPASS RICH-1. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 616(1):21 – 37, 2010.

[65] COMPASS collaboration: P. Abbon *et al.* Read-out electronics for fast photon detection with COMPASS rich-1. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 587(2–3):371 –387, 2008.

[66] Stefano Levorato. *Measurement of transverse spin effect in COMPASS on a transversely polarized proton target.* PhD thesis, COMPASS, 2009.

[67] Giulia Pesaro. *Messa a punto e caratterizzazione del RICH veloce dell'Esperimento COMPASS.* PhD thesis, COMPASS, 2007.

[68] T. Ypsilantis and J. Seguinot. Theory of ring imaging cherenkov counters. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 343(1):30 – 51, 1994.

[69] J. Beringer *et al.* Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012.

[70] J. Podolanski and R. Armenteros. Analysis ov v-events. *Philosophical Magazine Series 7*, 45(360):13 –30, 1953.

[71] F. Sozzi *et al.* A method to extract rich purity from data. *COMPASS Note*, 2006.

[72] S. Margulies and J. J. Phelan. On the parametrization of resonances occurring near threshold. *Il Nouvo Cimento*, LVIII(4):806 –814, 1968.

[73] Roger Barlow. Extended maximum likelihood. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 297(3):496 – 506, 1990.

[74] Giulia Pesaro. *Measurements at COMPASS of transverse spin effects on identified hadrons on a transversely polarized proton target.* PhD thesis, COMPASS, 2010.

[75] M. Lefebvre *et al.* Propagation of errors for matrix inversion. *arXiv:hep-ex/9909031*, September 1999.

[76] N. Make *et al.* Kaon and pion multiplicities from muon deep inelastic scattering on 6 LiD. *COMPASS Release Note*, 2010.

[77] Luis Miguel Faria Pereira Lopes da Silva. *Measurement of the Gluon Polarisation Through High p_T Hadron Production in COMPASS.* PhD thesis, COMPASS, 2011.

[78] B. Badalek *et al.* Radiative correction schemes in deep inelastic muon scattering. *Zeitschrift Für Physik C*, 66:591 – 599, 1995.

[79] Philip R. Bevington and O. Keith Robinson. *Data Reduction and Error Analysis for the Physical Sciences.* McGarw-Hill, 2003.

[80] Q. Curiel *et al.* Pion multiplicities from muon deep inelastic scattering on ^{6}LiD (2006 COMPASS data). *COMPASS Note*, 2014.

[81] Florian Thibaud. *Développement de détecteurs Micromegas pixellisés pour les hauts flux de particules et évaluation de la contribution diffractive à la lepto-production de hadrons à COMPASS.* PhD thesis, COMPASS, 2014.

[82] G Ingelman, J Rathsman, and G.A Schuler. AROMA 2.2 –a monte carlo generator for heavy flavour events in ep collisions. *Computer Physics Communications*, 101(1 –2):135 –142, 1997.

[83] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the lhc. *Eur. Phys. J.*, C63:189–285, 2009.

[84] A. Buckley. Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report. *arXiv:1405.1067*, May 2014.

[85] R. Brun *et al.* Geant: Simulation Program for Particle Physics Experiments. User Guide and Reference Manual. *CERN–DD–78–2–REV*, 2014.

[86] J. J. Sakurai. Vector-meson dominance and high-energy electron-proton inelastic scattering. *Phys. Rev. Lett.*, 22:981–984, May 1969.

[87] M. Diehl, W. Kugler, A. Schäfer, and C. Weiss. Exclusive channels in semi-inclusive production of pions and kaons. *Phys. Rev. D*, 72:034034, Aug 2005.

[88] A. Sandacz et P. Sznajder. HEPGEN –Generator for hard exclusive leptoproduction. *arXiv:1207.0333*, Jul 2012.

[89] M. Arneodo et al. Exclusive ρ^0 and ϕ muoproduction at large q^2. *Nuclear Physics B*, 429(3):503 –529, 1994.

[90] S.V. Goloskokov and P. Kroll. The role of the quark and gluon gpds in hard vector–meson electroproduction. *European Physical Journal C*, 53:367, 2008.

[91] John R. Taylor. *An introduction to Error Analysis. The study of uncertainties in physical measurements.* University Science Books, 1997.

[92] NNPDF Collaboration: Richard D. Ball *et. al.* Unbiased determination of polarized parton distributions and their uncertainties. *Nuclear Physics B*, 874(1):36 – 84, 2013.

[93] D. Hahne. Status report of k_s^0 multiplicities from 2006 at COMPASS. DPG-Physi Verhandlungen, Frankfurt, Germany, 2014.

[94] LHAPDF homepage. `https://lhapdf.hepforge.org/`.

[95] William H.Press, Saul A.Teukolsky, William T.Vetterling, and Brian P. Flannery. *Recipes in C (2Nd Ed.): The Art of Scientific Computing.* Cambridge University Press, 1992.