Quark Fragmentation and Hadron Formation in Nuclear Matter

Raphaël Dupré

Université de Lyon

Faculté des Sciences et Technologies

École Doctorale PHAST

THÈSE DE DOCTORAT

Présentée par

Raphaël DUPRÉ

pour obtenir le titre de Docteur ès Sciences de l'Université de Lyon

Specialité : PHYSIQUE NUCLÉAIRE

Sujet :

Quark Fragmentation and Hadron Formation in Nuclear Matter

Directeur de thèse: **Guy CHANFRAY**
Institut de Physique Nucléaire de Lyon, Université de Lyon

Encadrant local: **Kawtar HAFIDI**
Physics Division, Argonne National Laboratory

soutenue publiquement le **9 novembre 2011** à Villeurbanne

Jury :

Président :	Jean-Yves GROSSIORD	-	IPNL (Villeurbanne, France)
Directeur :	Guy CHANFRAY	-	IPNL (Villeurbanne, France)
Encadrant local:	Kawtar HAFIDI	-	ANL (Argonne, IL, USA)
Rapporteur :	William BROOKS	-	UTFSM (Valparaíso, Chile)
Rapporteur :	Michel GUIDAL	-	IPNO (Orsay, France)
Examinateur :	François ARLEO	-	LAPTH (Annecy-le-Vieux, France)

This thesis was prepared principally in the

Medium Energy Physics Group,
Physics Division

Argonne National Laboratory

9700 South Cass Avenue, Argonne, IL, USA

and occasionally in

Hall B,

Jefferson Lab

12000 Jefferson Avenue, Newport News, VA, USA

and

Institut de Physique Nucléaire de Lyon

Université de Lyon,
Université Claude Bernard Lyon 1,
IN2P3, CNRS

Domaine scientifique de la Doua
4, rue Enrico Fermi, Villeurbanne, France

[L]e savant ne choisit pas au hasard les faits qu'il doit observer. Il ne compte pas des coccinelles, comme le dit Tolstoï, parce que le nombre de ces animaux, si intéressants qu'ils soient, est sujet à de capricieuses variations. Il cherche à condenser beaucoup d'expérience et beaucoup de pensées sous un faible volume, et c'est pourquoi un petit livre de physique contient tant d'expériences passées et mille fois plus d'expériences possibles dont on sait d'avance le résultat. [...]

Le savant n'étudie pas la nature parce qu'elle est utile ; il l'étudie parce qu'il y prend plaisir et il y prend plaisir parce qu'elle est belle. Si la nature n'était pas belle, elle ne vaudrait pas la peine d'être connue, la vie ne vaudrait pas la peine d'être vécue. Je ne parle pas ici, bien entendu, de cette beauté qui frappe les sens, de la beauté des qualités et des apparences ; non que j'en fasse fi, loin de là, mais elle n'a rien à faire avec la science ; je veux parler de cette beauté plus intime qui vient de l'ordre harmonieux des parties, et qu'une intelligence pure peut saisir. C'est elle qui donne un corps, un squelette pour ainsi dire aux chatoyantes apparences qui flattent nos sens, et sans ce support, la beauté de ces rêves fugitifs ne serait qu'imparfaite parce qu'elle serait indécise et toujours fuyante. Au contraire, la beauté intellectuelle se suffit à elle-même, et c'est pour elle, plus peut-être que pour le bien futur de l'humanité, que le savant se condamne à de longs et pénibles travaux.

Science et méthode
Henri Poincaré

Fragmentation des Quarks et Formation des Hadrons dans la Matière Nucléaire

Résumé: La formation des hadrons est, dans le cadre de la théorie quantique de couleur (QCD), un processus non-perturbatif ; cette caractéristique entraîne d'importantes difficultés théoriques. C'est pourquoi, les mesures expérimentales de fragmentation dans différents noyaux sont une nécessité afin d'obtenir des progrès tangibles dans la compréhension des mécanismes de formation des hadrons. La thèse commence par les bases théoriques nécessaires à une telle approche, suivies des principaux modèles qui lui sont associés.

La thèse se poursuit par l'analyse de données de *Jefferson Lab* obtenues à l'aide d'un faisceau d'électrons de 5 GeV incident sur différentes cibles (^{2}H, C, Al, Fe, Sn et Pb). Les produits de la réaction sont mesurés avec le spectromètre CLAS. Les principaux résultats de cette expérience sont : (a) l'analyse multi-dimensionnelle des observables mesurées, qui permet une meilleure confrontation avec les modèles théoriques et l'extraction d'informations temporelles sur la fragmentation, et (b) l'observation d'une atténuation hadronique non-linéaire en fonction du rayon du noyau cible.

Dans une partie plus théorique, le générateur d'événements PyQM, développé dans le but de reproduire les données de la collaboration HERMES, est présenté. Les résultats sont mitigés, en effet la base théorique utilisée ne semble pas s'appliquer au cas étudié, néanmoins certaines caractéristiques des données sont reproduites permettant de comprendre leurs origines parfois inattendues. Enfin, les possibilités d'expériences futures, à *Jefferson Lab* et dans un collisionneur ion-électron (EIC), sont explorées.

Mots-clef: Fragmentation, hadronisation, QCD, Jefferson Lab, CLAS, noyau, Monte-Carlo, perte d'énergie des quarks, collisionneur électron-ion, EIC.

Quark Fragmentation and Hadron Formation in Nuclear Matter

Summary: The hadron formation is, in the framework of the quantum chromodynamics theory (QCD), a non-perturbative process; this characteristic leads to important theoretical challenges. This is why experimental measurements of fragmentation in nuclei are a necessity in order to obtain substantial progress in our understanding of the mechanisms of hadron formation. The thesis begins with the introduction of theoretical background, followed by an overview of theoretical models.

The thesis continues with the analysis of Jefferson Lab data obtained with a 5 GeV electron beam incident on various targets (^{2}H, C, Al, Fe, Sn and Pb). The reaction products are measured with the CLAS spectrometer of Hall B. The main results are: (a) a multi-dimensional analysis of the measured observables, which permits a better confrontation with theoretical models and the extraction of temporal information on fragmentation, and (b) the observation of a non linear hadronic attenuation as a function of the target's nuclear radius.

The PyQM event generator, developed to reproduce the data from the HERMES collaboration, is also presented. The results are ambivalent, the theoretical basis used does not seem to apply to the studied case, however, some characteristics of the data are reproduced allowing to understand their origin, which is sometimes unexpected. Finally, the possibilities for future experiments, at Jefferson Lab and at an Electron-Ion Collider (EIC), are explored.

Keywords: Fragmentation, hadronization, QCD, Jefferson Lab, CLAS, nuclei, Monte-Carlo, quark energy loss, Electron-Ion Collider, EIC.

Acknowledgments

A thesis work is not only a scientific work, it is also a human experience and I would like to thank here all those who made it possible and such a great experience!

I would like to thank first the many people who supervised me one way or another. Especially, Kawtar Hafidi who taught me day after day during these three years. But also Guy Chanfray for accepting to be my supervisor abroad, Alberto Accardi for his guidance on the PyQM and EIC projects and Stepan Stepanyan for sharing his knowledge on detectors during the preparation of the *eg6* run. I also would like to thank the members of my jury, especially W. Brooks and M. Guidal who accepted to be referee of my thesis.

A great part of this thesis is due to the great environment that is Argonne National Laboratory. Therefore, I would like to thank Roy Holt, Don Geesaman and all the MEP group for welcoming me and for their numerous contributions during my various presentations along the years.

I would like to credit the important participation of many other students and post-docs (in particular Aji, Hayk, Hyupwoo, Lamiaa and Taisiya) who worked on the *eg2* data and spent countless hours discussing analysis details. Special thanks to William Brooks who organized the meetings and who was endlessly patient with us.

I want to acknowledge the many members of the *université Claude Bernard de Lyon* who made my student years so enjoyable. Of course my many professors, of mathematics (T. Fack and M. Kibler) and quantum physics (S. Fleck and G. Chanfray) in particular. But also the professors, *BIATOSS* and students with which I worked in the various committees of the university, especially L. Collet who encouraged me to pursue my studies in the USA. I want also to credit all the people who make *Turbulence* (*spécialement Ahmed, Fabien, Kevin, Mirsal et Samia*) and the *BVE* (*spécialement Maryline et Patrice*) such great organizations, which made me appreciate fully the university.

I also need to express my gratitude to all the people who visited me in America and helped me keep the connection. *Samia, (Mirsal)^2, Thomas, Alexia et Marion merci d'être venu ! et merci aussi à mes nombreux correspondants : Samia (encore !), Aurélie, Anne-Ségo et tout particulièrement Kathalyne !* Also to the Moroccan team: Ahmed, Brahim, Lamiaa and Kawtar; who made the stay both at Argonne and JLab so enjoyable.

Finally, I would like to thank all my family members who supported me during these years. *Remerciement spécial à Alain, Anne, Christophe, Clara, Elisabeth, Enzo, Magalie et Nicolas qui se sont sacrifiés en traversant l'Atlantique !*

Contents

Introduction

Quarks and gluons, namely partons, are confined inside hadrons and cannot be found isolated. This experimental fact – known as confinement – is tightly related to hadronization, the process by which partons transform into hadrons. Indeed, the strong force provokes the production of new hadrons when one tries to isolate a parton. This force is described by quantum chromo-dynamics (QCD) and had great successes in the perturbative regime since the beginning of its development in the 1960's. However, at low energy, QCD cannot be treated perturbatively making reliable calculations very challenging, in particular for dynamics processes such as hadronization. In this regime, experimental measurements are, therefore, an important input to guide and constrain models.

Hadrons are formed on distances of few femtometers, making the nucleus the best tool to study the space-time properties of hadronization. Indeed, using electron deep inelastic scattering (DIS) on nuclei allows to produce quarks with known kinematics, in a static medium with well understood properties. Comparing hadron production in light and heavy nuclei is equivalent to compare hadronization in vacuum with hadronization in nuclear medium. As the struck parton evolves through different stages, its interaction with the medium changes, therefore, the hadronization dynamics can be deduced from the variations of the physical observables, as function of the kinematics and the nucleus size. However, the intermediate stages of hadronization are not known *a priori* and the interpretation of the data is, most of the time, a challenge.

Experimental investigations of hadronization, using DIS, started in the 1970's, but most of the early results lacked statistical precision and allowed only qualitative interpretations. During the last decade, the HERMES collaboration published results with significant improvements, both in term of statistics and hadron separation. Consequently, a clearer picture of the hadronization process emerged and many models were excluded, but this was not enough and some very different models remain. The CLAS collaboration data, presented in this thesis, offer great statistics and a large variety of nuclear targets. With this new experiment, the goal is to provide stringent tests of the models and, therefore, advance our understanding of hadronization dynamics.

Motivated by the recent improvements of the measurements quality, theorists improved their models by including new nuclear effects. This increasing sophistication favors the use of full Monte-Carlo simulation, to simplify the comparison with results presented in multi-dimensional bins. We developed such a simulation, called PyQM, the principal objective being to connect

traditional nuclear physics with relativistic heavy-ion collision physics. Such a comparison was made possible by recent theoretical developments, which link hadronization effects to the properties of the medium. Thus, we can make direct comparison between hot nuclear matter, such as Quark-Gluon Plasma (QGP), and normal nuclear matter, if we can provide a model fitting both kinds of data. Also, the effects observed in the PyQM simulation can help to interpret experimental data and understand the origin of certain observed features.

The thesis is organized as follow. In chapter 1, the necessary background and the physics motivations are introduced, in chapter 2, the theoretical models are reviewed, and, in chapter 3, they are confronted to published results. The Monte-Carlo simulation of hadronization in nuclei, PyQM, is presented in chapter 4. Chapter 5 is an overview of the apparatus used in the CLAS experiment, it is followed, in chapter 6, by the analysis of the data and, in chapter 7, by the results, which are presented and discussed. Perspectives for future experiments are discussed in chapter 8 and, finally, a conclusion will summarize the results disclosed in this thesis.

Chapter 1

Processes and Observables

1.1 Deep Inelastic Scattering

To trigger the hadronization of a quark, a hard QCD process needs to be involved, in this thesis we concentrate on deeply inelastic scattering (DIS). Its general form is $l(k) + n(p) \rightarrow l(k') + X$ with l a lepton and n a nucleon. Here, we will treat only the charged leptons and because the energy level is always much smaller than M_Z or M_W, photons are mediating the interaction[1]. Therefore, this is a pure electromagnetic interaction and only charged constituent of the hadron target – i.e. quarks – are probed. As a consequence, the DIS process treated in this thesis is of the form shown in figure 1.1.

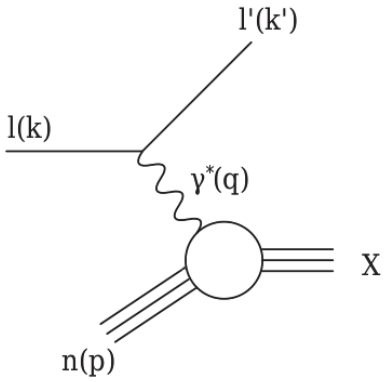

Figure 1.1: Leading-order Feynman diagram describing DIS on a nucleon with photon exchange.

The relevant variables for inclusive[2] measurements of DIS are:

- the 4-momentum transfer between the lepton and the nucleon squared $Q^2 = -q^2$,
- the energy transfer $\nu = p \cdot q / M_n$ ($= E_k - E'_k$ in the target rest frame),
- the Bjorken scaling variable $x_{Bj} = \frac{-q^2}{2p \cdot q} = \frac{Q^2}{2M_n \nu}$ (dimensionless),

[1] Very small effects remains from Z and W exchange, these are neglected here.

[2] Inclusive means that only the scattered lepton is detected. If the scattered lepton and one hadron are measured, the reaction is semi-inclusive. If all products are identified, the reaction is called exclusive.

- the ratio of the energy transferred to the total energy available $y = \frac{p.q}{p.k}$ ($= \frac{\nu}{E_k}$ in the target rest frame) (dimensionless),
- the mass of the total hadronic final state $W = \sqrt{M_n^2 - Q^2 + 2M_n\nu}$.

Physically, Q^2 is the scale probed by the photon and W indicates the inelasticity of the reaction ($W = M_n$ is elastic). In the target rest frame, ν is the virtual photon energy and, in the infinite momentum frame, the Bjorken variable is the fraction of the nucleon momentum carried by the struck quark.

By convention, in the case of nuclear targets, we calculate the kinematic variables considering that the target hadron is a nucleon (taken as the mean between proton and neutron). We do so because it facilitates comparisons and because, as DIS probes the nucleon constituents, it is more meaningful to keep it as reference. However, this choice is not perfect and several nuclear effects interfer, for example the EMC effect [Geesaman 1995], which introduces variations in the nucleon structure. The way around, in our hadronization studies, is to choose observables that are not sensitive to the modification of the initial state nucleon.

With semi-inclusive measurements, it is possible to get information on the struck quark. The flavor content of the produced hadron gives information on the flavor of the struck quark and the hadron 4-momentum p_h gives information on the quark kinematics and its hadronization. We list here the semi-inclusive kinematic variables used in the thesis (see figure 1.2 for a graphic representation):

- the fraction z of the virtual photon energy transferred to the hadron: $z = \frac{p_h.p}{q.p}$ ($= \frac{E_h}{\nu}$ in the target rest frame);
- the angle ϕ_h between the leptonic plane, defined by the virtual photon and the outgoing lepton, and the hadronic plane, defined by the virtual photon and the detected hadron;
- the transverse momentum $P_\perp$ of the hadron, defined in the target rest frame relative to the direction of the virtual photon;
- the Feynman scaling variable, x_F is the fraction of the maximum longitudinal momentum carried by the hadron $x_F = \frac{P_L}{P_L^{max}}$;
- the Mandelstam variable t is the square of the 4-momentum transferred to the hadron: $t = (q - p_h)^2$;
- the rapidity is defined as $y^* = \frac{1}{2} \ln \frac{E_h + P_L}{E_h - P_L}$.

Experimentally, the DIS events need to be separated from other processes like resonances and coherent production. In the case of resonances

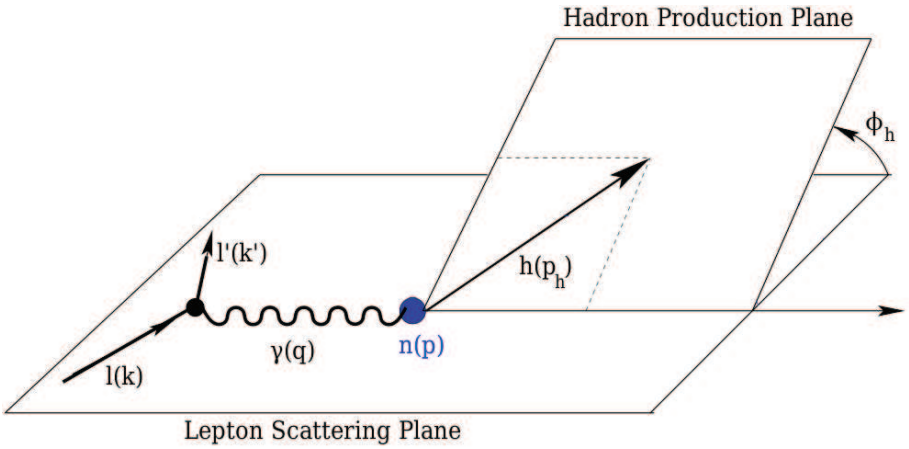

Figure 1.2: Semi-inclusive deep inelastic scattering on a nucleon.

the photon is absorbed by the nucleon as a whole, exciting it to a state which will eventually decay. This leads, as in DIS, to the emission of one or more hadrons in the final state, but these are produced by internal, collective effects of the hadronic target. The spectroscopy of these resonances is extensively studied [Aznauryan 2011] but is not relevant to the study of hadronization and could contaminate our DIS sample. The scattering is expected to occur from quarks when the momentum transfer exceeds the QCD scale ($Q^2 \gg \Lambda_{QCD} \simeq (300\text{MeV})^2$) and the energy of the final state exceeds all strong hadronic resonances ($W \gtrsim 2$ GeV). The quantum fluctuations of the virtual photon into hadronic states can also contaminate DIS samples. This is called diffractive production and comes mainly from vector mesons, it is concentrated at low x_{Bj} and high z (> 0.8).

Theoretically, one can describe the DIS process easily using the factorization theorem (see [Brock 1995] for a complete review of DIS theory). Factorization permits to separate the DIS cross section in three independent parts: (a) the hard scattering cross section between the lepton and the parton, which, as it is an electromagnetic process, can be exactly calculated using quantum electro-dynamics (QED); (b) the parton distribution functions ($f_{q,g}(x_{Bj}, Q^2)$), which contain all the information about the target and are, at leading order, the probability distribution to find a parton of fractional momentum x_{Bj} in the hadron; (c) the fragmentation functions $D_f^h(z, Q^2)$, which give, at leading order, the probability for a parton f to form a hadron h of fractional energy z[3].

The parton distribution functions and the fragmentation functions are expected to evolve with the scale Q^2, this is called the DGLAP evolution, for Dokshitzer-Gribov-Lipatov-Altarelli-Parisi [Dokshitzer 1977, Gribov 1972,

[3]Theorists usually define z in the same way than x_F and neglect both masses and transverse momentum in their fit of the fragmentation functions. This is not justified at low energy [Accardi 2009b].

Altarelli 1977]. This scaling violation is due to the nature of perturbative theory and the capacity to probe off-shell partons. This evolution is expected within QCD and is based on perturbative development and factorization. It can be exactly calculated and allows to extrapolate measured distributions to different Q^2.

1.2 The Hadronization Process

Because of confinement, quarks cannot be freed from hadrons; they have to hadronize, i.e. find partners and form new hadrons. This is a fundamental process of QCD for which precise calculation remains inaccessible. This is because only static systems can be approximated in the non-perturbative regime, using chiral theory, Dyson-Schwinger equations [Roberts 1999] or lattice calculation, for example, whereas at higher energies the perturbative QCD (pQCD) is successful for a wide range of processes. However, hadronization is, at its last stage, both a dynamics and a low energy process, therefore it cannot be precisely described from first principles. It is, therefore, described using models or phenomenological studies, which both need experimental guidance. In this section, we will review the basic information on hadronization that can be deduced from QCD principles.

After the hard scattering, during the production time, the struck quark propagates and, because it is a colored object, it emits gluons (even in vacuum). Then, the quark has to transform into a colorless object, referred to as pre-hadron, which will eventually fall in a given hadronic state after the formation time. The space-time characteristics of hadronization are believed to be dominated by these two scales, the production time and the formation time, illustrated in the figure 1.3.

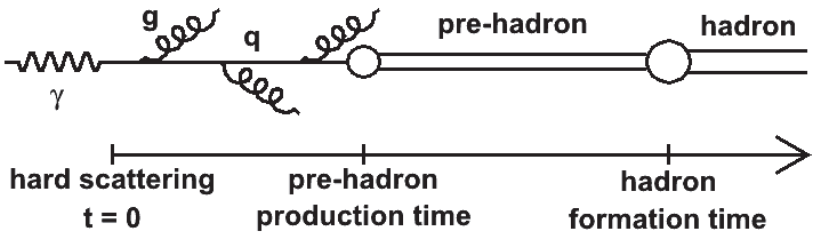

Figure 1.3: The time scales of the hadronization process.

One of the main goal of hadronization studies is to provide information on the dynamics scales of the process and because models give very different predictions, examples are shown in table 1.1, making experimental input is

$\langle t_h \rangle$	Kinematic	π	K	p	D
Basic QCD [Dokshitzer 1991]	HERMES	34 fm	8 fm	6 fm	1.9 fm
Basic QCD [Dokshitzer 1991]	RHIC	34 fm	8 fm	6 fm	1.9 fm
Lund Model [Andersson 1983]	HERMES	11 fm	9 fm	18 fm	
Lund Model [Andersson 1983]	RHIC	9 fm	8 fm	13 fm	
$\langle t_{preh} \rangle$					
Basic QCD [Adil 2007]	HERMES	28 fm	9 fm	3 fm	0.8 fm
Basic QCD [Adil 2007]	RHIC	18 fm	7 fm	3 fm	0.8 fm
Lund Model [Andersson 1983]	HERMES	4 fm	4 fm	6 fm	
Lund Model [Andersson 1983]	RHIC	2 fm	3 fm	1 fm	

Table 1.1: Table from [Accardi 2010b] showing typical hadron formation times $\langle t_h \rangle$ and prehadron production time $\langle t_{preh} \rangle$ from different models and kinematics.

crucial to evaluate them.

1.3 Hadronization in Vacuum

In the vacuum, hadronization was studied thoroughly using high energy e^+e^- collisions [Lafferty 1995]. These provide parton pairs, emitted back-to-back, leading to hadron production. The fit of the hadron spectra offers a good description of the fragmentation functions (see figure 1.4 for an example), for which universality[4] was then confirmed by comparing these results with e-p or hadron-hadron reactions. The fits of these data, all together, were performed by several theory groups [de Florian 2007b, Hirai 2007, Albino 2008] to obtain precise fragmentation functions.

An important success of QCD is that the fits, based on the Q^2 trend predicted by the DGLAP evolution expected from QCD, match properly the experimental data. Since the fitting procedure has attained, today, a good level of precision and covers many hadron flavors [de Florian 2007b, de Florian 2007a], this is an important success of QCD. However, the few theoretical studies attempting to reproduce the fragmentation functions from first principles remain inconclusive (see [Ito 2009] for example). This failure is mainly due to the difficulty to make calculations in the non perturbative domain. Nevertheless, the recent development of lattice QCD[5] (see [Nakamura 2005] for example) gave input on the nature of the quark-quark forces. However, this progress is limited to static forces, because lattice QCD is not able to describe evolving systems, yet this is a first strong theoretical input and, together with experimental input, it should help to construct mod-

[4] The fact that they are identical whatever the process.
[5] Computer based QCD calculations on a lattice.

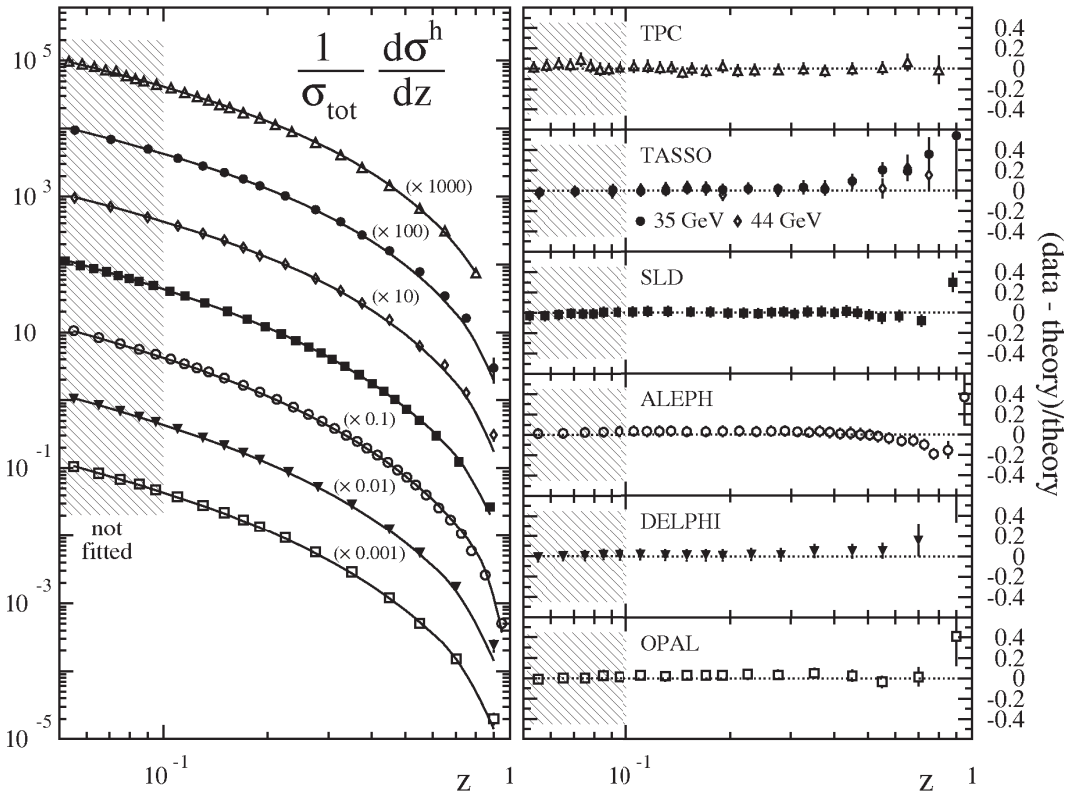

Figure 1.4: Fit, from [de Florian 2007b], of fragmentation functions of charged hadrons using e^+e^- data sets.

els.

1.4 Hadronization in Nuclei

A way to access information at the femtometer scale is to use nuclei. We do so using a hard process to free a quark from its initial hadron. Then, the quark will propagate in the nuclear medium and, as a colored object, emit gluons. It is believed that its interaction will drastically change when it becomes a prehadron, i.e. its color neutralizes. Therefore, the variations of the nuclear matter effects on hadronization, as a function of the nuclear size and the quark kinematics, should give information on the distance needed by the quark to neutralize its color and by the prehadron to expand into a hadron. However, studies on the prehadron need large nuclear size to ensure that the formation occurs inside the nuclei. Such studies can be simplified producing directly a color neutral object by diffractive process [Brodsky 1988].

The main complications to study hadronization with nuclei is the difficulty to evaluate the interactions of quarks and prehadrons with QCD matter. Indeed, it should be possible to calculate the quark interaction with the medium, within pQCD, and to relate the prehadron interactions to its size [Strikman 2007]. However, many different theoretical approaches exist to these problems. These are treated in chapter 2 and remain to be seriously challenged with experimental data.

1.5 Observables

To study hadronization, we want our observables to be independent of the initial state effects, sensitive to final state effects and easily interpretable. The observables presented in this section are well established and commonly used in the literature. They use a reference nuclei, which is in the following formulas deuterium (noted D), but it can sometimes be replaced by heavier nuclei, such as carbon, for experimental convenience.

The multiplicity ratio represents the relative production rate of a hadron h in a nuclear target A; it is defined as

$$R_A^h(Q^2, \nu, z_h, P_\perp^2) = \frac{N_A^h(Q^2, \nu, z_h, P_\perp^2)/N_A^e(Q^2, \nu)}{N_D^h(Q^2, \nu, z_h, P_\perp^2)/N_D^e(Q^2, \nu)} \tag{1.1}$$

with N_A^e and N_A^h, respectively, the number of electron and of semi-inclusive hadron h produced on a target A. The attenuation of the hadron production on a nuclear target, relatively to deuterium, is given by $1 - R_A^h$.

The transverse momentum broadening represents the increase of hadronic transverse momentum in a nuclear target A and is defined as

$$\Delta\langle P_\perp^2\rangle = \langle P_\perp^2\rangle_A - \langle P_\perp^2\rangle_D \tag{1.2}$$

with $\langle P_\perp^2\rangle_A$ the mean transverse momentum measured for a target A.

In the case of the multiplicity ratio, it is the normalization with the numbers of electron which permits to cancel the initial state effects. For the transverse momentum broadening, the initial state interaction is reduced because the transverse momentum is relative to the virtual photon whatever its kinematic. Finally, these observables are complementary; R_A^h probes the reduction of hadron production and $\Delta\langle p_\perp^2\rangle$ probes its modification.

1.6 Motivations

Hadronization is a common feature of high energy (> GeV) interactions and is present in many processes, such as e^+e^- annihilation, lepton-nucleon interactions and nucleon-nucleon interactions. In these, the nucleon can be replaced by a nucleus in order to study its properties or to increase the production yield.

Nuclear effects are numerous, they are often referred as initial and final state interactions depending on their nature. The main initial state effect, in nuclear lepton scattering, is the EMC effect [Geesaman 1995]. The final state effects in nuclear DIS are the main focus of this thesis and will be detailed in chapter 2. Inclusive DIS measurements have the particularity to be only sensitive to initial state effects, therefore, final state effects can only be observed using semi-inclusive and exclusive reactions. Such measurements are often used to study nuclear effects but, for many other experiments, these are important sources of uncertainty.

In the case of nuclear DIS, the initial state effects are mostly studied with inclusive measurements. Semi-inclusive measurements could provide more detailed information, but our limited knowledge of the final state effects introduces important systematic errors. In the case of neutrino scattering, nuclei are used to increase rates of the experiments. The correction from nuclear effects is necessary [Bleve 2001] and an input from electron or muon scattering helps to confirm and tune the models used for this correction [Manly 2006]. In heavy ion collisions, the picture is even more complicated, indeed initial and final state effects are entangled. Moreover, the soft QCD effects, tightly linked to the comprehension of hadronization, are involved in the background of all events and lead to difficulties in the interpretation of the data. At both RHIC (Relativistic Heavy Ion Collider at the Brookhaven National Laboratory)

and LHC (Large Hadron Collider at the *Centre Européen pour la Recherche Nucléaire*) the jets reconstructions have to be done on top of this background with the necessity to understand it.

1.7 Conclusion

It is clear that the importance of the hadronization process does not lie only in its links with the confinement phenomenon. It is also a central problem in many measurements where it sets important limits on experimental precision. But, while the hadronization process was studied by numerous experiments in vacuum, providing precise fits of the fragmentation functions, our fundamental understanding of hadronization did not progress much, especially, when considering in-medium hadronization, for which many pieces are still missing. Deep inelastic scattering offers a unique opportunity to study these final state effects and probe the hadronization dynamics. Eventually, such progress would have a large impact on many other QCD studies, in particular the ones involving the non-perturbative regime and the nuclear medium properties.

Chapter 2

Theoretical Models

2.1 Introduction

Because of its non-perturbative nature, hadronization cannot be studied directly from first principles. Therefore, only a development of models, based on experimental data, can help to describe it. Most of those models have the same basic assumptions about hadronization in vacuum (see section 1.2), but beyond that simple picture many differences appear. This can lead to different explanations for the phenomenons observed. This is introducing serious doubts on the validity of any given model and makes it difficult to establish strong ground for further progress.

The models describing SIDIS on nuclei are based on the following processes: hadron absorption, parton energy loss and medium modified fragmentation functions. Hadron absorption models consider the absorption of a prehadron with properties inspired from hadrons. Parton energy loss models rely on calculations (usually pQCD) of the colored parton interactions with the nuclear medium. Medium modified fragmentation functions models are based on calculations or phenomenological studies leading to a modification of the usual fragmentation functions due to the surrounding medium. In this chapter, we will review these processes independently, to facilitate comparisons. However, one has to keep in mind that some models are pure, i.e. consider only one effect, when others mix two of them. To conclude the chapter, a few examples of mixed models are given with a brief explanation of how the different processes are articulated.

The confrontation of the models with data will be treated in details in chapter 3. But it is important to point out in advance that all the models presented in this chapter can reproduce the main features of the published data. However, none is either in perfect agreement with the most recent data from HERMES. Therefore, work on the theoretical side is still necessary to describe properly the existing data and work on the experimental side is in order to provide measurements with more discriminating power.

2.2 Hadron Absorption

To evaluate the attenuation of the hadron production in nuclei, it is natural to think about the usual hadron-nucleon interaction. Especially because it can be studied independently from fragmentation with an hadronic beam. However the problem is not as simple as it seems, indeed, many models use hadron absorption to explain the suppression observed in experiments, but they usually adjust the effect and complement it with other features.

The common part of all hadron absorption models is the hadron-nucleus cross-section. Those cross sections are known from measurements [Nakamura 2010] (figure 2.1) and can be used directly to evaluate the hadron absorption in the medium. However, it appears that such a method does not give good results when compared with data [Gyulassy 1990]. The reason is believed to be linked with dynamical effects. Indeed, the hadron takes some time to be produced and might not fully interact immediately with the medium. Therefore, the properties of the prehadron, such as its life time and its evolving cross section, need to be modeled.

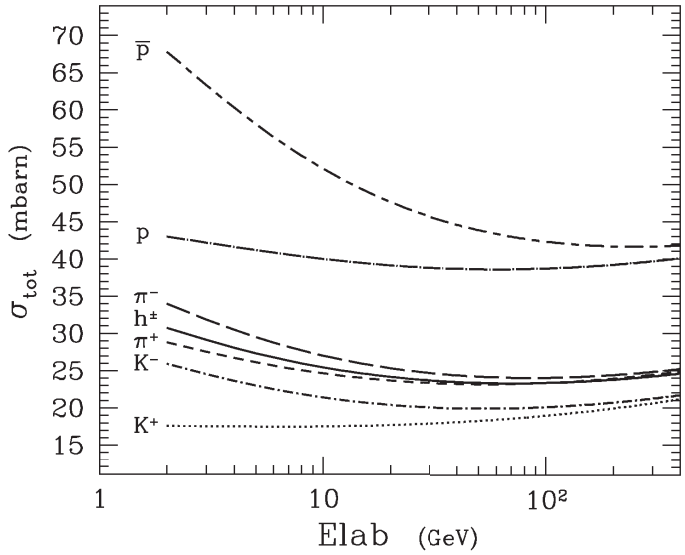

Figure 2.1: Hadron-nucleon cross section (from [Accardi 2003]).

The Monte-Carlo event generator called GiBUU[1] [Falter 2004, Gallmeister 2008] is based on a pure absorption model. They use PYTHIA [Sjostrand 2006] for the hard interaction and the fragmentation and they implement, in addition, several phenomena to describe nuclear reactions. In the initial state, the Fermi momentum of nucleons is implemented and the principal feature of the generator is to treat outgoing hadrons with the BUU transport equation [Teis 1997].

[1] The Giessen Boltzmann-Uehling-Uhlenbeck project.

PYTHIA produces the hadrons based on the Lund fragmentation model [Andersson 1983]. This is a string model, the strings are representing color fields and quark anti-quark pairs are produced when the string breaks. However, in PYTHIA, the Lund model is set for fragmentation in vacuum, by using it directly the authors assume that this process is not modified in the nucleus.

In the Lund model scheme, shown in figure 2.2, the hadron produced from the first hard quark is called a rank 1 hadron (like P_1 in the figure) and then the hadron to its left is of rank 2 (like P_2) and so on. Here, the current and target fragmentation regions can be redefined. The hadrons produced on the hard quark side are in the current region and, by opposition, the target fragmentation is composed of the hadrons produced on the side of the remnant diquark of the target. This theoretical definition of the fragmentation regions is very important and it will be clear in the chapter 3 that the experimental definition based on the sign of x_F is too simplistic. However, we can already see that hadrons produced at low energy accumulate at $x_F \sim 0$, where both regions are mixed, leading to possible confusion. It is interesting to note that, unlike most other models, GiBUU can describe both regions. This is one clear advantage when using the Lund model instead of fragmentation functions, which describe only the current region.

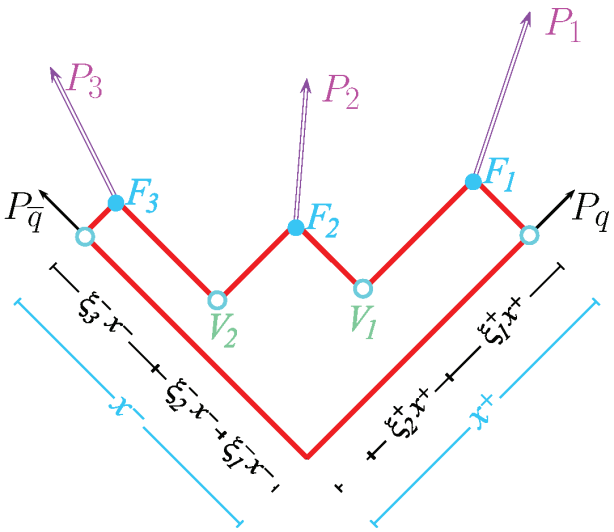

Figure 2.2: Sketch of the Lund model in the $q\bar{q}$ hadronization case. The vertical axis represents time and the horizontal one is the x-axis. The V_i vertexes are the breaking points of the strings, where new $q\bar{q}$ pairs are produced. The F_i, where the $q\bar{q}$ encounter, are the formation point of the hadrons P_i (figure from [Gallmeister 2005]).

The GiBUU absorption model also follows the results provided by the Lund model to evaluate the hadron production dynamics. The production time is set at the creation time of the first constituent quark of the hadron and the formation time is set at the junction point of its constituent quark and anti-

quark (or quark and diquark). A remarkable consequence is that the hadron of rank 1 has a production time of 0, i.e. is produced at the interaction point.

As previously mentioned, using cross section of hadrons on their whole path in nuclei gives a very poor description of hadronization in nuclear matter. One of the reasons is the need to determine a production time, another is that the prehadron does not necessarily behave like the final produced hadron. This last phenomenon is known as color transparency [Brodsky 1988] and is an analog of a classical feature of electric dipoles. Indeed, when the electric dipole size is small, its electro-magnetic field cancels at long distance leading to a vanishing cross-section. Within QCD, color dipoles are expected to behave similarly and these small size color dipoles can be theoretically treated [Kopeliovich 2007, Kaskulov 2009]. We can experimentally probe them at high energies [Mardor 1998] and it was shown that the color transparency effect is seen at energies as low as $Q^2 = 1$ GeV^2/c^2 [Clasie 2007, El Fassi 2012]. Therefore, this effect might be important for hadron production on nuclei at all energies considered in this thesis.

The color transparency is usually studied with diffractive processes, where Q^2 gives indication on the size of the prehadron. In the case of DIS, the effect is harder to estimate because the size of the prehadron is not known directly and needs to be modeled. To solve this problem, the GiBUU model was tested with different prehadron cross section behavior [Gallmeister 2008]. The best agreement with data from HERMES and EMC was found using a linear growth of the cross section, from null, at the production time, to the full hadronic cross-section, at the formation time.

Then hadrons and prehadrons, with their respective cross-sections, are introduced in the BUU transport model. This model is taking into account the usual absorption of hadrons by the nuclear medium but also the coupled channels, including side-feeding[2]. The GiBUU simulation takes also into account the Pauli exclusion for low energy baryonic decays. Overall, GiBUU is a very complete model for hadron absorption and is valid for comparison with data on a very large energy scale and for a very large kinematic range. Eventually the final results [Gallmeister 2008] (figure 2.3) give an overall good description of the HERMES and EMC data. However, some particles are badly reproduced, like protons and K^-. As it will be seen in chapter 3, these difficulties are experienced by most models.

Several other hadron absorption models exist, however they can, sometimes, give very different results. Those other models are generally not based on event by event simulations like PYTHIA. Instead they use fragmentation functions, parton distribution functions and cross sections. Therefore, they do

[2] Side-feeding meaning that the nature of a hadron can change and that new hadrons can be produced during the transport.

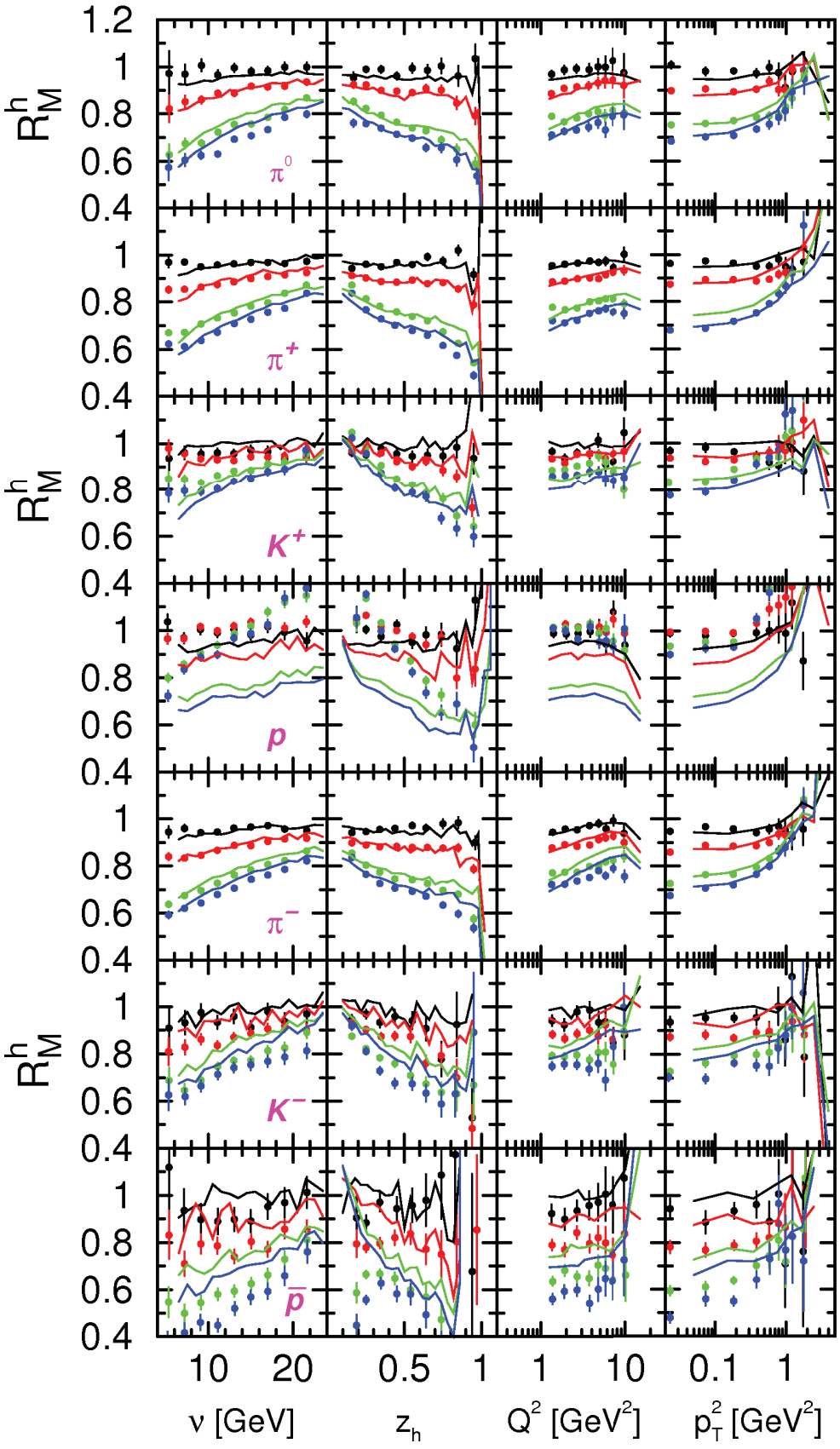

Figure 2.3: Multiplicity ratios measured by HERMES [Airapetian 2007] (points) and calculations from GiBUU [Gallmeister 2008] (lines). The colors represent the different targets, black for ^{4}He, red for ^{20}Ne, green for Kr ($A \sim 84$) and blue for Xe ($A \sim 131$).

not include the target fragmentation region and ignore it to focus only on the current region and its attenuation. However, even if experiments usually focus on the current region, the separation with the target region is not always very clear and the need for high statistics often wins over strict cuts to significantly reduce the target fragmentation contamination in the data. This can have a significant impact when comparing experimental results with theoretical models. On the other hand, the target region could also provide useful information by itself, but this part is not well studied yet.

Kopeliovich *et al.* [Kopeliovich 2004] propose to consider the prehadron as a $q\bar{q}$ pair and evaluate its cross section using a dipole-nucleon cross section, which is studied extensively for diffractive processes at low x_{Bj}. In this specific case, the authors calibrated their calculation using nuclear shadowing data ($x_{Bj} \gtrsim 0.02$) and real photon ($Q^2 = 0$) experiments. The size of the $q\bar{q}$ pair is estimated, at its origin, from its transverse momentum. Then, integrating over all possible paths leading to the observed hadron, they calculate the medium transparency as seen by this prehadron.

The hadron formation time is calculated using the uncertainty principle,

$$t_f = \frac{2 z_h \nu}{m_{h'}^2 - m_h^2} \tag{2.1}$$

with h' the first excitation of the hadron h. In this model, most of the attenuation is due to prehadron absorption, but other effects, in particular parton energy loss, are also included for the complete description of the process.

Accardi *et al.* [Accardi 2003] use two different models, Bialas-Gyulassy [Bialas 1987] and Bialas-Chmaj [Bialas 1983], which are mostly similar except for timing details. In these models, the absorption is treated with two parameters, the cross section of the prehadron and the production time. It is a standard treatment of the absorption of hadrons, the difficulty lying in the choice of the parameters. The original works just fitted these parameters to the data, in their work, Accardi *et al.* use, for the production times, a mean value given by the Lund model and, for the prehadron cross-section, a mean value between several hadron species ($h^{\pm}$ in figure 2.1). These estimations can seem rough, but they are strongly constrained compared to the former free parameters. However, it is important to note that this is not a pure model and modification of the fragmentation functions are also involved to described the data.

2.3 Parton Energy Loss

Most of the published work on parton energy loss is focused on heavy ion collisions and the characterization of the quark-gluon plasma. But, usually, the parton energy loss calculations can also be applied to the cold nuclear matter. This is especially interesting in the nuclear DIS case, as a better determination of the initial kinematic allows a precise comparison with experiments. However, some authors, focusing on heavy ions collisions, use assumptions in their calculations which limit their validity for cold nuclear matter. In this section, only models that were designed or successfully tested with cold nuclear matter are presented.

It is also important to mention another process, usually described within the same theoretical framework, and even, sometimes, used to calibrate parton energy loss models. The Drell-Yan reaction ($q\bar{q} \rightarrow \gamma^* \rightarrow l^+l^-$) is obtained using an hadronic beam sent on an hadronic target and observing lepton pair production. Indeed, because of their electromagnetic nature, the final state interactions are strongly suppressed, the hadron absorption and fragmentation function modification are, therefore, not relevant in this process. In nuclear targets, the quark from the beam loses energy by going through the nuclei before its interaction and the shift in distribution of the final state energy gives directly the parton energy loss [Johnson 2001]. Therefore, Drell-Yan experiments are mainly sensitive to initial state interaction. Hence, making a consistent picture of partonic energy loss for both DIS and Drell-Yan is a good indication that both initial and final state interactions are understood.

Most of the parton energy loss calculations presented here are made using the Glauber approximation, i.e. the parton interacts on independent scattering points. This has several implications, one being that the medium is static during the interaction. This assumption is motivated by the speed of most particles which is close to the speed of light. However, this might not apply, especially at lower energies and for heavy quark production. In the Glauber approximation, multiple soft interactions in the medium are often considered, but calculations for a single hard interaction can also satisfy the hypothesis.

Calculations are possible, within pQCD, for the parton energy loss in the Glauber approximation. We should highlight first the upper boundary found by Brodsky and Hoyer [Brodsky 1993]. They constrain the quantity of momentum transferred that can be resolved between two scattering points using the uncertainty principles. This boundary reflects the time necessary for the gluon field to be generated and gives a constrain on the energy loss as a function of the mean induced transverse momentum. Work from [Baier 1997], often referred as BDMPS (for the authors Baier, Dokshitzer, Mueller, Peigné and Schiff), goes beyond this simple picture. They carry out a calculation of

the parton energy loss process and find the total energy loss to be

$$-\Delta E = \frac{\alpha_s C_R}{8} \frac{\mu^2}{\lambda_g} \tilde{\nu}(B^2)\, L^2, \tag{2.2}$$

with μ the impact parameter and with $\frac{\mu^2}{\lambda_g}\tilde{\nu}(B^2)$ that can be related to the gluon distribution in the nuclei $G(x_{Bj}, Q^2)$ by

$$\frac{\mu^2}{\lambda_g}\tilde{\nu}(B^2) = \frac{4\pi^2\alpha_s C_R}{N_c^2 - 1}\, \rho x\, G(x, \mu^2/B^2). \tag{2.3}$$

This relation is very interesting as it allows to relate the energy loss, in a given nucleus, to its gluon distribution.

Furthermore, the authors of [Baier 1997] refined the result from [Brodsky 1993] mentioned before. Indeed, they directly relate the energy loss to the transverse momentum broadening with

$$-\frac{dE}{dz} = \frac{\alpha_s N_c}{8}\, \langle k_\perp^2 \rangle. \tag{2.4}$$

This result links the parton energy loss to the induced parton transverse momentum. This relation is particularly useful because it provides a way to access indirectly the quark energy loss, which is very difficult to measure directly.

Salgado and Wiedemann [Salgado 2002] extended the BDMPS calculations, in particular their emitted gluon spectrum, to obtain the quenching weights in finite medium. These are used by several authors and are the probability distributions for a quark to lose a given energy at given kinematics and medium properties. The calculation from [Salgado 2002] will be reviewed in more detail and applied to nuclear DIS in chapter 4.

Kopeliovich *et al.* [Kopeliovich 2004] also use BDMPS results (eq. 2.4) to estimate the impact of energy loss in DIS data from HERMES. They use values of $\langle p_\perp^2 \rangle$ measured with Drell-Yan reaction to calibrate their predictions. They also include some modification of the energy loss due to the modification of the hard scale: $Q^2 \to Q^2 + p_\perp^2$. This last modification also enhances the suppression of hadrons in the final state. However, this model leads to a suppression from parton energy loss much smaller than the one observed, it includes prehadron absorption effects in order to match the experimental multiplicity ratios.

Wang and Wang [Wang 2002], use slightly different assumptions. Landau-Pomeranchuck-Migdal (LPM) effect is used to take into account destructive interferences in the gluon emission. They eventually find the energy loss $dE/dz \simeq 0.5$ GeV/fm by fitting the data (their results are presented in fig-

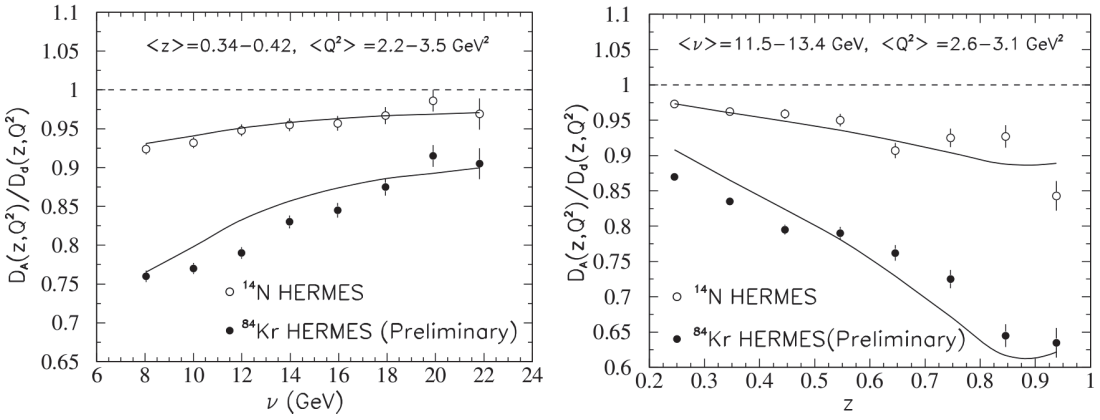

Figure 2.4: Multiplicity ratios measured by HERMES [Airapetian 2007] (points) and model from Wang and Wang [Wang 2002] (lines).

ure 2.4). They link directly the energy loss with the gluon distribution in nuclei and argue that the mismatch between Drell-Yan measurement and HERMES comes from the high Q^2 dependence of the gluon distribution $G(x_{Bj}, Q^2)$ at low x_{Bj}. Another interesting feature of this model, is its ability to describe heavy ion data from PHENIX using the same framework. They find a much larger energy loss, $dE/dz \simeq 7.3$ GeV/fm for hot nuclear matter. This result gives an interesting comparison of the nuclear matter in extreme conditions, and eventually of the quark gluon plasma, with the cold nuclear matter.

2.4 Medium Modified Fragmentation Functions

There are many reasons to imagine that fragmentation functions are modified in nuclei. Work by Accardi *et al.* [Accardi 2003] is based on models explaining the EMC effects by rescaling Q^2. In these models, the parton distribution functions $f_q(x_{Bj}, Q^2)$ are rescaled by a factor Λ as $f_q(x_{Bj}, \Lambda Q^2)$. As the fragmentation functions are driven by a similar DGLAP evolution, it is natural to extend the rescaling to them. They [Accardi 2003] found that this effect alone can explain hadronization data at high energy (EMC) but is insufficient for lower energy (HERMES) where prehadron absorption need to be included. Also, as it is unclear if Q^2 rescaling is the right way to model the EMC effect, therefore this model is not build on a strong ground. But it has the merit to offer a consistent picture for two apparently unrelated phenomena.

Similarly to energy loss models, the modified DGLAP models can be applied to both heavy ion collisions and nuclear DIS, moreover the effect is also linked with the gluonic properties of nuclei. In fact modifying the DGLAP equations can be seen as a parton energy loss model in which radiated gluons are reintroduced in the parton shower. Indeed, Deng *et al.* [Deng 2010] make a

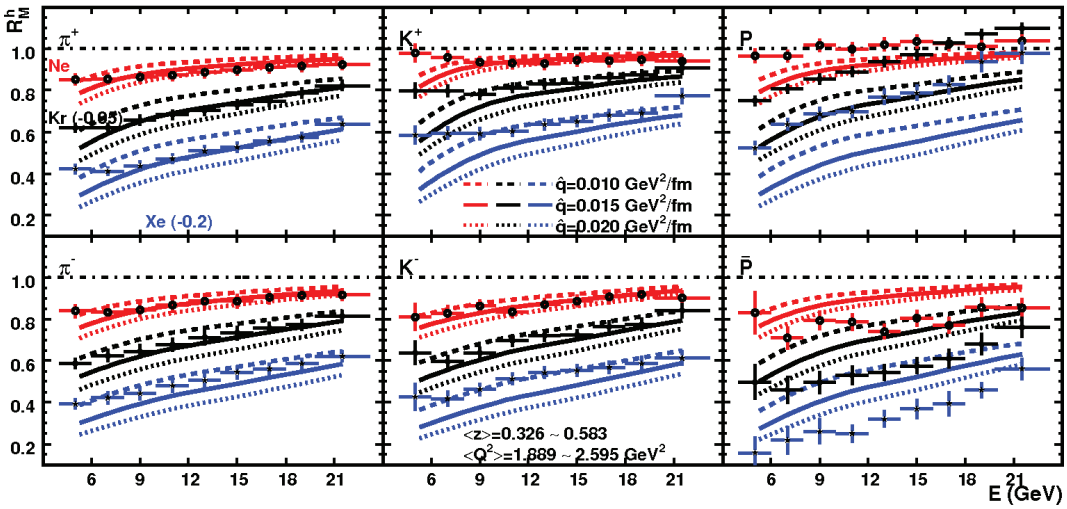

Figure 2.5: HERMES multiplicity ratios results [Airapetian 2007] (points) compared to a medium modified fragmentation calculation from Deng *et al.* [Deng 2011] for different transport coefficient of the medium $\hat{q}$ (lines). The colors represent the different targets (with some offset for clarity) red for ^{20}Ne, black for Kr ($A \sim 84$) and blue for Xe ($A \sim 131$).

coherent picture of the two processes and obtain a good description of the HERMES pion and kaon data (figure 2.5).

Similarly, a way to see medium modified fragmentation functions is to extend a standard parton energy loss model and account for the gluons emitted by the struck parton. In this case the fragmentation functions in vacuum are not modified but the increase of gluons can affect the produced flavor. In [Hwa 2002] it is shown that gluon emissions can significantly enhance the number of strange particles compared to light flavors. This picture is supported by a global fit of nuclear fragmentation functions [Sassot 2010] using both HERMES lepton-nuclei and RHIC nuclei-nuclei data sets. They find an enhancement of the gluon contribution in nuclei (figure 2.6). This phenomenological study can be interpreted in many ways but gives interesting perspectives on the possible effects of partonic energy loss on fragmentation functions. This question is especially important because of the recent observation of flavor effects and two hadrons production by HERMES [Airapetian 2006, Airapetian 2007] (see chapter 3 for detailed discussion).

Finally, a more standard way to look at fragmentation in nuclei is presented in [Gyulassy 1990]. They consider that the struck quark reinteracts similarly with each nucleons on its path. In this model, called string-flip model, a usual hadronization takes place each time the parton encounters a nucleon. The figure 2.7, sketching the process, shows that, compared to the classical figure in vacuum, the current hadrons are here produced by a string with lower energy and that more hadrons are produced in the target fragmentation region in this

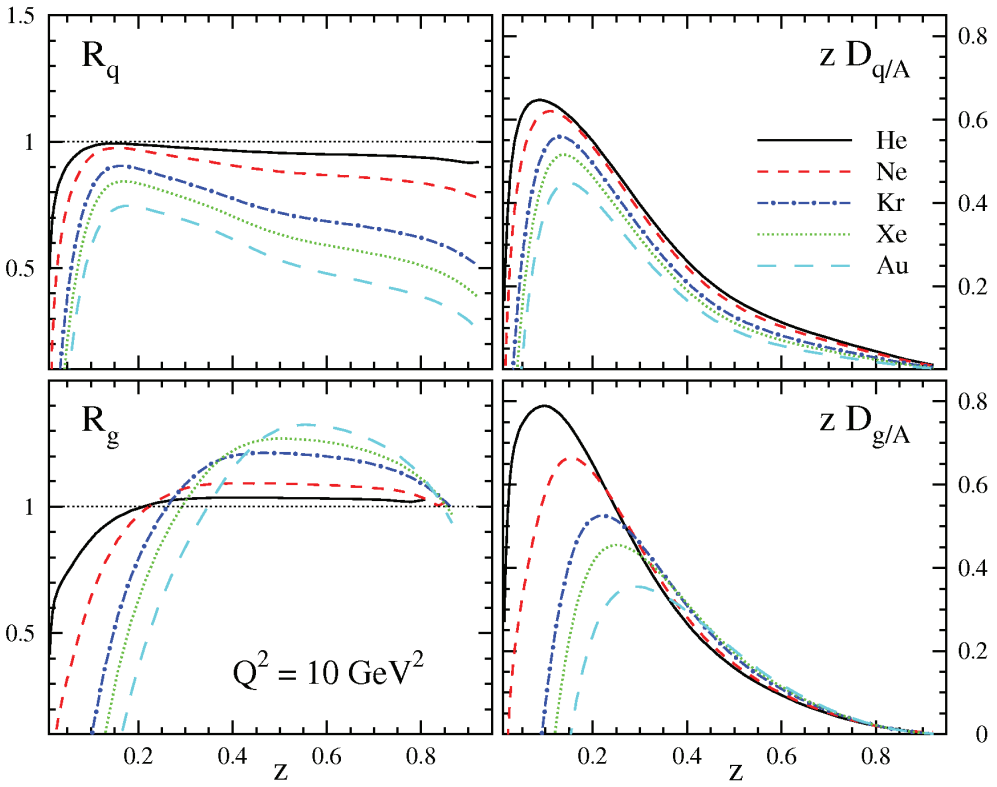

Figure 2.6: The right panels show the fragmentation functions for π^0 production from quark (top) and gluon (bottom) in various nuclei, the left panel shows the ratios of these functions with the usual fragmentation functions on nucleon (figure from [Sassot 2010]).

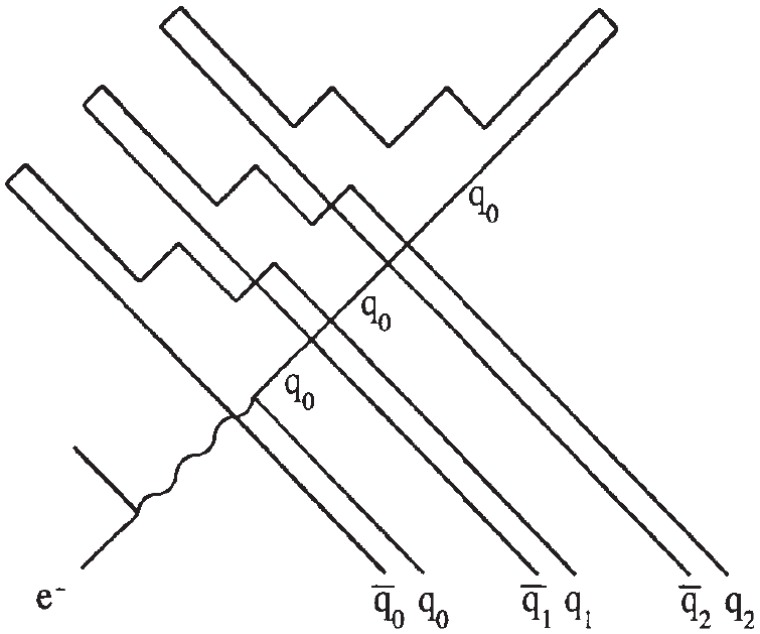

Figure 2.7: Representation of the string-flip model from [Gyulassy 1990], q_0 is the struck quark and the $\bar{q}_i$ are diquarks. The $\bar{q}_0 q_0$ represents the nucleon interacting with the virtual photon, the others pairs represent nucleons encountered by the leading quark (q_0) before exiting the nucleus.

case. This old and rather simple model lacks of comparison with recent data, but the global features seen in data are reproduced.

2.5 Conclusion

We have seen that many models describe the data with many different effects. To conclude this chapter, the hybrid models previously mentioned are briefly reviewed in order to clarify their global mechanisms.

Model from Accardi *et al.* [Accardi 2003] was mentioned several times before and is interesting because of its originality. The medium modified fragmentation are generated in a consistent way with the Q^2 rescaling model for the EMC effect. However this effect alone does not explain the attenuation observed by HERMES [Airapetian 2007] and hadron absorption effects have to be taken into account. The hadron absorption in this work is based on previous work from Bialas and collaborators [Bialas 1983, Bialas 1987].

Work from Kopeliovich *et al.* [Kopeliovich 2004] provides an original way to handle the prehadron as a color dipole, similarly to the one produced in diffractive processes. In this work, parton energy loss is a small correction applied using BDMPS results calibrated with Drell-Yan data.

Table 2.1 summarize briefly the various models mentioned in this chapter. It is important to point out that many of these used only a little part of the existing data for their comparisons, moreover these comparisons are sometimes based on preliminary data which are slightly different from the published one.

Authors	References	Model Description	Multiplicity Results	$p_\perp^2$ Broadening Results
Accardi *et al.*	[Accardi 2003]	Q^2 rescaling of FF + hadron absorption	Few comparisons with HERMES & EMC	None
Arleo	[Arleo 2003b]	BDMPS based parton energy loss (quenching weight calculation)	Scarce comparison with HERMES	$\hat{q} = 0.75$ GeV/fm^2 too large for HERMES
Deng *et al.*	[Deng 2010] [Deng 2011]	Modified DGLAP evolution	Few comparisons with HERMES	$\hat{q} = 0.015$ GeV^2/fm coherent with HERMES
Falter *et al.* (GiBUU)	[Falter 2004] [Gallmeister 2005] [Gallmeister 2008]	Pure hadron/prehadron absorption	Extensive comparison with HERMES & EMC	None
Gyulassy and Plümer	[Gyulassy 1990]	Medium modified FF using string-flip model	Comparison with old data (EMC & SLAC)	None
Kopeliovich *et al.*	[Kopeliovich 2004] [Domdey 2009] [Ciofi degli Atti 2005]	Q^2 rescaling of FF, energy loss and prehadron absorption	Extensive comparison with HERMES & EMC	Extensive comparison with HERMES
Salgado and Wiedemann	[Salgado 2002] [Salgado 2003]	BDMPS based parton energy loss (quenching weight calculation)	Few comparisons with HERMES See chapter 4	Extensive comparison with HERMES See chapter 4
Wang and Wang	[Wang 2002]	Pure parton energy loss	Few comparisons with HERMES	None

Table 2.1: Summary of the models discussed in chapter 2.

Chapter 3

Overview and Interpretation of Existing Data

3.1 Introduction

In this chapter, we will present an overview of the existing measurements and their implications. First, the early nuclear SIDIS data, which lack of statistical precision and are usually integrate over all hadron species. Little can be learned from these, yet a global picture is emerging from them and some of their ideas might be worth pursuing in the analysis of recent data sets. Second, the results published in the last decade, especially from the HERMES collaboration, which greatly improved our understanding of hadronization and revealed many astonishing features. Even if their statistical precision is not good enough to constrain all the models features, significant results emerge and help to direct the scope for future experiments. These future programs will be briefly mentioned in the conclusion of this chapter and detailed in chapter 8.

3.2 Early Results

3.2.1 SLAC Results

The first results on hadron production from nuclei were provided by Osborne *et al.* [Osborne 1978] from an experiment at the Stanford Linear Accelerator Center (SLAC). They used a 20.5 GeV electron beam on various nuclear targets: deuterium ($A = 2$), beryllium ($A = 9$), carbon ($A = 12$), copper ($A \sim 64$) and tin ($A \sim 119$). Because they had experimental limitations and a limited knowledge on electron scattering at that time, these early data have several intrinsic problems. First, the measurement of the hadrons is summed over all species. As the protons might not come from the fragmentation of a parton, because of their presence in the target, one may prefer to look at the negative hadrons for an easier interpretation in terms of hadronization. Another issue

is the normalization of their multiplicity ratios, they used

$$R_A^h(z_h, P_\perp^2) = \frac{N_A^h(Q^2, \nu, z_h, P_\perp^2)/\rho_A}{N_D^h(Q^2, \nu, z_h, P_\perp^2)/\rho_D} \tag{3.1}$$

with ρ_T the density of the target T. Unlike the usual multiplicity ratio (equation 1.1), this observable is directly sensitive to initial state nuclear effects, such as the EMC effect. However, even if, because of these problems, a quantitative statements cannot be made, a global picture is still emerging from their data and gives valuable information.

Abstract from [Osborne 1978]

We have measured the electro-production of hadrons from nuclei and compare it to the electro-production from deuterium. We find an attenuation of the forward component which increases with A. The attenuation is less for lower hadronic momenta, but is not a strong function of the other electro-production variables.

This very short abstract summarizes very well their experiment and their results. The first result is the attenuation of forward hadrons and that this effect increases with A. This indicates that the produced particles interact with the medium and somehow get suppressed. Figure 3.1 shows their main results, the multiplicity ratios as a function of x_F[1] for various nuclei. The suppression observed also increases with x_F. The effect in x_F is, in absorption models, related to the fact that high momentum hadrons are produced faster and, therefore, have to go through more material. This time dependence, well reproduced by the Lund model [Gallmeister 2005] (see figure 3.2), is due to the gluon emission during the vacuum fragmentation. In order to produce a hadron containing a large part of the momentum of the initial quark, it has to be produced fast, to limit the energy lost by gluon radiation in vacuum. In the case of parton energy loss models, the increased effect at high x_F is explained by the behavior of the fragmentation functions at high z. As the fragmentation functions are shifted in nuclear matter by the energy lost by the parton, the vanishing trend of fragmentation functions at high z leads to an increased relative suppression.

Finally, according to the authors of [Osborne 1978], the other variables do not affect significantly the hadron suppression. This feature of the data is important because later work will show significant dependencies on ν and $P_\perp^2$. One can guess that proton contamination and low statistical precision of

[1] In this paper [Osborne 1978], results are shown as a function of x_F (noted z_{CM}), while recent results are usually presented as a function of z. At high energy and relatively high z, these are equivalent, therefore all statements about x_F in this section are valid for z as well.

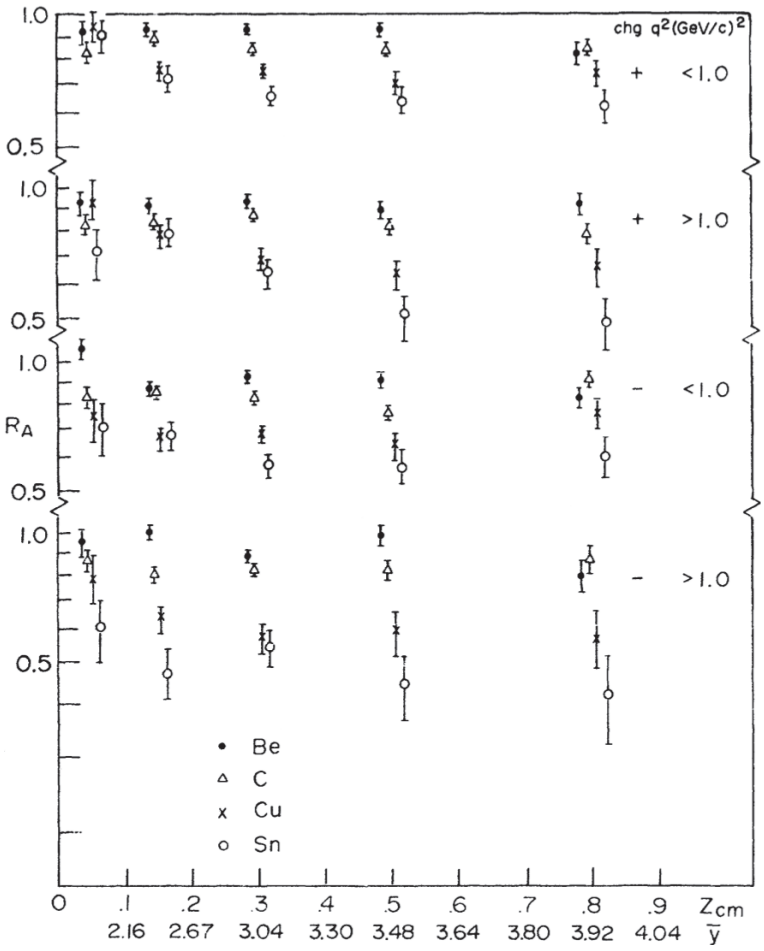

Figure 3.1: Multiplicity ratios from SLAC [Osborne 1978] shown as a function of $z_{CM} \equiv x_F$ for positive and negative hadrons and for $Q^2 > 1$ GeV^2/c^2 and $Q^2 < 1$ GeV^2/c^2.

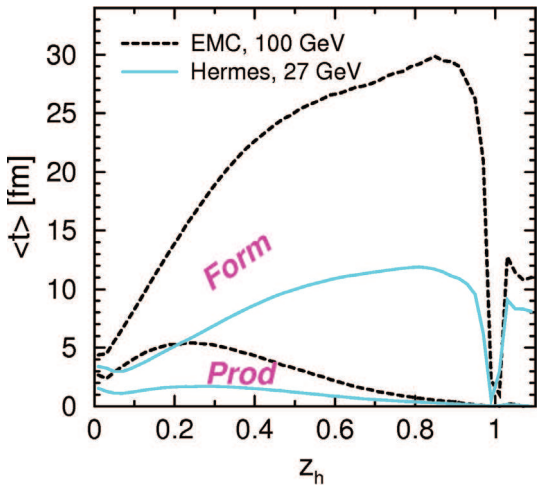

Figure 3.2: Production time and formation time extracted from the Lund model by Gallmeister *et al.* [Gallmeister 2005, Gallmeister 2008].

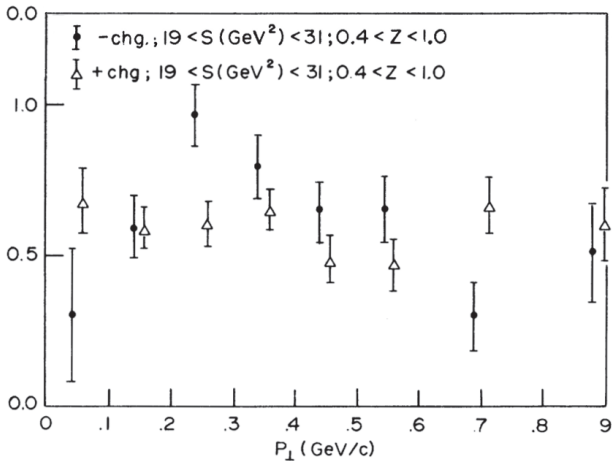

Figure 3.3: Positive and negative hadron multiplicity ratios from SLAC [Osborne 1978] shown as a function of $p_\perp$.

the SLAC experiment may have hidden the effect on ν, but the transverse momentum measurement shown in figure 3.3 is in apparent disagreement with HERMES data. Moreover, in HERMES data, protons are also subject to important variation at high $P_\perp^2$, so their presence in the data can not help to solve this problem. This question will be more discussed later but still remains problematic.

3.2.2 EMC Results

The European Muon Collaboration (EMC) at the CERN, is well known for the EMC effect, but it also provided interesting results on hadron production on nuclei, which are compiled in two papers [Arvidson 1984, Ashman 1991]. The first paper shows results which are difficult to interpret, probably because of the presence of protons in the hadron sample. The later publication is certainly more useful to learn about hadronization in nuclei. In both papers, all hadrons are detected without complete identification, as with SLAC results, causing troubles for the interpretation. Worse, the EMC results are also integrated over charge and, therefore, proton contribution cannot be avoided. As EMC effect was discovered at the time of these publications, the multiplicity ratio is defined as in equation 1.1 as it will be in all subsequent papers.

In [Arvidson 1984], the results presented were obtained using a 200 GeV muon beam on carbon (A =12) and copper ($A \sim 64$) targets. Instead of deuterium, carbon is used for observable normalizations, therefore the relative nuclear effects are slightly reduced. Strong cuts are applied to select the DIS region ($Q^2 > 5$ GeV^2/c^2, $x_{Bj} > 0.02$ and $50 < \nu < 180$ GeV), but a loose cut is applied for the current fragmentation region ($z > 0.1$). The cut on current fragmentation is always difficult to choose, especially because it reduces

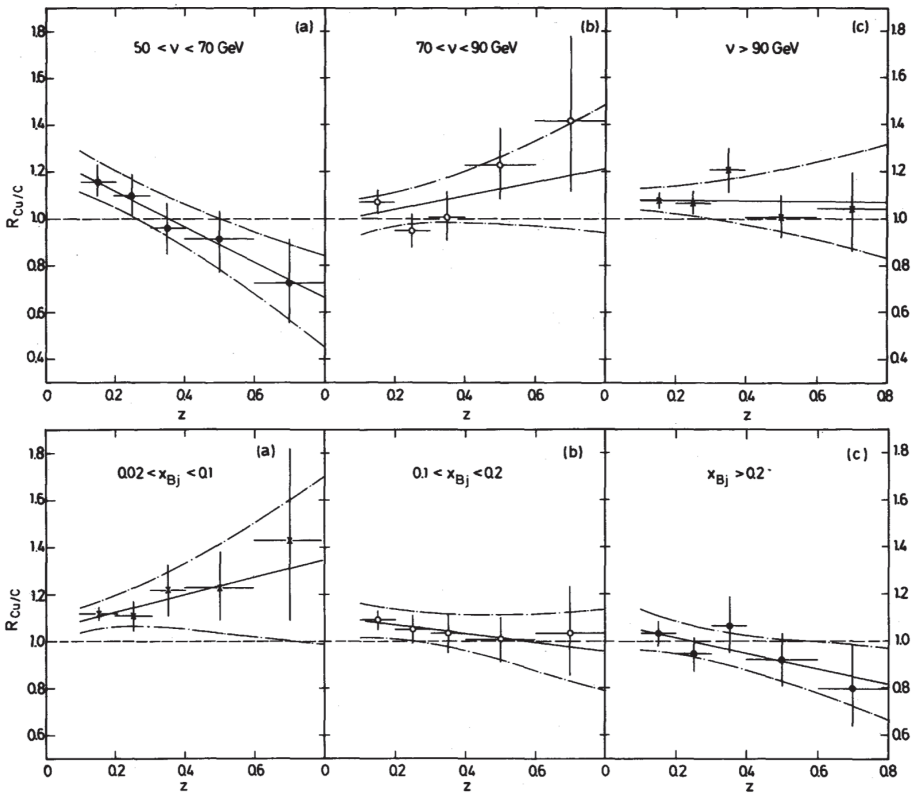

Figure 3.4: Multiplicity ratios from EMC [Arvidson 1984]. Top panels are binned in ν and bottom ones in x_{Bj}.

drastically the statistics available. Theoretical work on current fragmentation [Mulders 2001] seem to indicate a safe cut around $z \gtrsim 0.15$ for pions and $z \gtrsim 0.3$ for protons[2]. Therefore, the actual cut seems relatively low, but the subsequent contamination, due to this low cut, is difficult to estimate.

Results presented as a function of z and x_{Bj} or z and ν (figure 3.4) show new behavior. At high x_{Bj} and low ν the expected pattern is observed, the attenuation increases with z. However, in other cases, we observe a completely different picture where, instead of an attenuation, there is an enhancement, which is flat or increases with z. This is difficult to understand otherwise than to be an effect from the protons and the diffractive processes. However, the contamination from protons should be concentrated in the target fragmentation region and, therefore, lead to a stronger decreasing slope, as we observe in the low ν bin. Also, the EMC collaboration provides measurements of the number of proton and anti-proton compared to other charged hadrons. It appears that their number does not increase with A, making even more difficult to interpret these data this way. The diffractive processes contribution might be the key to understand these data, but it is difficult to make a quantitative statement. In conclusion, the lack of precision, especially in the particle iden-

[2] Calculated for $p_\perp < 0.5$ GeV/c and $\Delta\eta > 2$ at EMC energy with the eq. 17 of [Mulders 2001].

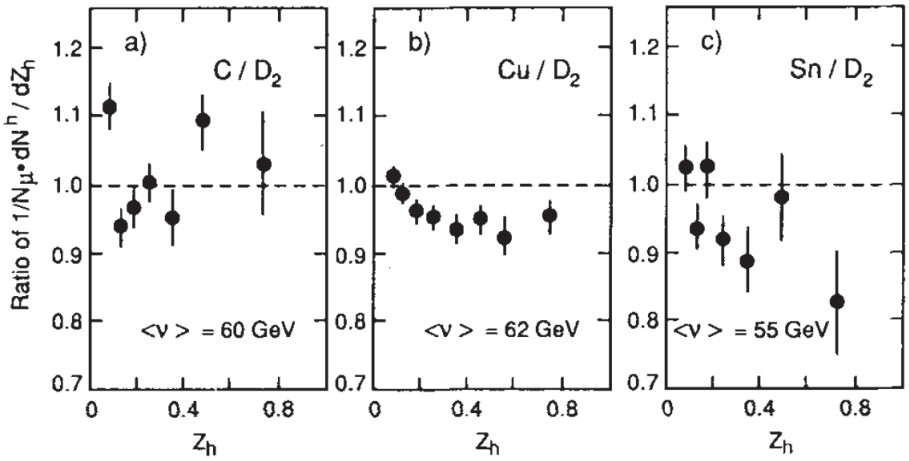

Figure 3.5: Multiplicity ratios as a function of z from EMC [Ashman 1991].

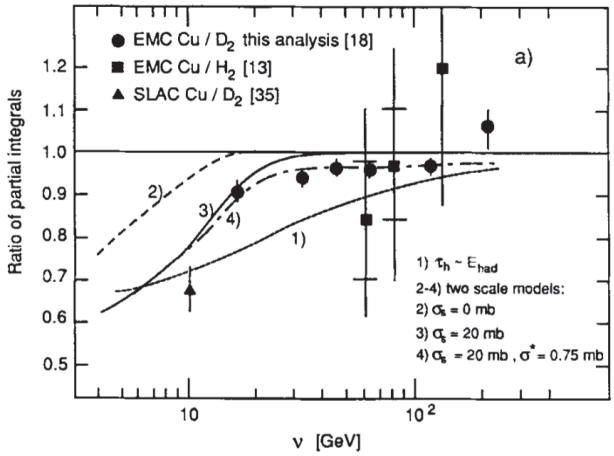

Figure 3.6: Multiplicity ratios as a function of ν from EMC [Ashman 1991].

tification, and the strange behavior of the distributions, compared to other measurements, does not allow a clear interpretation.

The EMC collaboration took more data using various muon energies, from 100 to 280 GeV, and various nuclear targets: deuterium, carbon, copper and tin. With much more statistics and deuterium as basis, these results [Ashman 1991] are easier to interpret. The z dependence of the multiplicity ratio, shown in figure 3.5, is much more coherent with the previous SLAC results, but many new features emerge. A dependence of the multiplicity ratio was found as a function of ν (figure 3.6); the attenuation is stronger at low ν and it even tends to disappear at the highest ν. No significant dependence is found in either x_{Bj} or Q^2, therefore, ν seem to be the only inclusive variable with a nuclear effect dependence. This behavior is interpreted by most of the models as a simple Lorentz boost compressing the size of the target nucleus. For hadron absorption models, this leads to a prehadron production time pushed outside of the nucleus and, therefore, less interaction. For par-

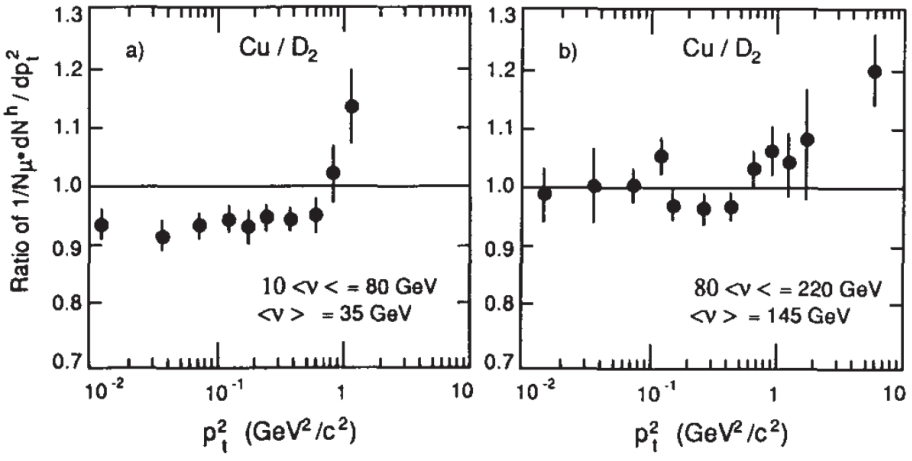

Figure 3.7: Multiplicity ratios as a function of $p_\perp^2$ from EMC [Ashman 1991].

ton energy loss or modified fragmentation models, the reduced interaction is due to the shorter time spent in the nuclei, leaving less time for gluons to be exchanged with the medium.

The $p_\perp^2$ effect was also studied in [Ashman 1991], they observed (see figure 3.7) a fast increase of the multiplicity ratio at high $p_\perp^2$. This was not seen in SLAC data, however the change begins to appear here around $p_\perp^2 = 1$ GeV^2/c^2, the limit of the SLAC measurement. This increase is due to the very small cross section at high $p_\perp^2$, which leads to a relatively more important effect. This is often referred as Cronin effect, which is, originally, a heavy ion collision effect named after James Cronin. Using another observable, $\langle p_\perp^2\rangle_A/\langle p_\perp^2\rangle_D$, EMC data seems also to lead to a suppression of $p_\perp^2$ effect at high ν. This feature is contradictory with most of the models using parton energy loss[3], but the size of the error bars is important and the choice of observable not optimal. No strong conclusion can be made from this result and more precise data are needed to interpret the $p_\perp^2$ behavior. In most models the extra $p_\perp^2$ is generated from multiple soft scattering in the nucleus; these scatterings occur at the partonic level. However, $p_\perp^2$ can be studied in more details using the variable $\Delta\langle p_\perp^2\rangle$, which was used in more recent experiments and has the advantage of a simpler interpretation.

Finally, the EMC collaboration studied the dependence of the attenuation in nuclei as a function of ϕ_h. They basically found no effects (figure 3.8), but their precision is not very good. It is interesting to point out here the link between hadronization and EMC effect. Indeed, models describing the EMC effect by a change of radius of the in-medium nucleons are leading to a change in the transverse momentum of their constituent quarks, i.e. increase of $\langle k_\perp\rangle$. One can explore $\langle k_\perp\rangle$ by measuring $\langle\cos(\phi_h)\rangle$, as shown by

[3] Other models do not describe that kind of $p_\perp^2$ effects.

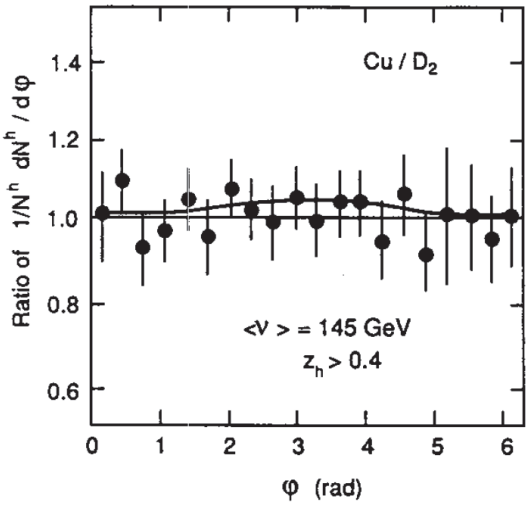

Figure 3.8: Multiplicity ratio as a function of ϕ_h from EMC [Ashman 1991].

[Cahn 1978, Konig 1982], and, therefore, access a change of the nucleon size.

3.2.3 Fermi Lab Results

Fermi National Accelerator Laboratory's (FNAL) studies on hadronization began very early with [Hand 1979]. However, because of the very low luminosity, it was difficult to compete with SLAC or EMC. Results of interest for us were published later by the E665 collaboration in [Adams 1995]. Though the focus was on diffractive processes and the shadowing effect[4], it provides with results on both shadowing and non shadowing regions.

The reported results were taken using a 490 GeV muon beam on deuterium and xenon ($A \sim 132$). The main observable is the number of Grey tracks. Grey tracks, are named after their experimental signal. Back in the time where events were recorded in photographic emulsions, these characterized slow moving protons. More specifically, in [Adams 1995], they are protons with momentum between 200 and 600 MeV and $x_F < -0.2$. These kinematic cuts ensure the target fragmentation nature of these protons. This is confirmed by the important number of Grey tracks observed in Xenon compared to Deuterium.

Figure 3.9 (left) shows very interesting features of the E665 data. The shadowing and non-shadowing regions give very similar results, that legitimate the use of these data in our context. One can associate the number of Grey tracks to the strength of the interaction. Indeed, when few or no Grey tracks are measured, the Xenon data are pretty much like Deuterium ones (figure 3.9 (right)). When more interaction occurs, the forward part is slightly suppressed and, as

[4]The shadowing effect is the reduction, at low x_{Bj}, of the structure function F_2 of nuclei compared to free nucleon.

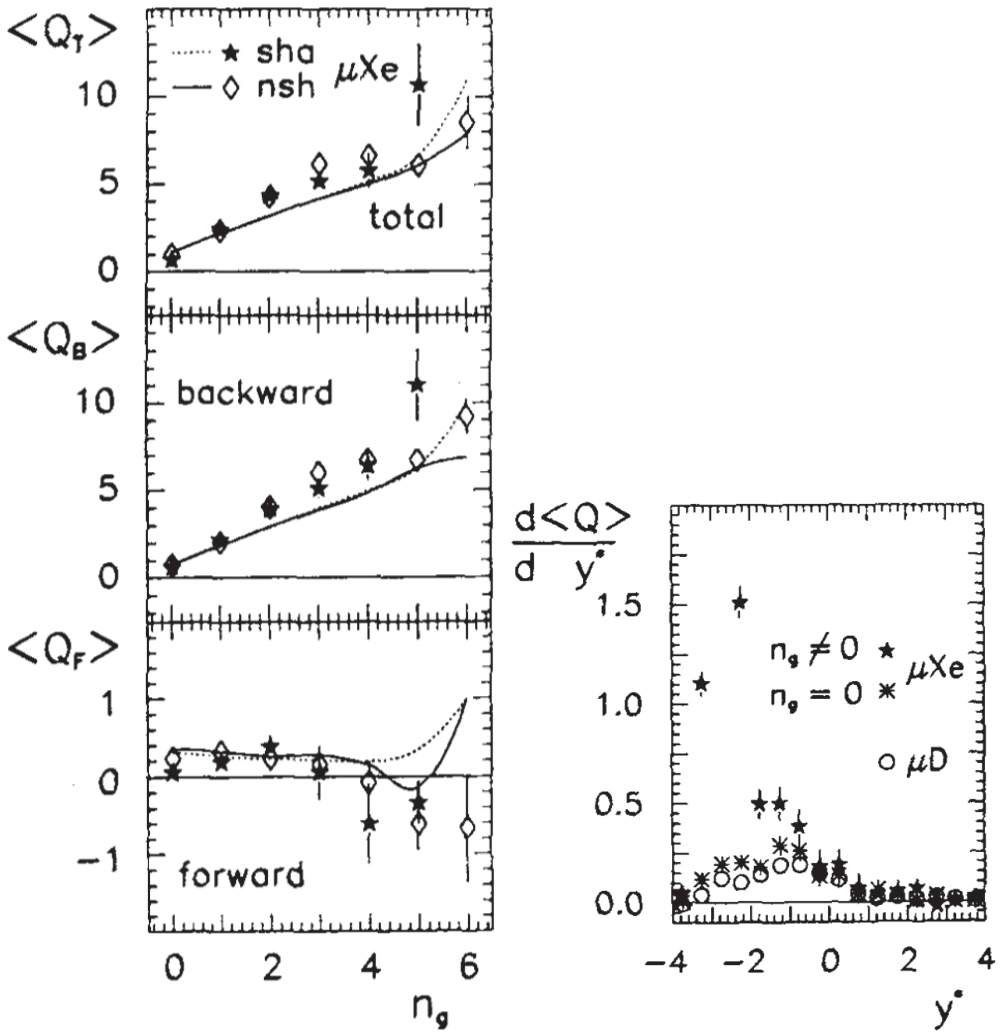

Figure 3.9: (Left) Charge measured in μD and μXe as a function of the number of Grey tracks. In top panel the total charge observed in the shadowing and non shadowing regions (respectively $x_{Bj} < 0.02$ and $x_{Bj} > 0.02$). In the middle panel only charge measured in the backward hemisphere (correspond to negative x_F) and in the bottom panel only charge in the forward hemisphere ($x_F > 0$). (Right) Charge as a function of rapidity for deuterium target and xenon target, for the later with and without observation of Grey tracks. (Figures are from [Adams 1995].)

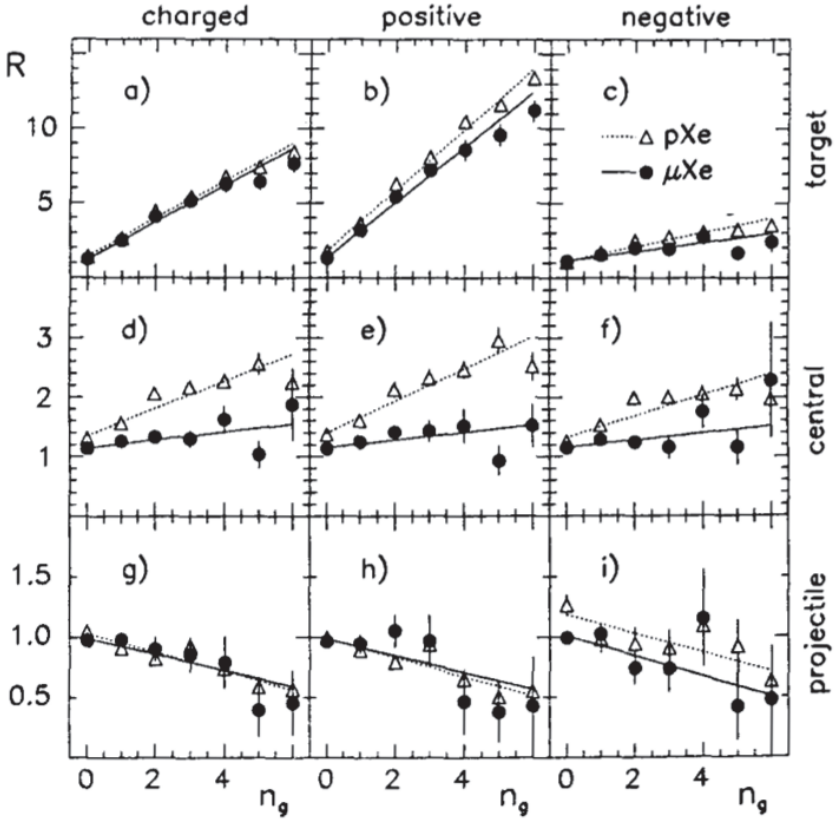

Figure 3.10: Full markers are multiplicity ratios of charged hadrons on xenon as a function of the number of Grey tracks from [Adams 1995]. Target region is defined by $y^* < -1$, central by $-0.5 < y^* < 0.5$ and projectile $y^* > 2$.

the lost energy is deposited in the nuclei expelling fragments, the backward region is enhanced. This result confirms our interpretation of previous data, but most importantly demonstrates the importance of a good separation of target and current regions as they behave differently. Moreover, effects in the current region are smaller and can be easily washed out by target fragmentation contamination. This feature could explain partly the early results from EMC [Arvidson 1984] and is illustrated in figure 3.10 where the central part, defined by $-0.5 < y < 0.5$, is dominated by target effects.

Theoretical interpretation of these data is very scarce in the context of hadronization, but globally the number of Grey tracks can be associated with centrality similarly as in heavy ion collisions. Nevertheless, these data were confronted to theory in [Ciofi degli Atti 2005], they lead to reasonable results within the model from Kopeliovich *et al.* [Kopeliovich 2004]. However, the measurement is not precise enough to obtain strong conclusions, thus more precise data on this topic would be valuable.

3.2.4 Conclusion

The results from SLAC, FNAL and CERN provide a coherent global picture of the hadronization process in nuclei. In the current fragmentation region, the produced hadrons are suppressed and this effect increases with z, x_F and A and decreases with ν. No sensitivity is observed, within error bars, to Q^2, x_{Bj} and ϕ_h and, finally, the multiplicity ratio appears to increase rapidly at high $p_\perp^2$. In the target fragmentation region, slow protons are the signature of an important interaction with the target nuclei. Subsequently the multiplicity in the target region is increased as the current fragmentation energy is reduced. This inversed effect demonstrate the importance of properly separating the two regions to interpret properly the data.

The global picture observed in the early measurements is driven by the geometry of nuclei and basic physics features, thus many models can describe easily these data. One need more precise results to discriminate and understand the processes involved. This is the goal of two major subsequent experiments, one by the HERMES collaboration, extensively described in the next section, and one by the CLAS collaboration described in details in this thesis.

3.3 Recent Results

Over the years, the HERMES collaboration, at *Deutsches Elektronen-SYnchrotron* (DESY), has published many results on hadronization in nuclei [Airapetian 2001, Airapetian 2003, Airapetian 2006, Airapetian 2007, Airapetian 2010, Airapetian 2011]. They used a 27.5 GeV positron beam for most of their data, but there is also some data taken at 12 GeV. Targets are made of deuterium, helium ($A = 4$), neon ($A \sim 20$), krypton ($A \sim 84$) and xenon ($A \sim 132$). Their results, are reviewed here by topic, instead of publication order, for clarity. Finally, very recent results from Jefferson Lab Hall C, where the beam energy is 5.5 GeV, will also be mentioned.

The two major improvements, over the previous measurements, are the increase of statistics, obtained by accumulating data over several years, and the complete particle identification. Indeed, HERMES [Airapetian 2007] provided results for the following particles species: π^+, π^0, π^-, K^+, K^-, p and $\bar{p}$. These results allow to test each particle's behavior and, subsequently, offer possibilities to test the parton flavor dependencies of hadronization. Also, as K^- is a purely sea object, it is composed of $s\bar{u}$ quarks, it is mainly produced as hadrons of rank 2, or more[5], but K^+s, composed of $u\bar{s}$, are mainly produced

[5] As defined in the Lund model (see section 2.2 and 4.2.4).

out of u fragmentation, like $\pi^+ s$, and can easily be of rank 1. These properties give access to effects linked to the rank of the hadrons. Finally, protons allow to access information about target fragmentation effects. In particular, the comparison with anti-protons, which production should be free of most of the target fragmentation contamination, should help understand both the target region and the baryon production.

3.3.1 HERMES Multiplicity Ratios

The results [Airapetian 2007] for multiplicity ratios, from the HERMES collaboration, are shown in figures 3.11. One can see that the general picture described previously is respected by all particles except protons, this might indicate that protons have played an important role in the puzzling results from EMC [Arvidson 1984].

As expected from most models, the results for all three pions (π^+, π^- and π^0) are very similar. However, the behavior of kaons is surprising, as K^-s behave like pions but K^+s are less attenuated. One can argue that positive kaons have a smaller cross section than other hadrons (see figure 2.1), leading to a reduced absorption by the medium. Indeed GiBUU model, which takes into account the hadron cross section, describe well K^+, but have a significant normalization error for K^- (see figure 2.3), making their result on this specific question not completely convincing. Another explanation, raised by Kopeliovich *et al.* [Kopeliovich 2004], is that $\pi + p \rightarrow K + \Lambda$ leads to an important contamination of positive kaons from the target fragmentation. This assumption leads them to consider only the high z component of kaons which they describe well[6] (see figure 3.12). Pure parton energy loss scenarios lack of quantitative explanation for the kaon enhancement, but one can refer to [Arleo 2003a] for a prediction of the difference between K^+ and K^- linked to the fragmentation functions in a parton energy loss model. The generation of gluons, by parton energy loss or by modified DGLAP evolution, might also lead to an enhancement of the strange sector [Hwa 2002]. This description is not completely coherent with the difference between the two charged kaons, but if the statement from Kopeliovich *et al.* is right this difference might have an unrelated origin. Furthermore, results on protons show that target fragmentation can reach relatively high z, making the difference between charged kaons even more questionable. To conclude on this, data at higher energy might be the key to separate clearly the target and current regions [Mulders 2001] and obtain results easier to interpret.

To understand the target region problems, the comparison between protons

[6] They still find a slight difference between K^+ and pions explained by the cross section argument.

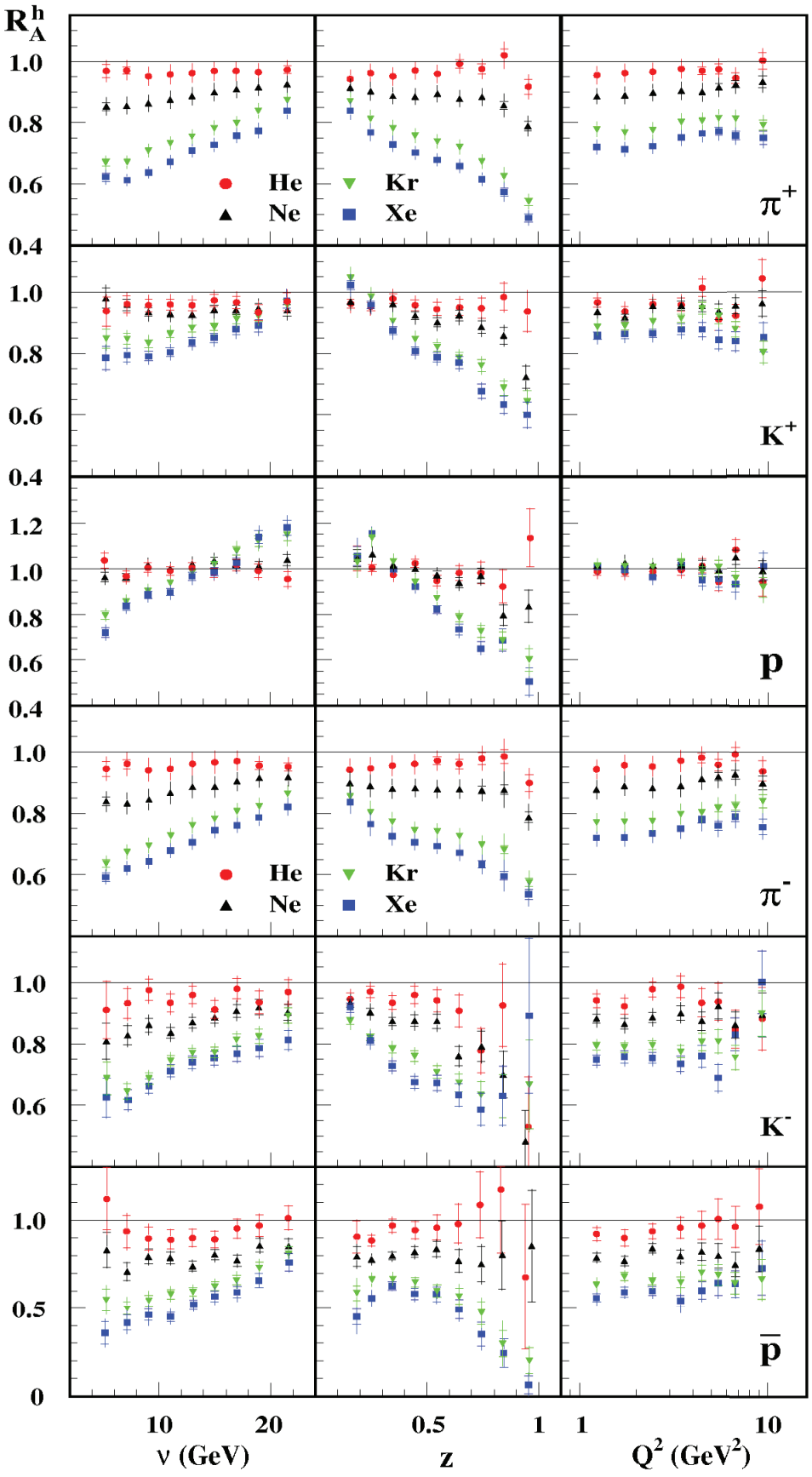

Figure 3.11: Multiplicity ratios of charged hadrons as a function of ν, z and Q^2 from HERMES [Airapetian 2007].

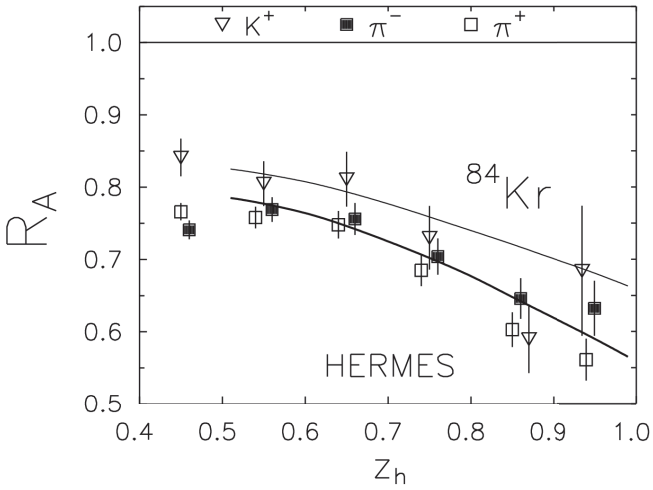

Figure 3.12: Multiplicity ratios of charged hadrons as a function of z, points are HERMES results [Airapetian 2003] and lines are from the Kopeliovich *et al.* model [Kopeliovich 2004] for pions (thick) and K^+s (thin).

and anti-protons, made available in the HERMES data, is very interesting. Because anti-protons are not present in the target and are not likely to be decay products of the target remnants, their contribution to the target fragmentation region should be small, especially compared to protons. Indeed, anti-protons behave more or less like mesons with two important differences (see figure 3.11). First, they are more suppressed over all; second, at low z the usual increase is replaced by a drop. Interpretation is, for the latter, that at very low z the absorption is large and that, if we do not see the same behavior for other particles, it is only because of the target fragmentation contribution. Indeed, the energy deposited in the nuclei leads to the production of an enhanced target region and the drop observed for anti-protons at low z is due to the absence of this compensation from target. The overall bigger suppression can be explained in various ways. In absorption models, it is obvious that anti-protons have a huge interaction cross section in nuclear medium (see figure 2.1). Also, one could argue that three quarks prehadron expand faster into hadron and are less likely to lead to color transparency effects. Another way to look at the increased suppression of anti-protons is its production threshold. As it is at higher energy than for pions, it leads to a more important suppression after the parton lost energy in the medium.

Protons, in their own way, also show an expected pattern in the figure 3.11. The multiplicity ratio becomes greater than one at low z, passing unity around $z = 0.3$. However, the ratio also becomes greater than 1 at high ν. This is a little more puzzling, but can be understood by taking into account the correlation in acceptance. Because protons are detected from 2 to 15 GeV/c, the high z events are suppressed at high ν and, similarly, low z is suppressed

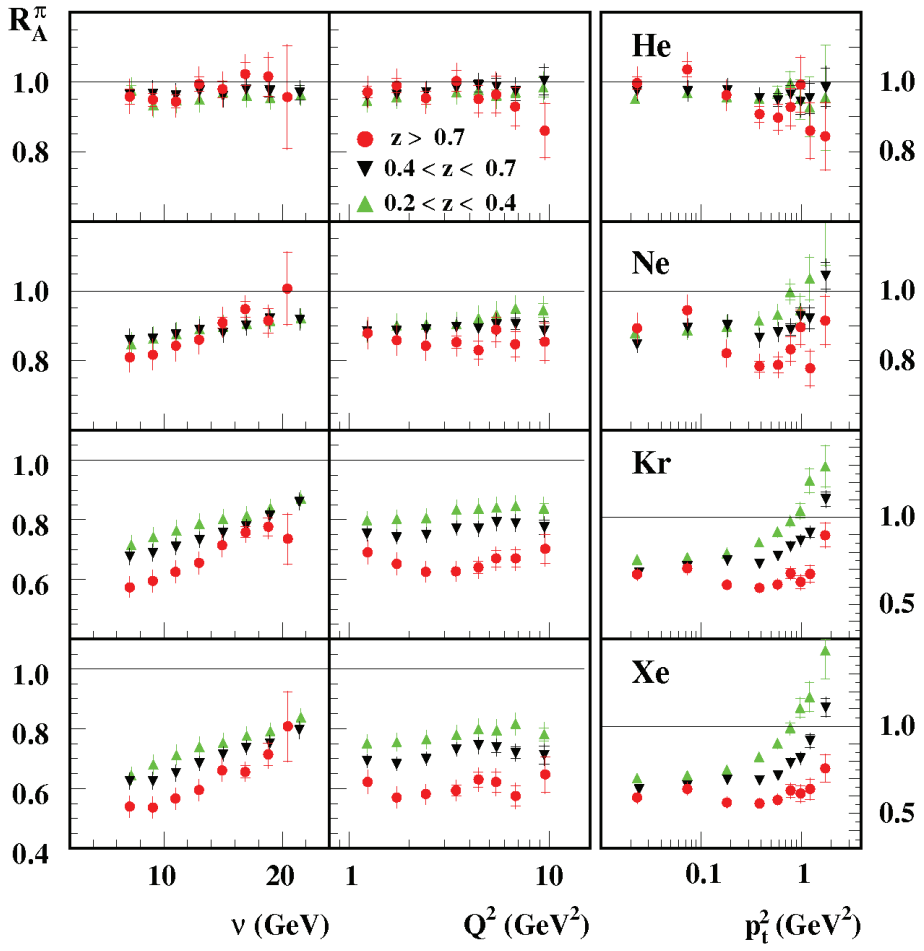

Figure 3.13: Multiplicity ratios of charged hadrons as a function of ν, Q^2 and $p_\perp^2$ in various z bins from HERMES [Airapetian 2007].

at low ν. This might accentuate the influence of target contamination at high ν and lead to the inversion of the ratio. In any case, the behavior at high z is similar to pions and this is a good indication that protons, in the current fragmentation region, behave similarly to other particles. Those data might indeed permit to identify the regions where target fragmentation has effects and where it eventually has none, helping to interpret the data properly. In this respect, it is interesting to see that no clear transition between target and current fragmentation is observed. Nonetheless, a subtle inflexion is observed for pions around z of 0.6 and is not present for protons. This might be an indication that target remnants contaminate the measurement at least up to this value, this would be coherent with theory [Mulders 2001].

Another tool, in the tracking of the target contamination, might be the multiplicity ratio as a function of $p_\perp^2$. At high $p_\perp^2$, most particles have an important increase of their multiplicity ratio. This Cronin effect is stronger for particles suspected of important target contamination (K^+ and p) and is reduced for

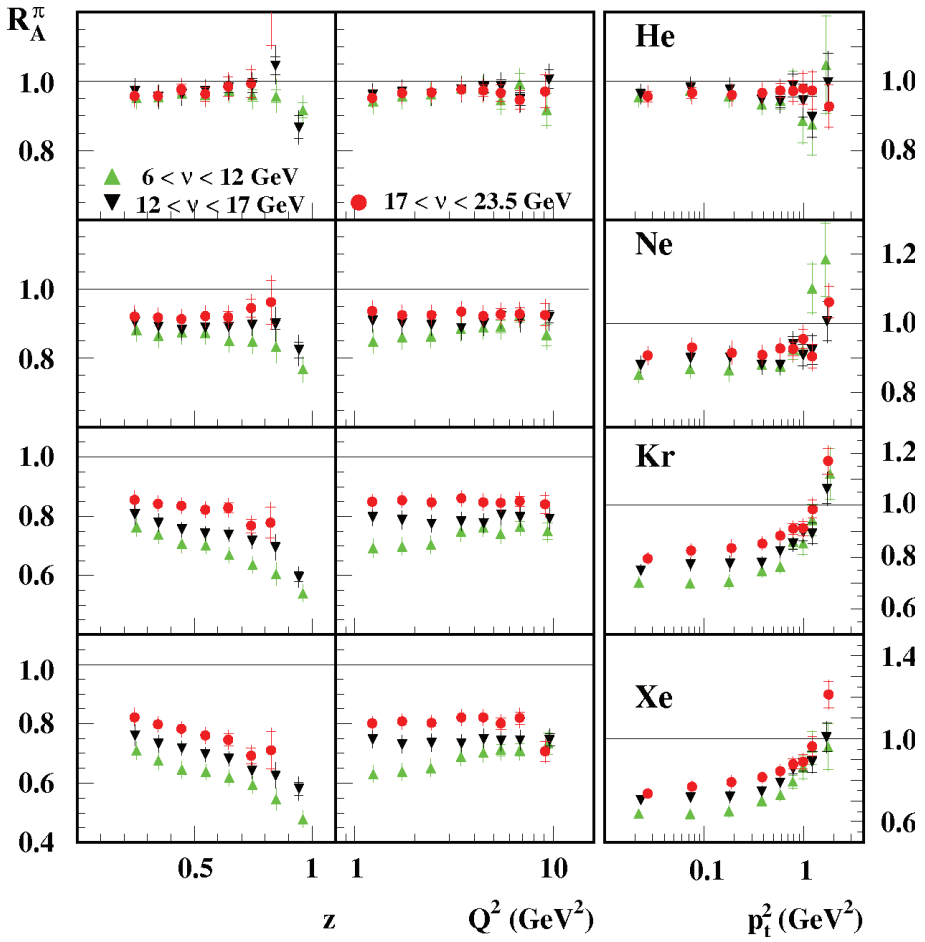

Figure 3.14: Multiplicity ratios of charged hadrons as a function of z, Q^2 and $p_\perp^2$ in various ν bins from HERMES [Airapetian 2007].

safer particles in this regard ($\bar{p}$). Therefore, the target region might contribute to the Cronin effect. This hypothesis can be related to the work on heavy ion collisions [Fries 2003, Hwa 2004] explaining the Cronin effect by a recombination of the constituents of the nuclei. To test this contribution from the target to the Cronin effect, we can use the multi-dimensional results of the multiplicity ratio from HERMES (figure 3.13). We observe a stronger Cronin effect for pions at low z, where we also expect more target fragmentation contamination, this is a confirmation of our hypothesis. This feature is important, because most models do not describe the target fragmentation region, therefore, they will not be able to reproduce this kind of measurement properly.

Data were also plotted for different ν bins (figure 3.14). In this case no surprise, all the distributions look similar with only a slight difference on the overall suppression. Still, it is interesting to look at the Q^2 dependence in these ν bins. In the first bin ($6 < \nu < 12$ GeV), a variation in Q^2 is visible, but none in the two other bins. It is difficult to interpret such a slight modification. A clearer signal can be seen in figure 3.15, where high $p_\perp^2$ events have a strong

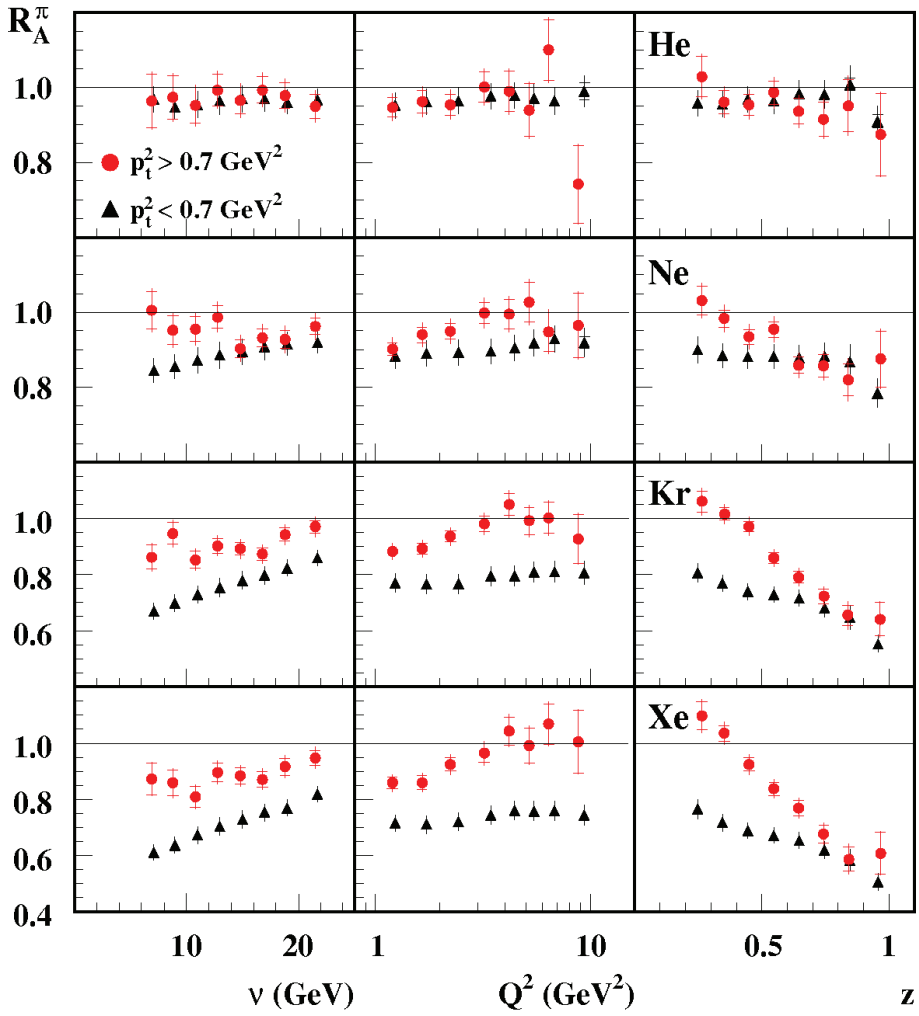

Figure 3.15: Multiplicity ratios of charged hadrons as a function of ν, Q^2 and z in two $p_\perp^2$ bins from HERMES [Airapetian 2007].

dependence in Q^2. These high $p_\perp^2$ events, as described before, are probably more contaminated by target fragmentation, but they also characterize events in which the interaction is important. Therefore, selecting these events might be a way to accentuate nuclear effects or to enhance the target region contribution. As some models, based on quark energy loss [Kopeliovich 2004], predict a slight increase of the multiplicity ratio with Q^2, this is an important question to solve. The present data shows a hint for a signal, but we need to be careful. Indeed, either a parasitic process or a kinematic correlation might lead to a wrong interpretation of this measurement.

Finally, HERMES tried to use a new variable to explore the nature of the interaction experienced during hadronization in nuclei, the formation length. They parametrize it using a Lund model formula as

$$L_c = z^{0.35}(1-z)\frac{\nu}{\kappa} \tag{3.2}$$

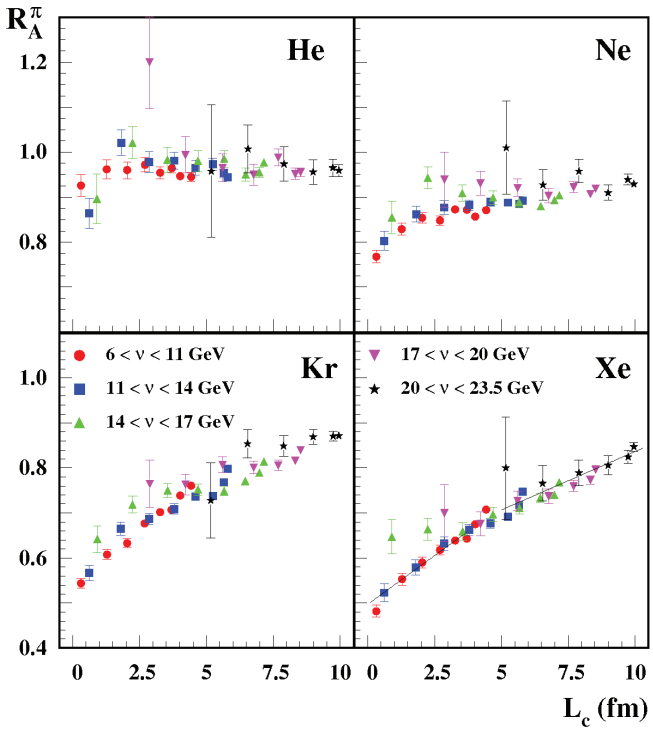

Figure 3.16: Multiplicity ratios of charged hadrons as a function of L_c (equation 3.2) from [Airapetian 2007].

with the string tension $\kappa = 1$ GeV/fm. Results are shown in figure 3.16. For helium and neon nuclei, a significant inflexion is visible at low L_c, which is expected for prehadron absorption models. For larger nuclei the inflexion is smaller and the slope extend to very high L_c, that could be an indication of energy loss effect. Also, the distribution in light nuclei does not seem to have a plateau at multiplicity ratio 1 and this could be another indication of parton energy loss. In [Airapetian 2007], the authors interpret these results as an indication that both parton energy loss and prehadron absorption are involved. However, this interpretation is model dependent and the way L_c is calculated may not be correct. Moreover, one of the results presented in chapter 4 shows a similar feature obtained with a pure energy loss model, invalidating the initial interpretation.

3.3.2 HERMES Transverse Momentum Broadening

The HERMES collaboration also provided results using the $\Delta\langle p_\perp^2 \rangle$ observable [Airapetian 2010], which is of main interest for parton energy loss models. In the pQCD framework, the **partonic** transverse momentum broadening can be directly related to the energy lost by the initial parton (see section 2.3, equation 2.4). However, what is observed is the **hadronic** transverse momentum broadening and one has to be careful how to relate it to the parton. Domdey *et*

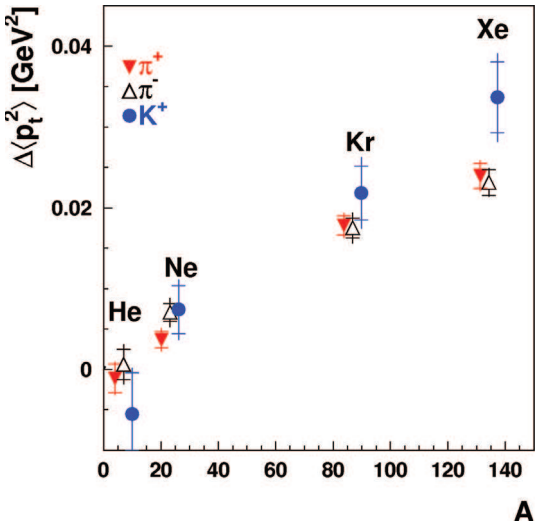

Figure 3.17: Integrated $\Delta\langle p_\perp^2\rangle$ of charged hadrons as a function of A from HERMES [Airapetian 2010].

al. [Domdey 2009] use a simple kinematic factor $\langle z^2\rangle$ to recover the partonic effect, but this might be over simplified. Indeed, a bias might be introduced by any nuclear effects, for example, hadron absorption. This will be studied in more details using a Monte-Carlo simulation in chapter 4. It is important to note that, except for some recent unpublished attempts [Gallmeister 2010], no prediction or description of the transverse momentum broadening is available for the models not including parton energy loss.

The figure 3.17 shows the integrated transverse momentum broadening obtained for π^+, π^- and K^+. At first sight, we see that both pions give very similar results. To interpret further this result, it is better to show it as a function of $A^{1/3}$, which is proportional to the radius of the nuclei and, therefore, to the path length L in the nuclear medium. $A^{2/3}$ is also of interest as BDMPS formula (equation 2.4) leads to a L^2 dependence of the energy loss and a L dependence of the $\Delta\langle p_\perp^2\rangle$. In figure 3.18, $\Delta\langle p_\perp^2\rangle$ is plotted as a function of $A^{1/3}$ and $A^{2/3}$, but it is difficult to discriminate between the two scenarios and more data are needed to settle this question. The second interesting feature is the behavior of K^+, which seem slightly different from the pions. Though, the error bars does not allow to be definitive about this measurement, if right, such a result could indicate that strange quarks behave differently than u and d quarks in medium. Of course, as it is suspected to impact multiplicity ratios, one should take into account the possibility that target fragmentation also plays a role in this feature.

Results are also available for $\Delta\langle p_\perp^2\rangle$ as a function of the usual kinematic variables (figure 3.19). In these results, several features emerge. First, $\Delta\langle p_\perp^2\rangle$ is constant with ν, which is coherent with parton energy loss calculations predict-

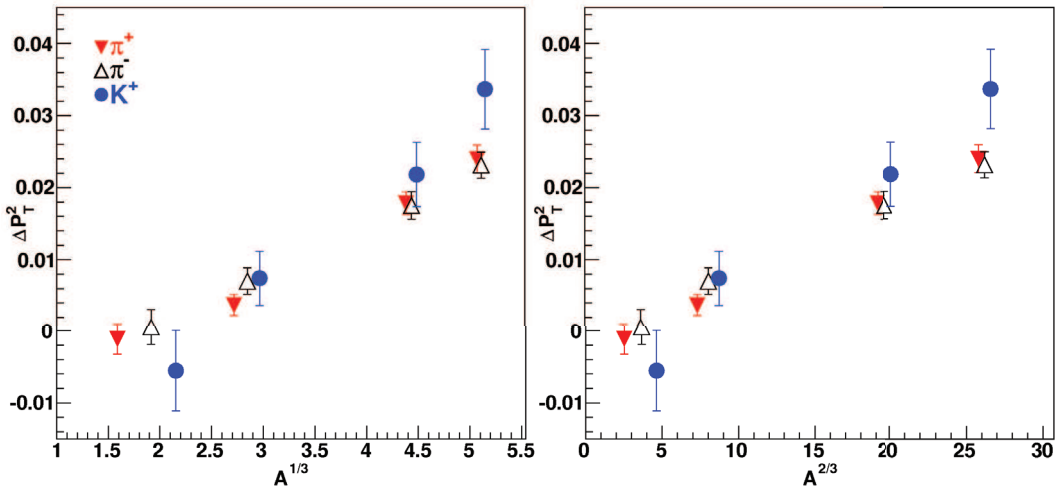

Figure 3.18: Integrated $\Delta\langle p_\perp^2\rangle$ of charged hadrons as a function of $A^{1/3}$ (left) and $A^{2/3}$ (right), adapted from [Airapetian 2010].

ing an energy loss independent of the initial energy [Brodsky 1993]. Second, $\Delta\langle p_\perp^2\rangle$ raise with Q^2, which is coherent with Domdey *et al.* [Domdey 2009] calculation, based on Kopeliovich *et al.* model [Kopeliovich 2004]. Finally, $\Delta\langle p_\perp^2\rangle$ is going to zero at high z indicating that $\Delta\langle p_\perp^2\rangle$ is not, as expected, generated by prehadron interaction. Indeed high z events in absorption models are produced rapidly and would be more subject to broadening, if it was generated during the prehadronic phase. It is also possible that initial state effects lead to an increased $p_\perp^2$ broadening at all zs, but the resolution of present results might not permit to access such an effect. To conclude, all these measurements confirm the assumption that $\Delta\langle p_\perp^2\rangle$ reflect mainly the partonic stage of hadronization, but the extrapolation to the partonic level, necessary for quantitative statement is still theoretically challenging.

3.3.3 HERMES Two Hadrons Multiplicity Ratio

HERMES also provided data on pion pair production [Airapetian 2006], their results are shown in figure 3.20. Here, z_2 is defined as the z of the second higher energy pion and it is requested that the leading hadron has $z > 0.5$. In prehadron absorption models, each hadron should behave independently leading to a scaling of R_{2h} similar to what is observed for the usual multiplicity ratio. We note that the important increase at low z_2 generated by the target fragmentation is observed. However a clear A dependence is expected but not observed and the results are similar for all targets. This is difficult to conciliate with the absorption models. In parton energy loss models, the energy lost by the leading hadron should impact all hadrons of a given event leading also to a clear difference between targets. However, the energy lost by the parton

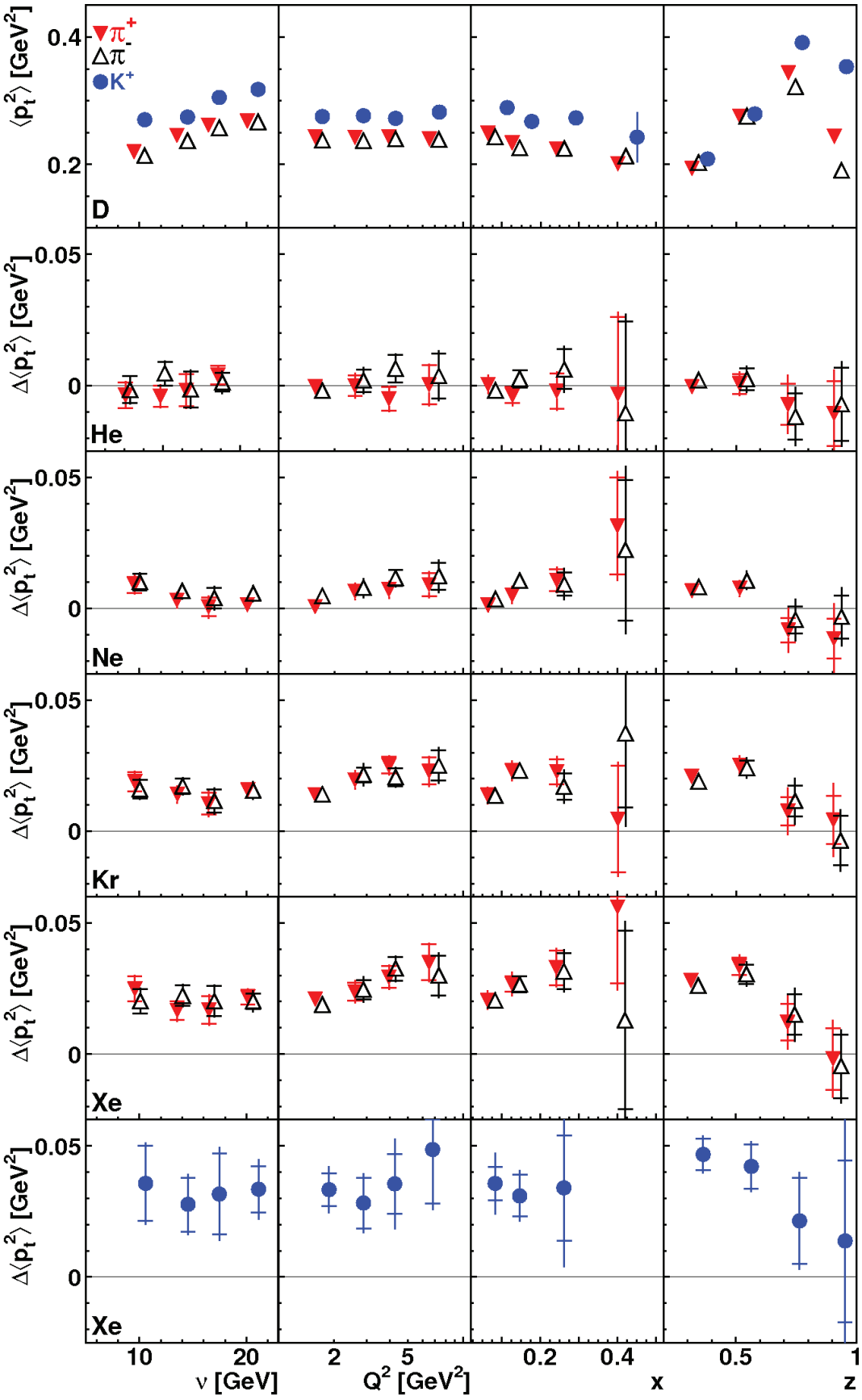

Figure 3.19: Transverse momentum broadening $\Delta\langle p_\perp^2 \rangle$ of charged pions as a function of ν, Q^2, x_{Bj} and z from HERMES [Airapetian 2010].

might remain, at least partly, in the current fragmentation, but most of the existing calculations do not keep track of this energy. Therefore, this result, where almost no A dependence is observed, needs elaborated models taking into account the remnants left by the leading hadron. These could then contribute to the second leading hadron production and compensate partly for the leading hadron attenuation. More precise data, and eventually different observables, are needed in order to interpret properly this result from HERMES. It is also unfortunate that only very few models are compared to these surprising results.

3.3.4 JLab Hall C Results

Finally, to complete this review, we have to mention the results from the Hall C at Jefferson laboratory [Asaturyan 2011] which uses a beam of 5.5 GeV electron. They obtained results relevant to our discussion using their empty target data, but, as the nuclear effects are not their primary focus, they have only results for deuterium and aluminum ($A = 27$). Their results (figure 3.21) are completely coherent with HERMES'. Nonetheless, we note the flat distribution of R^{π}_{Al} as a function of Q^2, but it is important to note that $p_{\perp}^2$ acceptance is limited to $p_{\perp}^2 < 0.2$ GeV^2/c^2. They also measure the multiplicity ratio as a function of $p_{\perp}^2$ and find a flat distribution (again $p_{\perp}^2 < 0.2$ GeV^2/c^2). These observations are coherent with the measurement from HERMES, respectively in figures 3.15 and 3.13.

3.3.5 Conclusion

The recent results, mainly from HERMES, greatly advanced our understanding of the processes involved in in-medium hadronization. However, no clear indication is provided on which effects dominates the hadronization in nuclei. Nevertheless, the new data are making more and more difficult for pure models to describe the whole picture. These data, using various observables, are a motivation for more complete models, which take into account both target and current fragmentation, with a global treatment of all the produced hadrons. To date, only the GiBUU Monte-Carlo simulation was designed in this spirit, a new one, based on parton energy loss is presented in chapter 4.

On the other hand, future experiments need to focus on clear aspects of the hadronization either with extreme kinematics or new observables. The goal being to provide directly interpretable results, leading to model independent interpretation. This is the objective for the CLAS experiment [Brooks 2002], presented in chapters 5 to 7 of this thesis (other preliminary results can be found in [Brooks 2009, Hicks 2009]). It is also possible by making experiments

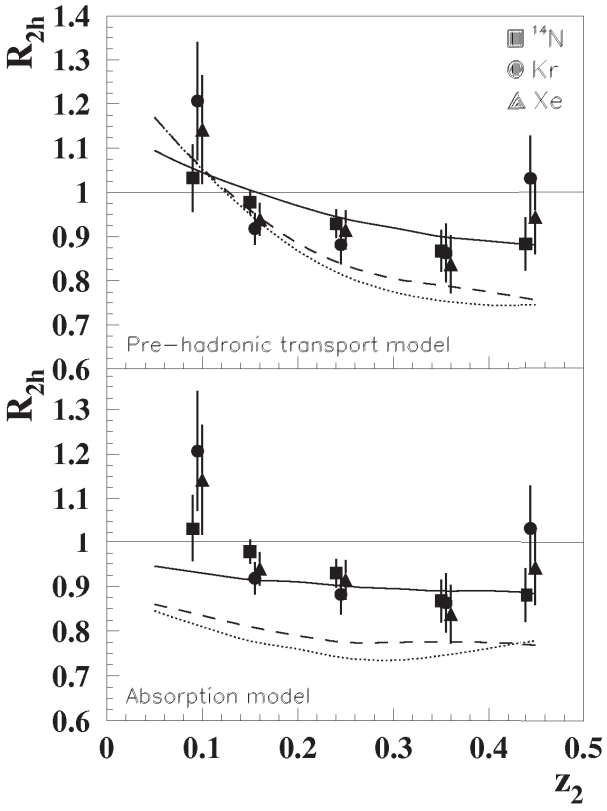

Figure 3.20: Two pions multiplicity ratios as a function of z_2 from HERMES [Airapetian 2006].

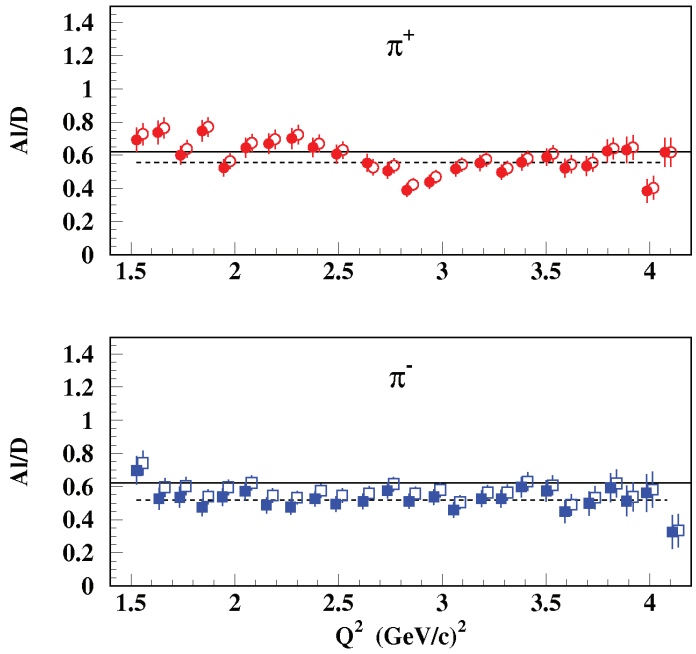

Figure 3.21: Multiplicity ratios of charged pions as a function of Q^2 from JLab Hall C [Asaturyan 2011].

at much higher energy, in the pQCD regime. Then, the parton energy loss is isolated from any possible prehadron absorption effects; this option is explored in chapter 8.

Chapter 4

PyQM Monte-Carlo Generator

4.1 Presentation

As pointed out in previous chapters, many existing models can claim some success in describing the data. Pure parton energy loss models have the specificity to be easily transposable between nuclear DIS and heavy ion collisions, but most of the existing calculations were applied only to hot QCD matter. The models by [Arleo 2002, Wang 2002, Arleo 2003b], for example, confront their calculations to both cold and hot nuclear matter. This chapter describes the adaptation of another calculation, from Salgado and Wiedemann [Salgado 2002, Salgado 2003], to cold nuclear matter. We implement their quenching weights into a Monte-Carlo event generator simulating both current and target fragmentations, PYTHIA. The quenching weights are an intermediate result of the parton energy loss calculation, which allows the rapid calculation of the energy loss for any given initial conditions.

The quenching weights can be used in a calculation using only the parton distribution functions and the fragmentation functions [Accardi 2007]. However, describing the most recent data can be difficult without a complete representation of the fragmentation process. Indeed, the fragmentation might have some effects on the transverse momentum, that would be difficult to explore using fragmentation functions, one of the reasons being that they include only the longitudinal description of hadron production. A complete Monte-Carlo simulation, including momentum conservation in a realistic fragmentation model, is a chance to get an insight on this problem. It might also be a chance to see effects not directly related to hadronization in nuclei, but constrained by the kinematic of the initial state. Finally, the target fragmentation region can be better understood in this framework, where it can be properly described.

The PYTHIA Monte-Carlo generator was chosen as basis for this work, because it is simple to use and it already includes many processes. PYTHIA (version 6.4 [Sjostrand 2006]) generates the lepton scattering, fragments the colored objects and allows unstable particles to decay. This chapter is about the implementation of nuclear effects in PYTHIA, especially the parton energy loss, which occurs after the DIS hard scattering and before the fragmentation.

The resulting generator is called PyQM for PYTHIA Quenching weight Model.

4.2 Technical Description of PyQM

4.2.1 The Hard Scattering

At moderate energies, for lepton-hadron interactions, the PYTHIA generator simulates two main processes: the deeply inelastic scattering and the scattering associated with vector meson dominance (VMD). In DIS the virtual photon directly interacts with a quark, see section 1.1; in VMD, the photon fluctuates into an hadronic state, that subsequently interacts with the target. The states easily accessible to the photon are the ones with similar quantum numbers, those are the vector mesons (ρ, ϕ, ω...). Then the virtual hadron can interact with the nucleon through different processes detailed in the PYTHIA manual [Sjostrand 2006]. However, this channel is problematic in PYTHIA because it leads to a too important diffractive ρ production (this was also noticed in the HERMES collaboration [Liebing 2004]). HERMES evaluated the correct impact of VMD on their multiplicity ratios to be only of few percents [Airapetian 2007]. Similarly, in JLab Hall C results [Asaturyan 2011], the associated correction is modest. However, this statement is not true at some extreme kinematics (high z, low x_{Bj} or low Q^2), but because these are not our regions of interest, this should not cause important contamination in our studies. Since VMD production and its potential effect on hadronization observables is not our focus and is expected to be small, we deactivate it in PYTHIA.

Many parameters can be used to adjust the simulation to the data, but the number of parameter and their intricate impact on observables make any adjustment difficult. Indeed, the two main processes, fragmentation and hard scattering are independent, this leads to ambiguities in any attempt to fit the simulation to data. Nevertheless, in order to describe data, many changes were tested, either using sets developed by HERMES or CLAS collaborations. However, some of these modifications are difficult to justify individually and provide only a limited improvement. Moreover, the hadronization observables are chosen to cancel these effects, especially the ones related to the initial hard scattering. For these reasons, no change on the hard scattering part is applied, only the kinematical limits of PYTHIA are lowered in order to allow the software to work in configurations down to CLAS energy (5 GeV electron beam).

An example of event is shown in table 4.1. PYTHIA is here stopped after the hard scattering and provides, in the final configuration, one scattered electron, a struck quark and a diquark remnant of the initial proton. The target can be changed to a neutron, but it always remains a free nucleon. In order to

I	Particle	Parent	p_x	p_y	p_z	E	Mass
1	(e^-)	0	0.000	0.000	27.500	27.500	0.001
2	(p)	0	0.000	0.000	0.000	0.938	0.938
3	(e^-)	1	0.565	-0.093	5.542	5.572	0.001
4	(γ)	1	-0.565	0.093	21.958	21.928	-1.275
5	(p)	2	-0.000	0.000	0.000	0.938	0.938
...	...	.	...	...	...	...	...
11	$e-$	3	0.565	-0.093	5.542	5.572	0.001
12	u $\uparrow$	5	-1.880	-2.031	17.553	17.770	0.000
13	ud_0 $\downarrow$	5	1.315	2.124	4.404	5.096	0.579
		sum:	0.000	-0.000	27.500	28.438	7.245

Table 4.1: Example of a DIS hard collision generated by PYTHIA 6.4. Energy, momentum and mass are given in GeV, GeV/c and GeV/c^2 respectively. Arrows indicate the beginning and end of color strings and particles in parenthesis are initial or intermediate particles not present at the end of the process.

simulate nuclei, we add several effects. First, we generate events on protons and neutrons according to the composition of the simulated nuclei. Second, we give to the nucleon some initial momentum to reproduce the internal Fermi motion of the nucleons. The Fermi momentum is randomly attributed, isotropically in space and according to [Ciofi degli Atti 1996] in magnitude (see figure 4.1). Relatively to deuterium, heavier nuclei look similar so, for simplicity, discussions about Fermi motion will be only about iron compared to deuterium.

4.2.2 Quenching Weights Calculation

The quenching weights are the results of calculations from Salgado and Wiedemann [Salgado 2002, Salgado 2003], which are accessible in the form of a FORTRAN routine provided by the authors. Their work is based on BDMPS calculations [Baier 1997] of the gluon emission spectrum in the multiple soft interaction approximation. They also provide results based on single hard scattering approximation using a different gluon spectrum, however this case will not be treated here.

The probability for a parton to lose an energy ΔE is given by

$$P(\Delta E) = \sum_{n=0}^{\infty} \frac{1}{n!} \left[\prod_{i=1}^{n} \int d\omega_i \frac{dI(\omega_i)}{d\omega} \right] \times \left(\Delta E - \sum_{i=1}^{n} \omega_i \right) \exp \left[- \int d\omega \frac{dI}{d\omega} \right], \quad (4.1)$$

with ω_i the emitted gluon energies and $dI/d\omega$ the gluon energy spectrum. This radiation spectrum usually depends on the length of the medium L and the transport coefficient $\hat{q} = k_\perp^2/L$, with $k_\perp^2$ the transverse momentum of the

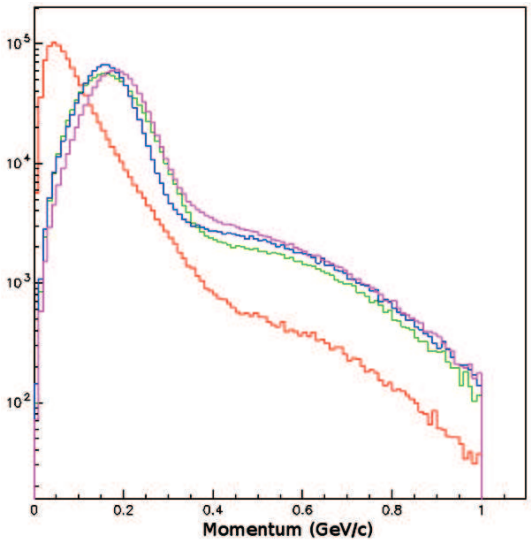

Figure 4.1: Fermi momentum distribution (normalization is arbitrary) for deuterium (red), carbon (green), iron (blue) and lead (purple).

quark induced by the medium. These two variables can be replaced by the characteristic gluon frequency $\omega_c = \frac{1}{2}\hat{q}L^2$ and a dimensionless density parameter $R = \omega_c L$. These variables are defined for both cold and hot nuclear matter, where, in the latter case, they are calculated by integrating over time as L and $\hat{q}$ vary.

BDMPS results are for $R \to \infty$; Salgado and Wiedemann extended the calculation of $\omega dI/d\omega$ for moderate R leading to a suppression of small energy gluon emission (figure 4.2 (left)). This distribution permits to evaluate easily the number of gluons emitted with energies higher than a given ω,

$$N(\omega) \equiv \int_{\omega}^{\infty} d\omega' \, \frac{dI(\omega')}{d\omega'}. \tag{4.2}$$

Their result for this variable is shown in figure 4.2 (right). The gluon radiation spectrum is also provided for the single hard scattering interaction approximation and compared to the work from [Gyulassy 2000], which gives qualitatively similar results but with slightly harder gluons.

The quenching weights calculated using equation 4.1 are provided in the form

$$P(\Delta E) = p_0 \, \delta(\Delta E) + p(\Delta E), \tag{4.3}$$

where p_0 represents the discrete probability for the parton to go through the medium unaffected and $p(\Delta E)$ the probability for the parton to lose an energy ΔE. Some of their results for $p(\Delta E)$ are shown in figure 4.3. We notice that, sometimes, the probability can go negative[1], this behavior reflects the

[1] As a probability $\int P(\Delta E) = 1$, the negative probability being compensated elsewhere.

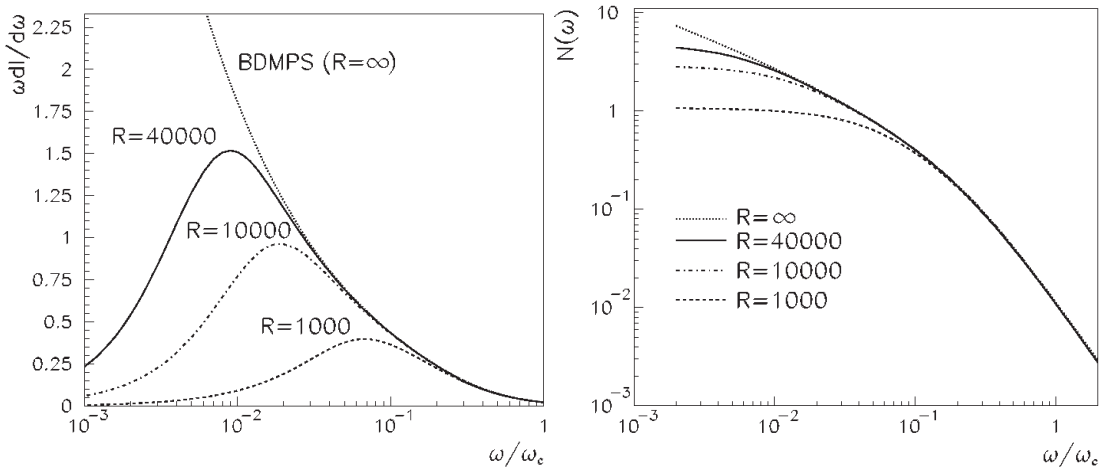

Figure 4.2: Gluon spectrum (left) and number of gluon emitted (right) as a function of ω/ω_c the fractional energy of emitted gluons calculated in the multiple soft interaction approximation [Salgado 2003].

possibility that a parton emits less gluons in medium than in vacuum (see [Salgado 2003] for more details).

In the Monte-Carlo implementation of the quenching weights, the angular dependence of the gluon emission is also of some importance. However, the full calculation was not done for this result and an exact calculation would be laborious. Using properties of the existing calculation, it is possible to obtain an approximated value (details in the next section) of the energy loss as a function of the gluon emission angle.

The main objective of the original calculation [Salgado 2002] was to test RHIC and eventually LHC data. However, in heavy ion collision, there is no experimental control of the initial hard scattering kinematic and only the highest $p_\perp$ jets are relevant. The contamination from the remnants of the hadronization process[2] might also influence the result as they cannot be separated from the main signal. Therefore, nuclear DIS is better suited to test this calculation.

4.2.3 Quenching Weights Implementation

To calculate the input variables needed for the quenching weights, we first determine the position of the interaction point within the nucleus. This is done by randomly selecting a position according to the density profile of the nucleus. One could use a simple hard sphere model, but previous result [Accardi 2007] indicates a better fit of the data using realistic density profile. We use the nuclear density profiles provided in [De Vries 1987]. Then, we calculate the in-medium path length, L_A, weighted by the nuclear density. For a parton

[2] This part is somewhat equivalent to the target fragmentation in e-A.

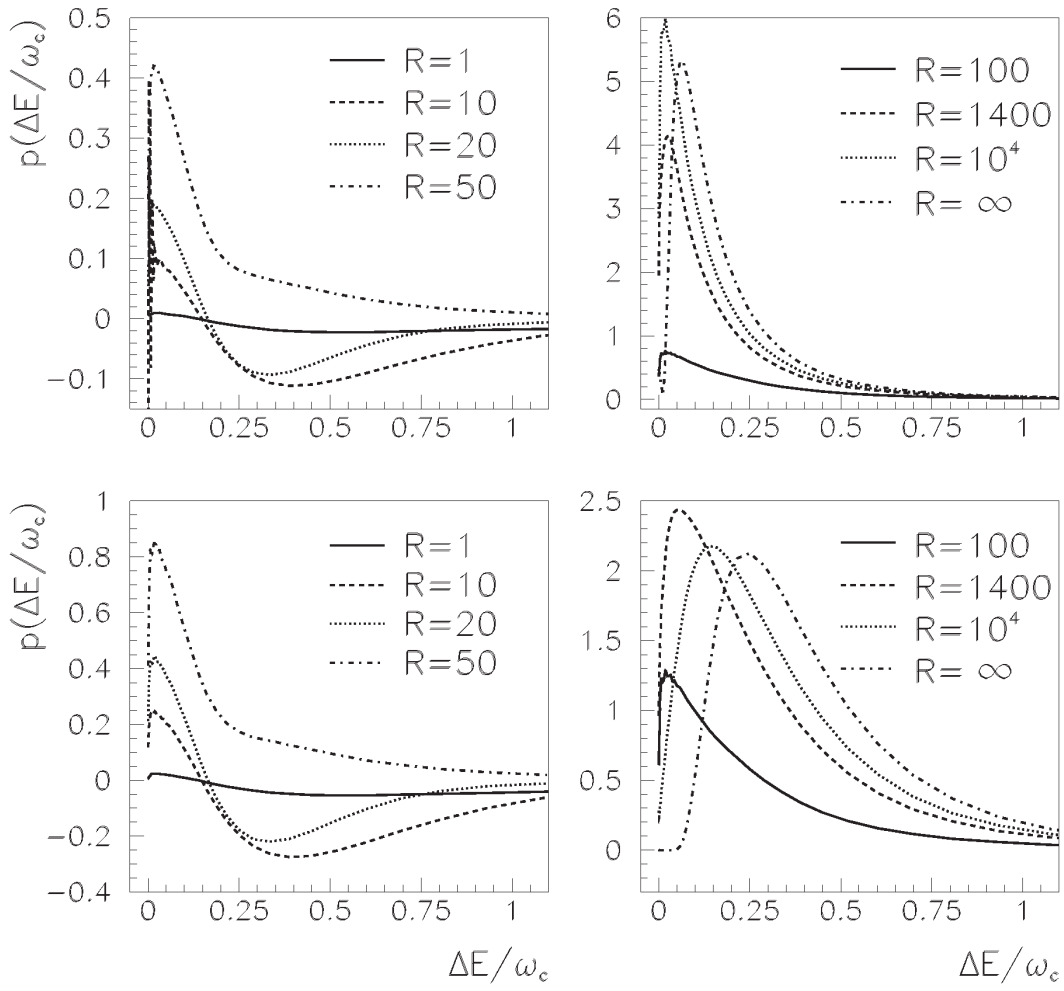

Figure 4.3: Continuous probability for a quark (upper plots) or a gluon (lower plots) to experience an energy loss ΔE in nuclear medium using the multiple soft interaction approximation [Salgado 2003].

struck at a position $(\vec{b}, y)$, propagating along y with an impact parameter $\vec{b}$,

$$L_A(\vec{b}, y) = 2\ \frac{\int_y^\infty dz(z-y)\rho_A(\vec{b}, z)}{\int_y^\infty dz\rho_A(\vec{b}, z)}, \tag{4.4}$$

assuming that the prehadron production always occurs outside of the nucleus. The average nuclear density on this path is

$$\bar{\rho}_A(\vec{b}, y) = \frac{\int_y^\infty dz\rho_A(\vec{b}, z)}{L_A(\vec{b}, y)}, \tag{4.5}$$

from which we can make an effective averaged transport coefficient

$$\bar{\hat{q}}_A(\vec{b}, y) = \frac{\hat{q}_0}{\rho_0}\ \bar{\rho}_A(\vec{b}, y), \tag{4.6}$$

where ρ_0 is the nuclear density at the center of the nucleus and $\hat{q}_0$ the transport coefficient at the center of the nucleus. The quenching weights input variables, which are the average characteristic gluon frequency and the average density parameter, can be calculated as

$$\bar{\omega}_c(\vec{b}, y) \equiv \frac{1}{2}\ \bar{\hat{q}}_A(\vec{b}, y)\ L_A^2(\vec{b}, y) = \int_y^\infty dz\ (z-y)\ \hat{q}_A(\vec{b}, z) \tag{4.7}$$

and

$$\bar{R}(\vec{b}, y) \equiv \bar{\omega}_c(\vec{b}, y)\ L_A(\vec{b}, y) = \frac{2\ \omega_c^2(\vec{b}, y)}{\int_y^\infty dz\hat{q}_A(\vec{b}, z)}, \tag{4.8}$$

where

$$\hat{q}_A(\vec{b}, y) = \frac{\hat{q}_0}{\rho_0}\rho_A(\vec{b}, y) \tag{4.9}$$

is the local transport coefficient. Finally, the two input variables of the calculation are based on only one free parameter, $\hat{q}_0$ the transport coefficient at the center of the nucleus, which is directly related to the transverse momentum induced by the crossing of the nuclear matter.

For each event the struck parton is used to generate a quenching weight distribution $P(\Delta E)$ used to get the quenching magnitude. At this stage we neglect any negative probability, this is probably not ideal but is difficult to avoid at this stage.

To determine the emission angle of the gluon, we define $R_\chi = R\sin^2\theta$, then

$$\langle\Delta E\rangle_\theta = \int dE'E'\left[P(E', \omega_c, R) - P(E', \omega_c, R_\chi)\right] \tag{4.10}$$

gives the average energy emitted by gluons outside of an angle θ [Salgado 2003]. Then, as we already determined the total energy loss, we can generate an angular distribution of averaged energy loss as a function of the emission angle. We use the averaged energy distribution over all angles to select randomly the gluon emission angle. This is an approximation, but, as we consider only single gluon emission, this can be justified, for moderate R, by the low number of emitted gluon (figure 4.2 (right)). Then the gluon kinematic is completed by an isotropic selection of a ϕ angle around the initial parton. The outgoing parton is reconstructed by subtracting the gluon momentum calculated to the initially generated parton. Sometimes, the propagating parton can be stopped in the medium, however this might block the fragmentation, which need a minimal energy to provide colorless objects. A lower limit of 400 MeV is applied on the parton energy in order to avoid this problem.

Finally, it is unclear how the gluons generated by the quenching should be treated. They can be directly added in the fragmentation chain, between the quarks, but they might also be absorbed in the nuclear medium, in which case their energy will dissipate in the nucleus in different ways (see [Ciofi degli Atti 2005] for example). We choose here to withdraw the emitted gluons and neglect this part of the process. This approximation might lead to a stronger attenuation, especially in the target region. However, there is no simple way to simulate this process, which need to be modeled.

4.2.4 The Fragmentation

Fragmentation is performed in PYTHIA using the Lund model [Andersson 1983, Sjostrand 1984], which is a string model. The idea of string models is based on color flux tubes – represented by strings – which generate the force between colored object when they are stretched. This force is linear and driven by the string tension $\kappa \sim 1$ GeV/fm. The Lund fragmentation is usually represented in the form of a diagram, like in figure 2.2. At the origin (bottom of the diagram) the two colored objects ($3\bar{3}$) move away from each other, the string accumulates energy which is released when it breaks producing a quark and an anti-quark or a diquark pair. These propagate away from each other, similarly to the initial partons. Eventually some quarks meet each other and form hadronic bound states. Gluons can also be introduced in the system, but they cannot serve as end points, so they are always connected to two strings. Therefore, within this string model, gluons are subject to twice the force applied to quarks, which is a good approximation of the charge ratio $2/(1-1/N_c^2) = 9/4$ [Sjostrand 1984] predicted by QCD theory.

The probability distribution for the string breaking has to respect many constraints. For the Lund model, the following form is used to determine the

position of the breaking in the string and the mass of the produced hadron[3]

$$f(z, m_\perp) = \frac{1}{z}\,(1-z)^a\ \exp(-bm_\perp^2/z) \tag{4.11}$$

with z the fraction of the string at which the breaking occurs and $m_\perp^2 = m^2 + p_\perp^2$ the transverse mass of the produced hadron calculated based on its mass and transverse momentum, a and b are parameters. The transverse mass factor leads to quark production proportion as follow $1:1:1/3:10^{-11}$ for $u:d:s:c$ respectively and, therefore, leads naturally to a suppression of heavy mesons and baryons compared to light mesons.

The time needed for the string breaking to occur is called τ and is usually characterized by $\Gamma = (\kappa\tau)^2$. Its probability distribution is given by

$$P(\Gamma) = \Gamma^a\ \exp(-b\Gamma) \tag{4.12}$$

using the same parameters a and b than previously. In PYTHIA, the default values for these parameters are $a = 0.3$ and $b = 0.58$ GeV^{-2}. Many other parameters exist to adjust the Monte-Carlo simulation to data. One parameter determines at which energy to stop the string breaking process, by default in PYTHIA, the threshold is set at 800 MeV. The remaining strings are then fragmented using a cluster fragmentation which is very simple at low energy. This parameter is the only one we decided to change. We lowered it to 300 MeV in order to increase the multiplicity. This modification is necessary because low multiplicity events lead to an overestimation of pion production at high z. This kinematic problem is due to the low energy of the experiments reproduced in this chapter, since PYTHIA was initially developed for much higher energies.

4.3 Results

4.3.1 Fermi-Momentum

Deuterium has much less Fermi momentum than nuclei (figure 4.1) and this might affect our observables and contaminate the signal from hadronization. To explore this problem, we simulate data from deuterium and iron implementing only Fermi motion effects in our Monte-Carlo. The main motivation for this study is to evaluate the effect of Fermi motion on the $\Delta\langle p_\perp^2\rangle$ observable. Indeed, the target nucleon have initial transverse momentum, which induce a broadening of the final state transverse momentum that we would like to quantify. Also, most of the kinematic variables are modified by the Fermi momentum, because, in their determination, we assume the nucleon to be at

[3] In the Lund model hadrons are produced at the same time than string breaking.

	CLAS	HERMES
ν	13%	14%
z	11%	8%
$p_\perp^2$	19%	14%

Table 4.2: Relative smearing of various kinematic variables due to the Fermi momentum of the struck nucleon in iron.

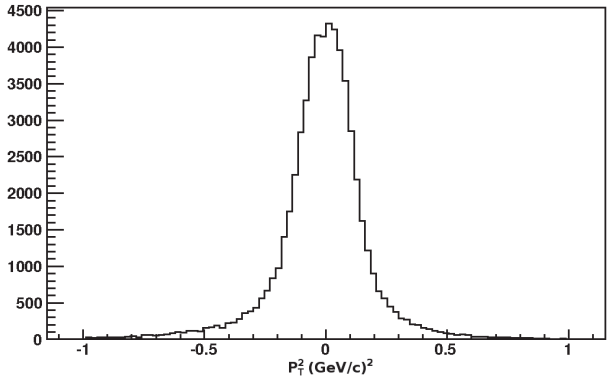

Figure 4.4: Smearing of the $p_\perp^2$ variable due to Fermi motion for pions produced on an iron target at CLAS energy.

rest. The smearing of the variables gives an indication on the bin size that can be resolved experimentally. However, as the Fermi motion magnitude is only of a few hundreds MeVs, it should affect mostly the lower energy experiments. Here, only CLAS and HERMES configurations, with 5 and 27.5 GeV beam respectively, will be studied.

To estimate the smearing of the kinematic variables due to Fermi motion, we compare the experimental variables, computed assuming the nucleon at rest, to the correct ones, knowing the nucleon kinematics. In the table 4.2 are the results for CLAS and HERMES kinematic and in figure 4.4 is a typical distribution of the smearing associated to the Fermi motion of the $p_\perp^2$ variable. These smearings are relatively important but must be put into perspective. First, ν is, in the hadronization studies, used to estimate the initial energy of the struck quark. The nuclear effects varying with ν depend in fact on the energy of the quark in the nucleus rest frame and, therefore, the measured ν is actually the relevant variable. The same is true for most of the variables and the error on these has only an impact on the expected cancellation of initial state effects in the multiplicity ratio. However, the induced broadening of $p_\perp^2$ might be more important, because of the rapidly falling distribution Fermi motion might populate significantly the higher $p_\perp^2$.

To evaluate the effects of the Fermi momentum on the observables, two sets of events are produced using Fermi motion of deuterium and iron. As

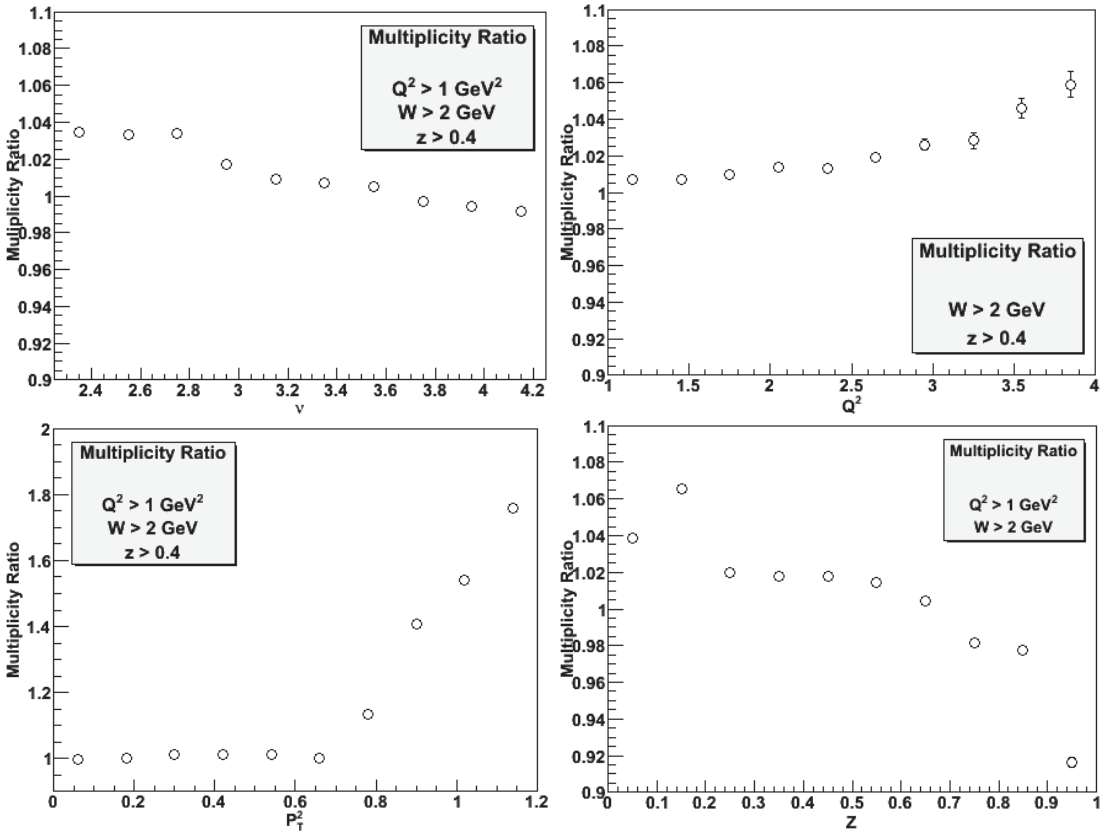

Figure 4.5: Multiplicity ratios of pions in iron, simulating only the Fermi motion, at CLAS energy.

the goal is to isolate the Fermi momentum effect, no other nuclear effects are included at this stage. Results for multiplicity ratios at CLAS energy are presented in figure 4.5. The effect is relatively small in most of the cases, except for $p_\perp^2$. Indeed, we observe a slope of ∼5% overall in ν which should attenuate the expected increase of R_{Fe}^{π}. The effect, as a function of z, is similar, but, this time, it will enhance the expected effect from hadronization. These are moderate effects but have to be taken into account for any quantitative study of the CLAS results presented in chapter 7. More important is the result as a function of $p_\perp^2$, because the Cronin effect is mimicked by a strong Fermi motion effect. The Cronin effect is generally believed to be a consequence of multiple scattering, but here the interpretation of CLAS multiplicity ratio, as a function of $p_\perp^2$, will be influenced by the Fermi motion. Another important impact is observed as a function of Q^2, the slope is relatively small (∼ 4% overall), but observed effects are also small! Therefore, to be definitive, an effect as a function of Q^2 will need to be substantial in order to rule out possible contribution from Fermi motion.

The impact on $\Delta\langle p_\perp^2\rangle$ is also interesting to observe. Figure 4.6 shows the effect as a function of various variables. Overall, the simulation indicates that

0.007 GeV^2/c^2 of transverse momentum is generated by Fermi motion. This effect has interesting dependencies in z and Q^2, which should be taken into account before interpreting experimental results. However, it is unclear how effects sum up between parton energy loss and Fermi momentum effects. This could be studied with a full generator such as PyQM, but is not part of this work.

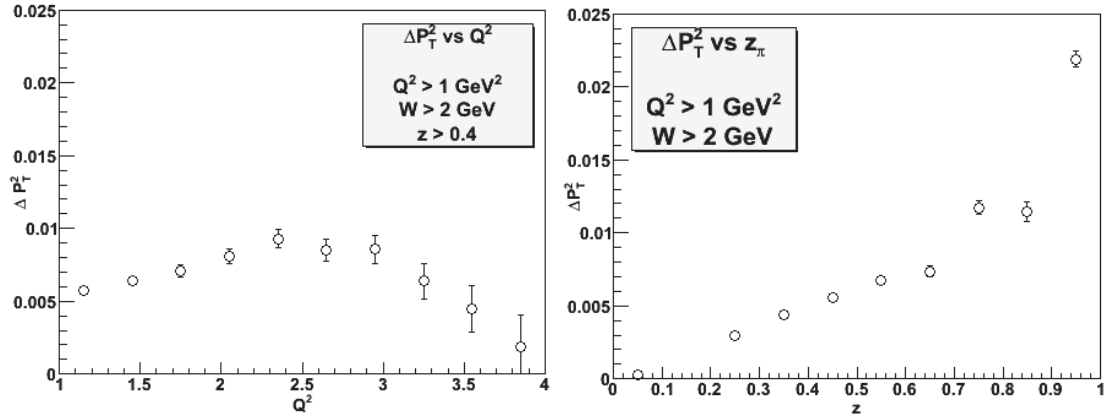

Figure 4.6: Transverse momentum broadening of pions in iron simulating only the Fermi motion effect at CLAS energy.

At HERMES, the Fermi motion impact is similar in trend, but much more limited in amplitude. As expected, with higher energies the Fermi momentum effects become relatively less important. In figure 4.7 and 4.8 are the results obtained for HERMES kinematics. We observe effects limited to a few percents for multiplicity ratios and we note the very different behavior as a function of $p_\perp^2$. The effect on $\Delta\langle p_\perp^2\rangle$ remains relatively important, in particular, we observe an increase of $\Delta\langle p_\perp^2\rangle$ as a function of Q^2. This feature might partly explain the increase observed in HERMES data, which, after correction, would come closer to no effect, but also closer to the pQCD expectation from [Domdey 2009]. However, the entanglement of the nuclear effects leading to $p_\perp^2$ broadening is complicated and it is not clear how to extract properly this contribution. The effect observed for $\Delta\langle p_\perp^2\rangle$ at high z was expected, but seems large compared to data. However, the fact that we are outside of the factorization region and also in the region of large diffractive ρ^0 contamination should lead to caution on interpreting this result in the high z region.

In conclusion, Fermi motion has important effects on CLAS data and especially on the Cronin effect, which is strongly enhanced. At both CLAS and HERMES energies, a small increase of $\Delta\langle p_\perp^2\rangle$ is to be expected overall. Globally all other variables are slightly affected and these effects should definitely be taken into account for any quantitative description of CLAS or HERMES data.

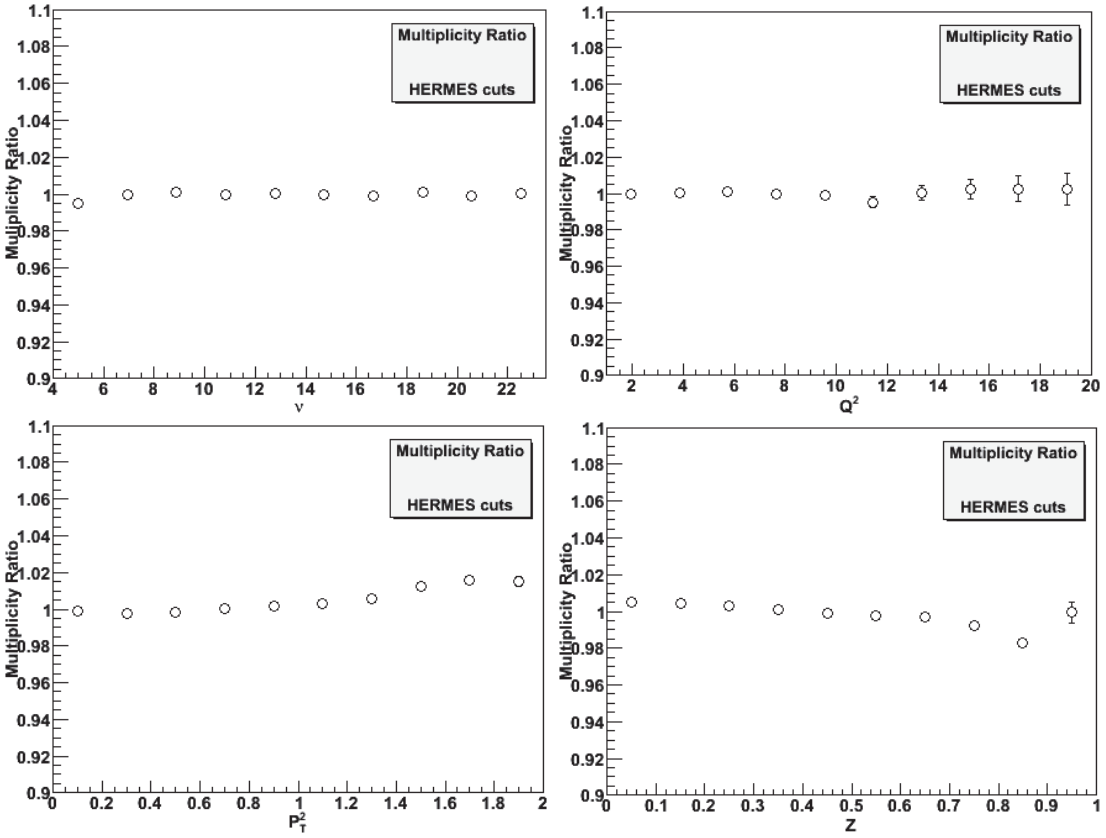

Figure 4.7: Multiplicity ratios of pions in iron, including only Fermi motion effect, at HERMES energy.

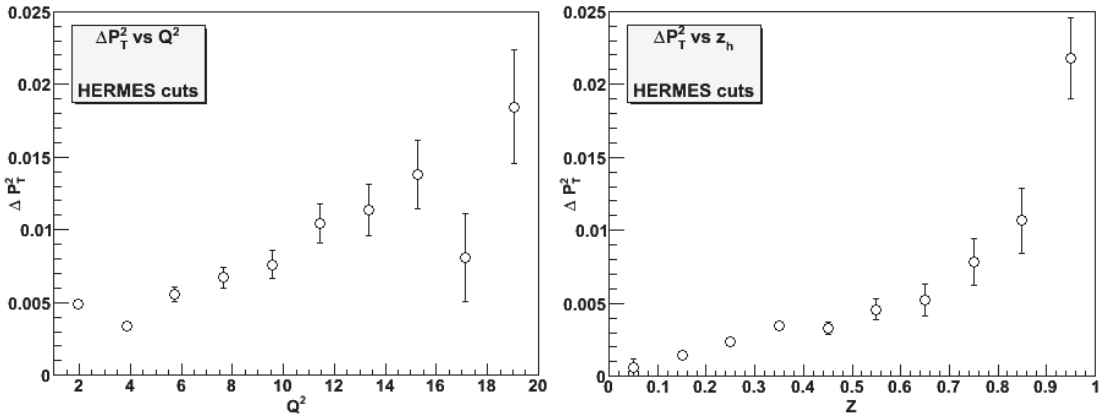

Figure 4.8: Transverse momentum broadening of pions in iron, including only Fermi motion effect, at HERMES energy.

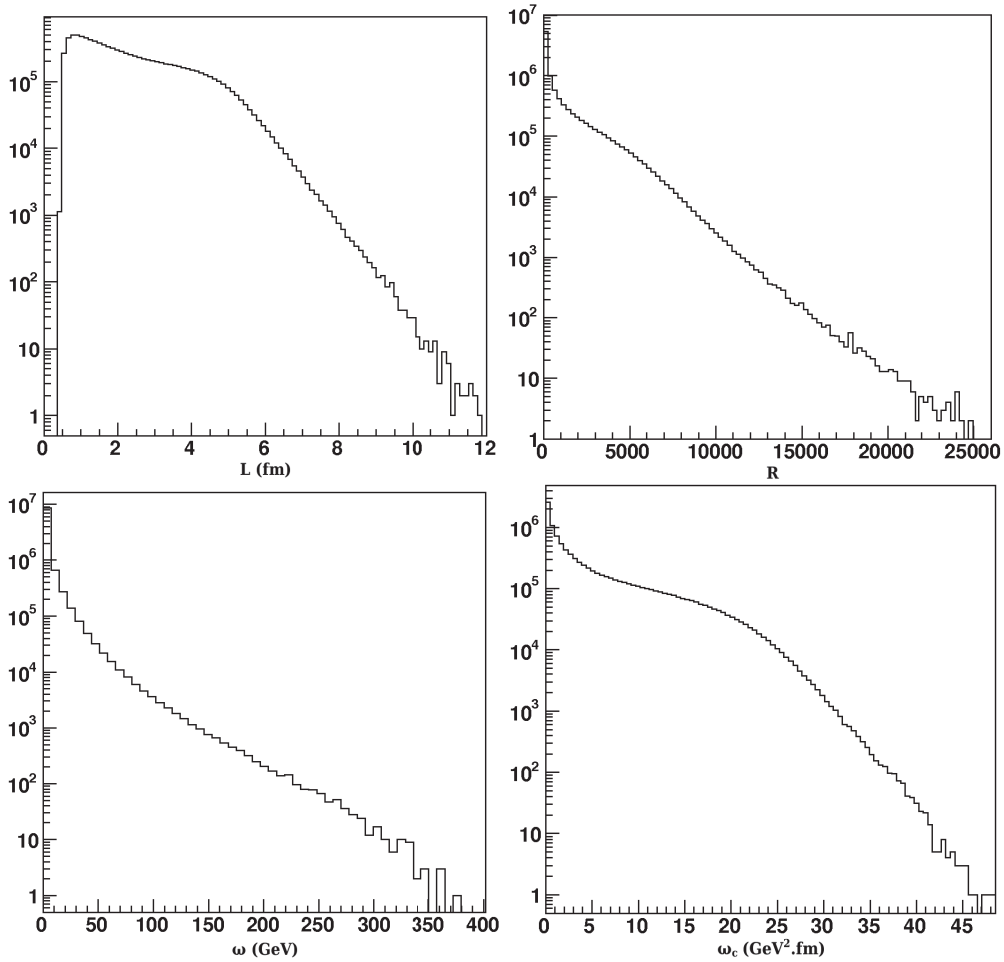

Figure 4.9: Characteristic variables of the quenching weight calculation at $\hat{q}_0 = 0.4$ GeV^{2}fm^{-1} using HERMES kinematic for xenon nuclei.

4.3.2 HERMES Description

In order to describe HERMES data with the PyQM code, $\hat{q}_0$ needs to be determined. We decided to adjust it to the multiplicity ratio of pions at $z = 0.5$. We choose this method instead of a fit because some regions are known to be poorly reproduced, low and high z typically, and they would affect the fit. The result obtained, which will be used in the rest of the section, is $\hat{q}_0 = 0.4$ GeV^{2}fm^{-1} with the quenching applied to both gluons and quarks. Figure 4.9 shows the distributions of various variables used or produced by the quenching weight calculation [Salgado 2003] for this value of $\hat{q}_0$ in xenon.

The result of the simulation, for π^+ as a function of z (figure 4.10 (left)), reproduces correctly the data, especially in the zone of best applicability $0.4 < z < 0.8$, where the fragmentation of a hard quark is dominating. However, because of the loose cut used in HERMES analysis ($z > 0.2$), the mismatch on the edges of the z distribution might affect the reproduction of other variables. The result as a function of ν (figure 4.10 (right)) has a slope sig-

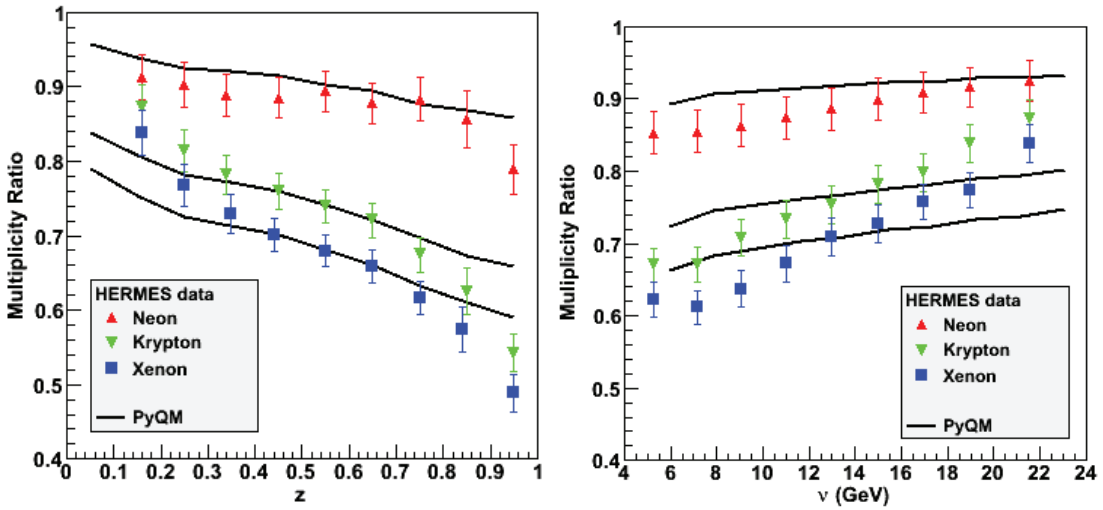

Figure 4.10: Multiplicity ratios from PyQM for π^+ as a function of z and ν (GeV) compared to HERMES data [Airapetian 2007].

nificantly off. This might be an indication that the calculation does not apply well to HERMES. However, it is important to mention that the multiplicity ratio trend in ν is enhanced by acceptance effects[4]. This was taken into account in our simulation but leads to more sensitivity to the absolute multiplicity, which can, eventually, influence results as a function of z.

The difference in slope for the ν dependence of the multiplicity ratio makes difficult to interpret other results linked to ν such as Q^2 or L_c (figure 4.11). However the global behavior can be interesting. There is no clear dependence of R_A^π with Q^2, which is natural, as Q^2 is not taken into account for the quenching weight calculation. As a function of L_c, R_A^π is slightly off, but seems to have some structure. The shallow inflexion in the HERMES data was initially interpreted [Airapetian 2007] as the indication that two processes were involved, i.e. parton energy loss and (pre)hadron absorption. The production of a similar behavior, in a pure energy loss model, is interesting because it is the demonstration that this feature can be generated in a pure model and, therefore, that the initial interpretation should be revised.

The results for $\Delta\langle p_\perp^2\rangle$ (figure 4.12) fail to reproduce the data. In particular, the description of the ν dependence is clearly off. The result for z follows the same trend as the data and the Q^2 dependence is flat. Despite of the fact that there is a global normalization mismatch, the reproduction of the observed trend in z shows that the attenuation effect on $\Delta\langle p_\perp^2\rangle$ is correctly handled and that the origin of the effect is consistent with a partonic process.

The question of the K^+ suppression, observed by HERMES

[4] HERMES detects only hadrons between 2 and 15 GeV/c, this lead to cuts in the ν-z plan (see [Accardi 2003] or [Airapetian 2007]).

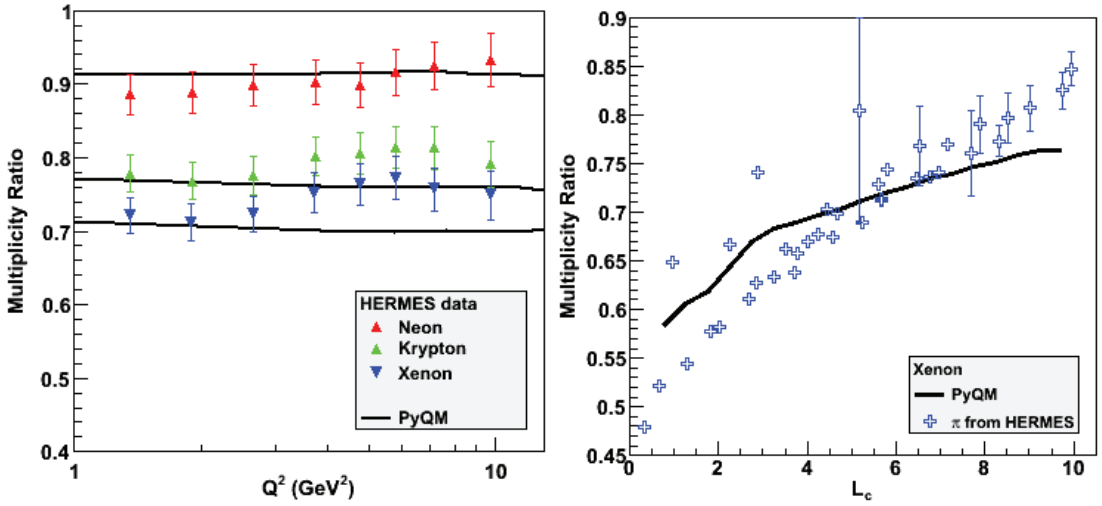

Figure 4.11: Multiplicity ratios from PyQM for π^+ as a function of Q^2 (GeV^2/c^2) and L_c (fm) compared to HERMES data [Airapetian 2007].

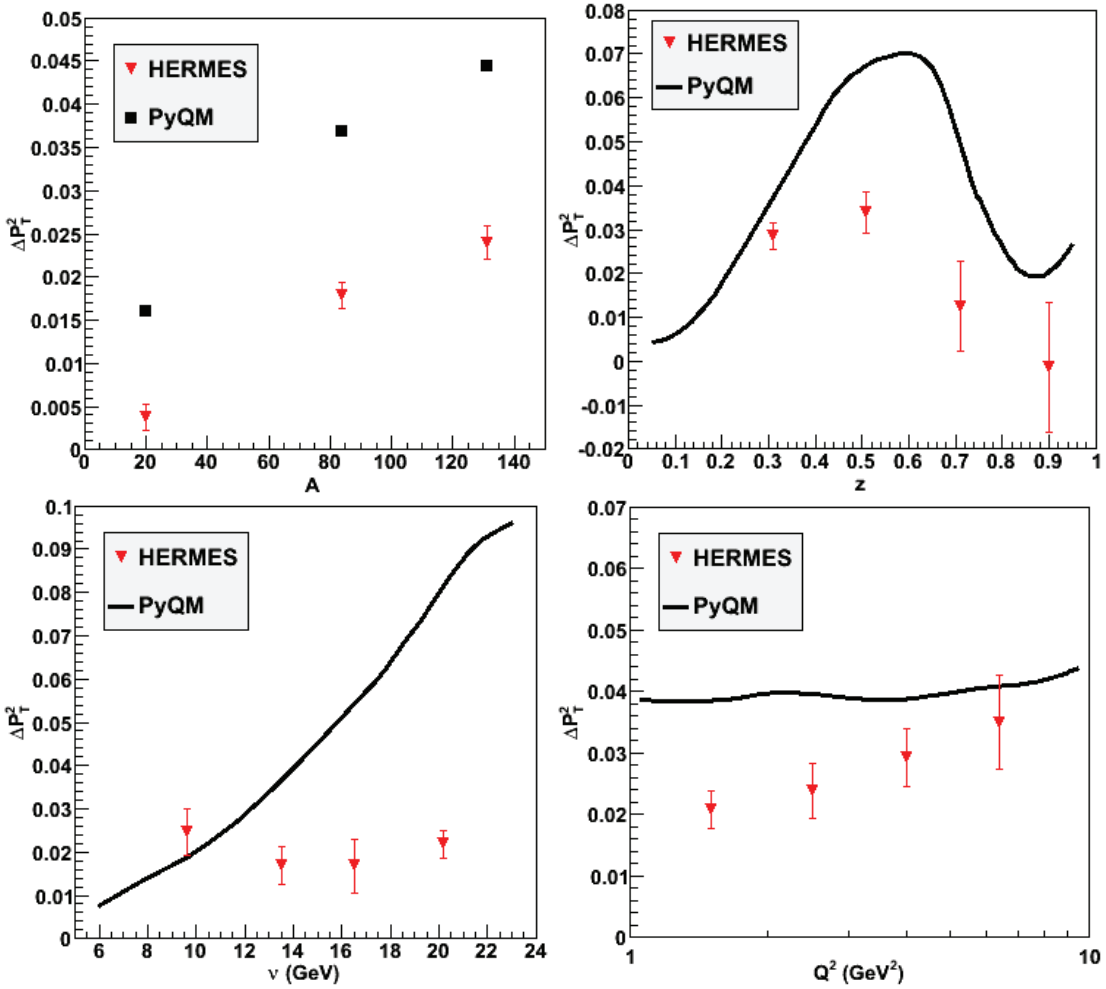

Figure 4.12: Transverse momentum broadening results of PyQM for pions as a function of A, z, ν (GeV) and Q^2 (GeV^2/c^2) compared to HERMES data [Airapetian 2007].

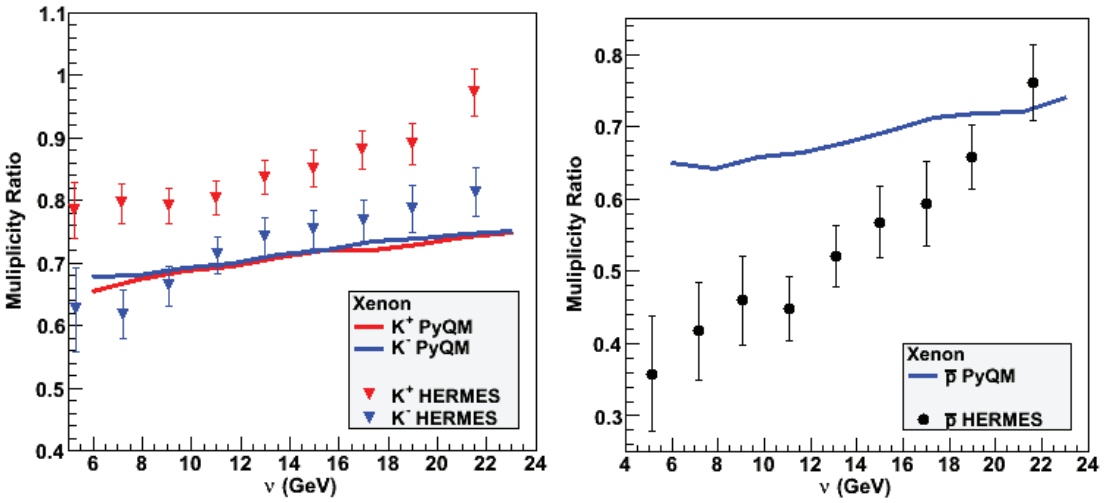

Figure 4.13: Multiplicity ratio results of PyQM as a function of ν (GeV) for K^+, K^- and $\bar{p}$ compared to HERMES data [Airapetian 2007].

[Airapetian 2007], can also be studied with PyQM; the relevant results are shown in figure 4.13 (left). Here both charged kaons give very similar results, this seems to eliminate the pure energy loss explanation based on fragmentation functions trend, raised in chapter 3. However, the target contamination and the strangeness enhancement, due to gluon emissions, are not simulated here and remain possibilities to explain this feature of the data in a pure energy loss scenario. Results for $\bar{p}$ give stronger attenuation than for pions, as expected in a pure energy loss model, but this is not enough to explain the data. Possibilities for the remaining discrepancy might be due to a problem with baryon production in PYTHIA or to hadron absorption. The latter, might be caused by a significantly shorter formation time or a significantly higher prehadron cross section than the one associated with pions.

4.4 Conclusions

The PyQM Monte-Carlo generator allows to simulate many basic features and study their effects on hadronization observables. It is found that the Fermi motion in nuclei leads to significant effects on hadronization observables at CLAS energy. The impact on HERMES results is smaller but is still to be taken into account for quantitative interpretation. As the effect is relatively reduced at higher energy, it should not be a concern beyond HERMES kinematics.

The full simulation results are not completely satisfactory, indicating that the Salgado and Wiedemann calculation do not describe properly the studied case. However, many interesting features are observed, for example, the

insufficient description of K^+ and $\bar{p}$ indicates some limit for pure energy loss models and the need for extra processes to be involved. On the other hand, $\Delta\langle p_\perp^2 \rangle$ as a function of A and z, in spite of a global overestimation, have a correct trend. This indicates that quark energy loss might be sufficient for a proper description of this observable. Also, the multiplicity ratio results have a correct trend as a function of z and L_c. The latter is especially interesting as the inflection observed by [Airapetian 2007] might have been wrongly interpreted.

Finally, the PyQM Monte-Carlo simulation was used in order to evaluate the possibilities for hadronization studies in a future electron-ion collider (presented in chapter 8). Obviously, predictions could also be made with PyQM, however as the shape found for ν is not matching the data, it is difficult to trust an extrapolation at higher energies. More work is needed before such predictions can be made. Improvements are also possible on the treatment of the emitted gluons and the simulation of the target fragmentation, but in this region, the lack of theory input is problematic.

Chapter 5

The Hall B of Jefferson Laboratory

This chapter is about the different apparatus used in the CLAS hadronization experiment. We briefly describe them, focusing on the essential information necessary to analyze properly the data. The experiment was run in the Hall B of the Thomas Jefferson National Accelerator Facility (TJNAF, also called Jefferson Lab or JLab)[1], which was founded in 1984 and started operations for physics in 1995. The main instrument of the laboratory is its electron accelerator and the experimental halls using its beam. The three different halls are complementary, the focus of the Hall B , where is CLAS (CEBAF Large Acceptance Spectrometer), is on multi-particle final state reactions and experiments which need a large acceptance or a photon beam. The Halls A and C are more focused on precision experiments, based on the very high luminosity available and their small acceptance, both statistical and systematic errors are drastically reduced.

The electron accelerator is described first, then the CLAS detectors will be reviewed. CLAS consists of an assembly of several detectors: drift chambers (DC), scintillator counters (SC), Cherenkov counters (CC) and electromagnetic calorimeters (EC). Because of the growing complexity of accelerators and detectors, collaborative work is necessary to successfully run major experiments. Numerous scientists and engineers of JLab are running the accelerator. CLAS management is organized in an eponymous collaboration composed of more than 200 physicists from around the world.

5.1 The Accelerator

A schematic of the JLab accelerator is shown in figure 5.1. The injector (the green box on the schematic) produces a low energy electron beam (45 MeV), which is accelerated in a first linear accelerator, the north linac. Then the beam is transported through a recirculation arc to the south linac to complete a first pass, the maximum number of passes is five. Each pass can accelerate

[1] Its former name was Continuous Electron Beam Accelerator Facility (CEBAF).

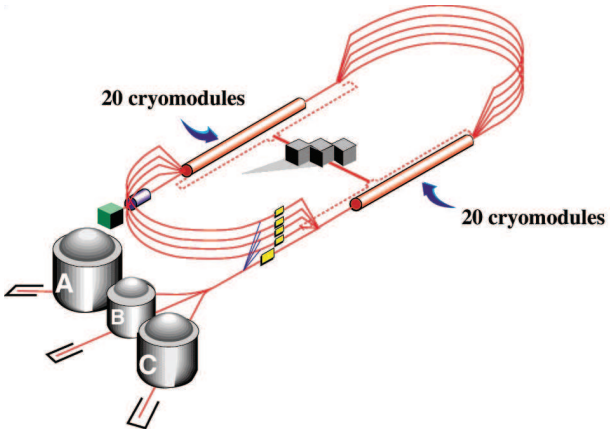

Figure 5.1: Schematic of the JLab accelerator.

the electrons up to 1.2 GeV, giving a maximum final energy of 6 GeV. At the end of the south linac, a distribution system is in place to send the beam in the three experimental halls.

In the injector, the electron beam is produced at radio frequency (RF), by lasers hitting a photo-cathode, in a -100 kV gun pushing the electrons in the beam line. There are three lasers, each associated with an experimental hall and pulsing at 499 MHz. Electrons for each hall are produced out of phase and kept that way in the accelerator, which is running at $3 \times 499 = 1497$ MHz. To improve this phase separation, a chopper cavity, which has an orthogonal field rotating at 499 MHz, spatially separates the three phases and a solid filter is used to select electrons with the correct phase. Then, a second chopper magnet recombines the beams. Finally, electrons are accelerated in a first series of superconductive cavities (in blue in the schematic) and leave the injector to the north linac with 45 MeV energy.

The electrons are accelerated by electromagnetic microwaves in superconductive cavities, which select a specific RF mode at 1497 MHz frequency. The electric field alternates at the same frequency as the electron injection. Therefore, all bunches see the same electric field and stay in phase (figure 5.2). The large amount of power needed for the accelerator to run continuously motivates the use of niobium, which is superconductive when cooled down to 2°K. This characteristic permits to reduce drastically the heat load and run the accelerator at ∼100% duty cycle. However, the two linacs require permanent cooling with liquid Helium and the 338 cavities of the accelerator are always kept immersed in a liquid helium bath.

At the end of each linac the beam is split by a magnetic field separating the electrons of various momentum. Each energy, corresponding to a number of pass, is kicked vertically and sent to a different recirculation arc. A second system (in yellow in the schematic), using the same principle, is installed at

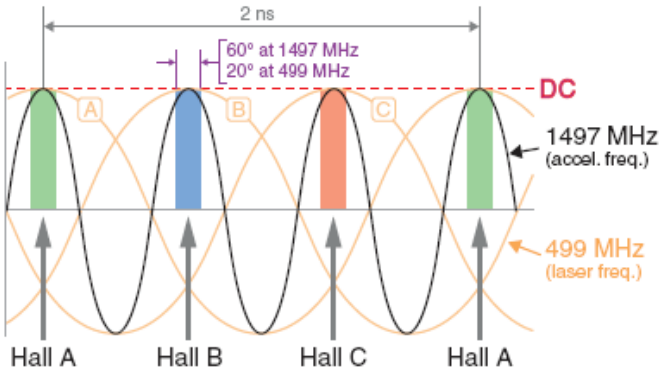

Figure 5.2: Schematic of the RF accelerating system.

the end of the south linac to send the beam in the experimental halls. Therefore, it is possible to send, at the same time, different beam energies to each of the three experimental halls.

5.2 The CEBAF Large Acceptance Spectrometer

5.2.1 Generalities

The CLAS [Mecking 2003] apparatus (figure 5.3) is a large assembly of detectors made of six similar sectors positioned around the beam line with the target at the center and torus magnets between them. Electrons, directly from the accelerator, or secondary photons, emitted in a radiator placed upstream of the detector, can be sent on the target. In the latter case, the electron beam is bent by a tagger magnet and sent to a beam dump before the detector. The beam going through the detector reaches a beam dump equipped with a Faraday cup, which allows the measurement of the total current received by the target.

The spectrometer (represented in figure 5.4) was built in order to study multi-particle final states. Those are only accessible using a large acceptance detector, because of the large available phase space. The detectors are organized in an onion like structure covering most of the solid angle. From the center to the periphery, we have three drift chamber regions [Mestayer 2000], a Cherenkov counter at forward angles [Adams 2001], scintillators for time of flight measurements [Smith 1999] and, at forward angle, an electromagnetic calorimeter [Amarian 2001].

The drift chambers are used for the tracking of the charged particles trajectory. In the magnetic field, produced by the superconductive torus magnets (in yellow on figure 5.4), the track curvature is directly linked to the momentum of the particle. Indeed, the radius R is given by $R = \frac{p}{q \times B}$ in which B is the

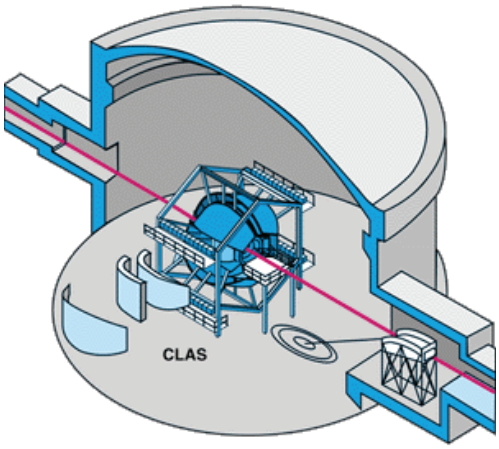

Figure 5.3: Schematic of the Hall B, with the beam entering from the right.

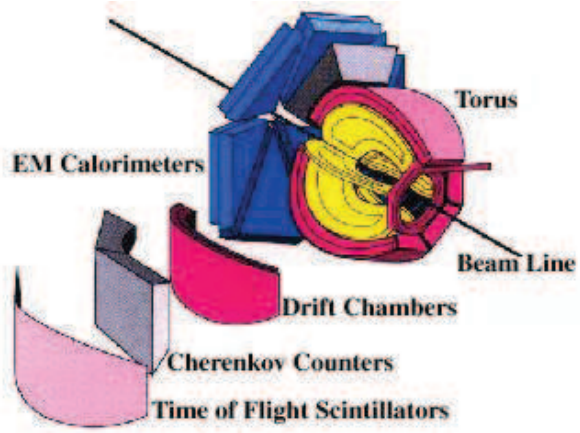

Figure 5.4: Schematic of CLAS.

magnetic field, p is the momentum and q the charge of the particle. The maximum field, 2 T.m, is at forward angles where the higher momentum particles are produced.

5.2.2 Drift Chambers

The drift chambers are organized in three concentric regions, each of them made of two superlayers, as depicted in figure 5.5 (left). In this geometry, the region 1 is the closest to the target and has important impact on the determination of the particle origin. The region 2 is placed in the high field region, in the middle of the toroid, and has a major impact on momentum reconstruction. The region 3 helps the global track reconstruction with other detectors, as well as the momentum determination. The goal of this system is to detect charged particles with momenta greater than 200 MeV/c on a large solid angle: from 8° to 142° and covering 80% of the azimuthal angle. To reach these objectives, all the electronic and heavy structure material is concentrated in

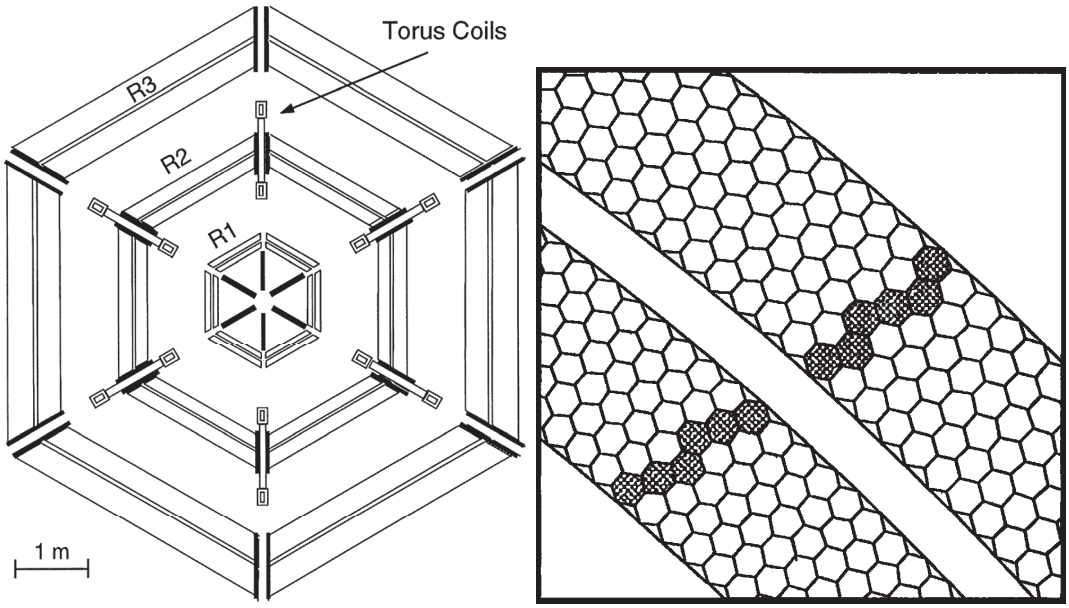

Figure 5.5: Vertical cut of the drift chambers of CLAS (left). Detail of the structure with a typical signal (right). (Both from [Mestayer 2000].)

the shadow of the torus magnets, because this region is already inefficient.

The structure of the superlayers is depicted in figure 5.5 (right). Hexagons corners represent the field wires with, in the center of each hexagon, a sense wire, on which the signal is detected. The hexagonal form of the cells mimics a circular geometry and gives a relatively regular field. The signal is produced by the avalanche of electrons triggered by the passage of a charged particle. In order to have avalanches in the chamber, we use Argon-CO_2 gas and strong difference of potential between the field and sense wires. Finally, each superlayer is surrounded by guard wires to avoid edge effects on the electric field.

Tracks are first reconstructed using only the wire positions, leading to a centimeter precision. Then, using the time delay of the signal, the position in each cell can be determined more precisely (∼300 μm precision). The calibration of the drift chambers is done in several steps. First, data is taken with no magnetic field for alignment calibration. Then, using the GARFIELD [Garfield 2010] simulation, the drift times are estimated and, eventually, they are refined using data. The calibration using data is based on the minimization of the χ^2 of reconstructed tracks. It allows to evaluate precisely the drift speed of the electronic cascades. However, even after the calibration, a left-right ambiguity remains for each cell. This is resolved by minimizing the χ^2 for each reconstructed track.

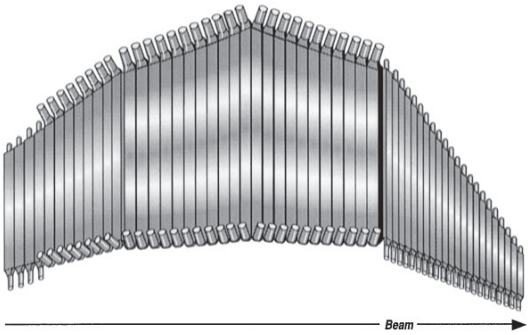

Figure 5.6: Schematic of a sector of the CLAS time of flight system (from [Smith 1999]).

5.2.3 Scintillator Counters

Scintillator counters are placed around the drift chambers to identify the different hadrons, namely π, K and p. Using the time of flight (TOF) in combination with the momentum known from the drift chamber, the mass, and therefore the nature of the particle, can be inferred. Moreover, because of the importance of multiple scattering, at low momentum (< 500 MeV) proton momentum can be determined with similar or better accuracy by the time of flight system than by the drift chambers. The scintillator counters are covering the same angular range as the drift chamber, from 8° to 142°, and similarly the electronics and the photo-multiplier tubes (PMT) are arranged to be in the shadow of the torus magnets. The structure of one sector is depicted in figure 5.6. The PMTs are at the two ends of each scintillator, allowing to use the time difference between them to find the hit position.

To reach optimal precision, a calibration system is installed directly on the scintillator. In order to synchronize the two PMTs on each side, laser pulses are sent to the center of each scintillator using optical fibers. This signal also permits to calibrate the gain after the analog-to-digital converters (ADC). The laser light is sent in synchronization to all scintillators giving a first estimation of the time response of each PMT. Delays between the various scintillators is then refined with data, using the $e^- + p \rightarrow e^- + \pi + X$ reactions. Pions can be identified, at low energy, even without good calibration of the time of flight system and, similarly, electrons can be identified using the Cherenkov counter and the calorimeter. Both electrons and pions momenta being known from the drift chambers, the time delay can be calculated and compared to the measurements. This allows to make the calibration in the exact condition of the data taking, improving the time resolution of the detector to ~ 160 ps.

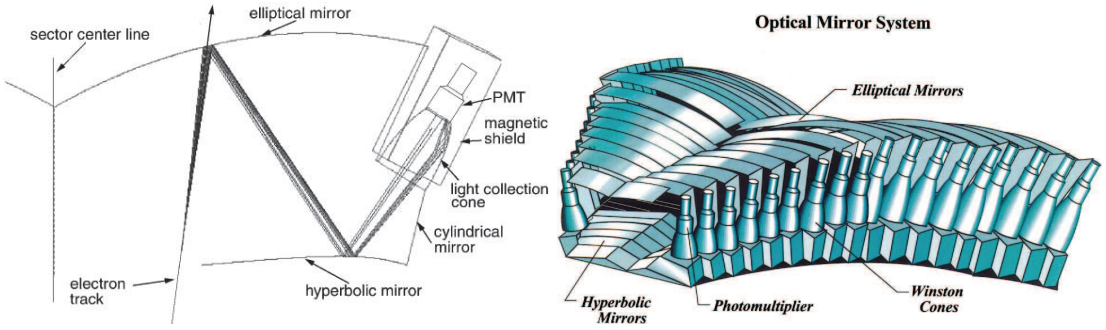

Figure 5.7: Transverse view of the CLAS Cherenkov counter and illustration of its light collection system (left). Schematic of a sector of CLAS Cherenkov counter (right). (Both from [Adams 2001].)

5.2.4 Cherenkov Counters

The Cherenkov light is emitted by particles going faster than the speed of light in a given material ($> c/n$ with n the refractive index). This phenomenon is very helpful for particle identification, especially at high energy, because the threshold for signal in a gas is very close to the speed of light, so only very light or very energetic particles will be detected. Detectors using this effect also permit to differentiate particles using the different momentum threshold, which depends on the mass of the particle.

The CLAS Cherenkov counters cover the forward angles up to 45°. The Cherenkov light is focused on the PMTs using a mirror system illustrated in figure 5.7 (left). As the sector size varies with θ, 18 different stripes were designed and assembled to form a complete sector (figure 5.7 (right)).

In CLAS, the Cherenkov detector is of major importance to separate electrons from pions. Indeed, it is used in the trigger system for electron scattering experiments. The Cherenkov counter needs to be very sensitive to avoid any gain issue at the trigger level. For this reason, PMTs were chosen to be sensitive to single photo-electron signal. The gas radiator is C_4F_{10} (perfluorobutane), which index of refraction is $n = 1.00153$ at atmospheric pressure. It corresponds to a threshold of ~ 10 MeV/c for electrons, ~ 2.5 GeV/c for pions and ~ 9 GeV/c for kaons.

Calibration of the PMT gains is obtained by recording signal while triggering on noise. Gains are adjusted with the single photon electron signal which, in those conditions, is clearly resolvable compared to other noise. Timing is calibrated using data. As the Cherenkov system is much less precise than the time of flight system, its timing is simply synchronized with time of flight using electrons.

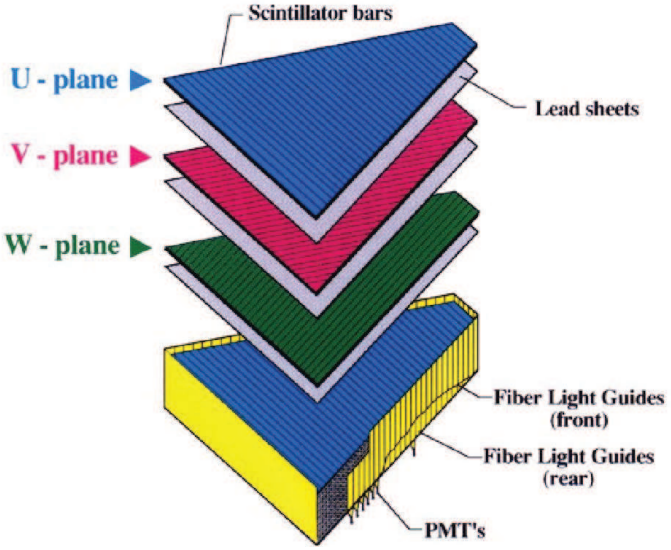

Figure 5.8: Schematic of the CLAS electromagnetic calorimeter (from [Amarian 2001]).

5.2.5 Electromagnetic Calorimeter

The electromagnetic calorimeter is an assembly of scintillators and lead, 16 radiation lengths thick. It is designed to collect the total energy deposited by an electromagnetic shower, i.e. deposited by an electron or a photon[2]. The calorimeter covers similar angles than the Cherenkov counters ($8° < \theta < 45°$) and is designed to differentiate electrons from pions at high momentum, in particular above the pion threshold in the Cherenkov counter (> 2.5 GeV). The calorimeter also detects neutral particles, such as photons and neutrons.

Each sector of the calorimeter is triangular and composed of scintillator strips organized in planes of 36 strips. There is 39 layers and each is rotated by 120° compared to each others, as shown in the figure 5.8. Those are called U, V and W planes, together, they give the position of the electromagnetic shower. The calorimeter signals are separated in two stacks: an inner composed of 5 groups of 3 planes (the closest to the target) and an outer composed of the remaining 8 groups. The scintillators are connected with optical fibers to the readout system placed on the back of the calorimeter.

GEANT simulation of the calorimeter showed that electromagnetic shower is at its peak between layers 6 and 12 for usual CLAS electrons ($0.5 < p < 4.5$ GeV/c) and leakage at the rear of the detector is of 1 to 2%. The energy deposited compared to the total energy of the incoming particle is called sampling fraction and is estimated to be 27.4% for electrons and photons. For comparison, charged pions, which are minimum ionizing particles, give only

[2] Pions and hadrons in general emit much less energy through EM processes because of their larger mass.

a signal of 75 MeV regardless of their momentum. However, the pions might interact hadronically in the calorimeter leading to a larger energy loss. Therefore, it is important to use together the Cherenkov counter and the calorimeter to differentiate the electrons from the pions.

Calibration of the gains is made using cosmic muons. Because they are also minimum ionizing particles, they deliver a very similar energy in all layers of the calorimeter independently of the position of the hit and of their energy. The time calibration is obtained using data with electrons in coincidence with the time of flight system. The calorimeter reaches a time resolution of ∼500 ps.

Chapter 6

Data Analysis

6.1 Introduction

This chapter treats of the analysis of the data taken in CLAS for the experiments E-02-104 [Brooks 2002] and E-02-110 [Hafidi 2002]. The code name for the run is *eg2*, it was composed of three phases labeled *a*, *b* and *c*. The analysis is performed on the data collected in the beginning of 2004 during the third phase (*eg2c*), for which the beam energy was 5.014 GeV[1].

As the two experiments, running simultaneously, aimed at comparing deuterium with heavier nuclei, it was decided to use a double target system [Hakobyan 2008]. The first target is filled with liquid deuterium and the second is a solid target. The latter can be made of carbon, aluminum, iron, tin or lead and changed remotely (the system is shown in figure 6.1). The two targets are separated by only 4 cm in order to reduce the acceptance differences between them. The advantage to have the two targets in the beam line simultaneously is that several systematic effects related to the beam and detector properties will cancel in the nuclear ratio.

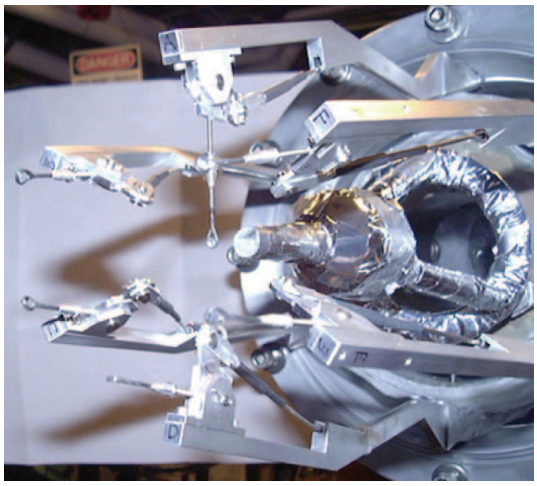

Figure 6.1: Picture of the target system of the *eg2* run. The cryogenic target is in the back, enveloped in aluminum foils. The solid targets are held by mechanical arms allowing to change targets remotely, in the picture the top arm is in the beam line.

[1] The other phases gave a small amount of data at a beam energy of 4 GeV.

The analysis of the experiment E-02-110, focusing on the study of color transparency effects, has been approved recently [El Fassi 2008]. As this analysis went through a careful review by the CLAS collaboration, we use similar analysis methods when possible. In particular, the electron selection presented here is very similar to theirs; the main difference being a new target determination method.

In the section 6.2, we identify the following particles e^-, π^- and π^+ using a series of cuts on the various detector outputs. For pions, we use the signal in the drift chambers (DC) and in the scintillator counters (SC), for which we require a positive status (global and DC) from the reconstruction code. For electron identification, we use signals from all detector parts DC, Cherenkov Counters (CC), SC and Electromagnetic Calorimeter (EC).

Once particles are identified, we can extract the observables (multiplicity ratio and transverse momentum broadening). The method is presented in the section 6.3, however these need several corrections presented in the section 6.4. The two main corrections are the acceptance of the detector and the radiative effects, both are evaluated using Monte-Carlo generator. The evaluation of the systematic errors is presented in the section 6.5. The analysis' final results are presented and discussed in the next chapter.

6.2 Particle Identification

6.2.1 Electron Identification

First, we apply a fiducial cut on the EC to remove the electrons detected on the edge of the calorimeter ($U_{EC} > 40$ cm, $V_{EC} < 360$ cm and $W_{EC} < 395$ cm in the calorimeter's coordinates). These are problematic hits because the generated electromagnetic shower might be partly outside the detector, leading to a wrong measurement of the energy deposited.

To reject pions, we apply cuts on the energy deposited in the EC using the measurements in the inner (E_{in}) and the outer part (E_{out}) of the calorimeter:

$$\mu\left[1-\frac{0.3}{\sqrt{a}}\right]-\frac{E_{in}}{p}\leq\frac{E_{out}}{p}\leq\mu\left[1+\frac{0.3}{\sqrt{b}}\right]-\frac{E_{in}}{p}, \tag{6.1}$$

where $\mu = 0.271$ is the mean of the fraction of the energy deposited in the calorimeter by electrons, the a parameter is set at 0.5 and the parameter b is a function of momentum given in table 6.1. The adjustment of b is motivated by the non-linear dependence of the energy deposited as a function of the particle's momentum (see figure 6.2). As pions are minimum ionizing particles, they are expected to lose a constant energy in the inner part of the calorimeter (around 30 MeV), regardless of their momentum. Therefore, by requesting

Momentum bin (GeV/c)	Parameter b
0.5 - 0.7	0.85
0.7 - 0.9	0.8
0.9 - 1.1	0.85
1.1 - 1.3	1.05
1.3 - 1.5	1.1
1.5 - 1.7	1.35
1.7 - 1.9	1.35
1.9 - 2.1	1.45
2.1 - 2.3	1.35
2.3 - 2.5	1.35
2.5 - 2.7	1.35
2.7 - 2.9	1.3
2.9 - 3.1	1.35
3.1 - 3.3	1.35
3.3 - 3.5	1.5
3.5 - 3.7	1.6
3.7 - 3.9	1.8
3.9 - 4.1	1.8
4.1 - 4.3	1.8
4.3 - 4.5	1.8

Table 6.1: Values of the parameter b used in equation 6.1 for different momentum ranges.

more than 60 MeV to be deposited, we efficiently cut the pion contamination (see figure 6.3).

In the CC, the mean number of photo-electrons[2] from a high energy electron is expected to be around 10. However, hadrons can generate noise due to δ electrons produced in the materials of the detector. This signal is expected around one photo-electron, to remove it, we keep only tracks with more than 2.5 photo-electrons (figure 6.4).

To select electrons, we also require that the particles are negatively charged, according to the bending direction of the tracks in the drift chamber. Positively charged particles are identified as positrons. In the rest of the analysis we use events with only one electron and no positron to avoid any confusion between the scattered electron and electrons from hadronic decays or photon conversion to e^+e^- pair, which often lead to the production of a positron.

[2] Photo-electrons are electrons produced in the front window of the Photo-Multiplier Tube (PMT) by single photons.

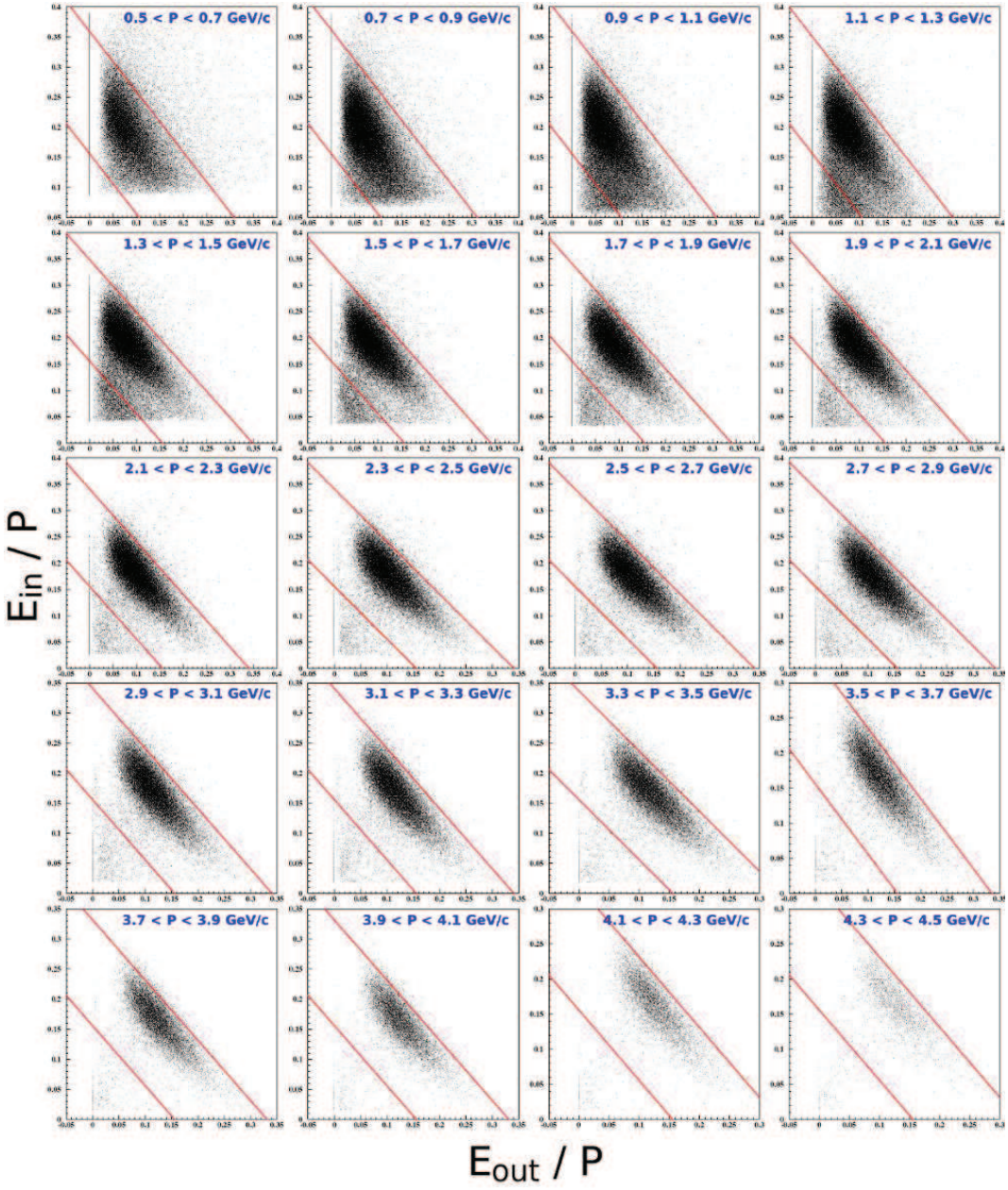

Figure 6.2: Energy deposited in the inner part of the EC (E_{in})) as a function of the energy deposited in the outer part (E_{out})), both divided by the momentum of the particle. Each panel is for a different momentum range, the red lines illustrate the cuts from equation 6.1.

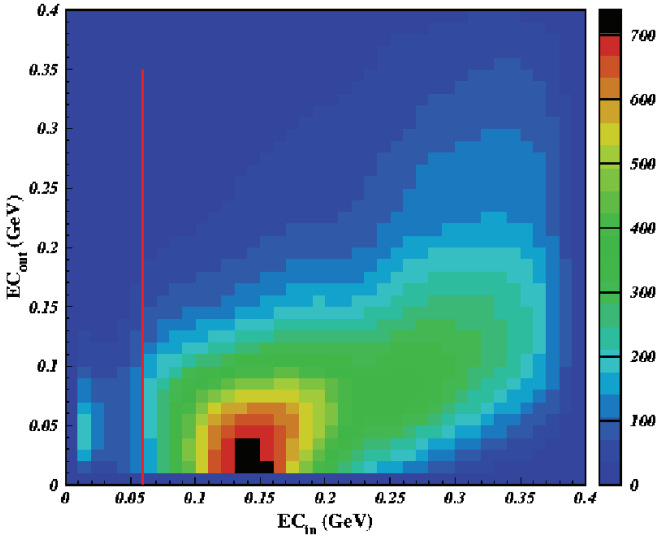

Figure 6.3: Energy deposited in the inner part of the EC as a function of the energy in the outer part. The red line illustrates the cut applied for electron selection.

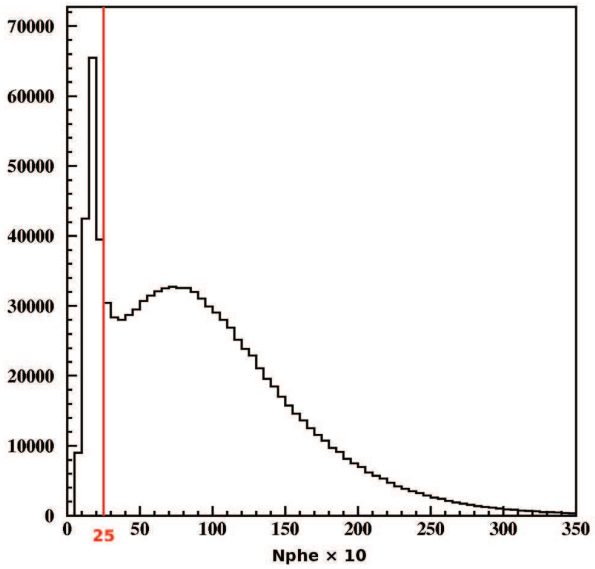

Figure 6.4: Number of photo-electrons (×10) per tracks, the red line shows the cut to remove pion contamination.

6.2.2 π^- Identification

To identify negatively charged pions, we select negative tracks, which are not identified as electron. Pions are detected in the angular range from ∼10 to ∼140 degrees using the DC and SC only.

The identification consists mostly of a time of flight (TOF) test. We define $\Delta\beta = \beta_{measured} - \frac{p}{\sqrt{p^2+m_\pi^2}}$ and request $\Delta\beta$ to be zero within ±0.03. Figure 6.5 illustrates the effect of this cut. We notice that there is not much negative kaons or anti-protons. Therefore, the contamination from these should be small (see section 6.5.1 for more detailed analysis).

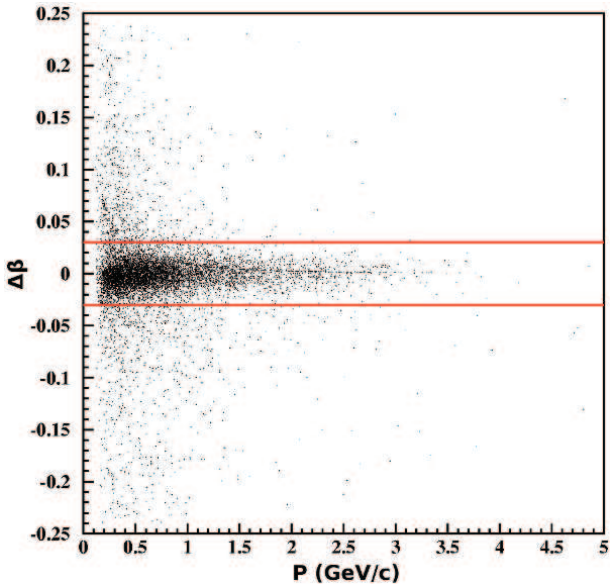

Figure 6.5: $\Delta\beta$ as a function of momentum (GeV/c). Only negative particles are plotted. The red lines are the cuts applied to select negative pions.

In principle, pion identification could be improved using the Cherenkov counter for momentum higher than 2.5 GeV/c. But, the low efficiency observed (figure 6.6), especially at momentum close to the threshold (∼25% at 2.5 GeV/c and ∼50% at 3 GeV/c), makes its use less compelling. Moreover, as only a very small amount of K^- and no $\bar{p}$ are present on the figure 6.5, we decided not to use the CC for pion identification.

6.2.3 π^+ Identification

The identification of positively charged pions is similar to that of negative pions. However, the time of flight plot is significantly more busy (figure 6.7), showing significant contamination from K^+ and protons at high momentum. As the CC is not efficient enough for hadron separation, the numerous kaons and protons should be removed by a tighter TOF cut:

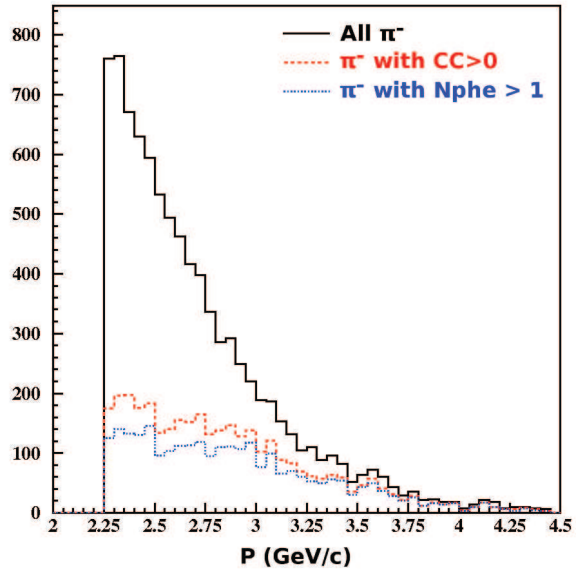

Figure 6.6: Histograms of momentum (GeV/c) for π^-. In black are all identified π^-, in red pions that also fire in the Cherenkov counter and in blue those that fire in the Cherenkov counter with more than 1 photo-electrons.

$$p_\pi \leq 3.375 \text{ GeV/c}: \qquad \Delta\beta > \max\left(-0.03, \frac{p}{\sqrt{p^2+0.4^2}}\right), \tag{6.2a}$$

$$p_\pi > 3.375 \text{ GeV/c}: \qquad \Delta\beta > \max\left(-0.02, \frac{p}{\sqrt{p^2+0.7^2}}\right). \tag{6.2b}$$

These cuts permit to minimize the kaon contamination, below 2.5 GeV/c, and the proton contamination for all momentum. The kaon contribution cannot be avoided, however it should remain a small contribution ($\sim$ 3% according to simulation) with only a small effect on the final results (see section 6.5.1 for details). Because protons could lead to even more contamination than kaons, a stricter cut is used at high momentum. This cut is also justified by HERMES data [Airapetian 2007], which show a very different behavior of the protons compared to the other hadrons (see chapter 3).

6.2.4 Target Determination

To differentiate the two targets and remove the background, we need to determine the origin of the particles. It appears that the different sectors are shifted in z (figure 6.8), because of a small misalignment of the beam with the detector. To correct this problem, we apply the modification shown in table 6.2 for the vertex determination. Those values were determined by fitting the

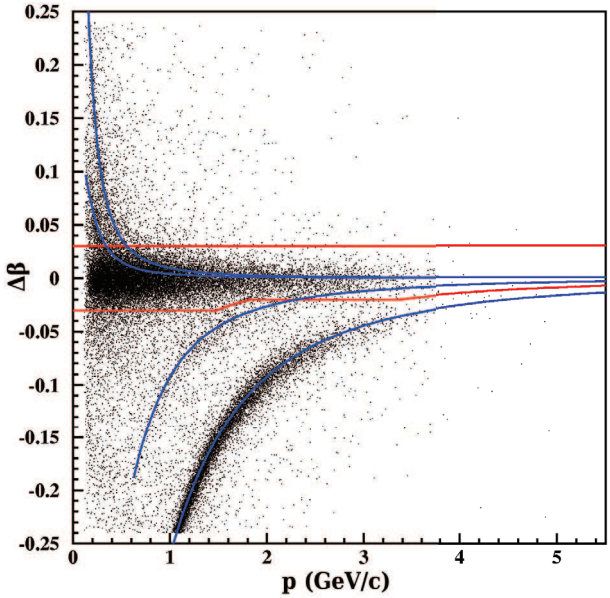

Figure 6.7: $\Delta\beta$ as a function of momentum (GeV/c). Only positive particles are shown, the red lines are the cuts applied to select positive pions. The blue lines indicate theoretical positions for other particle masses, are plotted from top to bottom the lines for positrons, muons, kaons and protons.

Sector	Shift (cm)
1	+ 0.1
2	- 0.4
3	- 0.6
4	- 0.1
5	+ 0.4
6	+ 0.6

Table 6.2: Values used to correct the vertex information for all sectors

solid target position viewed by each sector.

The position of the targets may also vary from one run to another, but this does not happen too often. Indeed, except for aluminum, the targets remain in the same position within one or two millimeters. The positions, in CLAS coordinates, are given in table 6.3.

The detected electrons are associated with the solid target if their vertex position is at less than 1.5 cm ($\sim 3\,\sigma$) from the value of table 6.3. For the liquid target, the cut is larger, 2 cm, in order to account for the size of the target (see figure 6.8). The vertex of the pions is checked against the electron one and we request that $|Vz^{e^-} - Vz^{\pi}| < 3$ cm (see figure 6.9).

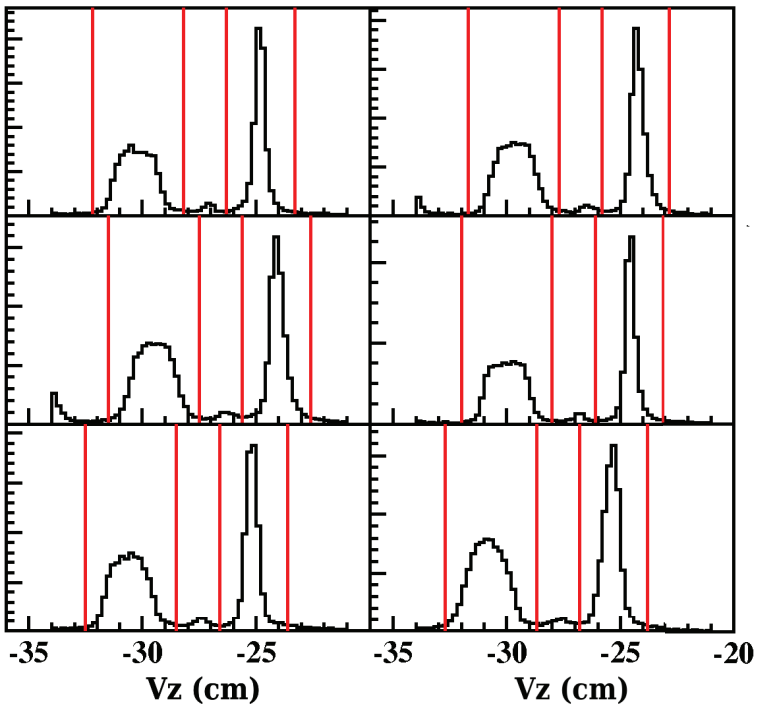

Figure 6.8: In each sector, the reconstructed vertex of electrons along the beam direction relatively to the center of CLAS. The red lines show the cuts to select the targets.

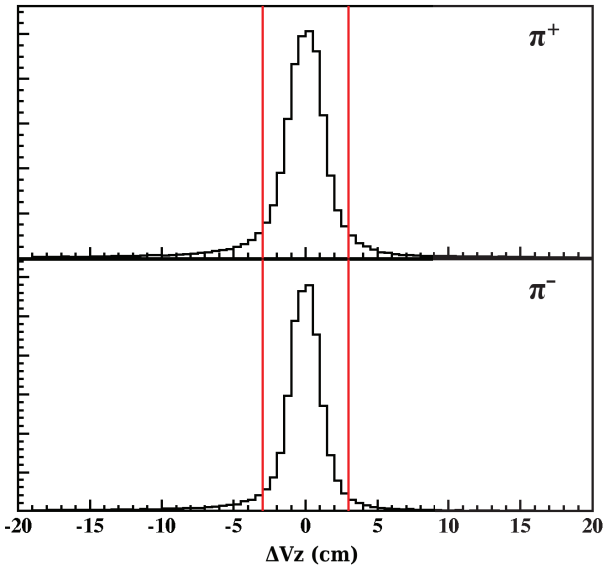

Figure 6.9: Distance along the beam axis between the electrons and pions vertexes. π^+ are plotted in the top panel and the π^- in the bottom. The red lines show the cuts applied to select pions.

Target	Carbon	Al (1)	Al (2)	Iron	Tin	Lead
Liquid	-30.1	n/a	n/a	-30.2	-30.1	-30.1
Solid	-24.7	-25.0	-23.8	-24.9	-23.8	-24.9

Table 6.3: Measured mean vertex position of targets, relative to the center of CLAS (in cm). (Aluminum data are separated in two sets because the target position appears to have changed at some point.)

6.2.5 Data Quality

To check the quality of the runs[3], we look at the ratio of the number of scattered electrons, between the liquid and solid targets. If the beam is hitting other materials than the targets or if the detector is not working properly this ratio can be off and indicates a problematic run. The figure 6.10 shows the values obtained; we fit the mean value for all runs of each target and eliminate runs away by more than 5σ. We note that the ratios are coherent with target thicknesses given in [Hakobyan 2008] except for carbon. Because of this, the density of the carbon target was remeasured recently and found to be coherent with the data[4].

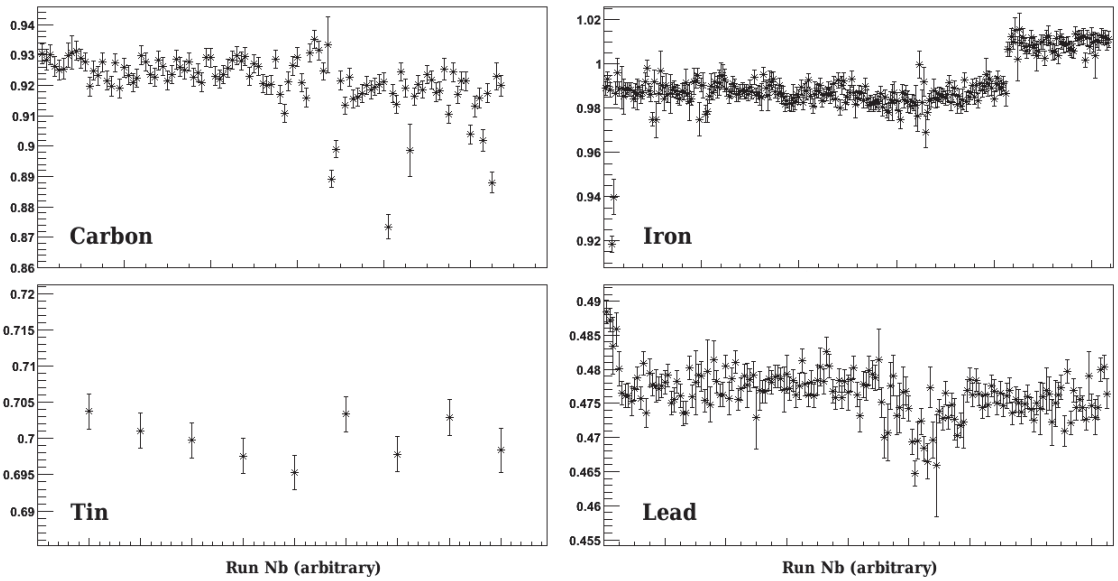

Figure 6.10: Ratio of the number of electrons scattered by the targets (solid over deuterium) for each run.

[3] Runs correspond roughly to two hours of data, they can be smaller in case of problem.
[4] i.e. (1.747+/-0.0007)g/cm^3 instead of 2.235 g/cm^3

6.3 Extraction of Multiplicity Ratio and $\Delta \mathbf{P}_{\perp}^2$

Since we are interested in deep inelastic scattering, we use the following cuts: $Q^2 > 1$ GeV^2/c^2 and $W > 2$ GeV. We also apply a cut $y < 0.85$ to reduce the impact of radiative effects, which will be discussed in more details in the section 6.4.2.

6.3.1 Method

Once the identification is done, the calculation of the observables is straight forward and the statistical errors are calculated with this expression

$$\frac{\delta\left(R_A^h\right)}{R_A^h} = \sqrt{1/N_A^h + 1/N_A^e + 1/N_D^h + 1/N_D^e} \tag{6.3}$$

and

$$\left(\delta\left(\Delta\langle P_\perp^2\rangle\right)\right)^2 = \left(\langle P_\perp^4\rangle - \langle P_\perp^2\rangle^2\right)_A / N_A^h + \left(\langle P_\perp^4\rangle - \langle P_\perp^2\rangle^2\right)_D / N_D^h. \tag{6.4}$$

The implementation of acceptance and radiative corrections is done through weights given to each particles depending on its kinematic. The multiplicity ratio becomes

$$R_A^h(Q^2, \nu, z_h, P_\perp^2) = \frac{\sum \omega_A^h(Q^2, \nu, z_h, P_\perp^2) / \sum \omega_A^e(Q^2, \nu)}{\sum \omega_D^h(Q^2, \nu, z_h, P_\perp^2) / \sum \omega_D^e(Q^2, \nu)}, \tag{6.5}$$

with ω the weights and all the sums running over all measured particles. The expression for the transverse momentum broadening remains

$$\Delta\langle P_\perp^2\rangle = \langle P_\perp^2\rangle_A - \langle P_\perp^2\rangle_D, \tag{6.6}$$

but with

$$\langle P_\perp^2\rangle = \frac{\sum P_\perp^2 \times \omega(Q^2, \nu, z_h, P_\perp^2)}{\sum \omega(Q^2, \nu, z_h, P_\perp^2)}. \tag{6.7}$$

This expressions leads to new expressions for the statistical error uncertainties:

$$\frac{\delta\left(R_A^h\right)}{R_A^h} = \sqrt{\left(\frac{\sum {\omega_A^h}^2}{\left(\sum \omega_A^h\right)^2}\right) + \left(\frac{\sum {\omega_A^e}^2}{\left(\sum \omega_A^e\right)^2}\right) + \left(\frac{\sum {\omega_D^h}^2}{\left(\sum \omega_D^h\right)^2}\right) + \left(\frac{\sum {\omega_D^e}^2}{\left(\sum \omega_D^e\right)^2}\right)} \tag{6.8}$$

and

$$\left(\delta\left(\Delta\langle P_\perp^2\rangle\right)\right)^2 = \left(\frac{\sum\omega_A^h P_\perp^4}{\sum\omega_A^h} - \left(\frac{\sum\omega_A^h P_\perp^2}{\sum\omega_A^h}\right)^2\right) \times \left(\frac{\sum\left(\omega_A^h{}^2\right)}{\left(\sum\omega_A^h\right)^2}\right) + \left(\frac{\sum\omega_D^h P_\perp^4}{\sum\omega_D^h} - \left(\frac{\sum\omega_D^h P_\perp^2}{\sum\omega_D^h}\right)^2\right) \times \left(\frac{\sum\left(\omega_D^h{}^2\right)}{\left(\sum\omega_D^h\right)^2}\right). \tag{6.9}$$

6.3.2 Preliminary Results

We present in figure 6.11 few preliminary results, before the application of any correction, with the goal to provide a first idea on data quality. The preliminary results will also be used to illustrate the effects of the corrections discussed below.

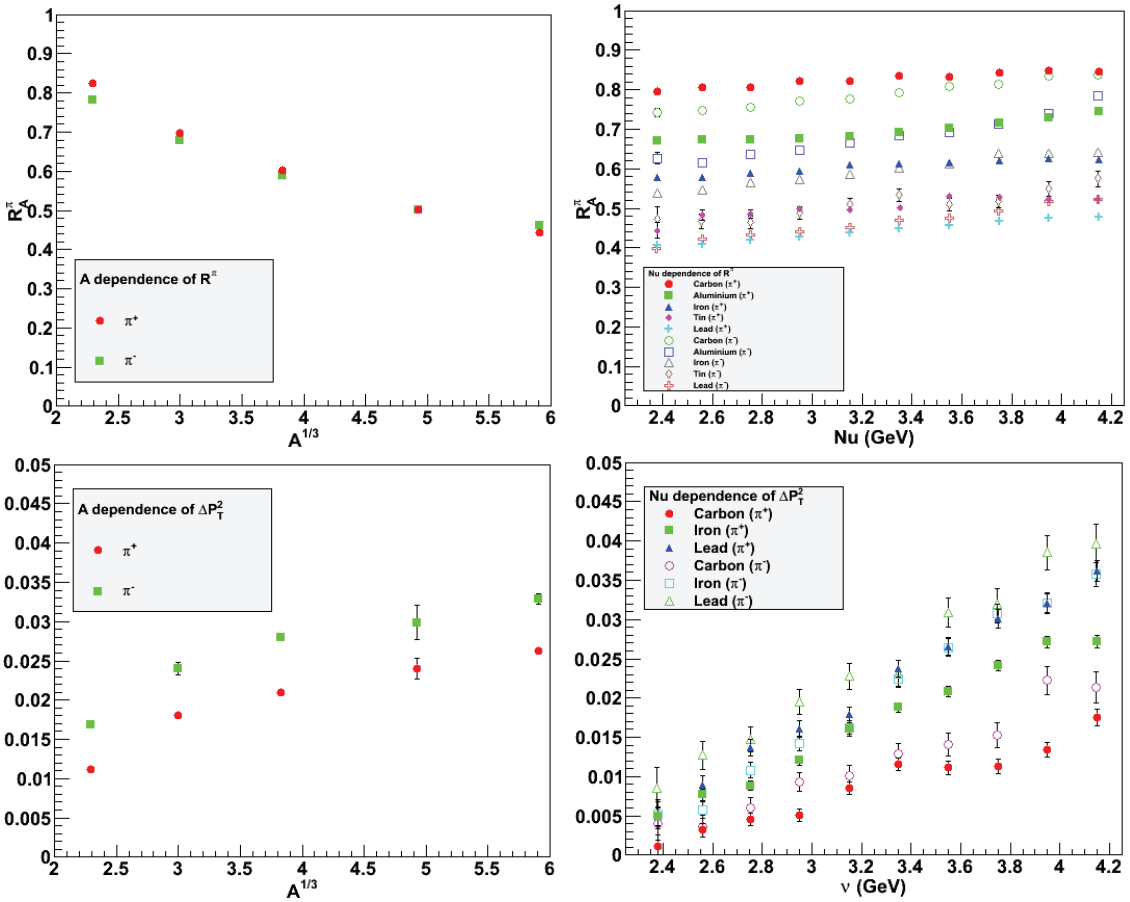

Figure 6.11: Results for multiplicity ratios (top) and transverse momentum broadening (bottom) without correction.

6.4 Corrections

6.4.1 Acceptance Correction

The acceptance correction consists of applying weights to the events in experimental data to correct for the inefficient parts of the detectors. Incidentally, it also corrects for small other issues of detection, such as misidentification and scattering on detectors materials. The quality of the correction depends on the ability of the simulation to reproduce the experiment, the statistics accumulated and the size of interfering effects, such as bin migration.

6.4.1.1 Simulation

To correct for acceptance effects, we simulated a total of a 100 million events per target (^{2}H, C, Fe and Pb) using the PYTHIA [Sjostrand 2006] event generator slightly modified to include Fermi motion effects. The generated events are processed by the CLAS software (GSIM, GPP and user_ana) to simulate the detector and the reconstruction process similarly to the one for experimental data.

Then the simulated data are processed in a similar way than the experimental data by applying the cuts described in the section 6.2. Overall the simulation reproduces quite well the detectors responses, yet two issues might affect us and have to be understood. First, the efficiency of the CC is overestimated in the simulation. On the electron side the signal is a little stronger in the simulation (11 photo-electrons) compared to experimental data (8 photo-electrons), but this feature should not affect us too much, because we are cutting only the tail of the distribution in both cases. For pion identification, as mentioned in section 6.2.2, we do not use the CC because of its unexpectedly low efficiency We observe, in the simulation, a much better detector response than in experiment (figure 6.12 compared to figure 6.6), this result also advocates against the use of the CC in the particle identification. Indeed, the poor reproduction of the experimental signal would introduce a bias in our acceptance correction. Second, in figure 6.8 the cuts on vertex are shifted from one sector to another, since the simulation have perfect alignment of beam with CLAS this feature is not present in the simulated data. Therefore, we do not apply the shift from the table 6.2 to the simulation; results are shown in figure 6.13.

The kinematic distributions from the simulation are compared to the experimental ones. This is important for the acceptance correction, to see whether or not, we can integrate over these variables for the correction. Comparisons between simulation and experiment are shown in figures 6.14 to 6.17.

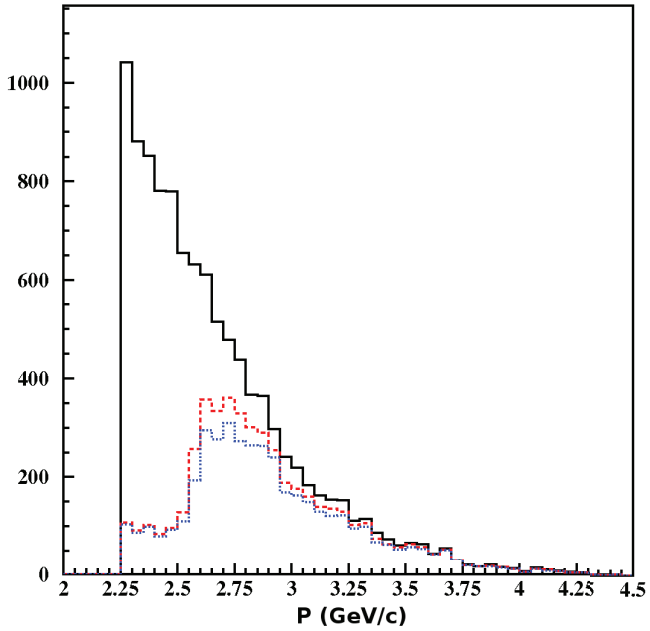

Figure 6.12: Histograms of momentum (GeV/c) of π^- for **simulated data**. In black are all identified π^-, in red pions that also fire in the Cherenkov counter and in blue those that fire in the Cherenkov counter with more than one photo-electron.

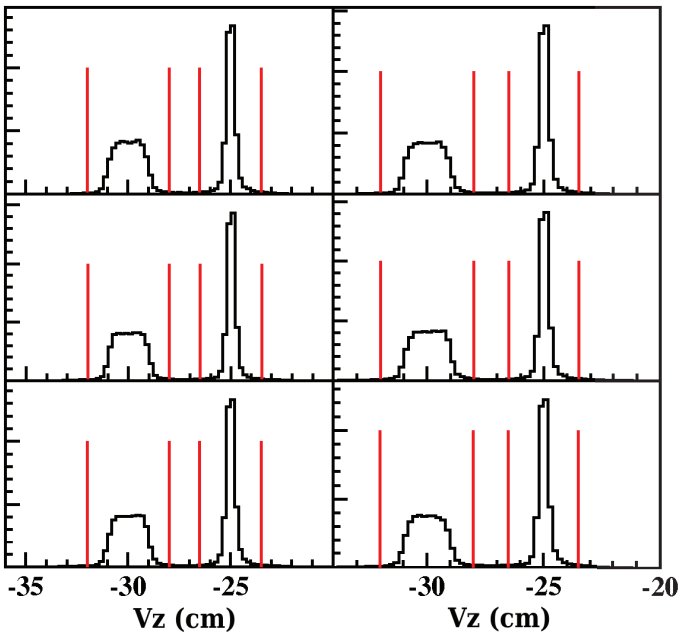

Figure 6.13: Reconstructed vertex position, along the beam direction, with values relative to CLAS center, for electrons from **simulated data** and each sector. The red lines show the cuts used to select the targets.

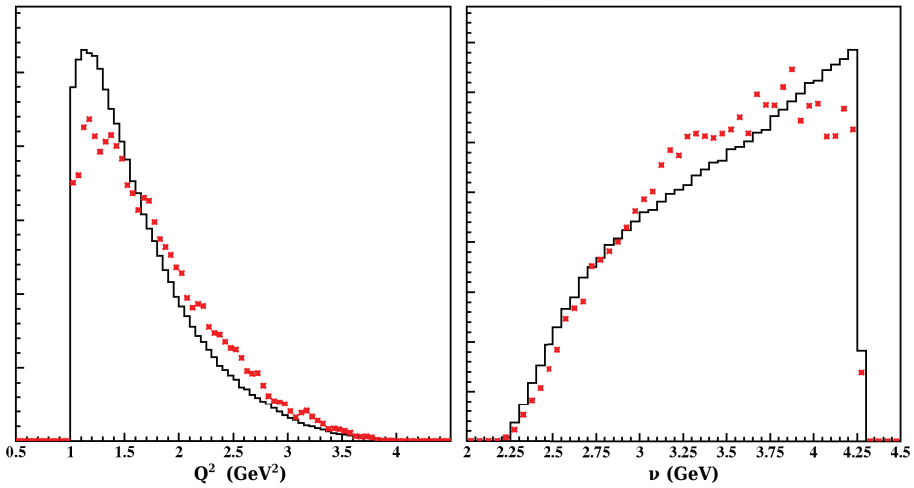

Figure 6.14: Comparisons, for Q^2 (GeV^2/c^2) and ν (GeV), of the distributions from simulated (red crosses) and experimental (histogram) data using deuterium target.

The agreement is reasonable, but not perfect. The differences are due to the PYTHIA simulation, which is not including some physical effects, such as radiative and diffractive processes.

6.4.1.2 Correction of the Data

The acceptance is defined as the ratio of reconstructed events over generated ones,

$$\text{Acc} = \frac{N_{rec}}{N_{gen}}. \tag{6.10}$$

The data is corrected using weights defined as $\omega = 1/\text{Acc}$. The ω coefficient are calculated in many bins, in order to be independent of the imperfections of the event generator. However, an excess of bin could results into a strong bin migration[5], which might introduce a bias if it becomes a large effect.

We use a 4-dimensional binning to divide the large phase space available to the two particles in the measured final state. To evaluate the systematic error associated to the correction, we use two different binning, which are presented in the table 6.4. The total number of bins is constrained by the amount of data generated and has to be maintained reasonably low in order to keep a good statistical precision and reasonable bin migration.

The weights ($\omega_{\pi^\pm}(\nu, x_{Bj}, p_h, t)$) are then extracted from the simulation using equation 6.10, however, it appears that these should not be used directly. Three factors are sources of problematic bins:

[5] Fraction of events not reconstructed in the bin they were produced in.

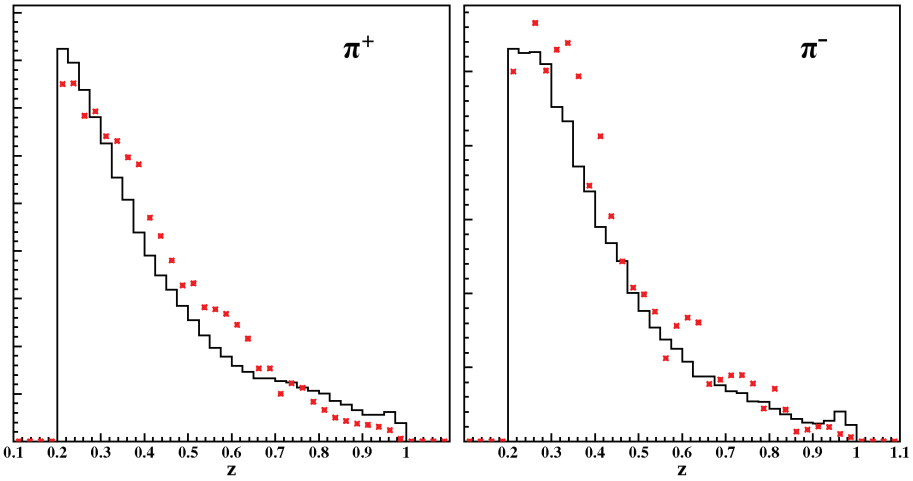

Figure 6.15: Comparisons, for z of charged pions, of the distributions from simulated (red crosses) and experimental (histogram) data using deuterium target.

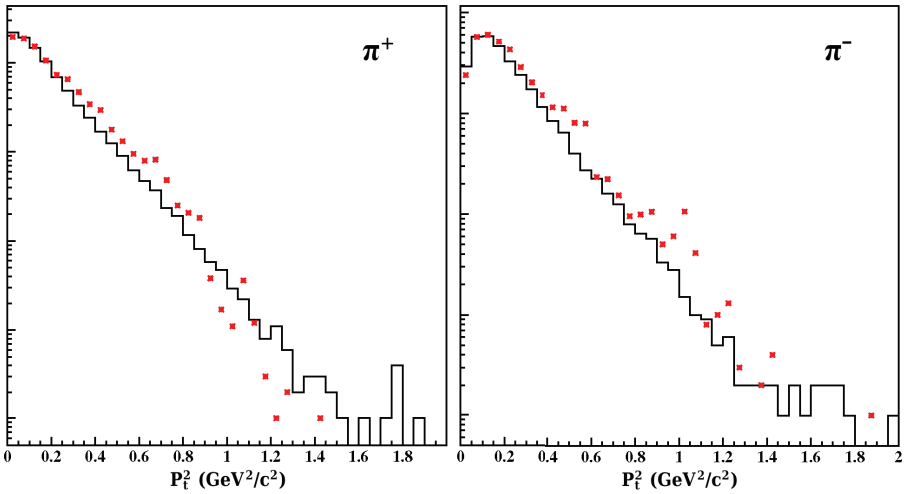

Figure 6.16: Comparisons, for $p_{\perp}^2$ (GeV^2/c^2) of charged pions, of the distributions from simulated (red crosses) and experimental (histogram) data using deuterium target.

Variable	Number of bins
ν	5
x_{Bj}	5
p_h	7
t	7
Total	1225

Variable	Number of bins
Q^2	5
ν	5
z_h	7
t	7
Total	1225

Table 6.4: Variables and their associated number of bins used for the multi-dimensional binning of the acceptance correction. The left panel shows the variables used for the correction and the right panel the variables used for the evaluation of the systematic error.

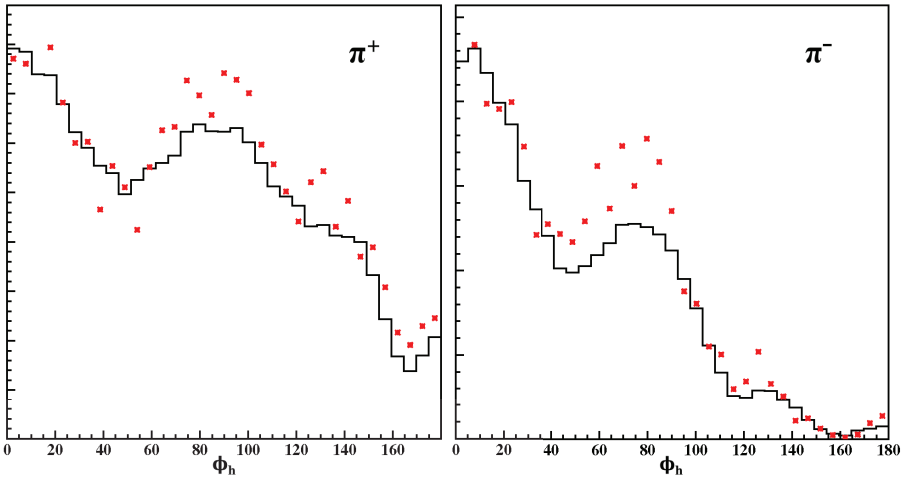

Figure 6.17: Comparisons, for ϕ_h of charged pions, of the distributions from simulated (red crosses) and experimental (histogram) data using deuterium target.

- too large weight values, i.e. very small acceptance;
- very few events reconstructed, leading to large statistical errors;
- many events that were generated in another bin (bin migration).

To remove these bins we apply the following cuts:

$$\left(\frac{\delta\omega}{\omega}\right)^2 \times \frac{R_c}{\omega} < 2, \tag{6.11}$$

and

$$N_{rec} > 4, \tag{6.12}$$

with $\delta\omega/\omega$ the relative error on the weight and R_c the proportion of events in the bin initially generated in another bin. Figure 6.18 shows the distributions of ω, after the cuts. Around 900 bins remain for both targets, but we notice that weights are much larger for π^-, indicating a larger correction, compared to π^+. In order to maximize the acceptance and reduce the sensitivity to the cuts applied on the weights, which are arbitrary, the bin weights are corrected. We apply a reweighting by calculating a new acceptance in a two dimensional binning:

$$\mathrm{Acc}_2(\nu, p_h) = \frac{\sum_{rec} \omega(\nu, x_{Bj}, p_h, t)}{N_{gen}(\nu, p_h)}. \tag{6.13}$$

with ω either the previously calculated weights or one for the excluded bins. One of the criteria for setting the limits in equation 6.11 and 6.12 is to keep the reweighting factors $\omega_2 = 1/\mathrm{Acc}_2$ at a few percents level.

The electron acceptance, necessary in order to correct the number of elec-

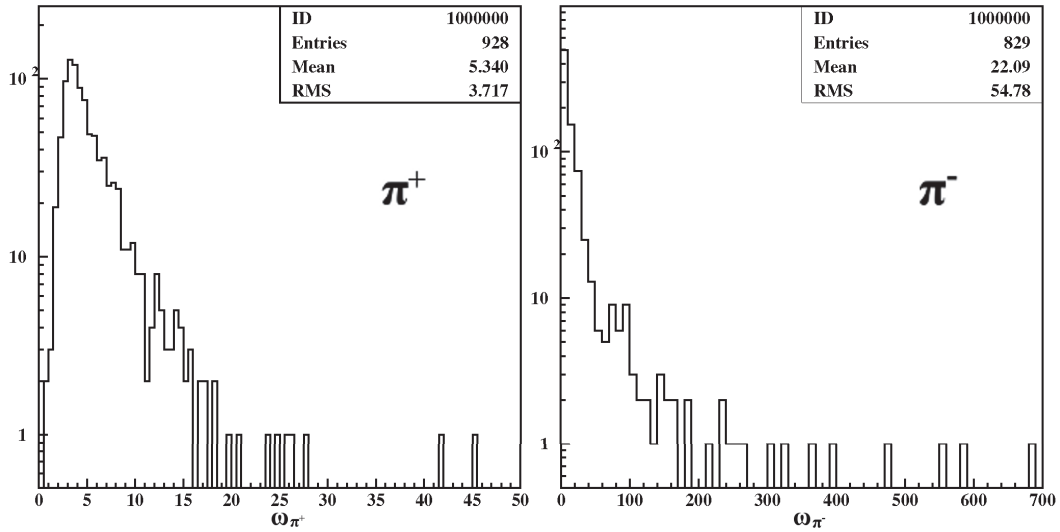

Figure 6.18: Acceptance weights for pions in deuterium (not reweighted).

trons in the multiplicity ratios, is done with the same method than in the semi-inclusive case, but only using the two first dimensions (ν and Q^2). It is also interesting to note that no reweighting is necessary for electrons as all non-empty bins pass the cuts 6.11 and 6.12.

Figure 6.19, where only statistical errors are represented, shows the new results after applying the acceptance correction. The acceptance effects appear to be of $\sim$ 10% for multiplicity ratios and $\sim$ 30% for transverse momentum broadening. This is a surprisingly significant correction for the few centimeters between the two targets. The π^-s are more affected because of CLAS' magnetic field. Indeed, the low $p_\perp^2$ π^-s have very low acceptance, this makes the errors on $\Delta\langle p_\perp^2\rangle$ very large and the related results difficult to exploit. Finally, because of some difference in position between the solid targets (table 6.3), some extra simulation is needed in order to exploit correctly the aluminum and tin results. Therefore, these two targets should not be considered for the final results.

6.4.2 Radiative Correction

Despite its lower magnitude, compared to the strong force, the electromagnetic force might have a non negligible impact on our results. The reason is that even a moderate energy photon emission can modify significantly the measured kinematic variables. To correct this effect, several simulation codes exist [Akushevich 2001], however, none permit to treat directly semi-inclusive DIS on nuclei. In order to make the correction, a dedicated Monte-Carlo simulation, based on existing codes, was developed.

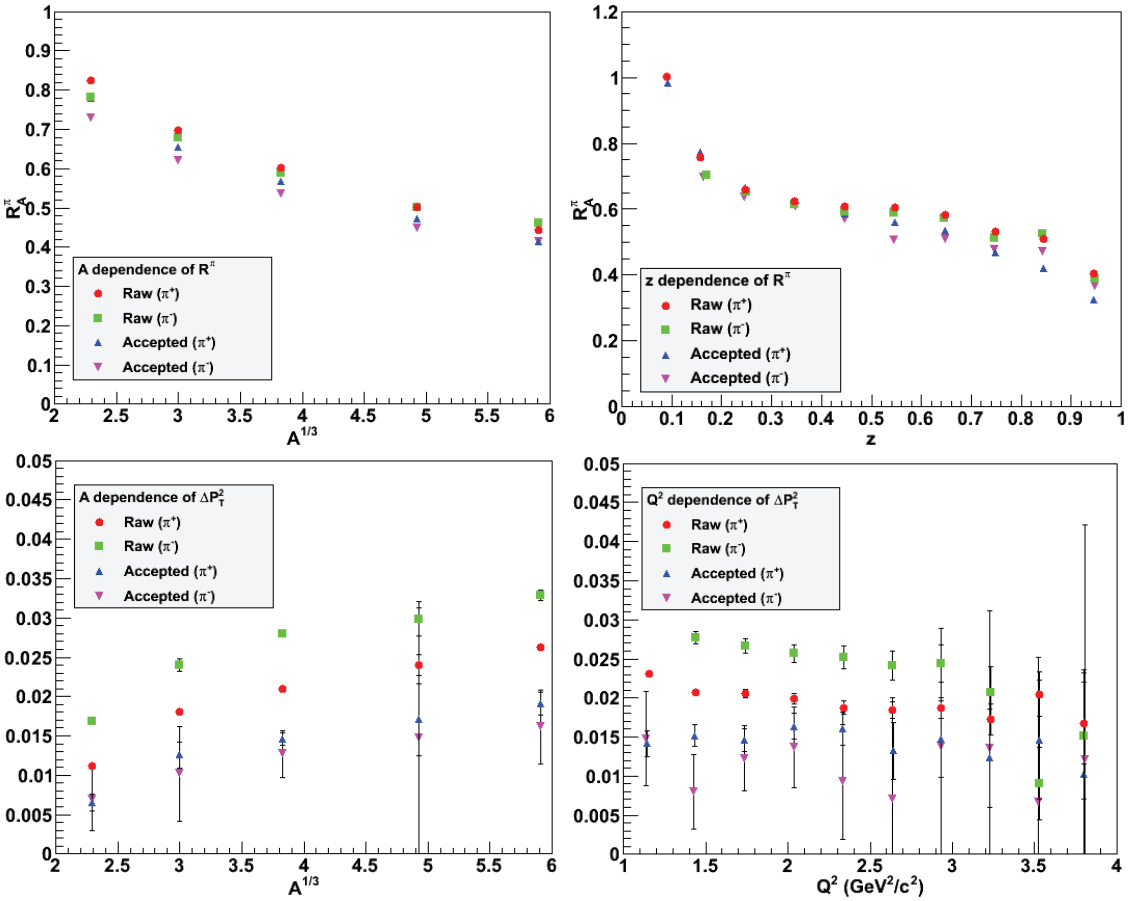

Figure 6.19: Preliminary results are compared with results corrected for acceptance, for multiplicity ratios (top) and transverse momentum broadening (bottom). Only statistical errors are shown.

6.4.2.1 Simulation

The inclusive radiative effects are generated using the RADGEN software [Akushevich 1998], which includes the Feynman diagrams shown in figure 6.20. This code evaluates directly the correction factors ($\delta = \sigma_{obs}/\sigma_{Born}$) to apply for inclusive measurements. As RADGEN is also a Monte-Carlo generator of photon emissions, it can be used to modify the virtual photon kinematic before the hadron production in any DIS generator. This allows to evaluate the radiative effects on semi-inclusive measurements by implementing RADGEN inside a full Monte-Carlo event generator.

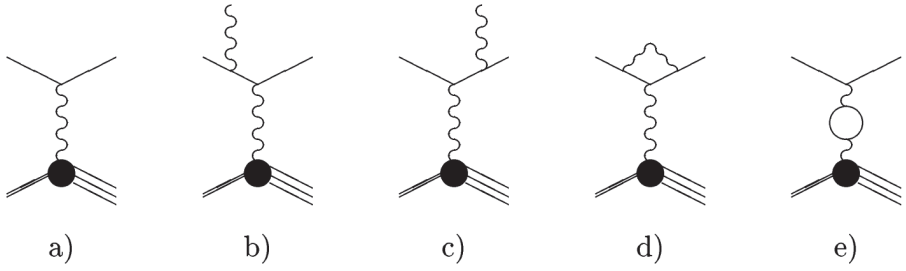

Figure 6.20: Diagrams taken into account in RADGEN [Akushevich 1998]. The diagrams b) to e) contribute to the first order calculation of the radiative effects on Born cross section (diagram a)).

The Monte-Carlo simulation we use is the GENIE software [Andreopoulos 2010], which describes well electron-nuclei scattering. Indeed, GENIE includes quasi-elastic scattering[6], which is the main channel contributing to the radiative events in the DIS region [Akushevich 2009]. The GENIE software includes also hadron cascades in the nuclei, which lead to the possibility to generate pions in quasi-elastic events. This is an important feature because the quasi-elastic cannot contribute directly to semi-inclusive measurements.

6.4.2.2 Correction of the Data

We generate 100 million events per target, with and without radiative effects. The RADGEN software gives direct indication for the inclusive correction and the comparison between the two data sets allows to extract factors for semi-inclusive correction.

The correction factors for inclusive ratio (figure 6.21) are relatively small (couple of percents maximum) except at low x_{Bj} and high ν. Semi-inclusive factors are shown in figure 6.22 as a function of z and $p_\perp^2$. However the statistical error on these is large, the problem being that only a small fraction of

[6] Elastic scattering on a single nucleon of a nuclei.

the events involved photon radiation.

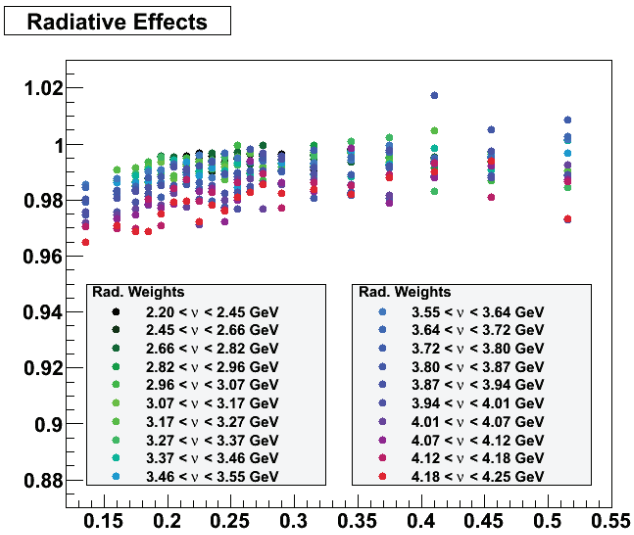

Figure 6.21: Ratios of the radiative correction factors $\delta_{Pb}/\delta_{^2H}$ as a function of x_{Bj} in various ν bins.

The radiative correction appears to be limited in amplitude in most of the kinematic, so we decided not to apply it before further studies. Improvement can be easily obtained by generating more statistics, in order to analyze the semi-inclusive effects in more depth. Also radiative effects from the target thickness are ignored in the current setup; these need to be evaluated and, eventually, corrected as well.

6.4.3 Isospin Correction

The important excess of neutron in heavy nuclei leads to a modification of the π multiplicity per DIS events. Using Hall C results [Asaturyan 2011], shown in figure 6.23, we evaluate correction factors for this effect. Our simple estimation is solely based on proton and neutron numbers; the results for π^+s are shown in table 6.5. The effect on π^- is found to be coherent with zero, therefore no correction is applied for them. We attribute errors of 10% of the effect for the isospin correction, this is chosen relatively to the precision of the Hall C measurement.

It is important to note that we correct isospin effects only for rates, thus we correct only the multiplicity ratios. Transverse momentum broadening could, in principle, also be affected, however, the results from [Asaturyan 2011] show no isospin effect in $p_\perp^2$. This gives good confidence that $\Delta\langle p_\perp^2\rangle$ is not affected by isospin effects.

We finally note that the A dependencies of the multiplicity ratios of the

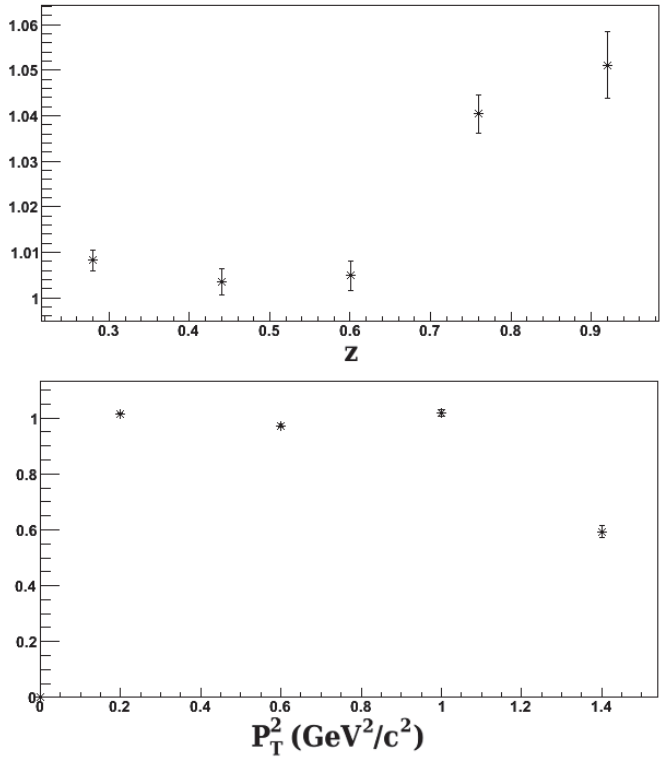

Figure 6.22: Ratios of the radiative correction factors $\delta_{Pb}/\delta_{^2H}$ as a function of z and $p_\perp^2$.

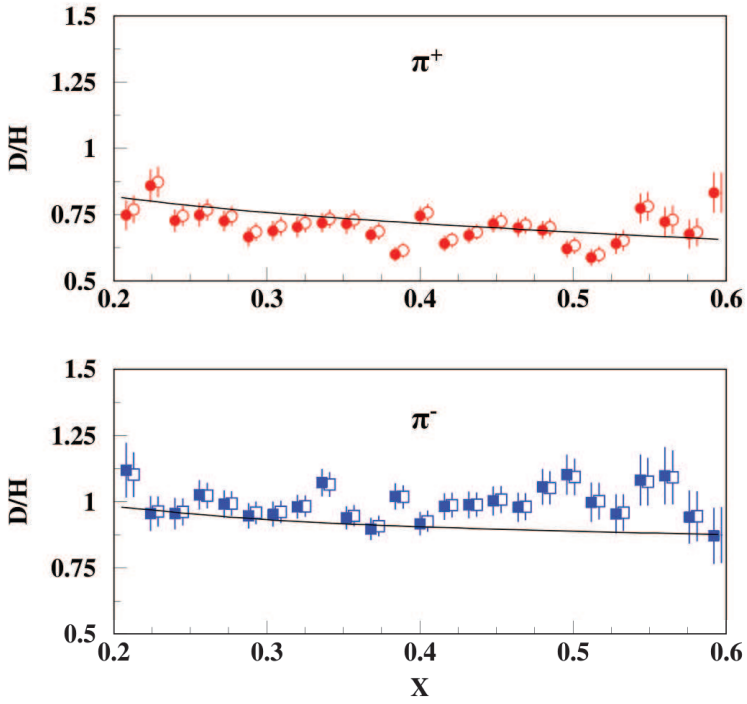

Figure 6.23: Ratios of deuteron over proton for π^+s and π^-s as a function of x_{Bj} at $z = 0.55$ [Asaturyan 2011].

Target	Isospin correction
C	0
Al	1.5%
Fe	3%
Sn	8%
Pb	10%

Table 6.5: Isospin correction applied to the π^+ multiplicity ratios for different targets.

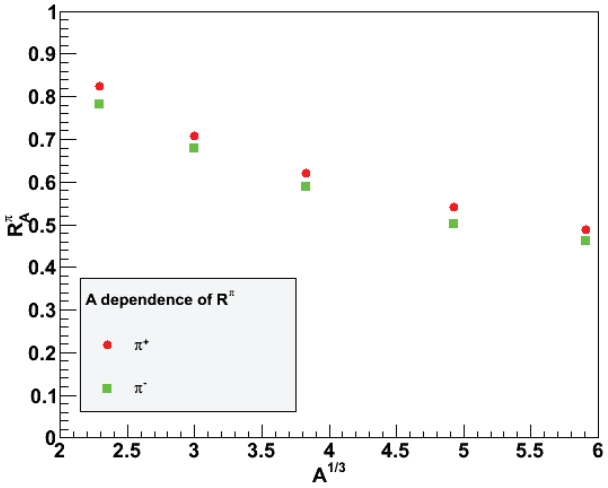

Figure 6.24: Multiplicity ratios as a function of $A^{1/3}$ with only isospin correction applied.

charged pions are similar after the isospin correction[7] (figure 6.24.). Considering previous measurements and existing models, this result was expected and confirm the need for this correction to be applied.

6.5 Systematic Uncertainties

In this section are summarized the systematic errors that we identified. The systematic point to point errors are calculated bin by bin for each results. They are caused by uncertainties on the acceptance and on the identification of particles. The normalization errors are attributed globally. They are due to acceptance effects, target misidentifications and isospin correction.

6.5.1 Quality of the Detection

The simulation of the CLAS detector, using GSIM package, is used to evaluate the systematic errors linked with:

[7] Within normalization errors presented in section 6.5

- experimental resolution of kinematic variables,
- particle misidentification,
- particle re-scattering in the detector,
- particle energy loss.

To evaluate those errors, we compare the reconstructed particles with the generated ones. Each reconstructed particle is associated with its generated parent using the similarity in momentum and angle. Distributions of Δp, $\Delta\theta$ and $\Delta\phi$ are defined as $\Delta x = \frac{\sum_n |x_{gen} - x_{rec}|}{n}$. They give the precision with which we measure these variables. Using the same simulation than for the acceptance, we find $\frac{\delta p}{p} = 0.03$, $\delta\theta = 3$ mrad and $\delta\phi = 10$ mrad. These are broader than announced in [Mecking 2003] but still reasonable. Then associated systematic errors can be easily evaluated for other variables: $\delta Q^2 \sim 0.013$ GeV^2/c^2, $\delta z \sim 0.4\%$, $\delta P_\perp^2 \sim 0.004$ GeV^2/c^2 and so on[8]. These are negligible because they are several times smaller than our usual bin sizes (see figures in section 6.3.2).

We also evaluate particle misidentification and particles originating from scattering in the detector or the coils. Only tails of distributions can lead to high contamination by misidentification, the most problematic being the tail of the $p_\perp^2$ distribution. This effect is due to the low probability to produce high $p_\perp^2$ events, therefore the contribution from accidental is relatively increased. We use the following cut $p_\perp^2 < 2.5$ GeV^2/c^2, which only removes a very little amount of data (~ 1 π in 30000). The misidentification of electrons is found to be of the order of 1 in a 1000 and does not contribute significantly to the uncertainty. For pions, the main contamination comes from kaons above 2 GeV/c ($\sim$3% of π^+ and $\sim$0.5% of π^-). Incidentally, protons also contaminates π^+ at high momentum (up to few %). The uncertainty related to misidentification is taken into account for the point to point systematic error evaluation, other effects are found to be negligible.

6.5.2 Target Reconstruction

Because of reconstruction errors or scattering on the detector materials, it is possible to associate a particle with the wrong target. To estimate this effect, we look, in the experimental data, at the number of events reconstructed upstream and downstream of the targets, where nothing should be detected. We define two test regions (see figure 6.25) to evaluate the contamination. The region 1 is upstream and chosen to be of the same size as the window used for solid target selection. We use it to evaluate contamination from the liquid target to the solid one. The region 2 is downstream and of the size of the

[8] These were evaluated for typical kinematics, the results can be significantly larger at extreme kinematics (large $p_\perp^2$ or Q^2).

window for the liquid target selection and allows to evaluate the solid target leak into the liquid one The distance between detection and test regions is identical to the distance between the two detection regions.

We find that the number of electrons in the test regions 1 and 2 represent 1 and 2%, respectively, of the total number of events. In the case of semi-inclusive measurement, where we request 2 particles in the final state, the number drops drastically and is of the order of 0.01%. In conclusion, only the number of electron is significantly affected by this problem leading to a normalization error of 1% for all multiplicity ratios.

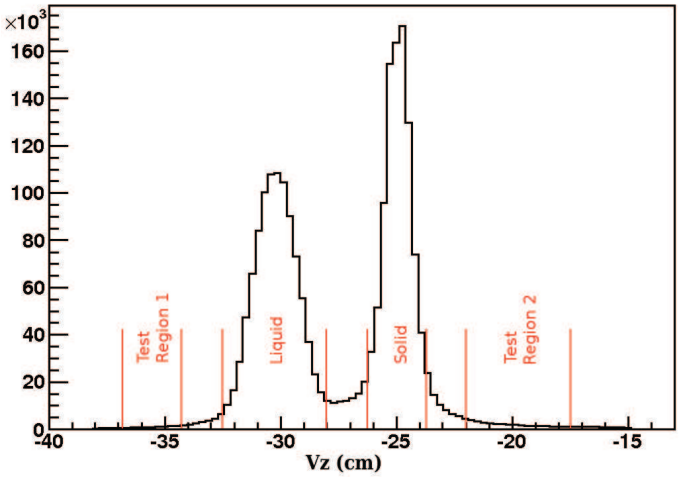

Figure 6.25: Distribution of the vertex positions along z (cm), i.e. along the beam line. In red are shown the cuts used to evaluate the leak from one target to another.

6.5.3 Acceptance

We apply two different acceptance corrections to the data using two different binnings (table 6.4). The differences between them give us a good indication on the systematic error associated with the method. Using a sample of results from iron, we find the systematic presented in the table 6.6. We note the significantly larger errors for π^-, which was expected because of the smaller acceptance and the larger weights. More significant are the large errors on the $\Delta\langle p_\perp^2\rangle$ measurements, this is mainly due to the nature of the observable; as a difference $\Delta\langle p_\perp^2\rangle$ enhances relative errors significantly. As was seen in figure 6.19, error bars are much larger for corrected $\Delta\langle p_\perp^2\rangle$ indicating an important statistical sensitivity introduced by the weighting. The errors on $\Delta\langle p_\perp^2\rangle$ are of the same order, or smaller, as the differences observed between the two sets. This is an indication that the uncertainties reported in table 6.6 are already taken into account in the statistical error.

Variable	Normalization errors	Point to point errors
$R_A^{\pi^+}$	1.2%	1.5%
$R_A^{\pi^-}$	2.5%	2.6%
$\langle \Delta P_\perp^2 \rangle_A^{\pi^+}$	5%	11%
$\langle \Delta P_\perp^2 \rangle_A^{\pi^-}$	5%	21%

Table 6.6: Relative errors on observables between the two weighting described in the text.

6.5.4 Normalization Error

The normalization errors – i.e. independent of kinematic variables – are due to acceptance effects, target identification and isospin correction. They are summarized in the table 6.7.

	$R_A^{\pi^+}$	$R_A^{\pi^-}$	$\Delta\langle p_\perp^2 \rangle^{\pi^+}$	$\Delta\langle p_\perp^2 \rangle^{\pi^-}$
Acceptance	1.2%	2.5%	5%	5%
Target id.	1%	1%	1%	1%
Isospin (Pb)	1%	1%	1%	1%
Total	1.9%	2.9%	5.2%	5.2%

Table 6.7: Normalization uncertainties.

Chapter 7

Results and Discussions

The results presented in this chapter are corrected for acceptance effects, but not for radiative ones. We use the following DIS selection cuts: $Q^2 > 1$ GeV^2/c^2, $W > 2$ GeV and $y < 0.85$. In order to ensure that factorization applies, we select $0.4 < z < 0.7$[1], this choice is based on experimental results from [Asaturyan 2011]. Moreover, this cut also ensures that we measure the leading hadron and that we avoid the high z region, which might be contaminated by the diffractive ρ^0 decay products. Finally, the cut $x_F > 0$ select the current fragmentation region.

7.1 Multiplicity Ratio

7.1.1 A Dependence

The figure 7.1 presents our result for the $A^{1/3}$ dependence of the multiplicity ratio. One can see a 5% normalization difference between pions. However, this difference is not significant as it represents only 1.5 standard deviation of our normalization uncertainty (see table 6.7).

The attenuation, presented in the figure 7.1, is not linear as a function of $A^{1/3}$ nor $A^{2/3}$. HERMES data has already showed some indication of this feature [Airapetian 2007, Airapetian 2010], but here, the nonlinearity, which has important implication on models, is clear. Indeed, it seems difficult to conciliate the prehadron absorption models with this result. Prehadrons are expected to have their cross section increasing with time and, therefore, distance. On the parton energy loss side, the BDMPS calculation gives a parton energy loss proportional to $L^2 \propto A^{2/3}$ also in contradiction with this result. However, this statement does not hold at low energy and might also be modified if the production time occurs inside the nuclei.

7.1.2 Cronin Effect

The Cronin effect is characterized by a large increase of the multiplicity ratio at high $p_\perp^2$ ($\sim$1 GeV^2/c^2), but it is a controversial measurement in hadronization studies. Indeed, whereas SLAC measurements [Osborne 1978] did not

[1] This cut is not applied for plots as a function of z.

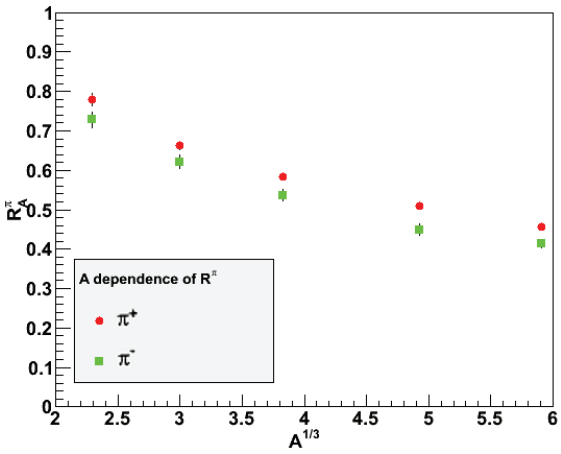

Figure 7.1: $A^{1/3}$ dependence of the multiplicity ratio. Normalization uncertainties are not shown.

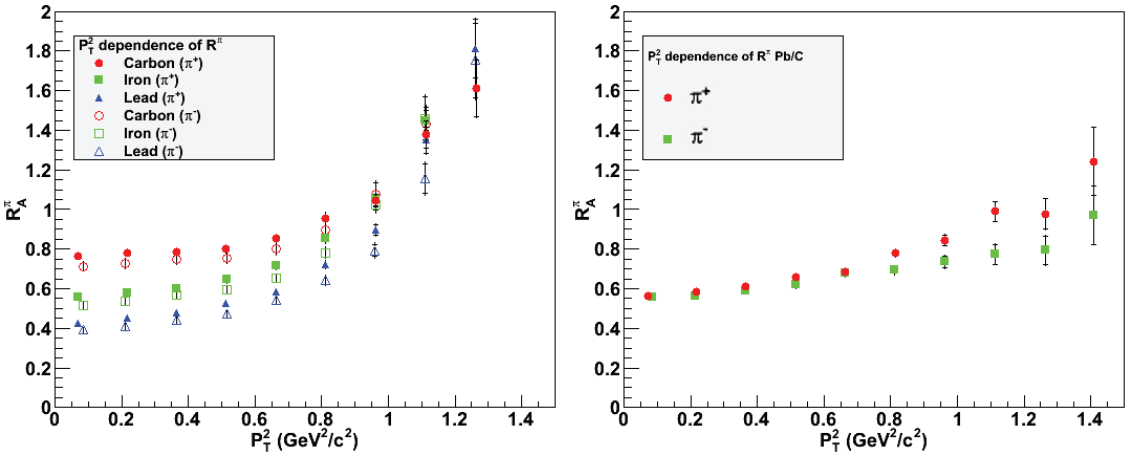

Figure 7.2: Multiplicity ratios as a function of $p_\perp^2$ (GeV^2/c^2) for charged pions. Left panel is the usual observable, right panel shows multiplicity ratio of lead normalized to carbon. Normalization uncertainties are not shown.

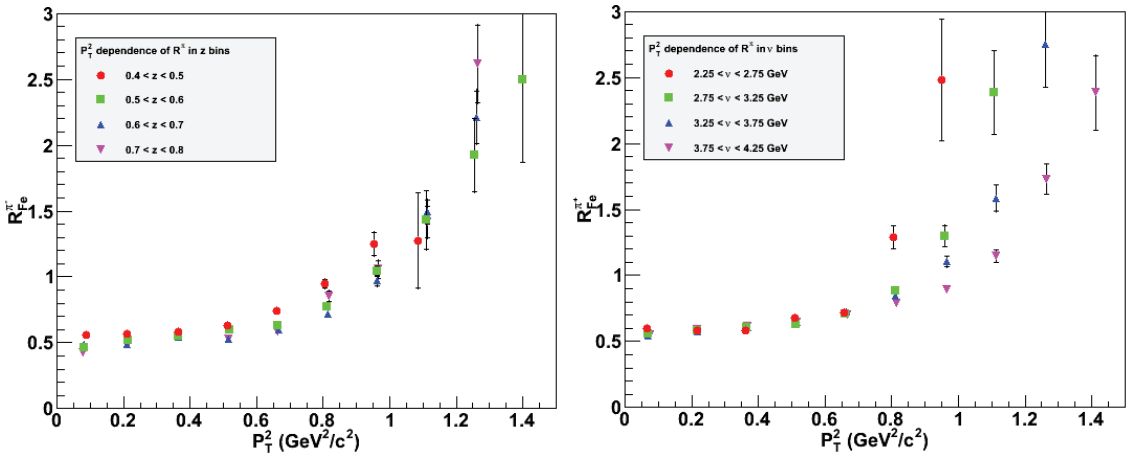

Figure 7.3: Multiplicity ratios as a function of $p_\perp^2$ (GeV^2/c^2) and z (left) or ν (GeV) (right) for positive pions. Normalization uncertainties are not shown.

show such an effect, HERMES [Airapetian 2007] measured a significant increase of the multiplicity ratio with $p_\perp^2$. Our result (figure 7.2 (left)), integrated over all other variables, shows a pattern similar to the HERMES measurement. However, there is potential contributions from the target fragmentation (see chapter 3) and the Fermi motion (see chapter 4).

Fermi motion effects can be significantly reduced by modifying our usual observables. Indeed, using carbon for normalization, instead of deuterium, cancels most of the effects of Fermi motion. Moreover, acceptance effects (section 6.4.1) are mostly canceled as well in such a ratio, reducing various systematic errors (including most of the normalization error). The figure 7.2 (right) represents the multiplicity ratio based on carbon. We observe an attenuation, at low $p_\perp^2$, similar to the usual multiplicity ratio of iron. This can be understood from the radius of the nuclei $R_{Pb} - R_C \sim R_{Fe}$. Nevertheless, the observed enhancement with $p_\perp^2$ is much more modest than in iron. At first sight the difference can be attributed to Fermi motion, which might affect the measurement based on deuterium. The target fragmentation effects might also lead to similar observation and more evidence are needed to confirm our interpretation.

In the figure 7.3, the multi-dimensionally binned multiplicity ratio are presented. HERMES found an important dependence of the Cronin effect with z [Airapetian 2007, Airapetian 2011] that we interpreted as a sign of an important target fragmentation contribution. The present result (figure 7.3 (left)) does not show this behavior, but it can be explained by our stricter cut on z ($z > 0.4$ whereas HERMES uses $z > 0.2$), which leads to a smaller target fragmentation contamination. The second result, binned in $p_\perp^2$ and ν (figure 7.3 (right)), shows an important dependence of the Cronin effect with ν. However, HERMES results did not reveal any similar ν variation. This is an evidence of the importance of the Fermi motion effect in our measurement. Indeed, in the chapter 4, it was shown that the impact of Fermi motion gets smaller at higher energy.

The result for the lead to carbon multiplicity ratio, figure 7.2 (right), has another interesting feature: the π^+s have a stronger Cronin effect. However, the signification of this observation is not clear. This could be a contribution from target fragmentation, as we might expect more positive particles than negative ones coming from the nuclei. This could also come from other sources such as isospin effect or factorization breaking at high $p_\perp^2$, but no test of these features exists in this kinematical range[2]. Last, it could be linked to the hadron rank and reflect some hadronization properties. Indeed, as we probe mostly valence quark at our energies, π^+s are more likely to be leading hadron com-

[2] [Asaturyan 2011] covers only $p_\perp^2 < 0.2$ GeV^2/c^2.

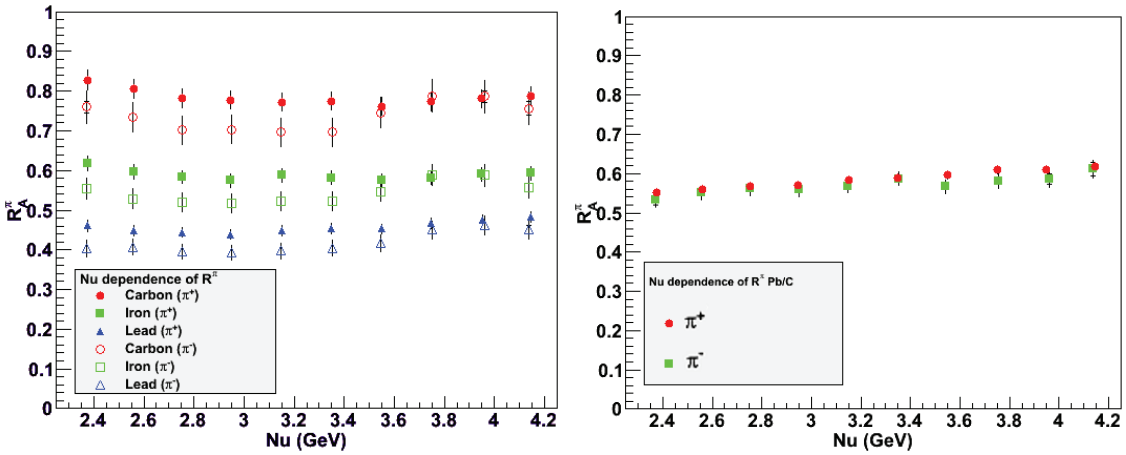

Figure 7.4: Multiplicity ratios as a function of ν (GeV). Left panel is the usual observable, right panel shows multiplicity ratio of lead normalized to carbon. Normalization uncertainties are not shown.

pared to π^-s.

In conclusion, our results indicate that the HERMES results were affected by target fragmentation and that our results are affected by Fermi motion. These effects can be easily controlled by selecting carefully our observables and kinematic coverage.

7.1.3 ν Dependence

The HERMES collaboration clearly observed a rise of the multiplicity ratio between ν of 4 and 22 GeV. However, at first sight, in our results (figure 7.4 (left)), no dependence is observed, except for a slight increase in lead. The results on Fermi motion together with the figure 7.4 (right) permit to reach a coherent explanation. The Fermi motion seems to cancel the hadronization effect almost completely and the systematic uncertainties linked with acceptance might wash out what remains of the effect. The multiplicity ratio based on carbon, however, offers a very neat slope consistent with what was observed on a much larger range by HERMES.

We also notice here that both pions give similar results in the figure 7.4 (right). This confirms our previous remark that the difference observed in the multiplicity ratio based on deuterium might come from the normalization uncertainty caused by the acceptance correction. Therefore, there is no clear difference between the charged pions for the integrated multiplicity ratios.

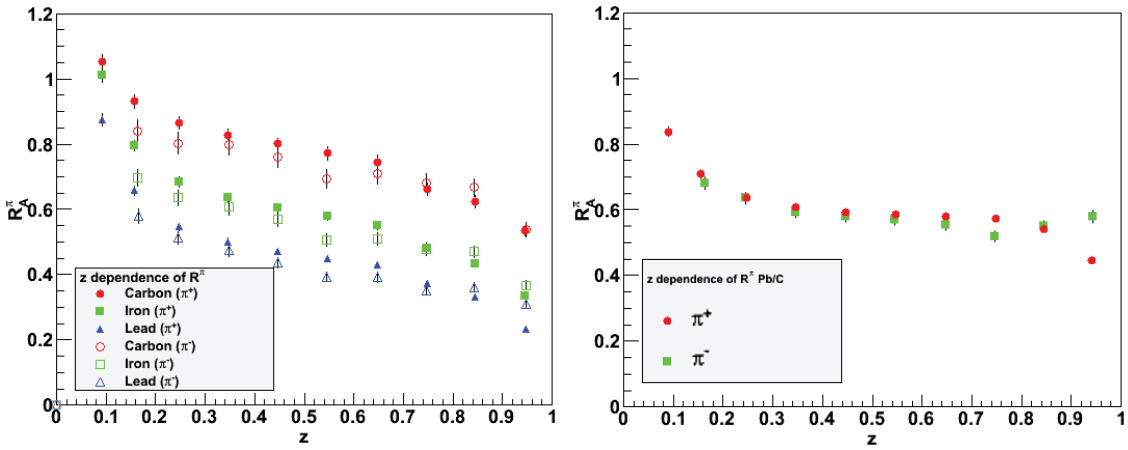

Figure 7.5: Multiplicity ratios as a function of z. Left panel is the usual observable, right panel shows multiplicity ratio of lead normalized to carbon. Normalization uncertainties are not shown.

7.1.4 z Dependence

The multiplicity ratio was observed to decrease with z in HERMES data, whereas this feature was not as significant in other experiments. Indeed the nature of this behavior is questionable, the target fragmentation also reduce with z and can mimic the effect. In figure 7.5 (left), as in HERMES, we see a clear slope even at values higher than 0.4, where target fragmentation effects are expected to be small. However, the ratio lead to carbon (figure 7.5 (right)) shows a much flatter behavior in the region of interest (0.4 to 0.7). This is coherent with what we showed in chapter 4, i.e. the expected z slope is enhanced by Fermi motion effects. Therefore, the situation seems to be similar to the Cronin effect one, where HERMES observation was enhanced by target fragmentation. The measurement using the carbon as basis is, therefore, more useful to isolate effects from the hadronization, but low z behavior remains driven by the target region. Another strange feature of the data is the behavior at high z (figure 7.5 (right)), where the two pions behave differently. However, we lack of solid theoretical grounds in this region to interpret this result.

7.1.5 Q^2 Dependence

The behavior of hadronization as a function of Q^2 is an important issue, that has direct implications on our understanding of nuclear matter properties in QCD. HERMES results, which covers $1 < Q^2 < 10$ GeV^2/c^2, give a hint for an increase of the multiplicity ratio with Q^2. Our result, in figure 7.6, does not allow to reach the same conclusion and indicates no significant dependence as a function of Q^2.

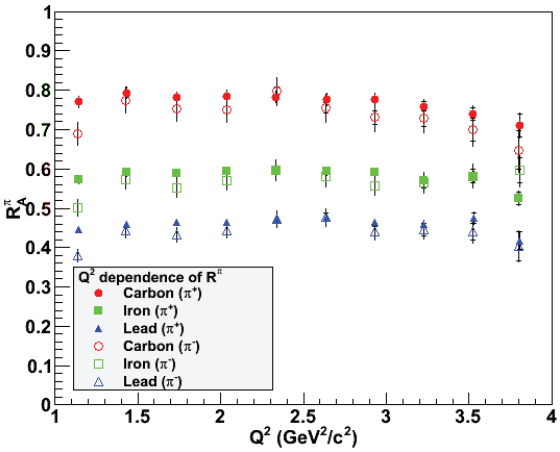

Figure 7.6: Multiplicity ratios as a function of Q^2 (GeV^2/c^2). Normalization uncertainties are not shown.

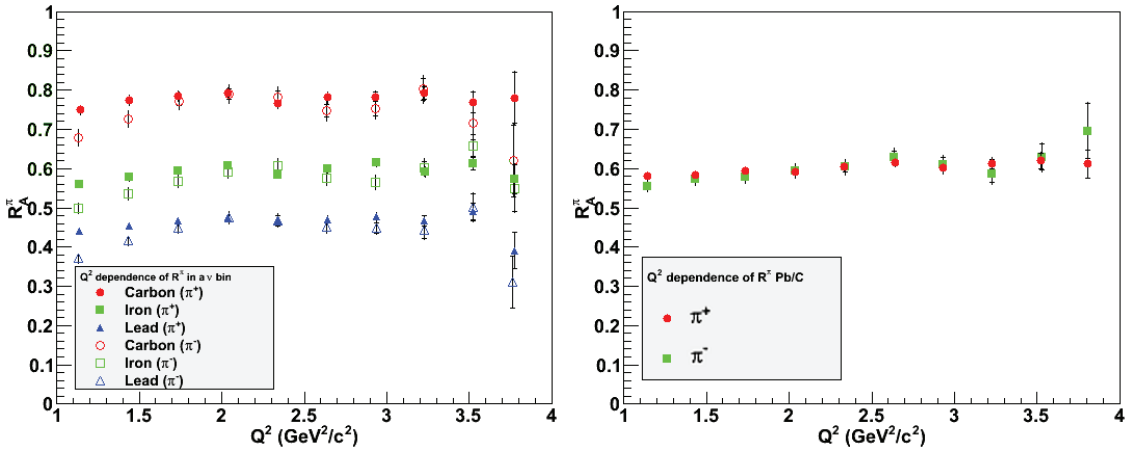

Figure 7.7: Multiplicity ratios as a function of Q^2 (GeV^2/c^2). In the left panel, the multiplicity ratio in a tight ν bin ($3.25 < \nu < 3.75$ GeV). In the right panel, the multiplicity ratio of lead normalized to carbon. Normalization uncertainties are not shown.

We can use our large statistics to get more information. As a dependence in ν is expected for the multiplicity ratio, it might be helpful to use a tight ν bin to plot the Q^2 dependence (figure 7.7 (left)), and remove any coupling between the two variables. We could also expect some effects from the Fermi motion, therefore, we show in figure 7.7 (right) the multiplicity ratio of lead to carbon as a function of Q^2. The two results of figure 7.7 show a slight increase with Q^2, but as for HERMES, no evidence is reached on this question. Our leverage on Q^2 appears to be too modest to make a clear measurement. Future programs to explore this question are discussed in chapter 8.

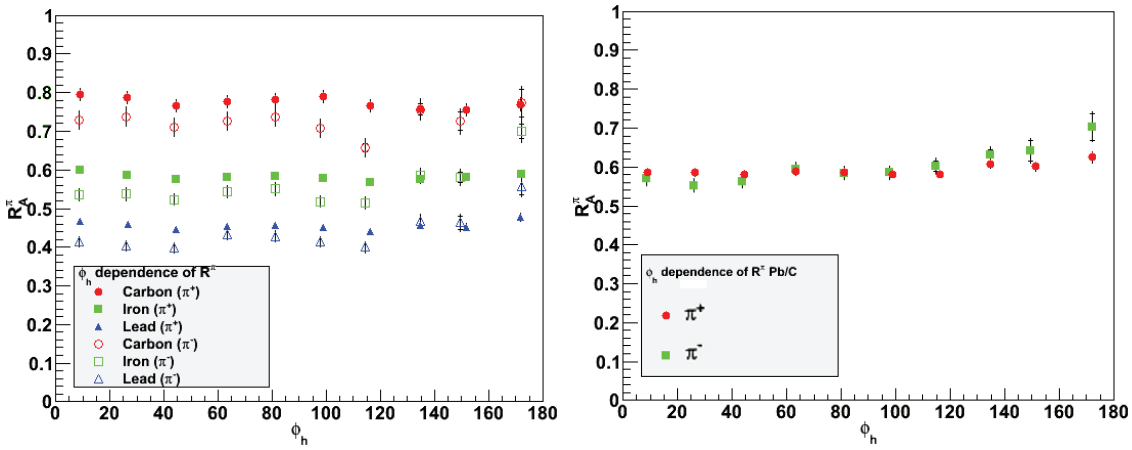

Figure 7.8: Multiplicity ratios as a function of ϕ_π. Normalization uncertainties are not shown.

7.1.6 ϕ_h dependence

The interest on the multiplicity ratio as a function of ϕ_h was pointed out by the EMC collaboration in [Ashman 1991]. They showed that a non-flat result could indicate a modification of the quark transverse momentum distribution in nuclei. However, EMC did not observe such an effect and our results in figure 7.8 do not show any clear variation either. Still some strange behavior is observed at large ϕ_h and could be explored in more depth. However, the complex acceptance in this region might affect this result.

7.2 Transverse Momentum Broadening

7.2.1 A Dependence

The A dependence of the transverse momentum broadening, presented in figure 7.9, is an important result. The large statistics coupled with the large coverage in A available to the CLAS experiment give precise indication on $\Delta\langle p_\perp^2\rangle$ as a function of $A^{1/3}$. Compared to HERMES [Airapetian 2010], we find smaller $\Delta\langle p_\perp^2\rangle$. This is coherent with theory, which predicts larger effects at larger energies. However, calculations of parton energy loss from BDMPS [Baier 1997] show that $p_\perp^2$ should depend on the nucleus radii (see chapter 2). However, this feature does not appear in our result. Of course, because of the low energy of the CLAS experiment, one might discuss the possibility to make a direct comparison with BDMPS calculations. Nonetheless, this result shows an unexpected pattern that remains to be explained. One possibility is that, as proposed for the multiplicity ratio as a function of $A^{1/3}$, the production time occurs inside the nuclei. In this case, the colored parton would not inter-

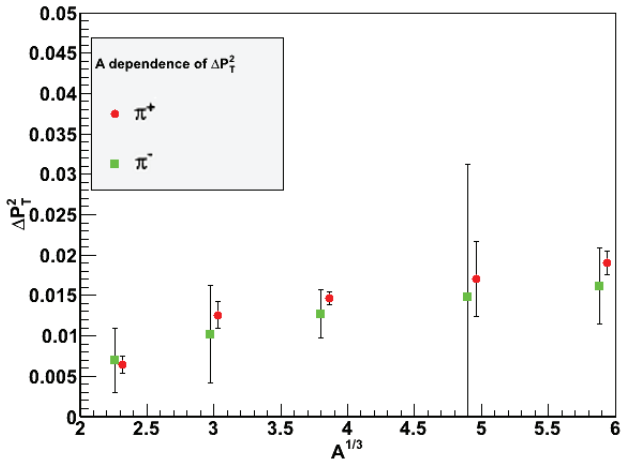

Figure 7.9: A dependence of $\Delta\langle p_\perp^2\rangle$. Points are slightly offset for readability, normalization uncertainties are not shown.

act with the whole nuclei and therefore changing its size would lead to limited effect.

7.2.2 ν Dependence

Interestingly, HERMES did not found any ν dependence of $\Delta\langle p_\perp^2\rangle$. Our results (figure 7.10) are very different as we see a clear raise with ν, but it is unclear why such a behavior appears. The high ν part of our $\Delta\langle p_\perp^2\rangle$ results seems to be flatter and might lead to a plateau or a smaller dependence as a function of ν. This might connect smoothly with the HERMES measurement and show a change of slope at few GeV. This result is surprising, yet theoretical input [Domdey 2009] have explicit dependence with ν. However, no calculation was made for CLAS energy.

7.2.3 z Dependence

The z dependence of $\Delta\langle p_\perp^2\rangle$ has some importance for parton energy loss models, because it is linked to the assumptions needed to extrapolate the hadronic $p_\perp^2$ to the partonic one. Moreover, usually, in parton energy loss models, $\Delta\langle p_\perp^2\rangle$ goes to zero at the highest z. Our results (figure 7.11) do not show such a pattern. However, our previous results showed the importance of Fermi motion in our kinematics, which is expected to enhance $\Delta\langle p_\perp^2\rangle$ at very high z. This is confirmed by the flattening observed for the carbon based $\Delta\langle p_\perp^2\rangle$, but the distribution is not clearly going down either at high z. This might be due to the size of the error bars and the potential contamination from diffractive ρ^0 production, both preventing us from any definitive statement on this feature of the data.

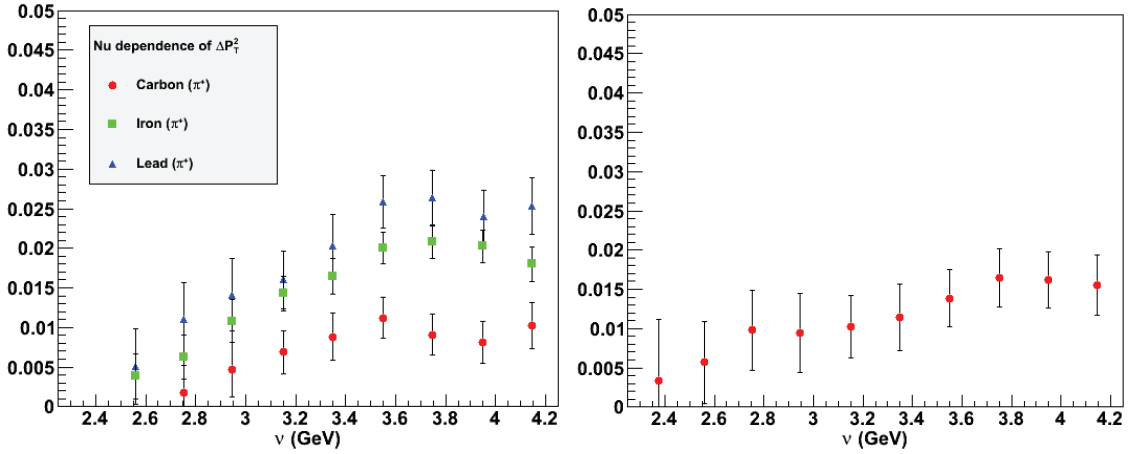

Figure 7.10: $\Delta\langle p_\perp^2 \rangle$ as a function of ν (GeV). Left panel is the usual observable. Right panel shows multiplicity ratio of lead normalized to carbon. Normalization uncertainties are not shown.

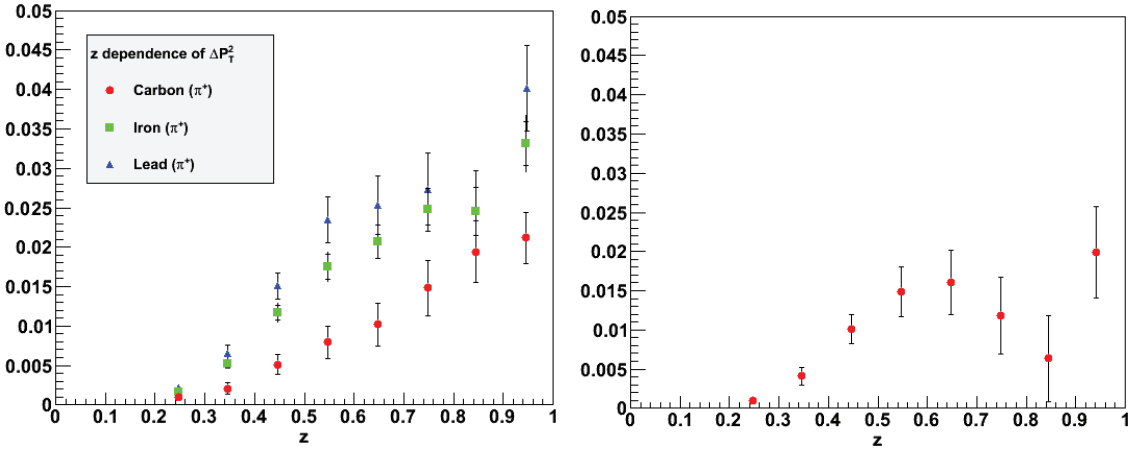

Figure 7.11: $\Delta\langle p_\perp^2 \rangle$ as a function of z. Left panel is the usual observable, right panel shows multiplicity ratio of lead normalized to carbon. Normalization uncertainties are not shown.

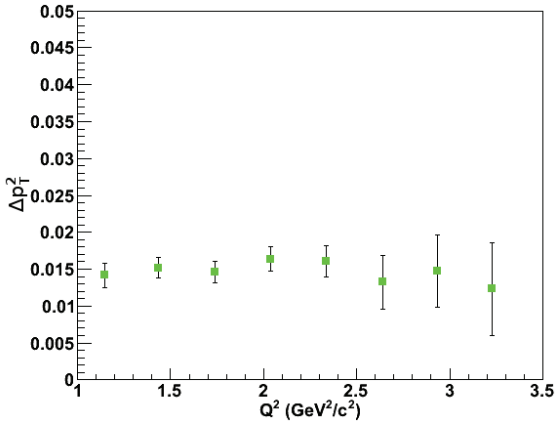

Figure 7.12: $\Delta\langle p_\perp^2 \rangle$ for iron as a function of Q^2 (GeV^2/c^2) with usual cuts. Normalization uncertainties are not shown.

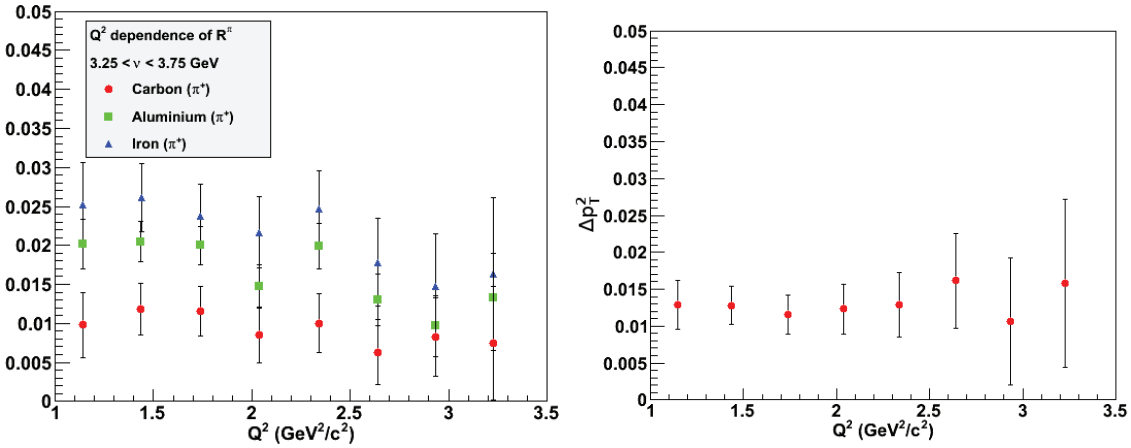

Figure 7.13: $\Delta\langle p_\perp^2 \rangle$ as a function of Q^2 (GeV^2/c^2). Left panel is the usual observable in a tight ν bin ($3.25 < \nu < 3.75$ GeV). Right panel shows $\Delta\langle p_\perp^2 \rangle$ of lead relative to carbon. Normalization uncertainties are not shown.

7.2.4 Q^2 Dependence

Finally, the Q^2 dependence of $\Delta\langle p_\perp^2 \rangle$ is an important result for the BDMPS based calculation from [Domdey 2009]. They expect a raise of $\Delta\langle p_\perp^2 \rangle$ with Q^2, which is not observed in the figure 7.12. Using binning in ν and carbon based $\Delta\langle p_\perp^2 \rangle$ (figure 7.13) gives a similar result. In conclusion, within error bars, no effect is observed for $\Delta\langle p_\perp^2 \rangle$ as a function of Q^2. However, as theoretical input is missing, it is not clear if we have the resolution to observe the effect expected within BDMPS based models.

Chapter 8

Future Experiments

8.1 Introduction

The past results, reviewed in chapter 3, and the CLAS results, presented in chapter 7, give us a coherent, but still incomplete, description of hadronization. In this final chapter, two proposed programs are presented. The first one, already approved to run by the Program Advisory Committee (PAC) of JLab, is using CLAS12, the upgrade of CLAS, in direct prolongation with the CLAS hadronization experiment presented in this thesis. The other one is based on the Electron Ion Collider (EIC) project, developed in the last few years, which is focusing on high energies. This setup, by pushing the prehadron production far outside the nuclei, allows a pure pQCD treatment of hadronization in nuclear medium. In the context of the EIC other programs, this is an original and independent way to access important nuclear properties, in particular, the gluons in nuclear matter and their saturation at low x_{Bj} (see chapter 2 for the associated theory).

8.2 The CLAS12 Experiment

The JLab accelerator, presented in section 5.1, will be upgraded in the next few years. After this upgrade, a 11 GeV beam will be available in the Hall B to be used with the CLAS12 detector. The whole project is planned to be completed mid 2015. Figure 8.1 shows the CLAS12 detector layout. However, the current project will not allow a good kaon identification above 2.5 GeV/c momenta. As a remedy, a Ring Imaging CHerenkov (RICH) detector is under consideration to replace the low threshold Cherenkov counter. The RICH detector will permit to separate properly protons, kaons and pions in the forward region for momenta up to 8 GeV/c.

The hadronization experiment using CLAS12 was initially proposed in 2006 [Hafidi 2006] and was updated in 2010 [Brooks 2010] for rating and beam time attribution by the PAC. It appeared in the initial proposal, based on the initial CLAS12 project, that the technical issue of particle identification was underestimated. As kaons and anti-protons hadronization are major questions, the

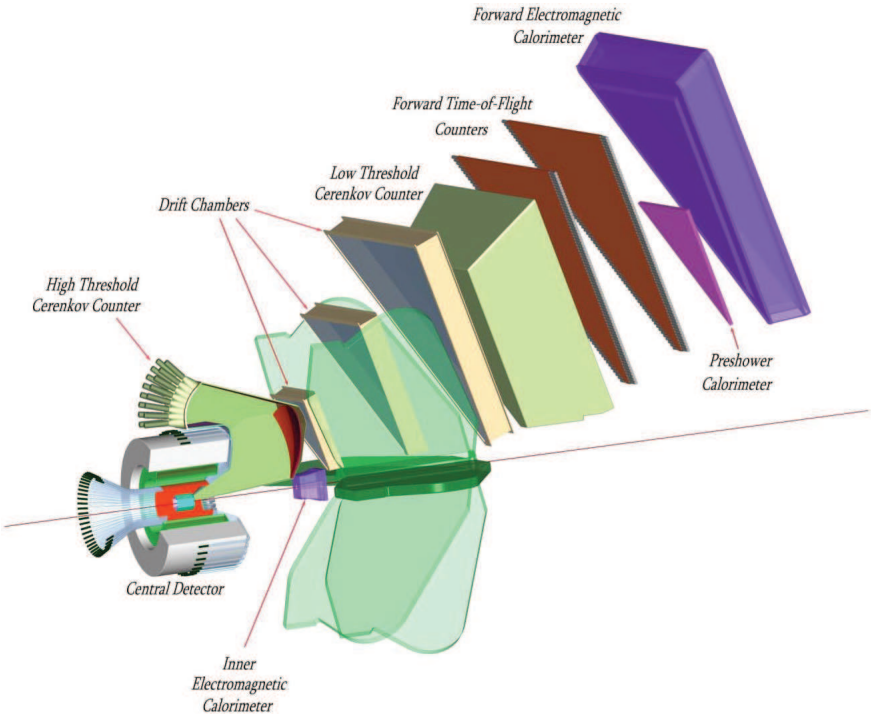

Figure 8.1: Exploded view of one sector of the CLAS12 spectrometer.

RICH detector seems to be necessary for the completeness of the experiment [Schoch 2006, Brooks 2010].

The authors propose to run on 5 different targets in order to map properly the A dependence. This is globally similar to the proposal for the CLAS experiment [Brooks 2002]. The results presented in section 7.1.1 show clearly the importance of the A dependence and are justifying the request for a wide set of nuclei.

There are many possibilities to better understand hadronization and nuclei using CLAS12. First, the $p_\perp^2$ broadening should be a more effective tool than it was in CLAS or HERMES. Indeed, the high energy, the high statistics and the access to several flavors will ease the extraction of the quark $p_\perp^2$ broadening, i.e. $\hat{q}$, from $\Delta\langle p_\perp^2\rangle$. Second, CLAS12 will have a major role to play in the understanding of the nature of the time scales involved in the in-medium hadronization. Using the larger energy coverage and statistics, the production time measurement will be greatly improved and eventually the information on formation time and prehadron characteristics might be accessible. Third, a multi-hadron measurement, similar to the one from HERMES [Airapetian 2006], was difficult to carry on at CLAS energy because the low energy lead to low multiplicity. Using CLAS12, the measurement should be possible; the high luminosity should even allow to go at higher z_2 than HERMES. This measurement is important to understand the puzzling results from HERMES with this observ-

able. Fourth, to access directly information on the hard struck parton, we can measure photons directly emitted by the quark in the medium. This is an usual observable in heavy ion collisions, but so far none were observed in nuclear DIS. However, quarks, as charged particles, have to emit bremsstrahlung photons in the nuclear media and these should be experimentally accessible. The spectrum of these photons would give direct information on the initial quark dynamics. Finally, an important part of the experiment plan is to explore the hadronization of various baryons (p, $\bar{p}$, Λ, Σ and Ξ). The goal is to look into production and formation time of baryons, because the information on baryon hadronization is even scarcer than for mesons, this input will be very valuable.

Even more topics can be explored, either by extending the proposed measurement or just by exploring different observables. For example, tagging strangeness in multi-hadron final states would extend the two pions measurement and could contribute to solve the puzzle of the K^+ multiplicity ratio. Another possibility, in the analysis of both mesons and baryons, is to try to find a recombination sector as in heavy ion collisions [Hwa 2010]. The signature for such an effect would be a plateau or a peak of the multiplicity ratio as a function of $p_\perp^2$. Baryons can also be detected in the target region in order to make analysis similar to the one from FNAL [Adams 1995] and see how theoretical description, like [Ciofi degli Atti 2005], can be scaled to much lower energy.

In conclusion, this experiment can make a great use of the very high luminosity and the high energy available at JLab12. With the addition of a RICH detector to CLAS12, it would be the ideal medium energy experiment to study hadronization dynamics. Using the many observables and particles accessible, it will provide many new information and a very stringent tests for the different descriptions of in-medium hadronization.

8.3 The Electron Ion Collider

8.3.1 Hadronization at EIC

A collaboration of nuclear physicists, mainly in the United-States, promotes the construction of an Electron Ion Collider as the next major facility for research on QCD [Aidala 2007]. Two projects are being developed, one at the JLab [Thomas 2009] (see figure 8.2), the other at the Brookhaven National Laboratory (BNL) [Otboev 2006] (see figure 8.3). These two projects have technical differences, but have similar goals in term of energy and luminosity. Their development is ongoing and the design parameters reevaluated regularly. In this section the energy, luminosity and other technical issues will not be discussed directly. The basic parameters, energy and luminosity, of both designs will be used here and the specificities attached to each project will be

ignored. The goal of this work is to find the main interests for hadronization studies at such a facility and, also, to raise the experimental challenges these measurements would face.

There is a growing interest for pure energy loss models, because of the opportunity to use them to learn about the properties of the nuclear media. Calculations of parton energy loss in QCD matter are now numerous and most of them show an underlying link between the gluonic content of the medium and the parton energy loss. These calculations are often made to describe heavy ion collisions but they are of particular interest in the e-A configuration. Indeed, because of the better knowledge of the initial kinematic and the reduced initial state interaction, it is possible to compare precisely experimental results with calculations and deduce specific information about gluons densities in nuclei [Baier 1997, Kopeliovich 2010a]. However, the theoretical uncertainties remain important for the recent low energy experiments, where significant prehadron absorption is possible and pQCD applicability is questionable. Results from the EMC collaboration [Ashman 1991] show that attenuation becomes small around energy transfer $\nu \sim 100$ GeV for tin. However, EMC results are not precise enough to allow a quantitative interpretation, but they can be used to determine an optimum energy for future high energy experiments.

By going at higher energies, we reach an energy level comparable with RHIC and allow for a better comparison between hot and cold nuclear matter. It also opens the possibility to measure hadronization observables on heavy quarks and jets. Heavy quarks permit to explore the mass dependence of quark energy loss, and generally of all hadronization characteristics, like time evolution, adding a new dimension to the measurement. Jets can give access to the usual hadronization variables, but independently of fragmentation functions. They are also the chance to access new observables typical of jets, like the jet radius, jet number or hadron multiplicity in the jet.

In the projections presented in this section, the error bars are statistical only and for a configuration of EIC, either $s = 200$ GeV^2 or $s = 1000$ GeV^2. The integrated luminosity used is 200 fb^{-1} per target. Using lead and deuterium target at the best configuration available in each EIC design [Accardi 2010a], this would correspond to a run of roughly 2 years per energy setting. We consider only data in the DIS regime ($Q^2 > 1$ GeV^2/c^2 and $W > 4$ GeV) where we can assume a single quark is hit ($x_{Bj} > 0.1$), and ensure that we are detecting the leading hadron in the factorization region ($0.4 < z < 0.8$). For experimental reasons, such as electron detection and radiative effects, we also apply a cut on the y coverage ($0.1 < y < 0.85$). The acceptance is assumed to be 50% for light mesons and 2% for heavy mesons (D and B) to account for

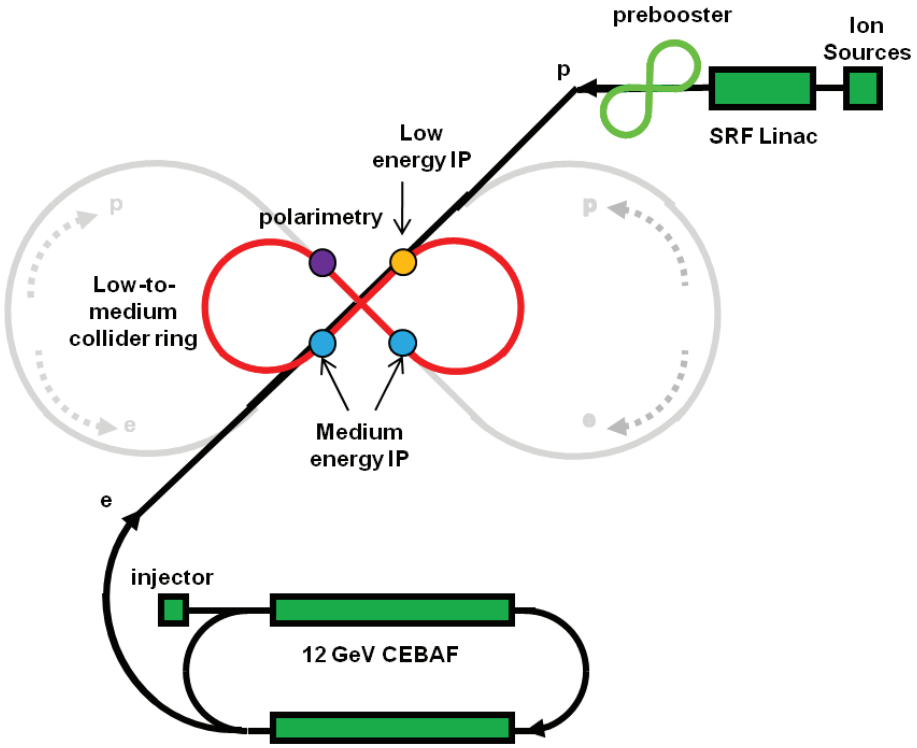

Figure 8.2: Design layout of the medium (Red) and full (Grey) energy projects for an EIC at JLab.

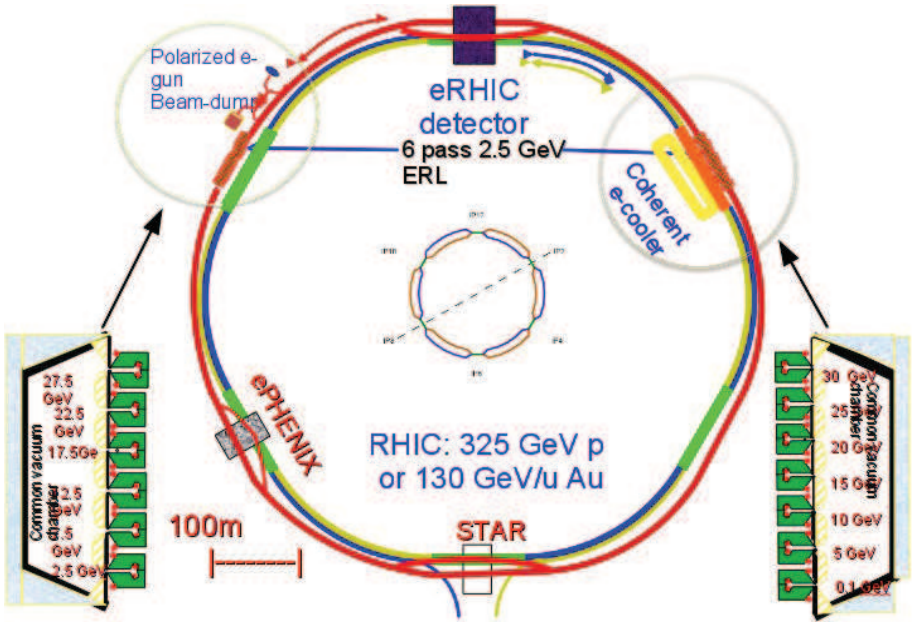

Figure 8.3: Design layout of the eRHIC project for an EIC at BNL.

their much more challenging detection. This global configuration is reasonably achievable by both EIC designs, indeed even higher energies might be available. However, it does not seem useful to push the energy parameter because this would be at the expense of the luminosity, which is a critical parameter for heavy quarks measurements and Q^2 leverage.

8.3.2 Parton Energy Loss

As highlighted several times in this thesis, the theory of parton energy loss is still progressing and needs experimental input. EIC is the perfect tool for a precise measurement of quark energy loss and transverse momentum broadening. It is important to point out that measuring directly the energy loss is challenging, because at high ν, the relative effect on the quark momentum gets smaller. The transverse momentum broadening observable allows to get around this problem, as very small effects can be resolved with this observable. Besides, the transverse momentum induced by energy loss has direct theoretical interpretations.

The energy loss effects can be isolated for $\nu \gtrsim 100$ GeV allowing a clean measurement by limiting the impact of hadron absorption. However, we expect a good control on models by the time the EIC will be running. Therefore, it should also be possible to exploit the lower ν data where the nuclear effects are relatively large.

The multiplicity ratio, as a function of z, will help to verify precisely the assumptions made at lower energy and, eventually, to make direct evaluation of the quark energy loss (figure 8.4). Moreover, measurements of different hadron flavors at moderate ν (figure 8.5) would give access to information on hadron absorption for hadrons heavier than the pions in a region much more safer in term of target fragmentation contamination. At high ν, $\Delta\langle p_\perp^2 \rangle$ of pions is an almost direct measurement of $\hat{q}$, and using dependencies in ν and z permits to make model independent extraction. This topic is considered one of the golden measurement for EIC [Boer 2011] because of the possibility to use it to access the saturation scale [Kopeliovich 2010a] independently from other measurements.

The Q^2 dependence of $\Delta\langle p_\perp^2 \rangle$ is also of interest to check the scaling between different flavors predicted, within pQCD, by [Domdey 2009] (figure 8.6). It can also be used to detect any modification of the Q^2 evolution of fragmentation functions (figure 8.7 (left)). To this end, the multiplicity ratio can also be used in low ν bins (figure 8.7 (right)). In this case the low energy part of the EIC program ($s = 200$ GeV^2) seems particularly interesting. The Q^2 dependence of in-medium hadronization is an open question and EIC offers unprecedented Q^2 coverage for this measurement!

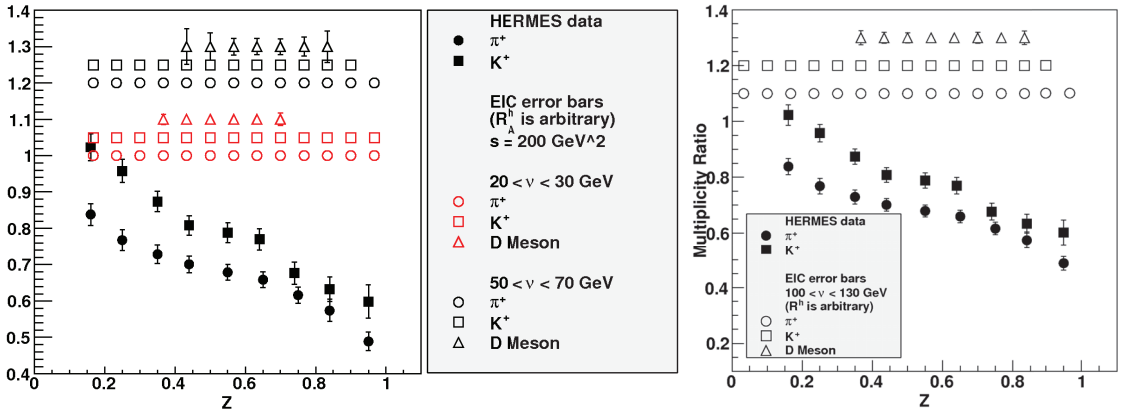

Figure 8.4: Projections for multiplicity ratio errors at EIC in various ν bins for pions, kaons and D mesons (R_A^h is arbitrary) compared with HERMES results [Airapetian 2007].

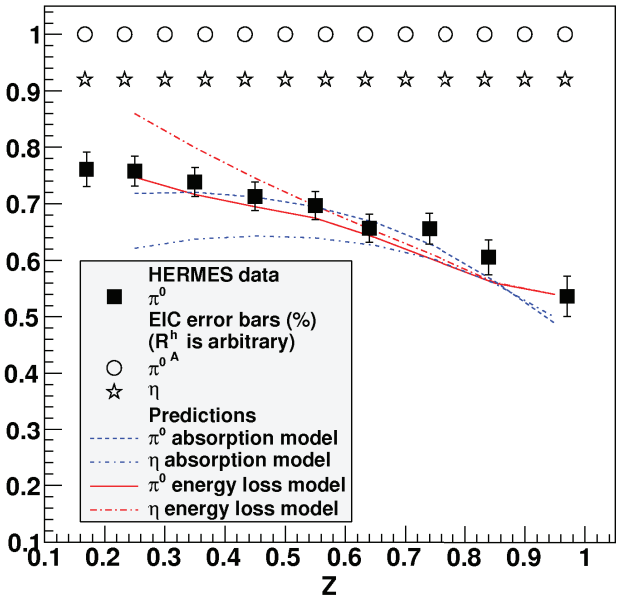

Figure 8.5: Results from HERMES [Airapetian 2007] for π^0 multiplicity ratio compared with predictions based on an absorption model and an energy loss model (both from [Accardi 2009a]). Circles and stars are the projections for the statistical errors at EIC ($20 < \nu < 30$ GeV).

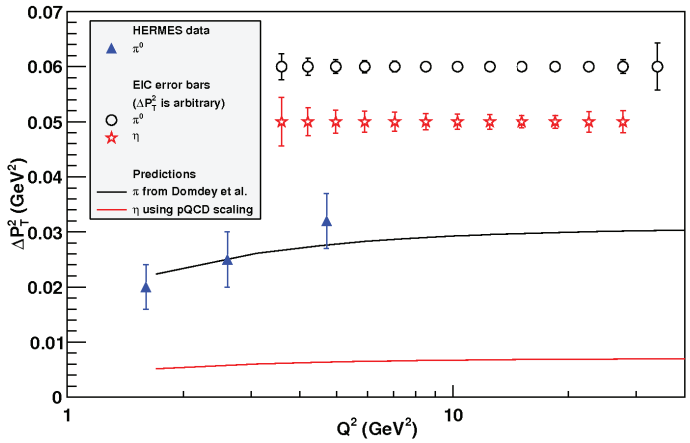

Figure 8.6: Preliminary results from HERMES for π^0 transverse momentum broadening compared with prediction based on a pQCD calculation (pion from [Domdey 2009] and η from [Accardi 2009a]). Circles and stars are the projections for the statistical errors at EIC ($20 < \nu < 30$ GeV).

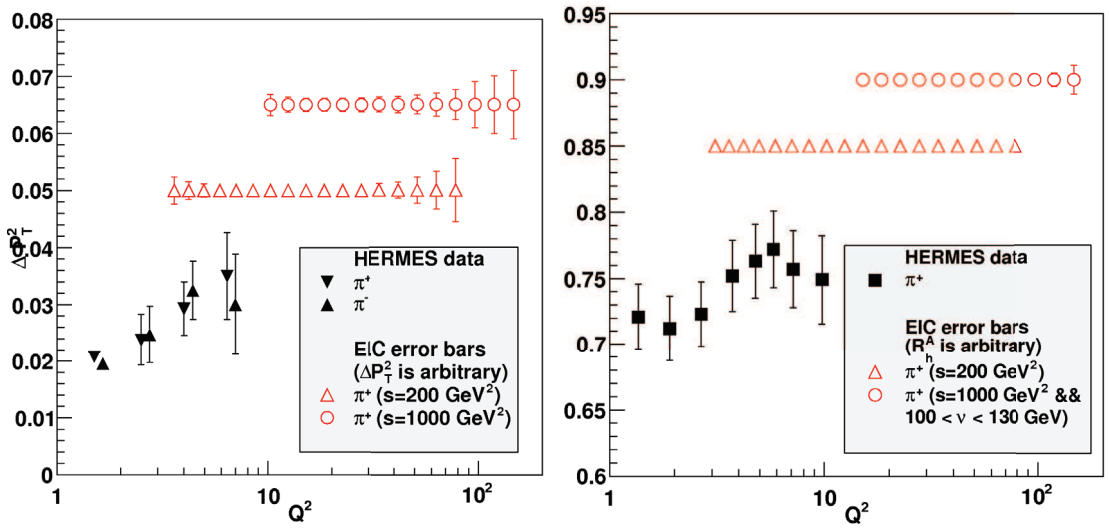

Figure 8.7: Projections of errors on transverse momentum (left) and multiplicity ratio (right) of pions achievable at EIC (constrained in a ν bin for ratio at high s). Compared with HERMES results [Airapetian 2007, Airapetian 2010].

8.3.3 Heavy Quarks

The possibility to measure heavy quarks at EIC is completely new for the e-A configuration giving access to the effect of the parton mass on hadronization. For the measurement of heavy mesons, the key points are the detector vertex resolution, which needs to be of the order of few tens of micrometer, and the luminosity, which needs to be at least 10^{33} $cm^{-2}s^{-1}$ for D mesons and, ideally, few times 10^{34} $cm^{-2}s^{-1}$ for B mesons. The observables are then similar to the ones for light quarks and allow to explore the hadronization time scale and the quark energy loss. The scaling of both production time and energy loss with the mass of quarks is an important question that can be directly confronted to theory. This measurement could, indeed, be used to reveal non perturbative effects in the QCD energy loss [Horowitz 2008, Kopeliovich 2010b]. Many measurements are possible to explore heavy quarks propagation and hadronization with an EIC, figures 8.4 and 8.8 illustrate some possibilities.

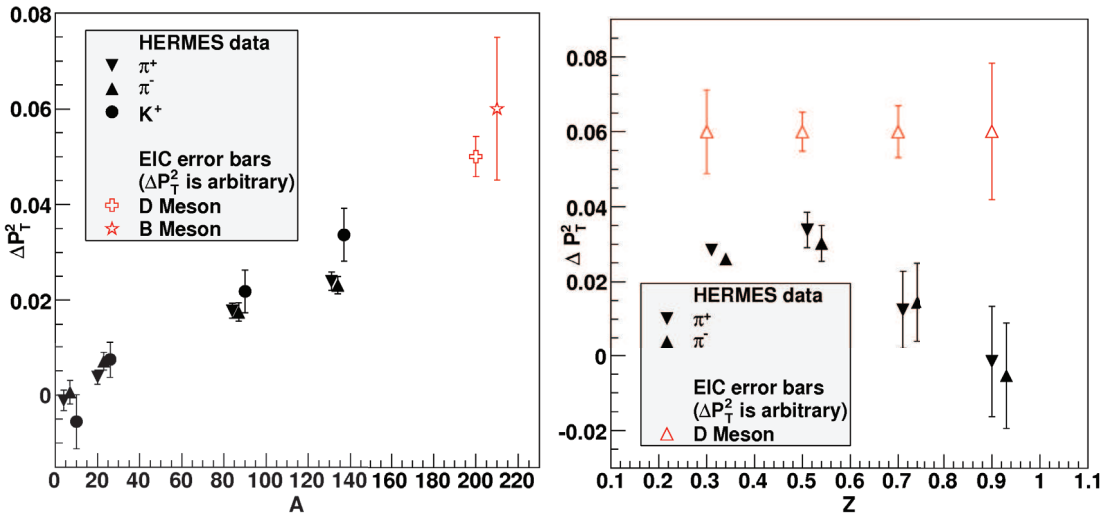

Figure 8.8: Projections for the transverse momentum errors at EIC for heavy flavors, compared with HERMES results for pions and kaons [Airapetian 2007, Airapetian 2010].

8.3.4 Conclusion

To complete the discussion on the EIC, it is important to mention jets, because the high energy available is the chance to measure them in e-A. Jets offer many opportunities; they are a new and independent way to access transport coefficient $\hat{q}$ and to confirm the other measurements mentioned for an EIC. One can also use the modification of the jet radius to explore the modification of fragmentation functions in nuclear medium. The modification of the number of jet per DIS events in nuclei can give information about hard gluon

emissions. Unfortunately, studies of jets at an EIC are still at their beginning and no quantitative evaluation of these measurements exists yet. Nevertheless, jets observables give a unique access to medium modified fragmentation functions and eventually medium modified DGLAP evolution. They would definitely be a major part of the scientific program of an EIC.

In conclusion, the possibility offered by an EIC to go at higher energy, while keeping a high statistics, is very promising for hadronization studies. Indeed, many important measurements can be achieved on the parton energy loss in nuclear medium, for both light and heavy quarks. These would help improve our understanding of the nuclear medium, the pQCD and the effects of QCD medium on DGLAP evolution. For an optimal study of hadronization at the EIC, the main concerns are the luminosity, the possibility to run lower energy than the standard setting and the detector capabilities for the vertex reconstruction.

Conclusion

To date, understanding the low energy component of the QCD theory remains a major challenge. One of the key processes to understand this regime of QCD is hadronization. However, the possibility to study hadronization in vacuum is linked with its understanding in nuclear medium. The harnessing of hadronization in nuclear matter has also other interests for nuclear physics; it would provide a new tool to study the nuclear matter itself in normal or under extreme conditions. To reach these goals, an important theoretical question has to be answered: during in-medium hadronization, are the interactions with the nuclear matter of partonic or hadronic nature?

After 35 years of measurements, we can draw a picture of the in-medium hadronization in which the forward hadrons are attenuated and their energy redistributed in the target fragmentation region. HERMES provided important measurements to establish it and, in particular, their $\Delta\langle p_\perp^2\rangle$ results are believed to give a direct insight on the parton energy loss. However, the measured multiplicity ratios, as a function of ν, Q^2, z and $p_\perp^2$, can still be described within many different theoretical assumptions and the question of the nature of the interactions driving hadronization in nuclei remains pending.

The results obtained with the PyQM event generator, confirmed most of the interpretation of the published data on $\Delta\langle p_\perp^2\rangle$, indicating that this variable can be linked to the partonic transverse momentum broadening. For multiplicity ratios, the results are not as decisive. However, by reproducing some behavior of HERMES data, we showed that their results are compatible with a pure parton energy loss model. Finally, the simulation appears to be of important help to understand the data as it allowed to emphasize the impact of Fermi motion on CLAS measurements.

In the analysis of the CLAS data, the acceptance was found to be the main source of correction and, subsequently, of systematic uncertainties. The contamination of the pion samples by kaons and the radiative effects were found to have only a small impact on our measurements. The use of carbon as a baseline for our observables was found to reduce the effect of Fermi motion and allowed to provide a cleaner signal with more valuable input for models. Indeed, the slope of the multiplicity ratios, as a function of z and ν in particular, are a key measurement that models need to reproduce properly. Many other results are also of interests. The Cronin effect was found to be of reduced size after being cleared from the target fragmentation and Fermi motion contamination. Also, we did not observe any effect with Q^2, giving some upper limit for such an effect, which is expected in models based on medium

modified fragmentation functions or parton energy loss.

Our most integrated results, as a function of $A^{1/3}$, might be the most interesting. Indeed, we noticed a saturation of the nuclear effects with $A^{1/3}$, which is very challenging to describe within existing theoretical models. The saturation of the nuclear attenuation with nuclear size indicates that attenuation is mainly caused in the first few femtometers. This is difficult to conciliate with (pre)hadron absorption models, because, in these models, the hadronic cross-section grows with time and leads to an important attenuation at longer distances. The results for $\Delta\langle p_\perp^2\rangle$ are similar and, therefore, in apparent contradiction with the BDMPS calculations, indicating that something is also wrong with the pure energy loss picture. A solution, to conciliate the $A^{1/3}$ dependence observed with the quark energy loss models, might be to consider a production of the prehadron, with reduced interactions, inside the nuclei. To conclude on the results shown as a function of $A^{1/3}$, we notice a similar behavior for both R_A^π and $\Delta\langle p_\perp^2\rangle$, which is consistent with a similar origin for both observations.

Two complementary experiments, in the future, should give a more detailed picture of the hadronization process in nuclei. The CLAS12 hadronization experiment will give an important insight on the dynamics of hadronization. Indeed, based on the present results, the relatively low energy appears as an advantage in order to explore this dynamics. On the other side, EIC, by pushing the energy in a completely new territory, will give a chance to access parton energy loss. There, the parton energy loss can be safely calculated within pQCD, making possible the precise measurements of nuclear properties using hadronization observables. Moreover, the measurement of heavy quarks and jets will give access to new tests for the high energy processes in QCD.

Bibliography

[Accardi 2003] A. Accardi, V. Muccifora and H.-J. Pirner. *Hadron production in deep inelastic lepton nucleus scattering*. Nucl.Phys., vol. A720, pages 131–156, 2003. (Cited on pages 26, 30, 33, 34, 37 and 77.)

[Accardi 2007] A. Accardi. *Quenching weights in DIS*. Working note, not published, 2007. (Cited on pages 63 and 68.)

[Accardi 2009a] A. Accardi. Private communication, 2009. (Cited on pages 133 and 134.)

[Accardi 2009b] A. Accardi, T. Hobbs and W. Melnitchouk. *Hadron mass corrections in semi-inclusive deep inelastic scattering*. JHEP, vol. 0911, page 084, 2009. (Cited on page 17.)

[Accardi 2010a] A. Accardi. *EIC Luminosity*. https://eic.jlab.org/wiki/index.php/EIC_luminosity, 2010. (Cited on page 131.)

[Accardi 2010b] A. Accardi, F. Arleo, W. K. Brooks, D. D'Enterria and V. Muccifora. *Parton Propagation and Fragmentation in QCD Matter*. Riv. Nuovo Cim., vol. 32, pages 439–553, 2010. (Cited on page 19.)

[Adams 1995] M. Adams *et al.* *Nuclear shadowing, diffractive scattering and low momentum protons in μXe interactions at 490-GeV*. Z.Phys., vol. C65, pages 225–244, 1995. (Cited on pages 46, 47, 48 and 129.)

[Adams 2001] G. Adams, V. Burkert, R. Carl, T. Carstens, V. Frolov *et al.* *The CLAS Cherenkov detector*. Nucl.Instrum.Meth., vol. A465, pages 414–427, 2001. (Cited on pages 84 and 87.)

[Adil 2007] A. Adil and I. Vitev. *Collisional dissociation of heavy mesons in dense QCD matter*. Phys.Lett., vol. B649, pages 139–146, 2007. (Cited on page 19.)

[Aidala 2007] C. Aidala *et al.* *A High Luminosity, High Energy Electron-Ion-Collider*. White Paper Prepared for the NSAC LRP 2007, 2007. (Cited on page 129.)

[Airapetian 2001] A. Airapetian *et al.* *Hadron formation in deep inelastic positron scattering in a nuclear environment*. Eur.Phys.J., vol. C20, pages 479–486, 2001. (Cited on page 49.)

[Airapetian 2003] A. Airapetian *et al.* *Quark fragmentation to $\pi^{\pm}$, π^{0}, $K^{\pm}$, p and anti-p in the nuclear environment*. Phys.Lett., vol. B577, pages 37–46, 2003. (Cited on pages 49 and 52.)

[Airapetian 2006] A. Airapetian *et al.* *Double hadron leptoproduction in the nuclear medium*. Phys.Rev.Lett., vol. 96, page 162301, 2006. (Cited on pages 34, 49, 59, 61 and 128.)

[Airapetian 2007] A. Airapetian *et al.* *Hadronization in semi-inclusive deep-inelastic scattering on nuclei*. Nucl.Phys., vol. B780, pages 1–27, 2007. (Cited on pages 29, 33, 34, 36, 49, 50, 51, 53, 54, 55, 56, 64, 77, 78, 79, 80, 97, 117, 119, 132, 133, 134 and 135.)

[Airapetian 2010] A. Airapetian *et al.* *Transverse momentum broadening of hadrons produced in semi-inclusive deep-inelastic scattering on nuclei*. Phys.Lett., vol. B684, pages 114–118, 2010. (Cited on pages 49, 57, 58, 60, 117, 123, 134 and 135.)

[Airapetian 2011] A. Airapetian *et al. Multidimensional Study of Hadronization in Nuclei*. Eur.Phys.J., vol. A47, page 113, 2011. (Cited on pages 49 and 119.)

[Akushevich 1998] I. Akushevich, H. Bottcher and D. Ryckbosch. *RADGEN 1.0: Monte Carlo generator for radiative events in DIS on polarized and unpolarized targets*. 1998. (Cited on page 110.)

[Akushevich 2001] I. Akushevich, A. Ilyichev and N. Shumeiko. *Radiative effects in scattering of polarized leptons by polarized nucleons and light nuclei*. 2001. (Cited on page 108.)

[Akushevich 2009] I. Akushevich, A. Ilyichev and M. Osipenko. *Complete lowest order radiative corrections to five-fold differential cross-section of hadron leptoproduction*. Phys.Lett., vol. B672, pages 35–44, 2009. (Cited on page 110.)

[Albino 2008] S. Albino, B. A. Kniehl and G. Kramer. *AKK Update: Improvements from New Theoretical Input and Experimental Data*. Nucl. Phys., vol. B803, pages 42–104, 2008. (Cited on page 19.)

[Altarelli 1977] G. Altarelli and G. Parisi. *Asymptotic Freedom in Parton Language*. Nucl.Phys., vol. B126, page 298, 1977. (Cited on page 17.)

[Amarian 2001] M. Amarian, G. Asrian, K. Beard, W. Brooks, V. Burkert *et al.* *The CLAS forward electromagnetic calorimeter*. Nucl.Instrum.Meth., vol. A460, pages 239–265, 2001. (Cited on pages 84 and 88.)

[Andersson 1983] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand. *Parton Fragmentation and String Dynamics*. Phys.Rept., vol. 97, pages 31–145, 1983. (Cited on pages 19, 26 and 70.)

[Andreopoulos 2010] C. Andreopoulos *et al. The GENIE Neutrino Monte Carlo Generator*. Nucl. Instrum. Meth., vol. A614, pages 87–104, 2010. (Cited on page 110.)

[Arleo 2002] F. Arleo. *Tomography of cold and hot QCD matter: Tools and diagnosis. ((V))*. JHEP, vol. 11, page 044, 2002. (Cited on page 63.)

[Arleo 2003a] F. Arleo. *Parton energy loss or nuclear absorption: What quenches hadron spectra at HERA?* pages 558–563, 2003. (Cited on page 50.)

[Arleo 2003b] F. Arleo. *Quenching of hadron spectra in DIS on nuclear targets*. Eur. Phys. J., vol. C30, pages 213–221, 2003. (Cited on pages 37 and 63.)

[Arvidson 1984] A. Arvidson *et al. Hadron production in 200 GeV μ-copper and μ-carbon deep inelastic interactions*. Nucl.Phys., vol. B246, page 381, 1984. (Cited on pages 42, 43, 48 and 50.)

[Asaturyan 2011] 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*. 2011. (Cited on pages 59, 61, 64, 112, 117 and 119.)

[Ashman 1991] J. Ashman *et al. Comparison of forward hadrons produced in muon interactions on nuclear targets and deuterium*. Z.Phys., vol. C52, pages 1–12, 1991. (Cited on pages 42, 44, 45, 46, 123 and 131.)

[Aznauryan 2011] I. Aznauryan, V. Burkert, T.-S. Lee and V. Mokeev. *Results from the N* program at JLab*. 2011. (Cited on page 17.)

[Baier 1997] R. Baier, Y. L. Dokshitzer, A. H. Mueller, S. Peigne and D. Schiff. *Radiative energy loss and p(T)-broadening of high energy partons in nuclei*. Nucl. Phys., vol. B484, pages 265–282, 1997. (Cited on pages 31, 32, 65, 123 and 131.)

[Bialas 1983] A. Bialas and T. Chmaj. *Leptoproduction of hadrons from nuclear targets and fragmentation of quarks into hadrons*. Phys.Lett., vol. B133, page 241, 1983. (Cited on pages 30 and 36.)

[Bialas 1987] A. Bialas and M. Gyulassy. *Lund model and an outside - inside aspect of the inside - outside cascade*. Nucl.Phys., vol. B291, page 793, 1987. (Cited on pages 30 and 36.)

[Bleve 2001] C. Bleve, G. Co, I. De Mitri, P. Bernardini, G. Mancarella *et al. Effects of nuclear reinteractions in quasielastic neutrino nucleus scattering*. Astropart.Phys., vol. 16, pages 145–155, 2001. (Cited on page 22.)

[Boer 2011] D. Boer, M. Diehl, R. Milner, R. Venugopalan, W. Vogelsang *et al. Gluons and the quark sea at high energies: Distributions, polarization, tomography*. 2011. (Cited on page 132.)

[Brock 1995] R. Brock *et al. Handbook of perturbative QCD: Version 1.0*. Rev.Mod.Phys., vol. 67, pages 157–248, 1995. (Cited on page 17.)

[Brodsky 1988] S. J. Brodsky and A. H. Mueller. *Using Nuclei to Probe Hadronization in QCD*. Phys.Lett., vol. B206, page 685, 1988. (Cited on pages 21 and 28.)

[Brodsky 1993] S. J. Brodsky and P. Hoyer. *A Bound on the energy loss of partons in nuclei*. Phys.Lett., vol. B298, pages 165–170, 1993. (Cited on pages 31, 32 and 57.)

[Brooks 2002] W. Brooks *et al. Quark Propagation Through Cold QCD Matter*. A proposal to PAC 22, 2002. (Cited on pages 62, 91 and 128.)

[Brooks 2009] W. Brooks and H. Hakobyan. *Experimental Studies of Hadronization and Parton Propagation in the Space-Time Domain*. Nucl.Phys., vol. A830, pages 361C–368C, 2009. (Cited on page 62.)

[Brooks 2010] W. K. Brooks *et al. Quark Propagation and Hadron Formation*. 10 pages update to PAC 36, 2010. (Cited on page 127.)

[Cahn 1978] R. N. Cahn. *Azimuthal Dependence in Leptoproduction: A Simple Parton Model Calculation*. Phys.Lett., vol. B78, page 269, 1978. (Cited on page 45.)

[Ciofi degli Atti 1996] C. Ciofi degli Atti and S. Simula. *Realistic model of the nucleon spectral function in few and many nucleon systems*. Phys.Rev., vol. C53, page 1689, 1996. (Cited on page 65.)

[Ciofi degli Atti 2005] C. Ciofi degli Atti and B. Kopeliovich. *Time evolution of hadronization and grey tracks in DIS off nuclei*. Phys.Lett., vol. B606, pages 281–287, 2005. (Cited on pages 37, 48, 69 and 129.)

[Clasie 2007] B. Clasie *et al. Measurement of Nuclear Transparency for the A(e,e' π^+) Reaction*. Phys. Rev. Lett., vol. 99, page 242502, 2007. (Cited on page 28.)

[de Florian 2007a] D. de Florian, R. Sassot and M. Stratmann. *Global analysis of fragmentation functions for pions and kaons and their uncertainties*. Phys.Rev., vol. D75, page 114010, 2007. (Cited on page 19.)

[de Florian 2007b] D. de Florian, R. Sassot and M. Stratmann. *Global analysis of fragmentation functions for protons and charged hadrons*. Phys.Rev., vol. D76, page 074033, 2007. (Cited on pages 19 and 20.)

[De Vries 1987] H. De Vries, C. De Jager and C. De Vries. *Nuclear charge and magnetization density distribution parameters from elastic electron scattering*. Atom.Data Nucl.Data Tabl., vol. 36, pages 495–536, 1987. (Cited on page 68.)

[Deng 2010] W.-T. Deng and X.-N. Wang. *Multiple Parton Scattering in Nuclei: Modified DGLAP Evolution for Fragmentation Functions*. Phys.Rev., vol. C81, page 024902, 2010. (Cited on pages 33 and 37.)

[Deng 2011] W.-T. Deng, N.-B. Chang and X.-N. Wang. *Modified DGLAP Evolution for Fragmentation Functions in Nuclei and QGP*. Nucl.Phys., vol. A855, pages 416–419, 2011. (Cited on pages 34 and 37.)

[Dokshitzer 1977] Y. L. Dokshitzer. *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.* Sov.Phys.JETP, vol. 46, pages 641–653, 1977. (Cited on page 17.)

[Dokshitzer 1991] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller and S. Troian. *Basics of perturbative QCD*. 1991. (Cited on page 19.)

[Domdey 2009] S. Domdey, D. Grunewald, B. Kopeliovich and H. Pirner. *Transverse Momentum Broadening in Semi-inclusive DIS on Nuclei*. Nucl.Phys., vol. A825, pages 200–211, 2009. (Cited on pages 37, 57, 74, 124, 125, 133 and 134.)

[El Fassi 2008] L. El Fassi *et al. Search for the Onset of Color Transparency via ρ^0 Electroproduction off Nuclei*. Internal Analysis Note, 2008. (Cited on page 92.)

[El Fassi 2012] L. El Fassi *et al. Evidence for the Onset of Color Transparency in ρ^0 Electroproduction off Nuclei*. To be published, 2012. (Cited on page 28.)

[Falter 2004] T. Falter, W. Cassing, K. Gallmeister and U. Mosel. *Hadron attenuation in deep inelastic lepton-nucleus scattering*. Phys.Rev., vol. C70, page 054609, 2004. (Cited on pages 26 and 37.)

[Fries 2003] R. Fries, B. Muller, C. Nonaka and S. Bass. *Hadron production in heavy ion collisions: Fragmentation and recombination from a dense parton phase*. Phys.Rev., vol. C68, page 044902, 2003. (Cited on page 53.)

[Gallmeister 2005] K. Gallmeister and T. Falter. *Space-time picture of fragmentation in PYTHIA / JETSET for HERMES and RHIC*. Phys. Lett., vol. B630, pages 40–48, 2005. (Cited on pages 27, 37, 40 and 41.)

[Gallmeister 2008] K. Gallmeister and U. Mosel. *Time Dependent Hadronization via HERMES and EMC Data Consistency*. Nucl.Phys., vol. A801, pages 68–79, 2008. (Cited on pages 26, 28, 29, 37 and 41.)

[Gallmeister 2010] K. Gallmeister. *eA collisions within GiBUU*. Talk at INT-10-3 workshop, 2010. (Cited on page 57.)

[Garfield 2010] Garfield. *http://garfield.web.cern.ch/garfield/*, 1984-2010. (Cited on page 85.)

[Geesaman 1995] D. F. Geesaman, K. Saito and A. W. Thomas. *The nuclear EMC effect*. Ann.Rev.Nucl.Part.Sci., vol. 45, pages 337–390, 1995. (Cited on pages 16 and 22.)

[Gribov 1972] V. Gribov and L. Lipatov. *Deep inelastic e-p scattering in perturbation theory*. Sov.J.Nucl.Phys., vol. 15, pages 438–450, 1972. (Cited on page 17.)

[Gyulassy 1990] M. Gyulassy and M. Plumer. *Jet quenching in lepton nucleus scattering*. Nucl.Phys., vol. B346, pages 1–16, 1990. (Cited on pages 26, 34, 35 and 37.)

[Gyulassy 2000] M. Gyulassy, P. Levai and I. Vitev. *NonAbelian energy loss at finite opacity*. Phys.Rev.Lett., vol. 85, pages 5535–5538, 2000. (Cited on page 66.)

[Hafidi 2002] K. Hafidi *et al.* *Q^2 Dependence of Nuclear Transparency for Incoherent ρ^0 Electroproduction*. A proposal to PAC 22, 2002. (Cited on page 91.)

[Hafidi 2006] K. Hafidi *et al.* *Quark Propagation and Hadron Formation*. A proposal to PAC 30, 2006. (Cited on page 127.)

[Hakobyan 2008] H. Hakobyan, W. Brooks, K. Bruhwel, V. Burkert, T. Carstens *et al.* *A double-target system for precision measurements of nuclear medium effects*. Nucl.Instrum.Meth., vol. A592, pages 218–223, 2008. (Cited on pages 91 and 100.)

[Hand 1979] L. Hand, D. Petersen, H. Scott, K. Gottfried, M. Atac *et al.* *Intranuclear cascading in deep inelastic scattering of 150-GeV muons in emulsion*. Z.Phys., vol. C1, pages 139–142, 1979. (Cited on page 46.)

[Hicks 2009] K. Hicks and A. Daniel. *K(s)0 Hadronization Following DIS at CLAS*. 2009. (Cited on page 62.)

[Hirai 2007] M. Hirai, S. Kumano, T. H. Nagai and K. Sudoh. *Determination of fragmentation functions and their uncertainties*. Phys. Rev., vol. D75, page 094009, 2007. (Cited on page 19.)

[Horowitz 2008] W. Horowitz and M. Gyulassy. *Testing AdS/CFT Drag and pQCD Heavy Quark Energy Loss*. J.Phys.G, vol. G35, page 104152, 2008. (Cited on page 133.)

[Hwa 2002] R. C. Hwa and C. Yang. *Strangeness enhancement in the parton model*. Phys.Rev., vol. C66, page 064903, 2002. (Cited on pages 33 and 50.)

[Hwa 2004] R. C. Hwa and C. Yang. *Final state interaction as the origin of the Cronin effect*. Phys.Rev.Lett., vol. 93, page 082302, 2004. (Cited on page 53.)

[Hwa 2010] R. Hwa. *Recombination in nuclear collisions*. PoS, vol. CERP2010, page 004, 2010. (Cited on page 129.)

[Ito 2009] T. Ito, W. Bentz, I. Cloet, A. Thomas and K. Yazaki. *The NJL-jet model for quark fragmentation functions*. Phys.Rev., vol. D80, page 074008, 2009. (Cited on page 19.)

[Johnson 2001] M. Johnson *et al.* *Energy loss of fast quarks in nuclei*. Phys.Rev.Lett., vol. 86, pages 4483–4487, 2001. (Cited on page 31.)

[Kaskulov 2009] M. M. Kaskulov, K. Gallmeister and U. Mosel. *Pionic transparency in semi-exclusive electroproduction off nuclei*. Phys. Rev., vol. C79, page 015207, 2009. (Cited on page 28.)

[Konig 1982] A. Konig and P. Kroll. *A realistic calculation of the azimuthal asymmetry in semiinclusive deep inelastic scattering*. Z.Phys., vol. C16, page 89, 1982. (Cited on page 45.)

[Kopeliovich 2004] B. Kopeliovich, J. Nemchik, E. Predazzi and A. Hayashigaki. *Nuclear hadronization: Within or without?* Nucl.Phys., vol. A740, pages 211–245, 2004. (Cited on pages 30, 32, 36, 37, 48, 50, 52, 54 and 57.)

[Kopeliovich 2007] B. Kopeliovich, J. Nemchik and I. Schmidt. *Color transparency in electroproduction of the ρ meson at low energies*. Phys.Rev., vol. C76, page 015205, 2007. (Cited on page 28.)

[Kopeliovich 2010a] B. Kopeliovich, I. Potashnikova and I. Schmidt. *Measuring the saturation scale in nuclei*. Phys.Rev., vol. C81, page 035204, 2010. (Cited on pages 131 and 132.)

[Kopeliovich 2010b] B. Kopeliovich, I. Potashnikova and I. Schmidt. *Why heavy and light quarks radiate energy with similar rates*. Phys.Rev., vol. C82, page 037901, 2010. (Cited on page 133.)

[Lafferty 1995] G. Lafferty, P. Reeves and M. Whalley. *A Compilation of inclusive particle production data in e+ e- annihilation*. J.Phys.G, vol. G21, pages A1–A151, 1995. (Cited on page 19.)

[Liebing 2004] P. Liebing. *Can the gluon polarization in the nucleon be extracted from HERMES data on single high-p(T) hadrons?* 2004. (Cited on page 64.)

[Manly 2006] S. Manly. *Recent electron scattering results from Jefferson Laboratory*. NUFACT 2006 presentation, 2006. (Cited on page 22.)

[Mardor 1998] I. Mardor, S. Durrant, J. Aclander, J. Alster, D. Barton *et al. Nuclear transparency in large momentum transfer quasielastic scattering*. Phys.Rev.Lett., vol. 81, pages 5085–5088, 1998. (Cited on page 28.)

[Mecking 2003] B. Mecking *et al. The CEBAF Large Acceptance Spectrometer (CLAS)*. Nucl.Instrum.Meth., vol. A503, pages 513–553, 2003. (Cited on pages 83 and 114.)

[Mestayer 2000] M. Mestayer, D. Carman, B. Asavapibhop, F. Barbosa, P. Bonneau *et al. The CLAS drift chamber system*. Nucl.Instrum.Meth., vol. A449, pages 81–111, 2000. (Cited on pages 84 and 85.)

[Mulders 2001] P. J. Mulders. *Current fragmentation in semi-inclusive leptoproduction*. AIP Conf. Proc., vol. 588, pages 75–88, 2001. (Cited on pages 42, 43, 50 and 53.)

[Nakamura 2005] A. Nakamura and T. Saito. *QCD color interactions between two quarks*. Phys. Lett., vol. B621, pages 171–175, 2005. (Cited on page 19.)

[Nakamura 2010] K. Nakamura *et al. Review of particle physics*. J.Phys.G, vol. G37, page 075021, 2010. (Cited on page 26.)

[Osborne 1978] L. Osborne, C. Bolon, R. Lanza, D. Luckey, D. Roth *et al. Electroproduction of hadrons from nuclei*. Phys.Rev.Lett., vol. 40, page 1624, 1978. (Cited on pages 39, 40, 41, 42 and 119.)

[Otboev 2006] A. Otboev, Y. Shatunov, L. Ahrens, M. Bai, J. Beebe-Wang *et al. eRHIC, Future Electron-ion Collider at BNL*. AIP Conf.Proc., vol. 842, pages 1046–1048, 2006. (Cited on page 129.)

[Roberts 1999] C. D. Roberts. *Nonperturbative effects in QCD at finite temperature and density*. Phys.Part.Nucl., vol. 30, pages 223–257, 1999. (Cited on page 18.)

[Salgado 2002] C. A. Salgado and U. A. Wiedemann. *A Dynamical scaling law for jet tomography*. Phys.Rev.Lett., vol. 89, page 092303, 2002. (Cited on pages 32, 37, 63, 65 and 68.)

[Salgado 2003] C. A. Salgado and U. A. Wiedemann. *Calculating quenching weights*. Phys.Rev., vol. D68, page 014008, 2003. (Cited on pages 37, 63, 65, 66, 67, 69 and 76.)

[Sassot 2010] R. Sassot, M. Stratmann and P. Zurita. *Fragmentations Functions in Nuclear Media*. Phys.Rev., vol. D81, page 054001, 2010. (Cited on pages 33 and 35.)

[Schoch 2006] B. Schoch *et al. PAC 30 Recommendations*. 2006. (Cited on page 127.)

[Sjostrand 1984] T. Sjostrand. *Jet fragmentation of multiparton configurations a string framework*. Nucl.Phys., vol. B248, page 469, 1984. (Cited on page 70.)

[Sjostrand 2006] T. Sjostrand, S. Mrenna and P. Z. Skands. *PYTHIA 6.4 Physics and Manual*. JHEP, vol. 0605, page 026, 2006. (Cited on pages 26, 63, 64 and 103.)

[Smith 1999] E. Smith, T. Carstens, J. Distelbrink, M. Eckhause, H. Egiian *et al. The time-of-flight system for CLAS*. Nucl.Instrum.Meth., vol. A432, pages 265–298, 1999. (Cited on pages 84 and 86.)

[Strikman 2007] M. Strikman. *Color transparency: 33 years and still running*. 2007. (Cited on page 21.)

[Teis 1997] S. Teis, W. Cassing, M. Effenberger, A. Hombach, U. Mosel *et al. Pion production in heavy ion collisions at SIS energies*. Z.Phys., vol. A356, pages 421–435, 1997. (Cited on page 26.)

[Thomas 2009] A. W. Thomas. *An Electron-Ion Collider at Jefferson lab*. vol. arXiv:0907.4785, 2009. (Cited on page 129.)

[Wang 2002] E. Wang and X.-N. Wang. *Jet tomography of dense and nuclear matter*. Phys.Rev.Lett., vol. 89, page 162301, 2002. (Cited on pages 32, 33, 37 and 63.)

Quark Fragmentation and Hadron Formation in Nuclear Matter

Raphaël Dupré

Summary: The hadron formation is, in the framework of the quantum chromodynamics theory (QCD), a non-perturbative process; this characteristic leads to important theoretical challenges. This is why experimental measurements of fragmentation in nuclei are a necessity in order to obtain substantial progress in our understanding of the mechanisms of hadron formation. The thesis begins with the introduction of theoretical background, followed by an overview of theoretical models.

The thesis continues with the analysis of Jefferson Lab data obtained with a 5 GeV electron beam incident on various targets (^{2}H, C, Al, Fe, Sn and Pb). The reaction products are measured with the CLAS spectrometer of Hall B. The main results are: (a) a multi-dimensional analysis of the measured observables, which permits a better confrontation with theoretical models and the extraction of temporal information on fragmentation, and (b) the observation of a non linear hadronic attenuation as a function of the target's nuclear radius.

The PyQM event generator, developed to reproduce the data from the HERMES collaboration, is also presented. The results are ambivalent, the theoretical basis used does not seem to apply to the studied case, however, some characteristics of the data are reproduced allowing to understand their origin, which is sometimes unexpected. Finally, the possibilities for future experiments, at Jefferson Lab and at an Electron-Ion Collider (EIC), are explored.

Keywords: Fragmentation, hadronization, QCD, Jefferson Lab, CLAS, nuclei, Monte-Carlo, quark energy loss, Electron-Ion Collider, EIC.