978-91-7473-142-2
LUNFD6/(NFFL-7233)2011

# Identifying Charged Hadrons on the Relativistic Rise Using the ALICE TPC at LHC

Philippe Gros

Lund University

Department Of Physics

Lund, 2011

# Identifying Charged Hadrons on the Relativistic Rise Using the ALICE TPC at LHC

Philippe Gros
Department of Physics
Lund University

Thesis for the degree of Doctor of Philosophy

| | |
|---|---|
| Faculty Opponent: | John Haggerty |
| Thesis Advisor: | Anders Oskarsson |
| Deputy Supervisors: | Peter Christiansen, Oxana Smirnova |

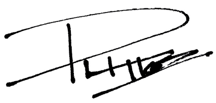
To be presented, with the permission of the Faculty of Science of Lund University, for public criticism in lecture hall A of the Department of Physics on Friday, the 10th of June 2011, at 10.15.

| | | |
|---|---|---|
| **Organisation**<br>**LUND UNIVERSITY**<br>Department of Physics<br>Lund University<br>Box 118<br>SE-221 00 Lund<br>SWEDEN | | **Document name**<br>**DOCTORAL DISSERTATION**<br>**Date of issue**<br>May, 2011<br>**CODEN**<br>LUNFD6/(NFFL-7233)2011 |
| **Author(s)**<br>Philippe Gros | | **Sponsoring organisation** |
| **Title and subtitle**<br>Identifying Hadrons on the Relativistic Rise Using the ALICE TPC at LHC | | |
| **Abstract**<br><br>The chain from hadron collisions to the physics results requires several important links. First the outcome of the collision is measured by the detectors. Then, the signal from the detector is processed and transformed into information relevant for the study of the physics processes. The data is made available to physicists to be analysed and used to improve theories. This thesis presents work done on no most of these steps for the ALICE experiment at LHC.<br>First a study of the main processes in the TPC detector for ALICE was done using simulation and test beam data. The results are shown in paper I. The study was deepened with the analysis of test beam data from a TPC prototype for the ILC, as shown in paper III.<br>Concurrently, a study on the Grid – computing framework for distributed computing and storage resources – was performed. This involved the development of an interface module between the ALICE software AliEn and the ARC software developped in the Nordic countries. This work is presented in paper II.<br>Finally, an analysis of the particle identification capabilities of the ALICE TPC was done. Paper IV shows the results obtain from proton collisions at 7 TeV. The analysis is completed here with preliminary results on heavy ion collisions. | | |
| **Key words:**<br>LHC, ALICE, TPC, ILC, Grid, pion spectra | | |
| **Classification system and/or index terms (if any)** | | |
| **Supplementary bibliographical information:** | | **Language**<br>English |
| **ISSN and key title:** | | **ISBN**<br>978-91-7473-142-2 |
| **Recipient's notes** | **Number of pages**<br>115 | **Price** |
| | **Security classification** | |

**Distribution by (name and address)**

Philippe Gros
Department of Physics
Div. of Experimental High-Energy Physics
Box 118, SE-221 00 Lund, SWEDEN

**I, the undersigned, being the copyright owner of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.**



**Signature** ________________________________ **Date** , 2011

*To My Mother and the memory of my Father —*

# Populärvetenskaplig Sammanfattning

# (Popularised Summary in Swedish)

En stor fråga för människor, och särskilt fysiker, är hur Universum skapades. Numera känner alla till Big Bang teorin, där Universum började i en mycket våldsam händelse för fjorton miljarder år sedan. I det unga Universum (tusentals år efter Big Bang), var materien så het att elektroner och protoner inte var bundna tillsammans i atomer. Detta kallas plasma, och är materiens tillstånd i solen. När Universum var mycket yngre och hetare (mikrosekunder efter Big Bang), var inte ens kvarkar och gluoner bundna inom protoner, och materia var ett så kallad kvark-gluon plasma (QGP efter engelskans Quark-Gluon Plasma).

Fysikerna vet väldigt lite om QGP. För att förstå det bättre, försöker vi att skapa det med kärnkollisioner vid acceleratorer. Det är huvudmålet inom fältet högenergetisk tungjonfysik. Kollisioner av blyjoner med högsta energi sker i experimentet ALICE på CERN i Schweiz. Energin är nästan femton gånger högre än i tidigare experiment. I dessa kollisioner kan QGP skapas för en kort ögonblick. QGP studeras då genom dem många partiklar som kommer från det. För att detektera alla dessa partiklar, och mäta deras egenskaper, behövs stora och komplexa detektorer, och kraftfulla datorer för analys av de enorma datamängderna.

Denna avhandling beskriver olika delar i analysen av blyjonkollisioner. Eftersom det bara fanns data från kollisioner tillgänglig från och med mitt tredje år som doktorand, ägnade jag mycket tid att förbereda för det. Första delen av avhandlingen handlar om en särskilt detektor i ALICE-experimentet: tidprojektionkammaren (TPC). TPCn är en elegant detektor som registrerar passagen av elektriskt laddade partiklar med god positionupplösning i tre dimensioner och hög effektivitet. Dessutom kan TPCn hjälpa till med att identifiera många av dessa partiklar. Arbetet som redovisas här syftar till att bättre förstå och förbättra denna partikelidentifierings förmåga genom tester med olika TPC prototyper och genom att modellera detektorns respons.

Den andra delen av avhandlingen handlar om Grid, ett system som gör det möjligt för forskare i hela världen att använda datorer på många olika institutioner för att bearbeta och analysera data från experimenten. Grid verkställdes med flera olika mjukvaror (s.k. mellanvara) för olika projekt. I synnerhet utvecklade ALICE och nordiska institutioner två mellanvaror kallade AliEn och ARC. För att ge så många forskare som möjligt tillgång till så många datorer som möjligt, försöker man bygga gränssnitt mellan mellanvaror. Målet med mitt projekt var att utarbeta gränssnittet mella AliEn och ARC. Grid består av ett antal noder (sk Tier-1) som ofta är stora samlade datorresurser i nationella datorcentra.

Till slut använde jag kunskapen om både TPC och Grid för att analysera data från ALICE. Efter en kort analys ser vi redan nu hur sammansättningen av partiklar från kollisioner ändras när QGP har skapats.

# Résumé Vulgarisé en Français

# (Popularised Summary in French)

Les origines de l'Univers représentent une question fondamentale pour l'humanité, et en particulier pour les physiciens. Tout le monde a entendu parler de la théorie du Big Bang, qui décrit le commencement de l'Univers par un violent événement il y a quatorze milliard d'années. Dans le jeune Univers (quelques milliers d'années après le Big Bang), la matière était si chaude que les électrons et les protons ne pouvaient pas se lier pour former des atomes. On appelle cet état de la matière un plasma, c'est l'état que l'on trouve par exemple dans le soleil. Quand l'Univers était encore bien plus jeune et plus chaud, les quarks et les gluons n'étaient pas non plus confinés dans des protons et neutrons. Cet état est appelé Plasma Quark-Gluon (QGP).

On sait très peu de chose sur le QGP. Pour essayer de comprendre cette phase de la matière, les physiciens essaient de la recréer dans des accélérateurs de particules. C'est le but principal de la physique des ions lourds à hautes énergies. Des collisions d'ions de plomb à très haute énergie sont observées par l'expérience ALICE au CERN en Suisse. L'énergie atteinte est près de quinze fois supérieure à celle des expérience précédentes. Au cours de ces collisions, un QGP peut être créé pour un court instant. On étudie alors le QGP à travers les nombreuses particules qui en sont émises. Pour observer ces particules et mesurer leurs propriétés, on utilise des détecteurs gigantesques et complexes et de puissants ordinateurs pour traiter les énormes quantités de données produites.

Cette thèse décrit différents aspects de l'analyse de collisions d'ions lourds. Comme les données expérimentales sur les collisions n'ont été disponible que dans la troisième année de ma thèse, j'ai pu passer beaucoup de temps pour préparer cette analyse. La première partie de la thèse traite d'un détecteur particulier de l'expérience ALICE: la TPC. La TPC est un détecteur élégant qui permet de mesurer la trajectoire de particules chargées électriquement avec un haute résolution tridimensionnelle et une bonne efficacité. De plus, la TPC permet d'identifier une grande partie de ces particules. Le travail présenté ici approfondit la compréhension de cet aspect pour améliorer la capacité d'identification de la TPC. Ceci est fait en analysant les données de tests réalisés avec différents prototypes de TPC.

La seconde partie traite de la "Grille", un systême informatique qui permet aux chercheurs du monde entier d'accéder à des ordinateurs dans de nombreuses universités et instituts de recherche afin de traiter et analyser les données des expériences. Le concept de la Grille a été implémenté dans différents logiciels (appelés intergiciels) pour différents projets. En particuliers, ALICE et la communauté scientifique des pays nordiques ont développé deux intergiciels appelés AliEn et ARC. Pour permettre à un maximum de scientifiques d'accéder à un maximum de ressources, il est intéressant de connecter les différents systèmes. Une telle interface entre AliEn et ARC était le but de mon projet.

Finalement, j'ai pu utiliser les connaissances que j'ai acquises aussi bien sur la TPC que sur la Grille pour analyser des données produites par ALICE. Après une rapide analyse, on aperçoit déjà l'influence du QGP sur la composition des particules émises au cours des collisions.

## Acknowledgements

There are many people I have to thank for helping me to finish this thesis. I will probably forget some, please forgive me.

I will start with the obvious, big thank you to my supervisors: Peter Christiansen for guiding me through these years, full of ideas and ready to help on every subject, and also making every stay CERN more enjoyable around a few beers; Anders Oskarsson for sharing generously his experience and wisdom in so many fields from electronics to Scanian weather; Oxana Smirnova for introducing me to the world of the Grid.

I would like to thank the whole High Energy Physics division for the friendly atmosphere. In particular Bozena for the help with every administrative question. Leif for inviting me in the LCTPC project. Also Evert for keeping me up to date on statistics and French sports, Ulf for computer help. All the people I shared my office with: Alex, for keeping a lively atmosphere and helping with so many questions; Lene, Tuva, Ryô and Martin for their great work on the LCTPC. I will not forget Alejandro, Weina and Nele for letting me bother them so many times when I wanted to have a break or a fresh view on a problem.

I will also thank all my friends in Lund, for creating the wonderful international atmosphere and joining me for so many walks, bike rides and other crazy activities whenever I did not feel like working. Here comes a long, incomplete list: Lisa, Iskra, Alejandro, Ilaria, Weina, Tina, Cristina, Nils, Simona, Florido, Antonio, Chris, Luciana, Stefano... And I will not forget my friends in France: Laurent, Bénédicte, Claire, Manu, Tiphaine and Mathilde.

Finally, I would like to thank my whole family, including the three little ones born during this thesis, for their love and support all these years.

Merci à tous!

## Preface

This thesis gathers the work I have done on different aspects of the full process of physics analysis at the early stages of the Large Hadron Collider (LHC). Over the four years of my doctoral work, we prepared for the analysis at LHC. After the incident at the first LHC startup in September 2008 and the following shut down for repairs, we saw the first proton collisions in December 2009. In 2010, we observed proton collisions at increasing energies, quickly up to 7 TeV, and finally in November the first lead ions were collided at LHC with an energy of $\sqrt{s_{NN}} = 2.76$ TeV. During the several years before the first collisions, I could develop and test different tools necessary in the full analysis process.

The chain from the hadron collision to the physics results requires several important links. First the outcome of the collision is measured by the detectors. Then, the signal from the detector is processed and transformed into information relevant for the study of the physics processes. The data is made available to physicists to be analysed and used to improve theories.

During my doctoral work, I had the opportunity to work on almost every link of the physics analysis chain. I have first studied details of a detector, the ALICE Time Projection Chamber (TPC). By creating a complete simulation of the detector, I studied the most important phenomena taking place in the TPC. The development of the TPC prototype for the International Linear Collider (ILC) gave me the opportunity to go deeper in this study. The more modern technology used in the prototype allowed to look at the processes in more details. The participation in the test with an electron beam gave us the possibility the perform specific tests for the effects we wanted to study. It also gave me the opportunity to participate in the actual running of the prototype.

I have also participated in the Grid project, which provides access to computing and data storage resources for processing and analysis of the data created by the LHC experiments. I worked on an interface between the two Grid middleware solutions from the ALICE experiment and from the Nordic community. This gave me an interesting perspective on the different Grid strategies, while setting up a dummy Grid for tests gave me a overview of the internal details of the Grid.

When collision were finally recorded in the LHC, I could use the experience I gained on the TPC and the Grid to start analysing data from the ALICE experiment. In the little time had left, I could begin to study the actual signal registered in the TPC during heavy ion collisions. Even from a quick study, I could observe clear effects of the size of the interaction on the composition of the outcoming particles.

This thesis is based on four publications:

**Paper I** was published in Nuclear Instruments and Methods A (NIMA) in October 2009[1]. It describes a study done on data taken in 2006 with an ALICE TPC test setup. Using a fully controlled simulation of the entire TPC process, we studied the influence of the different phenomena in the TPC.

**Paper II** was published as proceedings to the ISCG2010 conference in the book ”Data

[1]Papers I and II are reprinted with the authorisation of the copyright holder Springer

Driven e-Science: Use Cases and Successful Applications of Distributed Computing Infrastructures (ISGC 2010)"[1]. It describes the work I have done on the interface between the two Grid middleware AliEn and ARC.

**Paper III** was published as a memo in the EUDET project[2]. It describes the test setup of the TPC prototype in DESY, and the performance analysis done by the Lund group. It also shows a study of the space resolution and energy loss distribution in the TPC, with comparison with simulation. I developed the reconstruction and simulation software for these studies.

**Paper IV** will be published as proceedings to the Epiphany conference as the July 2011 issue of Acta Physica Polonica B[3]. It describe the first results on the performance of the charged pion identification method on the relativistic rise using the ALICE TPC.

[2]Paper III is a public note for the Eudet project, available at http://www.eudet.org/e26/e28/

[3]Paper IV has been approved by the ALICE collaboration and submitted for publication

# Contents

# Chapter 1

# Heavy Ion Collisions and the Quark-Gluon Plasma

## 1.1 A Story of Origins

The Big Bang is a common and widely accepted theory to describe the origin of the Universe. All matter is created in a violent event. At its origin, the Universe had an extremely high temperature. It then cooled down and condensed into matter as we know it: the atoms that build up stars and planets. This evolution followed some steps as shown in Fig. 1.1. When looking back in time, or as matter gets hotter and hotter, the molecules first decouple themselves from each other to form a gas. Then energy from the temperature becomes higher than the electromagnetic energy that binds the electron to the nucleus, and we have an "electromagnetic" plasma of nuclei and electrons. Finally, at high enough temperature, the quarks and gluons, which are the constituents of the protons and neutrons, cannot be bound together by the strong interaction, and "strong" plasma is formed. It is called Quark-Gluons Plasma (QGP).

The transition from QGP to hadronic matter (where quarks and gluons are confined inside protons and neutrons, which are hadrons) is estimated to have happened some microseconds after the Big Bang. It is not possible to get direct observations related to that state of the Universe, because matter was then opaque to light. However, we can now observe the Cosmic Microwave Background (CMB) radiation, which is the remnant of the photons emitted by the Universe at its origin. From that, we can estimate how matter was distributed when the Universe became transparent to light (a few hundred thousand years after the Big Bang). This distribution reflects the dynamics of matter from its very origin, and therefore the thermodynamic properties of the different phases of matter. Therefore, to understand this astrophysical distribution, we need to understand the dynamics of matter at the highest temperature, i.e. at the largest energies. In particular, the early phases of matter, like the QGP.

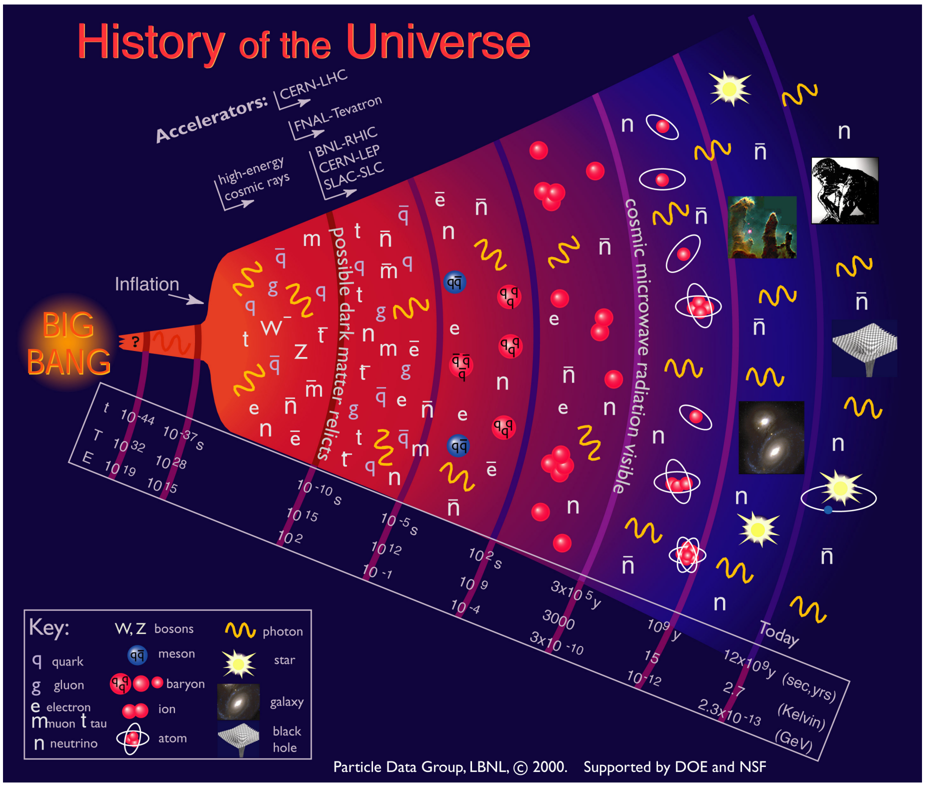



Figure 1.1: Cartoon of the evolution of matter in the Universe since the Big Bang.

## 1.2 The Quark-Gluon Plasma

The physics at very high energies, or very short distances, is described by the Standard Model of particle physics. The Standard Model is a very successful theory to describe matter at the shortest distances, and highest energies. The Standard Model describes matter with 30 fundamental particles (6 leptons and their 6 anti-particles, 6 quarks and their 6 anti-particles, and 6 bosons mediating their interactions, see Fig. 1.2).

Quantum ChromoDynamics (QCD) is the branch of the Standard Model that deals with quarks and gluons, and their interactions.The quarks and gluons carry a charge called "colour", that can take three values, usually referred as "red", "blue" and "green". Like electrically charged particles, coloured particles are attracted to each other through the so-called "strong" interaction to create a colour neutral ("white") environment. In QCD, the strength of the interactions increases linearly at large distances ($> 1$ fm) and disappears at short distances, much like the tension of a spring. This implies the two phenomena called confinement and asymptotic freedom. Since the strength increases with distance, quarks and gluons stay confined in colour neutral hadrons: baryons (e.g. the proton) with three quarks of the three different colours, and mesons (e.g. the pion) with a quark and an anti-quark of the anti-colour. At short distances (e.g. within a proton), quarks do not interact and can move freely.


Figure 1.2: Illustration of the fundamental particles constituting the Standard Model, with the three groups: The quark – which interact with the strong force – and the leptons – interacting with the weak force – are fermions. The gauge bosons are the carriers of the interactions: the gluons for the strong force, the photon for electromagnetism and the $Z$ and $W^{\pm}$ for the weak force. The Higgs boson $H^0$ is predicted in the Standard Model, but has not yet been observed. The pictures are there only for aesthetic purposes (see Appendix A).

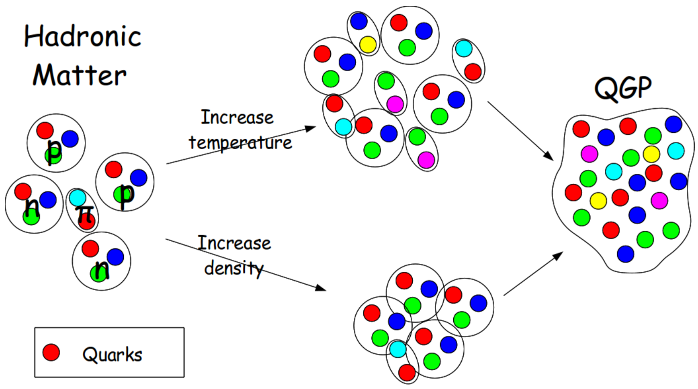



Figure 1.3: Sketch of the transition from normal hadronic matter to the QGP. At high enough temperature, quark pairs are created. When the quark density is large enough, the long distance strong potential – that confines the quarks in hadron – is screened by the other quarks. The quarks are deconfined in a "coloured" medium, the QGP.

When the temperature or the quark density increases, i.e. at large energy density, quarks might not be confined in hadrons any more, but are expected to move freely in a weakly interacting medium (see Fig. 1.3). This is the phase of matter called *Quark-Gluon-Plasma* (QGP). An estimated representation of the QCD phase diagram is shown on Fig. 1.4. The QGP can be created at very high densities (which probably happens in neutron stars) or very high temperatures (like in the early Universe, or in collisions in accelerators). For low densities, it seems the transition is not sharp, but happens smoothly ("cross-over").

Though the existence of a new phase of matter at very high energy density is accepted and has been observed in accelerator experiments, the nature of the QGP is still unclear. Different experiments try to probe the different regions of the phase diagram by producing very hot and dense matter. They then gather information about the medium created, but also about its evolution in the phase diagram. One surprising results of experiments was that instead of a weakly interacting medium, similar to a gas, the QGP created in accelerators seemed to behave like a liquid, where the components are strongly coupled.

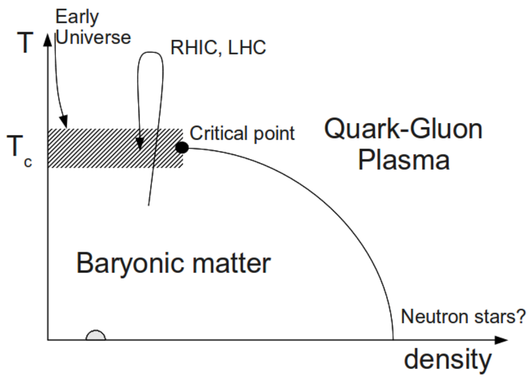



Figure 1.4: Sketch of the QCD phase diagram. The line represents the phase transition between normal baryonic matter and the QGP. In the hatched region, the transition is only a cross-over. The early Universe, neutron stars and the accelerator experiments are indicated. The arrows represent a time evolution.

## 1.3 Heavy Ion Collisions

In order to study the QGP, two main conditions are required. The first is to reach very high energy density. The second is to have a large enough sample to observe collective effects. The energy density required to create a QGP exists naturally only at the centre of neutron stars, or at the midst of violent astronomical events. To recreate them in a laboratory, we need particle accelerators. However, particle accelerators usually use protons or electrons, which are simple objects convenient to study the fundamental processes in collisions. To get the large samples and observe collective effects, we use instead heavier nuclei, typically gold (Au) or lead (Pb).

Heavy ion accelerator physics at high energies is relatively recent, and started with fixed target experiments at AGS[1] at Brookhaven National Laboratory (BNL) and SPS[2] at CERN from the late 1980s, reaching centre of mass energies of a few tens of GeV per nucleon. The first heavy ions collider was RHIC (Relativistic Heavy Ion Collider) at BNL, which observed the first collisions of gold ions at 200 GeV in 2000. The LHC at CERN has collided lead ions with a record centre of mass energy per nucleon pair of 2.76 TeV in 2010. At SPS, indications of a phase transition of matter have been observed. RHIC experiments established the existence of QGP, and some of its properties have been observed to be consistent with the hydrodynamical model of a perfect fluid [1, 2]. LHC should give new information and improve the understanding of the QGP and its evolution back to normal hadronic matter on the low baryon density side of the phase diagram (see Fig. 1.4).

It is unfortunately not possible to observe directly the medium created in the collisions. In each collision event, it can only exist for some extremely short time (a few $10^{-23}$s [3]) before it "freezes" out (hadronises) to normal hadronic matter. The only things that can be measured are the outgoing particles. Their nature and kinematic properties give indications about the dynamics of the medium, but also about the hadronisation processes where quarks and gluons regroup into hadrons. In order to understand the global characteristics of the medium, a single occurrence of the very short lived QGP is not enough. Besides, quantum phenomena at this scale are statistical in nature. Therefore, the collisions are reproduced many times and their outcome is studied statistically.

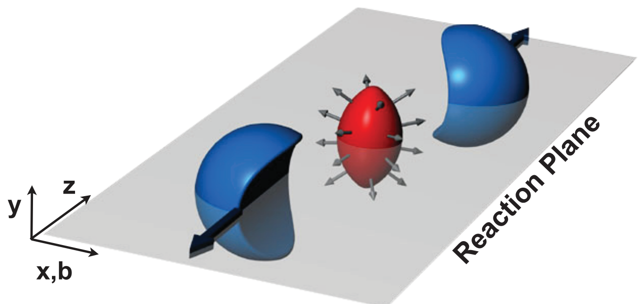



Figure 1.5: Sketch of the collision of two nuclei [4]. In blue, the spectator nucleons follow the beam directions. In red, the almond shaped interaction volume.

To describe the collective behaviour of the QGP in many different events, we need to define some references. First, to describe space orientation, we define a reference that is called the reaction plane (see Fig. 1.5). With the simplified assumption that the colliding nuclei are spherical, the reaction plane is defined as the plane containing the beam axis and the centre of the two nuclei. The region of interaction between the two nuclei has, in this simplified picture, an almond shape perpendicular to the reaction plane. Another important parameter is the amount of particles (and therefore energy) participating in the collision. This is related to the number of nucleons in each nucleus that will interact with those of the other nucleus (so called number of participants). A corresponding variable called "centrality" is often used, which is defined such that a "central" collision

[1] Alternate Gradient Synchrotron
[2] Super Proton Synchrotron

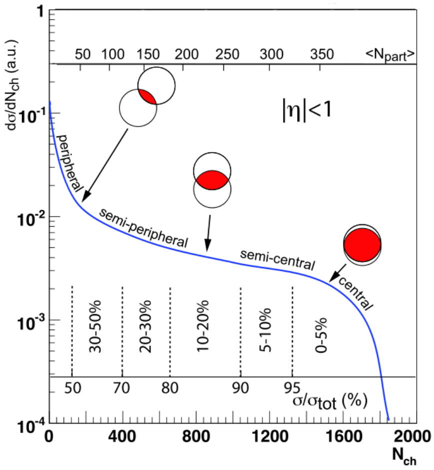



Figure 1.6: An illustration of the correlation of the final state observable $N_{ch}$ (number of charged particles) with Glauber calculated quantities $N_{part}$ (number of participants) and centrality [5]. The plotted distribution and various values are illustrative (based on LHC energies) and not actual measurements.

(with low centrality) has almost all the nucleons of the two nuclei interacting, while a "peripheral" collision (centrality approaching 100%) has only a few nucleons on the edge of the nuclei interacting. These value are evaluated from the number of produced charged particles using the so-called Glauber model as illustrated in Fig. 1.6.

## 1.4 QGP Experimental Observations

It is impossible to observe the QGP directly. However, the indirect measurements in heavy ions collisions have given convincing evidence of the creation of QGP, and shown some interesting properties. I will give an overview of some of the main results from RHIC, and expectations from LHC. I have concentrated on those observations at RHIC, where identified hadrons at high transverse momentum are of particular interest. Therefore, the method deployed in this thesis are expected to play an important role in these investigations. One should bear in mind that what is considered high $p_T$ at RHIC is rather intermediate at LHC.

### 1.4.1 Particle momentum spectra and nuclear modification factor

To check if the matter created in heavy ion collisions (AA) presents new properties or even a new phase, the first step is to compare the observables with those in proton collision (pp). This is done using the nuclear modification factor $R_{AA}$, which is the ratio of the particle yields in AA and pp, where the latter has to be scaled with the number of nucleon-nucleon collisions (also called *binary collisions*) in the AA collision. If AA behaves exactly the same as pp, $R_{AA}$ is 1. Of course, the binary collisions in AA and pp should have the same energy for the comparison to be relevant. Rare processes can be expected to scale with the number of binary collisions. Such rare processes can be hard scattering leading to jets or other high $p_T$ production, heavy quark production etc. Soft processes are rather expected to scale with the number of participants.

The most basic observable in particle physics is the momentum of the particle, or even the transverse momentum $p_T$ (i.e. the projection of the momentum on the plane transverse to the beam axis). It is the most important quantity to measure in an experiment. It is then natural to use the $R_{AA}$ vs $p_T$ as a main observable for heavy ion collisions. Besides, measurement of identified particles up to the highest possible $p_T$, is crucial to constrain theoretical models. It can help to separate contributions of initial state – from the collision itself – and final state – from the hadronisation stage – effects from the energy loss mechanism in the QGP.

Measurements of $R_{AA}$ for identified particles for relatively high $p_T$ at RHIC is shown in Fig. 1.7. One important feature of the $R_{AA}$ is the strong suppression of high $p_T$ particles in central events, where a value of one would be expected. At even lower collision energies, at the SPS, the corresponding ratio was even in excess of 1. This was then attributed to rescattering in the nuclear material. The observed suppression here is thus a large effect. It disappears and the ratio approaches 1 if peripheral collisions are selected. This suppression can be associated with the phenomenon of jet quenching, described in the following section.

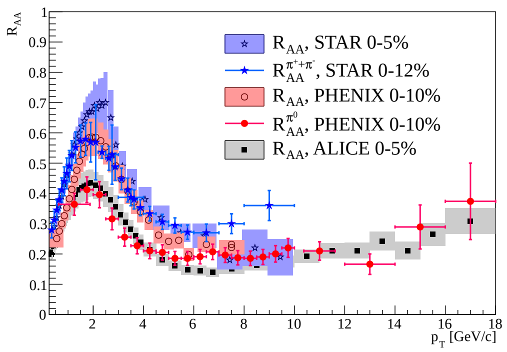



Figure 1.7: $R_{AA}$ measured for unidentified hadrons in ALICE, PHENIX and STAR and for identified pions in PHENIX and STAR. The data come from [6], [7] and [8].

Another observation in this context is the relative yields of baryons and mesons as shown on Fig. 1.8. There is a clear difference in the relative abundance of the protons and pions at intermediate $p_T$ in the case of central Au–Au collisions. This points to effects due to the creation of a QGP in these collisions. The data can be described by different models using recombination or coalescence. The hadronic composition at high $p_T$ is associated with the nature of the jets (whether they are very high $p_T$ quarks or gluons), and the energy loss process in the medium. On Fig. 1.8 we see that the composition at high $p_T$ does not seem to vary with centrality. This indicates that both quark and gluon jets are affected in the same way by the QGP. The energy loss processes should therefore be similar for quarks and gluons in the QGP [9].

### 1.4.2 Jet Quenching

Interactions with large momentum transfer (hard processes) in hadron collision can produce very high momentum partons (quarks and gluons). The partons are observed in

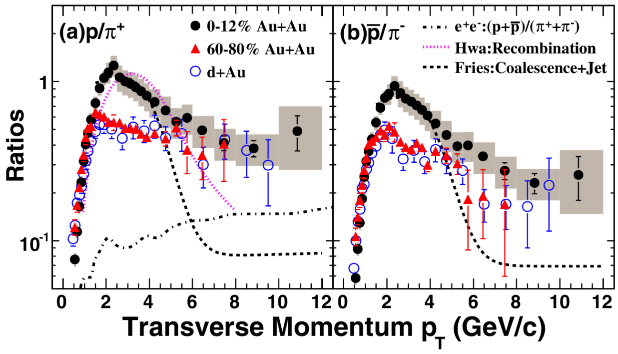



Figure 1.8: The $p/\pi^+$ and $\bar{p}/\pi^-$ ratios from d–Au and Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV. The $(p+\bar{p})/(\pi^+ + \pi^-)$ ratio from light quark jets in ee collisions at $\sqrt{s} = 91.2$ GeV is shown as a dotted-dashed line. The dotted and dashed lines are model calculations in central Au–Au collisions. Protons and antiprotons are less suppressed than pions at intermediate $p_T$. At higher $p_T$, the weak centrality dependence indicate that the partonic sources of $\pi^{\pm}$, $p$, and $\bar{p}$ have similar energy loss when traversing the nuclear medium [9].

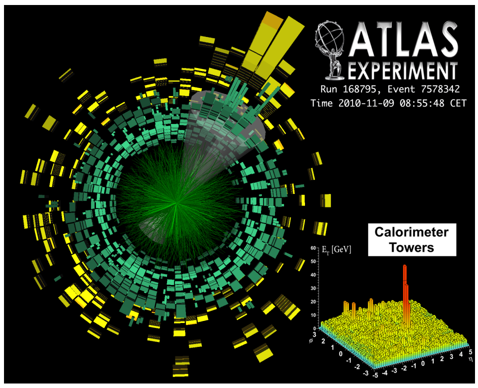



Figure 1.9: Event display of a Pb–Pb collision with jet quenching in ATLAS [10]. A jet is visible only on one side of the detector. The opposite side shows an overall higher activity corresponding to the scattered other jet.

the detector as "jets" (group of high momentum hadrons going in the same general direction). In a dense coloured medium, the high momentum partons are expected to lose a lot of energy through radiative effects, this is called "jet quenching". Due to momentum conservation, the jets are usually produced by pairs, back-to-back. Therefore, if the jet pair is produced on the outer part of the QGP region, one jet will escape and be observed, and the other one will be absorbed or softened in the medium.

Jet Quenching was a convincing evidence of the creation of QGP at RHIC. At LHC, ATLAS and CMS have observed the jet quenching on single events from the very first hours of the Pb–Pb collisions (Fig. 1.9). The study of jet quenching is expected to give very interesting information about the energy loss processes in the QGP produced at LHC.

### 1.4.3 Elliptic flow

One of the main goals in the study of QGP is to understand the collective behaviours of the medium. Since the interactions between quarks and gluons are very fast, the extremely hot and dense medium created in the collisions will thermalise into a QGP. If the constituents interact very little, as expected from the idea of asymptotic freedom, the system would expand like a gas of quarks and gluons and eventually condense into hadrons at a phase transition back to normal hadronic matter. What has been observed at RHIC is instead that the expansion has a collective hydrodynamic component. A collective expansion in the direction of the largest pressure gradient. This expansion is called flow. Due to the asymmetry of the almond shape reaction volume (see Fig. 1.5), we observe a so-called anisotropic flow (as described in Fig. 1.10). This anisotropic flow is one of the clearest signature of collective effects in the QGP.

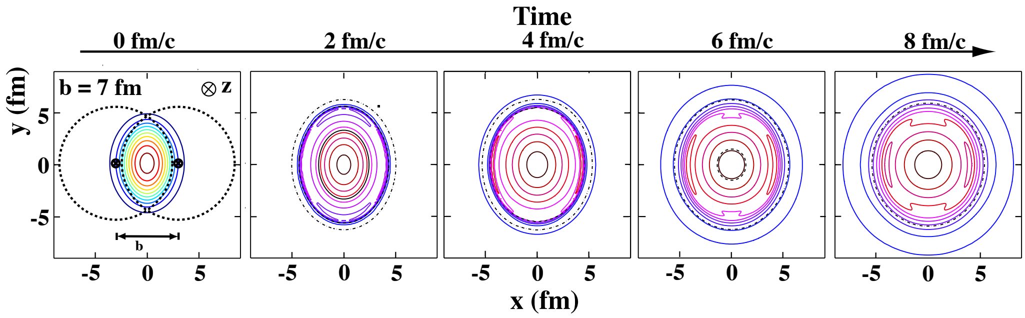



Figure 1.10: The created initial transverse energy density profile in space and its evolution in time [4]. The $z$-axis is along the colliding beams.

The flow is thus directed into the reaction plane. The experimental observable related to flow is a directionality of soft particle emission relative to plane. One can observe it either as azimuthal angular correlations to this plane (if it is known) or as the second component of the Fourier expansion of the anisotropic flow. This so-called elliptic flow ($v_2$) can be associated to multiparticle azimuthal correlations in the event. The large elliptic flow measured at RHIC has been interpreted as the signature of a perfect liquid [1,2].

When looking at the elliptic flow for identified hadrons in the STAR experiment at RHIC (Fig. 1.11), we observe that the mesons ($\pi$, $K$ and $\Xi$) and the baryons ($p$, $\Lambda$ and $\Omega$) are following two different curves. With a scaling to the number of constituent quarks $n_q$ (two for mesons, three for baryons), all the point can be made to align (Fig. 1.12). This observed $n_q$-scaling can be interpreted as evidence that the freeze out process can be described using coalescence models [11].

Due to the high energy, collective flow effects could also be present in high multiplicity $pp$ collisions at the LHC, as is studied in for example [13].

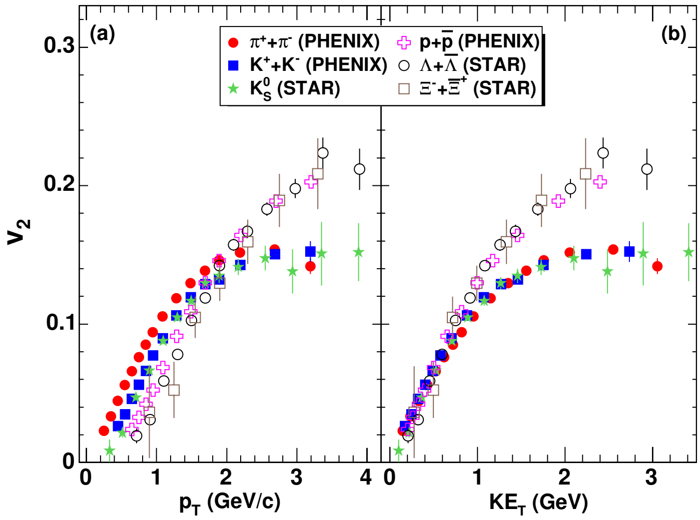



Figure 1.11: (a) $v_2$ vs $p_T$ and (b) $v_2$ vs transverse kinetic energy $KE_T$ for identified particle species obtained in minimum-bias Au–Au collisions at RHIC [12].

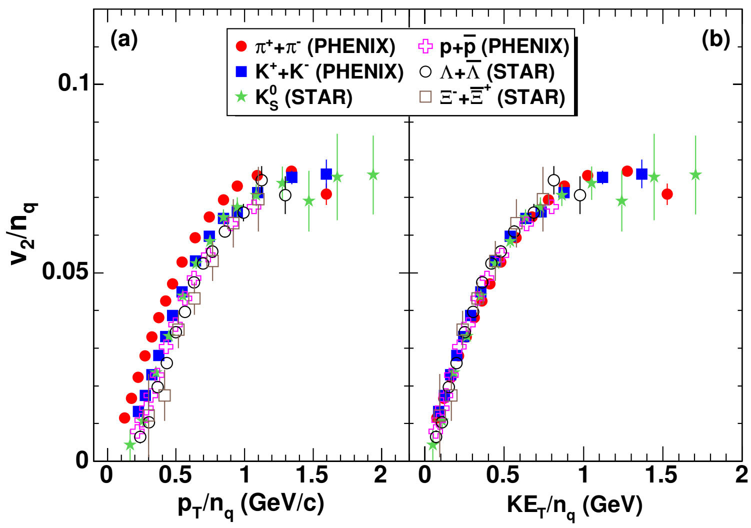



Figure 1.12: Same spectra as Fig. 1.11, scaled to the number of constituent quarks $n_q$ [12]. With this scaling, all the curves align.

# Chapter 2

# The ALICE Experiment

## 2.1 The Large Hadron Collider at CERN

The Large Hadron Collider (LHC) was built between 2002 and 2009 at the CERN laboratory near Geneva. It is the largest and most powerful particle accelerator ever built. It is installed in the 27km long circular underground tunnel across the border between France and Switzerland that hosted its predecessor the LEP (Large Electron–Positron collider). 16 radio-frequency (RF) accelerating cavities and over 1600 superconducting magnets allow it presently to accelerate, store and collide protons with a centre-of-mass energy up to 7 TeV, more than three times larger than the previous record held by the Tevatron at Fermilab, near Chicago. Future improvement should allow to reach the design energy of 14 TeV. Besides, it is also used for shorter period to accelerate and collide lead ions (Pb) to an energy per nucleon pair of 2.76 TeV, 13 times higher than at RHIC (5.5 TeV after future improvement).

The LHC was designed and commissioned to address several question that still puzzle physicists:

- The existence of the Higgs boson, predicted by the current and very successful theory of the Standard Model.
- The nature of Dark Matter and Dark Energy, which constitute most of the Universe.
- The reason for the dominance of matter over antimatter in the Universe.
- The nature and properties of extremely hot matter (such as the QGP) that constituted the early Universe.

To address and study these different issues, 4 detectors were designed with different focus, and installed in 4 interaction points on the LHC ring as shown on Fig. 2.1.

**ALICE [14]** A Large Ion Collider Experiment, is designed to study the QGP in heavy ion collisions. ALICE is the main focus of this thesis.

**ATLAS [15] and CMS [16]** A Toroidal LHC ApparatuS and Compact Muon Solenoid, are general purpose detectors, which will try to gather as much information as

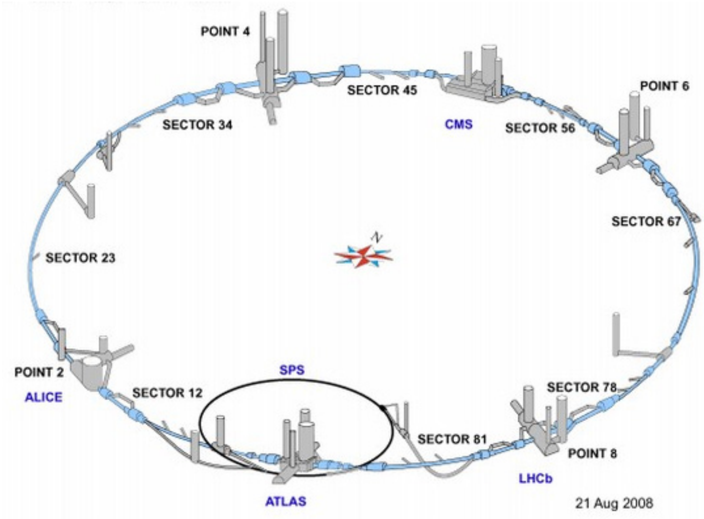



Figure 2.1: The LHC ring layout

possible from hadron collisions. Their main focus is proton-proton collisions and the fundamental questions of the Standard Model, and expected new phenomena beyond the Standard Model.

**LHCb [17]** which will study the asymmetry between matter and antimatter through the precise observation of the physics of the bottom quark (*b-quark*)

Besides the main questions described above, exotic and new physics beyond the Standard Model might be unravelled by LHC. The theory suggesting a new symmetry between fermions and bosons called SuperSymmetry (SuSy) [18] shows a lot of promises, but extra dimensions and black holes [19] or generic approaches like secluded sectors [20] have been investigated in the past years. The high sensitivity data in a new energy range that the LHC is providing will allow to confirm or exclude many of these hypotheses.

## 2.2 The ALICE Experiment

ALICE (A Large Ion Collider Experiment) is a large experiment placed at one of the four collision points of LHC. The collaboration involves over a thousand scientists and engineers from 116 institutes in 33 countries. It was designed to study the strong interaction sector (QCD) of the Standard Model and to characterise the Quark-Gluon Plasma (QGP) It is the only experiment at LHC which was optimised for the heavy ions (Pb+Pb) collisions.

The detector is placed in the solenoid magnet from the old LEP experiment L3. This provides a relatively low magnetic field of 0.5 T, which allows to measure with high performance low momentum particles corresponding to the so-called soft QCD, as well as more energetic particles from hard processes. Because of the extremely large number

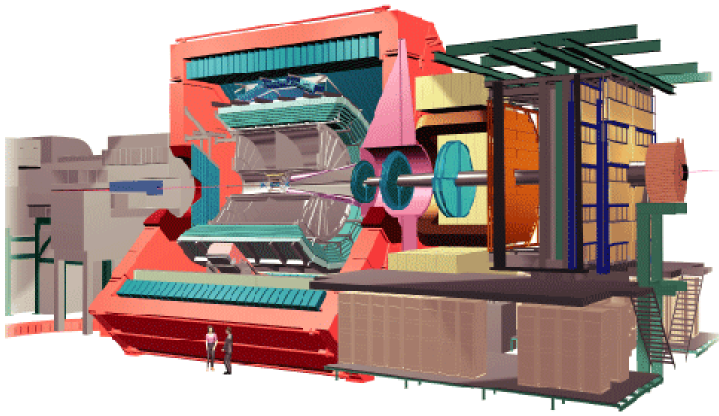


Figure 2.2: The ALICE detector setup

of created particles (*multiplicity*) expected in central nucleus-nucleus collisions at LHC energies, the design of ALICE was optimised for a multiplicity $dN_{ch}/d\eta$=4000, but tested by simulation up to $dN_{ch}/d\eta$=8000. The first results showed a lower $dN_{ch}/d\eta$ $\approx$1600 at half energy [21]. ALICE has an efficient and robust tracking system over a large momentum range, from tens of MeV/c (soft physics) to over 100 GeV/c (jet physics). A specificity of the ALICE detector over the other LHC experiments is its large focus on hadron and lepton identification (PID). It is achieved over much of the momentum range using most known PID techniques: specific ionisation energy loss $dE/dx$, time-of-flight, transition and Cherenkov radiation, electromagnetic calorimetry, muon filters, and topological decay reconstruction.

### 2.2.1 The detectors

The detectors in the ALICE experiment are arranged in a classical layered structure around the interaction point as shown on Fig. 2.2. Here is a short description of the main detectors:

**Inner Tracking System (ITS),** placed closest to the interaction point. It consists in 6 cylindrical layers of silicon detectors at radii from 3.9 to 43 cm, and covering the pseudorapidity region $|\eta| < 0.9$. It allows tracking and identification of low momentum particles, detection of secondary vertexes, and provides a reference to improve the tracking resolution of the TPC. Combined with the TPC, it can identify vertexes with a 14 $\mu$m resolution in the transverse plane [22].

**Time Projection Chamber (TPC),** large cylindrical gas detector. It has an inner (resp. outer) radius of 85 cm (resp. 250 cm), and covers $|\eta| < 0.9$. It is described in detail in the next chapter.

**Transition Radiation Detector (TRD),** from 2.9 to 3.8 m from the interaction point, $|\eta| < 0.84$. It discriminates electrons from pions with high efficiency for momenta above 1 GeV/c

**Time Of flight (TOF),** from 3.99 to 7.41 m from the interaction point, $|\eta| < 0.9$. By measuring the time particles take to reach it, and combined with the tracking information of the TPC, it allows to identify pions, kaons and protons at low and intermediate momentum ($p_T < 2.5$ GeV/c).

**High Momentum Particle Identification Detector (HMPID),** does not have a full azimuthal coverage. The HMPID discriminates pions and kaons in the range $1 < p < 3$ GeV/c and protons and kaons in the range $2 < p < 5$ GeV/c by means of their Cherenkov rings.

**PHOton Spectrometer (PHOS),** has a limited acceptance at central rapidity, at the bottom of the setup. It is a high resolution electromagnetic crystal calorimeter for low $p_T$ photon measurement, notably from $\pi^0$ decays.

**ElectroMagnetic Calorimeter (EMCal),** has a limited azimuthal coverage, opposite of PHOS. Like PHOS, it provides photon measurements with a large $Pb$-scintillator array. The position back-to-back of the two detectors allows dijet studies.

**Muon Spectrometer,** situated in the forward region ($-4 < \eta < -2.5$), on one side of the experiment. It offers tracking of muons with high momentum ($p > 4$ GeV/c, having passed the large concrete absorber). It provides measurement of the different quarkonia ($J/\psi$, $\psi'$, $\Upsilon$, $\Upsilon'$, $\Upsilon''$) decaying into the $\mu\mu$ channel.

**Forward Multiplicity Detector (FMD),** covering the forward regions $-3.4 < \eta < -1.7$ and $1.7 < \eta < 5$. It is a silicon strip detector which measures charged particle multiplicities in the forward direction, mainly to evaluate the reaction plane of Pb–Pb collisions.

### 2.2.2 Computing framework

The ALICE experiment, like all the LHC experiments, will yield enormous amounts of data. Sharing and processing it requires large resources, which are provided through the Grid, as explained in chapter 4.

Another effect of the large community using the data, is the necessity for a unified computing framework. Every scientist in the ALICE collaboration needs to rely on the work of many others to extract the relevant information from the experiment. Detector signal reconstruction and interpretation, data quality, event selection as well as physics and detector simulation have to be handled with a common system to enable the end result of physics analysis.

In ALICE, the *AliRoot* framework is based on the ROOT [23] software. The ROOT system can be seamlessly extended with user classes and libraries that become effectively part of the system. These libraries are loaded dynamically and the classes contained share the same services of the native ROOT classes. This allows all the contributors to the ALICE experiment to develop and then share code for detector simulation, reconstruction and optimisation, as well as for physics analysis.

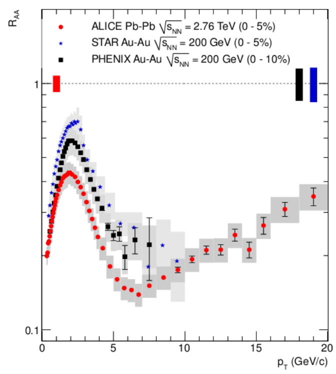



Figure 2.3: $R_{AA}$ in ALICE for Pb–Pb at $\sqrt{s_{NN}} = 2.76$ TeV, compared to the values obtained with Au–Au at $\sqrt{s_{NN}} = 200$ GeV at RHIC [6]. At higher collision energy, the $R_{AA}$ measured in ALICE has a similar $p_T$ dependency as the one measured at RHIC.

## 2.3 First Results from ALICE

In November 2010, Pb–Pb collisions have been recorded in ALICE, and the data has been readily analysed. Results have been published on several different observables. I will mention here first the measurement of the nuclear modification factor $R_{AA}$, shown in Fig. 2.3. The data in ALICE shows an agreement in the shape with the $R_{AA}$ measured in STAR and PHENIX at 200 GeV. This seems to contradict expectations from recombination models [6]. At high $p_T$, the behaviour can be associated with jet quenching.

In absence of pp data at matching energy (2.76 TeV), these results were obtained using an interpolation of the $p_T$ spectra for pp from the data at 900 GeV and 7 TeV using Monte Carlo simulations. In March 2011, protons were collided at 2.76 TeV that should provide a better reference for $R_{AA}$ measurements. More work is required to obtain $R_{AA}$ for identified particles. The analysis shown in Paper IV shows a method to obtain it for pions, kaons and protons.

ALICE already obtained also measurements of elliptic flow that show remarkable agreement with the previous experiments, over a very wide range of collision energies (see Fig 2.4).

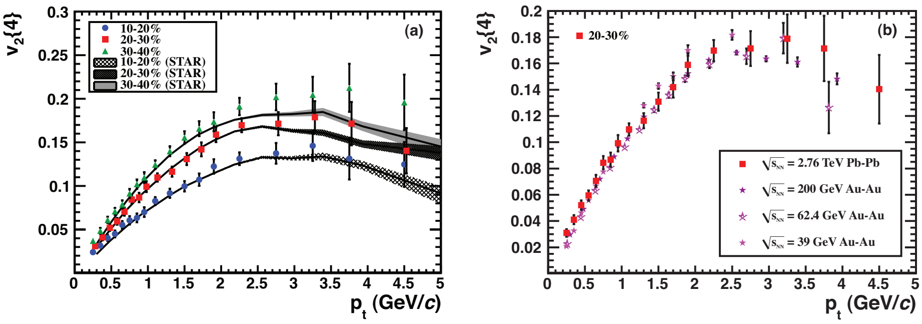



Figure 2.4: Elliptic flow in ALICE compared to previous experiments. (a) compare with STAR at $\sqrt{s_{NN}} = 200$ GeV for different centralities. (b) compared, at same centrality, with measurements for a collision energy range covering almost two orders of magnitude [4]. The agreement between all the results is remarkable.

# Chapter 3

# The Time Projection Chamber

The Time Projection Chamber (TPC) is a type of detector invented in the 1970s and designed to perform tracking of charged particles in three dimensions. It was used in many high energy physics experiments, notably ALEPH [24]. NA49 at SPS and STAR at RHIC have demonstrated the possibilities of TPC detectors for high energy heavy ions physics. The largest TPC in the world was built and used for the ALICE experiment. The same type of detector is considered for experiments at the International Linear Collider (ILC), and prototypes are already being tested. Here I will describe the principle and main features of the TPC, and I will then show the specific characteristics the ALICE TPC and the large TPC prototype for ILC, which I worked on.

## 3.1 Particle Tracking with a TPC

### 3.1.1 Principle

The basic principle of the TPC is as follows: a charged particle passing through a gas will interact with the gas and thus free electron along its trajectory through ionisation. In the TPC, we detect these electrons to reconstruct the particle's trajectory.

More specifically, the TPC is a gas volume (chamber), with a uniform electric field. The electrons freed along an ionisation track will drift along the field line to the end plate of the chamber, where 2-dimensional electron avalanche detectors are installed (see fig. 3.1). This 2-dimensional detector allows to measure the projection of the track on the end plate (thanks to the straight field lines). Besides, since the drift velocity of the electrons can be determined, the time of arrival of the electrons on the end plate will give the distance of the track to the end plate at each $x$-$y$ position (this assumes of course that we know the time when the track passed through the chamber). This way, we finally get the 3-dimensional trajectory of the particle. If we add a uniform magnetic field, the particle will follow a helical trajectory, and the curvature will be proportional the momentum $p$ of the particle.

The characteristics of the energy loss process of relativistic charged particles in a gas gives an extra information. The energy that the particle loses in the gas (which is on average proportional to the number of ionisation electrons) depends only on $\beta\gamma = p/m$

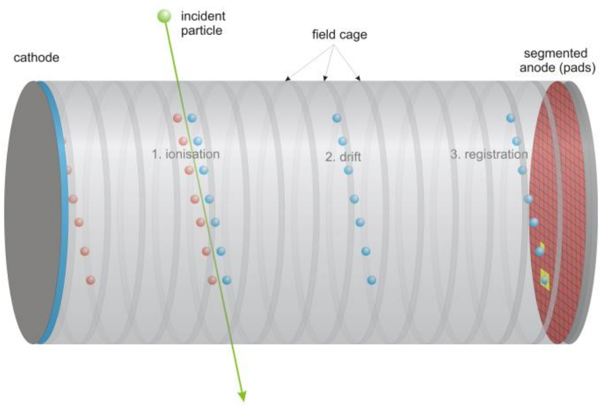



Figure 3.1: Working principle of a TPC [25]. A charged particle passes through the gas producing ionisation electrons. The electrons drift along the electric field to the anode and are readout on the end plate.

(more details in the following section). Therefore, by using the number of ionisation electrons measured and the momentum $p$, we can find the mass $m$ of the particle.

### 3.1.2 Energy loss of charged particles in a gas

The TPC is a very elegant detector which offers great possibilities for precise measurement. It is however a very dynamic device whose response depends on many parameters, in particular temperature and pressure. For this reason, it has to be calibrated at all times. To optimise and automatise this calibration, a good understanding of the different mechanisms is required.

The energy loss and ionisation of the relativistic particle in the gas is a complex process which is to date not very well described by models. It can be divided in two aspects: the mean energy loss $\langle dE/dx \rangle$ of the particle and the fluctuations around this value. The mean energy can be calculated, with a few approximations. It gives the separation between the different particles. The fluctuations are more difficult to model. They are the main component in the detector's $dE/dx$ resolution and should be understood as well as possible to improve this resolution. A higher resolution allows to distinguish particles even when the difference in the $\langle dE/dx \rangle$ gets small. Besides, the understanding of these processes is important for applications in many different fields, in particular for ion therapy in medicine.

#### Mean energy loss: Bethe-Bloch function

Relativistic charged particles passing through a gas lose their energy mainly by ionisation or excitation of electrons of the atoms in the gas. The mean rate of energy loss of a particle with charge $Z$ is given by the Bethe-Bloch formula [26]:

$$-\left\langle \frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta} [\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}] \qquad (3.1)$$

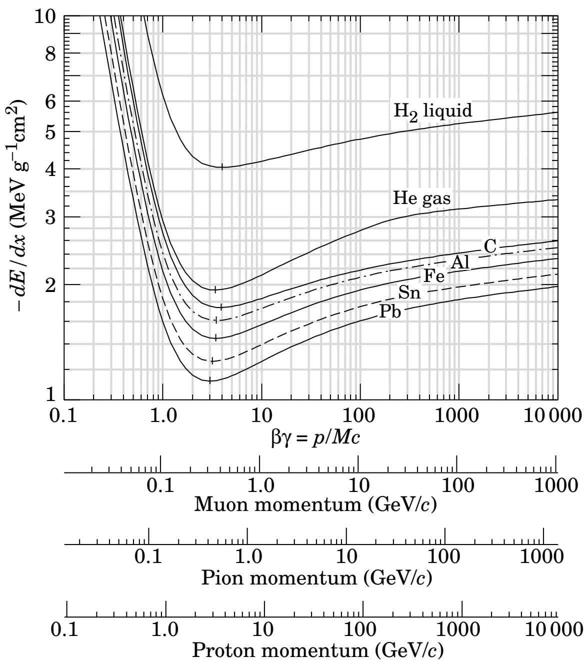



Figure 3.2: Mean energy loss rate in liquid (bubble chamber) hydrogen, gaseous helium, carbon, aluminium, iron, tin, and lead [26].

where $K = 4\pi N_A r_e m_e c^2$ , $r_e$ and $m_e$ are the classical radius and mass of the electron, $N_A$ Avogadro's number, $Z$ and $A$ the atomic number and atomic mass of the material, $z$ and $\beta$ the electric charge and the relativistic velocity of the incident particle, $\gamma = 1/\sqrt{1-\beta^2}$. $I$ is the mean ionisation potential of the material, which has been measured for several materials, and is usually extrapolated from these. $T_{max}$ is the maximum kinetic energy that can be given to a free electron in a single collision which, for a particle of mass $M$, is given by

$$T_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2} \tag{3.2}$$

and becomes $T_{max} = 2m_e c^2 \beta^2 \gamma^2$ for $M \gg m_e$ (the so-called "low-energy" approximation). The term $\delta$ represents a density effect correction and is relevant for particles with $\beta\gamma \gg 3$.

Even though a small dependence on $M$ is introduced through $T_{max}$ at high energies, for all practical purposes the $dE/dx$ given by Eq. 3.1 depends on $\beta$ only. Conventionally it is a function of $\beta\gamma = p/Mc$. Equation 3.1 is valid for $0.1 \lesssim \beta\gamma \lesssim 1000$ since at the lower limit the velocity of the incident particle becomes comparable with the atomic electron "velocity" and at the upper limit radiative losses start to be important (1% at around 10 GeV/c for muons) [26].

A plot of the Bethe-Bloch curve for different materials is shown is figure 3.2. At low $\beta\gamma$, the function dives following the component $1/\beta$. It reaches a minimum at $\beta\gamma \simeq 3.6$, which is referred to as the Minimum Ionising Particle (MIP) region. Then it rises following a more or less logarithmic slope. This is called the relativistic rise. It is associated with the maximum recoil energy of the electrons in the Rutherford scattering of the particle. Finally, the function reaches a plateau for very high $\beta\gamma$ (ultrarelativistic plateau). The polarisation of the medium by the charged particle shields the atoms from the electric field at larger distances.

The relativistic rise is a region of particular interest for particle identification. The steady increase provides a difference in $dE/dx$ for particles with same momentum, but different mass over a large momentum range. However, the slow logarithmic evolution makes this difference small, and requires high precision measurements to perform particle identification. This implies a good resolution on the *average* value of $dE/dx$. In practise the TPC is the only detector where this can be attempted since one performs large number of $dE/dx$ measurements (159 in the ALICE TPC).

**Energy loss fluctuations: Photo-Absorption Ionisation model**

The particle will of course not lose energy uniformly along its trajectory. The energy loss occurs by interaction with atoms in the gas. This can be modelled with three steps:

1. The interaction between the particle and the gas has a certain cross-section. We assume the interactions are not correlated and therefore happen with a Poisson distribution along the trajectory. The parameter depends on the characteristics of the gas (composition, pressure, temperature).

2. Each interaction will transfer a certain amount of energy. The probability density for this process is not very well understood and requires a model. A very popular model for this is called Photo Absorption Ionisation (PAI) [27].

   From Maxwell's equations in the gas, the energy loss of a charged particle through its interaction with the electric field it induces can be calculated. The result depends on the oscillator strength density of the atoms in the gas. This can only be calculate rigorously for hydrogen. Using a dipole approximation, it can be written for other elements in term of the so-called photo-absorption cross-section. Since this cross section has been measured and tabulated for many materials, this model can be used for most situations.

   One characteristic feature of this model is that the shell structure of the atom will enhance the probability to have energy losses corresponding to the energy level of these shells, as shown in fig. 3.3

3. The energy deposited by the particle is usually transferred to an electron of the atom which will then sometimes liberate a number of secondary electrons in the region. A simple approximation will give a number of secondary electrons proportional to the energy transferred.

### 3.1.3 Electron detection

The electrons produced by ionisation have to be detected with high precision. A first effect to take into account is the diffusion of the electrons while they are drifting to the end plate. To minimise this effect, one could reduce the drift time. For this, one should either reduce the drift length, or increase the drift velocity. Increasing the drift

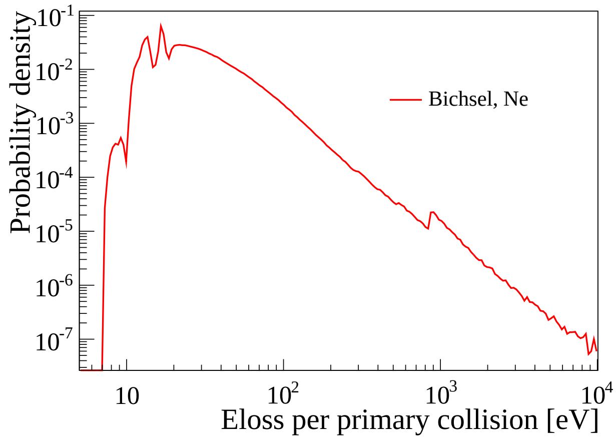



Figure 3.3: Probability for a 3 GeV/c proton to lose an energy $E$ over 1 cm of gas obtained from calculations by H. Bichsel [28].The effect of the shell structures of Neon is apparent.

velocity can be done by changing the gas mixture, but also requires electronic readout with matching time characteristics to keep the same space resolution in the drift direction. The electron diffusion is therefore an important factor in the choice of the gas for the TPC. The magnetic field used to measure the particle momentum has a beneficial side effect to reduce the lateral diffusion of the electrons by forcing them into a helical trajectory. So increasing the magnetic field will reduce the spatial effect of the diffusion.

Another effect to avoid is the absorption of the electrons in the drift volume. This is done by removing as much as possible some molecules from the gas, in particular oxygen.

Finally, when the electrons reach the end plate, they have to be amplified to create a signal that can be sensed by electronics. Different electron avalanche systems can be used. For many years, only MultiWire Proportional Chambers have been used. More recently, new technologies such as Micro Pattern Gas Detectors (MPGDs), of which the Gas Electron Multiplier is an example, have been developed and can be used in TPCs. The electron avalanche amplifiers I worked with in this thesis are described in the next sections.

### 3.1.4 TPC dE/dx resolution

The very long tail of the energy loss distribution shown in Fig. 3.3 makes it impossible to measure the average energy loss with a finite number of samples. To optimise the resolution, we must evaluate the most probable value of $dE/dx$ in a track. Moreover, the limited space resolution of the detector does not allow to measure the energy loss in single interactions with the gas, but rather the integrated ionisation over segments of the track. The resulting distribution can be described with a Landau distribution (see Fig. 3.4). Because of the long tail of the Landau distribution, we can approach the most probable value by discarding the sample with the largest signal in the track, and only take the average of the 60% lowest values. This so-called truncated mean is used as the

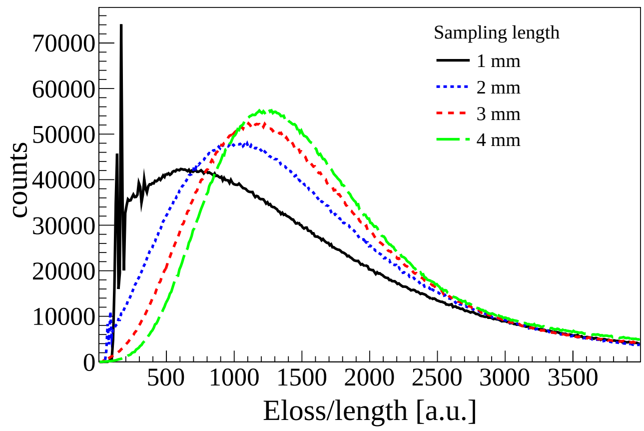



Figure 3.4: Energy loss for a minimum ionising particle in an Argon based gas mixture (P10) over different sampling lengths using a model calculated by H. Bichsel [29]. The energy loss is normalised to the sampling length so that the mean value is constant. For very short samplings (1 mm), the details of the energy loss are visible. For long samplings (4 mm), the shape is described by the smooth Landau distribution, with a long tail. The most probable value (peak position) depends on the length.

TPC signal for a track, and often referred to as the $dE/dx$ of the track. Unlike the average, the truncated mean (as the most probable value) does not scale with the sample length.

The Landau distribution is an approximation which is valid when the sampling length is long, so that the signal is the average of many interactions. When we improve the resolution of the detector and reduce the sampling length, so that each hit corresponds only to a few interaction of the particle with the gas, this model is not valid any more. Details of the energy loss process can be observed in the distribution.

In order to optimise the TPC particle identification capacity, it is fundamental to have a good description of these distributions, and their dependence on parameters such as the track angle (related to the sample length). For high precision detectors like ALICE or the new TPCs at linear colliders, this requires to understand many details of the whole TPC system. Since many effects are entangled like energy loss and ionisation, and the details are smeared out by diffusion and gain variations, a simulation is the best way to get a clear view of the different processes.

I have developed a simulation program which was used for the work leading to paper I and III. The simulation (described in detail in these papers) takes into account all the main effects inherent to the TPC functioning. It treats all the effects individually, so that each can be modified or removed independently to see its influence on the final distributions.

## 3.2 The ALICE TPC

The ALICE TPC [30] is the main tracking detector of the ALICE experiment. It has been part of the ALICE project since its very beginning in 1993. It was designed to cope with the highest conceivable charged particle multiplicities predicted, reaching 20 000 tracks in one interaction in the TPC acceptance. It includes an electron readout system of MWPCs which is described below, followed by an overview of the technical characteristics of the detector.

### 3.2.1 MultiWire Proportional Chambers (MWPC) readout in ALICE

The MWPC was invented in 1968 by Georges Charpack and has been used in TPCs as standard device for gas amplification. In this technique, two or three grids of parallel wires are set to different electrical potentials, and the electronics is placed on a readout plane behind them. The cathode wire grid is set to a potential matching the drift field in the chamber. The anode wire grid, between the cathode and the readout plane, is set to a higher potential so that the electrons are accelerated towards the anode and create an amplification avalanche (see fig. 3.5). A third wire grid called "gating grid" is added on the side of the drift volume. By setting the wires on the gating grid to alternate potentials, a special field configuration is created that prevents ions to drift into the gas volume when the TPC is not active, and electrons from the drift volume to enter the MWPC and develop avalanches when the TPC is not being readout.

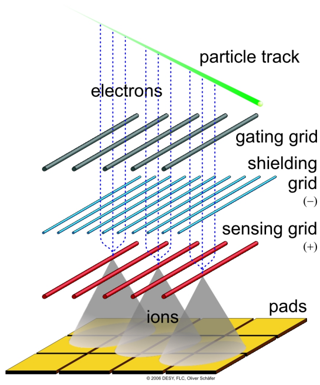



Figure 3.5: Layout of gas amplification and read-out with proportional wires [25]. The ionisation electrons from the tracks drift to the cathode wires (shielding grid) through the open gating grid. The potential difference between the cathode and the anode accelerates them so they create an avalanche close the the anode wire. The moving ions created induce a signal in the readout pads.

The readout plane is made as copper pads etched on a circuit board, with connections through the board to the Front End Electronics (FEE). The produced ions drift back towards the TPC volume and induce a signal on the pad plane. The induced signal is wide enough to be sensed on several pads. The ions are stopped by the gating grid from getting into the TPC volume where they could distort the electric field if the rate is high. The induced ion signal on the pads is further amplified by a charge sensitive preamplifier, followed by a shaper. Due to the slow drift of the ions, the induced signal has a long receding time, appearing as a long tail in the pulse. The signal is then digitised and and can be digitally corrected for tails in the FEE, as described in paper III, section 6.

The induced signal reaches over several pads which allows a high position resolution with a simple centre of gravity reconstruction. However, the width of the signal limits the separation power for nearby tracks in the TPC. For high precision detectors, the physical limit for the distance between two wires can also be an issue for this type of detector. But since high momentum particle are basically perpendicular to the wires, the wire spacing has minor influence on the momentum resolution. Moreover, the wire spacing is smaller than the pad size. Due to the diffusion a point-like information at the track will give avalanches on several wires, thus allowing a centre of gravity calculations also in this direction.

### 3.2.2 Design data of the ALICE TPC

The ALICE TPC is a 5 m long cylinder with an 80 cm inner and 2.8 m outer radius, containing about 90 $m^3$ of gas volume. It is separated in the middle by a 30 $\mu$m cathode plane with 100 kV potential creating the 400 $Vcm^{-1}$ drift field. Each end is equipped with MWPC readout organised in 18 trapezoidal sectors on each side. The MWPC readout is highly segmented with a total of around 560,000 pads.

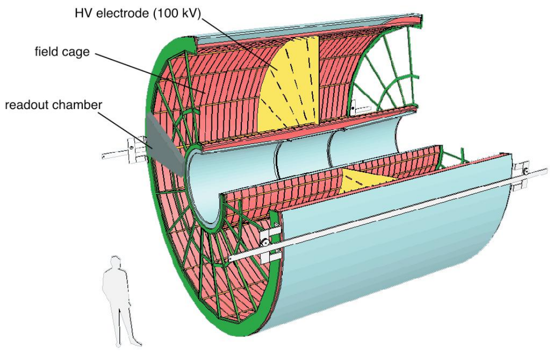



Figure 3.6: The ALICE TPC. The drift volume is a 5 m long and 2.5 m radius cylinder, cut in its middle by a 100 kV electrode creating the electric field. Each end cap is equipped with MWPC readout chambers installed in 18 sectors.

## 3.3 The ILC TPC Prototype (LCTPC)

The International Linear Collider (ILC) is a collider project started in the early 2000s, and still at a preliminary development stage. If accepted for funding, the ILC will accelerate and collide electrons with a centre of mass energy of about 500 GeV/c using several thousands superconducting accelerating cavities over about 30 km. The plans the LCTPC development is such that, awaiting final decisions about the ILC, the development is also adequate for e.g. CLIC and other high precisions collider experiments.

Two detector concepts for ILC have been selected: ILD (International Large Detector) and SiD (Silicon Detector). Research and development on all the specific detectors goes on under the EUDET project (European Detector R&D towards the ILC). The ILD concept includes a large TPC, and a large prototype has been built and tested with dif-

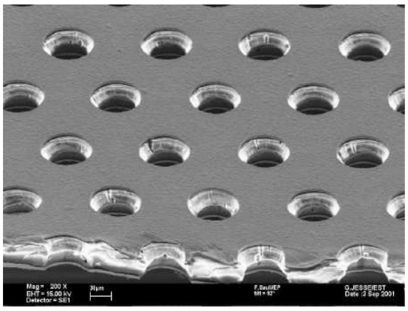


Figure 3.7: Image of the GEM structure taken with an electron microscope [25]

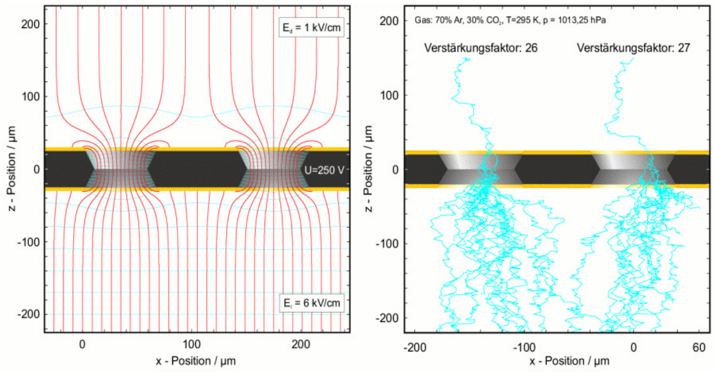



Figure 3.8: Illustration of the working principle of a GEM [25]. Left side shows the field lines (red) in a GEM hole. The right side shows the multiplication of a single electron -turquoise lines show electron paths- in GEM holes

ferent electron readout systems at DESY, as is explained in paper III. The work carried on in this thesis involved an GEM readout developed by a Japanese group [31].

### 3.3.1 Gas Electron Multipliers (GEM)

The GEM (Gas Electron Multiplier) technology, which has been introduced by F. Sauli in 1996, is in use in high energy and medical physics. The standard CERN GEMs consist of an insulator made of a thin Kapton foil (about 50 $\mu$m) which is coated on both sides with copper layers (about 5 $\mu$m). This structure is perforated with holes that typically have a diameter of 70 $\mu$m and a pitch of 140 $\mu$m (see fig. 3.7).

The working principle of the GEMs is shown in fig.3.8. Between the two copper coatings a voltage of a few 100 V is applied. Since the field lines are focused in the holes, there the resulting electric field strength is in the order of some 10 kV/cm which is high enough for the avalanche amplification. It is possible to achieve an amplification up to ten thousand, but usually a setup consists of two or three successive GEMs which are operated at a lower gain and therefore run more stable.

Most electric field lines end on the side towards the cathode while on the other side most lines go into the direction of the anode. The ions from the gas amplification are

| | |
|---|---|
| Size | $\phi$ =3.6 m, $L$ =4.3 m outside dimensions |
| Momentum resolution (3.5T) | $\delta(1/p_T) \sim 9 \times 10^{-5}$ / GeV/c TPC only ($\times$0.4 if IP incl.) |
| Momentum resolution (3.5T) | $\delta(1/p_T) \sim 2 \times 10^{-5}$ / GeV/c (SET+TPC+SIT+VTX) |
| Solid angle coverage | Up to $\cos\theta \simeq 0.98$ (10 pad rows) |
| TPC material budget | $\sim 0.04X_0$ to outer field cage in $r$ |
| | $\sim 0.15X_0$ for readout end caps in $z$ |
| Number of pads/timebuckets | $\sim 1 \times 10^6/1000$ per end cap |
| Pad size/no.pad rows | $\sim$ 1 mm $\times$ 4 $-$ 6 mm/ $\sim$ 200 (standard readout) |
| $\sigma_{\text{point}}$ in $r\phi$ | $<$ 100 $\mu$m (average over $L_{sensitive}$ , modulo track $\phi$ angle) |
| $\sigma_{\text{point}}$ in $rz$ | $\sim$ 0.5 mm (modulo track $\theta$ angle) |
| 2-hit resolution in $r\phi$ | $\sim$ 2 mm (modulo track angles) |
| 2-hit resolution in $rz$ | $\sim$ 6 mm (modulo track angles) |
| dE/dx resolution | $\sim$ 5% |
| Performance | $>$ 97% efficiency for TPC only ($p_T$ >1 GeV/c), and |
| | $>$ 99% all tracking ($p_T$ >1 GeV/c) |
| Background safety factor | Chamber will be prepared for 10 $\times$ worse backgrounds |
| | at the linear collider start-up |

Table 3.1: Goals for performance and design parameters for an LCTPC with standard electronics [32].

thus pulled to and collected on the GEM surface while most of the electrons are extracted out of the GEM holes. The intrinsic ion back-drift suppression is one of the main advantages of the GEMs and may make a gating grid unnecessary. Another advantage of the GEM amplification is the fast signal: the electron signal is measured directly on the pad plane, which leads to a better time (consequently $z$) resolution, and two-track resolution in the time dimension as there is no slow ion tail.

The GEM signal on the pad plane is narrower than in the case of proportional wires. This is an advantage in the two-track separation, but poses new challenges to fully use the potential $r-\phi$ resolution. In order to achieve a space resolution better than the size of the pads, they have to be smaller than width of the electron signal. However, the number of pads is limited by the feasible number of readout channels, so the challenge becomes to develop electronics that can readout many channels on a small space.

### 3.3.2 Design goals for the ILC TPC

The performance and characteristics of the TPC concept for ILD are summarised in table 3.1

A large prototype for this TPC (LPTPC) was built for testing different readout technologies. It can accommodate GEM or MicroMeGA readout systems developed by different groups. The prototype is essentially a cylindrical field cage, where the anode end cap can be instrumented with the different readout systems. It is 60 cm long, with a 73 cm diameter, and placed in a 1 T superconducting magnet originally designed for the BESS-Polar balloon experiment [33]. The electrodes provide a 230 V/cm electric field, and the chamber is filled with the so called T2K gas mixture (95/3/2 % Ar/$CF_4$/Isobutan).

A detailed description of the LPTPC is given in paper III.

## 3.4 Comparison between the ALICE and ILD TPC designs

The two TPC concepts, for ALICE and ILD, are very similar. They have however differences in some characteristics which are dictated partly by the available technologies, and partly by the physics goals of the two experiments.

The major difference between ALICE and ILD is the strength of the magnetic field (0.5 T for ALICE and 4 T for ILD). ALICE has a focus on QCD, and in particular "soft" QCD. For this it is important to measure low momentum particles, which can be difficult if a very strong magnetic field confine them close to the beam. Another argument for using a 0.5 T magnetic field in ALICE was the possibility to reuse the magnet from the L3 experiment. On the other hand, ILD will focus on high momentum processes in electron-positron collisions. For this, a strong magnetic field is required to bend the particles trajectories even at very high momentum.

The choice of readout technology is partly related to the magnetic field. In ALICE, the low magnetic field allows a rather large transverse diffusion in the drift volume. The resulting signal on the readout plane can get large, and it is therefore unnecessary to have very fine pad size to measure it. The well understood MWPCs offered a fine enough readout. The GEM technology was not yet considered mature enough, and would not have offered a strong enough improvement to the resolution. For ILD, the low transverse diffusion should allow a much finer resolution. Besides, for the high momentum physics considered at ILC, the separation between two very close tracks is important. The wide readout signal of the MWPCs cannot offer this feature. This is why the highly segmented readout associated to MPGD like the GEMs is a much better candidate in that case.

Despite these differences, it could be interesting, now that MPGD technologies have improved to be manufacturable on large scale, to consider the possible improvement of ALICE with new readouts. In particular, the high segmentation in the radial direction could improve the $dE/dx$ resolution of the TPC, and therefore improve the particle identification capabilities of ALICE. However, with the weak magnetic field in ALICE the potential improvement probably does not justify a long study.

# Chapter 4

# The Grid

Across the twentieth century, experimental science, and physics in particular has evolved in the complexity of the experiments to perform, and the number of scientists involved in each of them. In the beginning of the century and before, a handful of scientist in one laboratory would make discoveries with table top experiments, writing the data and calculations on pieces of paper. In the middle of the century, larger and larger experiments required to gather scientists in large research facilities. As computers arose and communication tools improved, the end of the century would see international collaborations of several hundred people, recording data and performing calculations using mainframe computers. At the beginning of the new century, a project like LHC involves several thousands of scientists from many countries in the world, and produces tremendous amounts of data. This new situation requires new tools, such as the Grid.

## 4.1 Computing Challenge for LHC

### 4.1.1 Requirements

An experiment like ATLAS at LHC should observe at full luminosity $4 \cdot 10^9$ collision events per second. The size of such events is of the order of a Megabyte (MB = $10^6$ B). So if every event were recorded, we would expect several Terabytes (TB = $10^{12}$ B)per second. This number is far to large for present computers to handle. A selection is therefore done at different levels in the event recording process, and only a few hundred events per second are saved. This number is the same for ALICE, even though the event rate is much smaller. This results in hundreds of MB stored every second, several Petabytes (PB = $10^{15}$ B) every year.

The data recorded by the LHC experiments have then to be processed and accessed by the thousands of members of the collaborations, all over the world. Such a task requires not only extremely large amounts of storage space and computation power, it also requires a network infrastructure to get access to it. Such an infrastructure can hardly be supported by a single institution, for technical, financial and security reasons. The solution adopted to face this computing challenge is the *Grid*.

### 4.1.2 Distributed solution: the Grid

The Grid is a concept for distributed data and computing. Distributed computing had already been developed and used since the early 1980s, with for example CONDOR [34], but it required a lot of work from the user point of view. The term "Grid" was coined in the late 1990s in analogy with the electric grid. The idea was to provide "pervasive, dependable, consistent and inexpensive access to advanced computational capabilities, databases, sensors and people" [35]. The Grid was meant to provide computing power and storage in a transparent way to the user. The Grid is particularly suitable for the trivially parallel type of computing in High energy physics experiments. Both simulations and analysis of experimental data are handled event by event. Thus the complete processing of a single event can be kept in one processor. Only the results of the lengthy event computations need to be pooled from the large number of processors involved.

The concept of Grid evolved a lot, and different solutions and philosophies were adopted by different user communities and projects. Part of the work done in this thesis was to integrate two very different implementations of the Grid idea:

**AliEn [36]** a software specific to the ALICE experiment, offering to the scientists in ALICE access to a pool of dedicated computer clusters and storage systems, through a central database at CERN.

**ARC [37]** a software developed in the Nordic countries, offering a flexible and very distributed system to many user communities, from biologists to big CERN experiments.

In practise, from the user point of view, the Grid typically works in the following way: The user develops a piece of code to produce or analyse data. They send the code and a list of files to access to the Grid interface. The computing jobs are sent to computer nodes on different clusters, along with the data to be processed, in a completely transparent way to the user. The user retrieves the outputs of the jobs.

## 4.2 The ALICE Grid Environment: AliEn

The AliEn framework works on three different levels:

**Central Services** Services running on a set of computers at CERN, managing and optimising the jobs and data distribution on the Grid.

**VO box** Services running on a front end machine in each cluster providing resources to AliEn, managing the jobs on the cluster.

**Job Agent** Service started by the VO-box on individual computing nodes, to run and manage single jobs.

### 4.2.1 Central Services

The core of the AliEn system is a set of services running on powerful machines at CERN. They provide a single entry point for the users and resources. They provide global job and data transfer management and optimisation.

The key components are three databases:

**Configuration LDAP** AliEn relies on a small LDAP database to store the relevant configuration information. It is mostly manually filled and provides the information about the characteristics of the clusters (size, operating systems, interfaces...), and configurable parameters of the full system.

**Task Queue** A large database registers all the jobs submitted to the ALICE Grid, including their description, the path to the associated files and the current status. It is a large central queue where the thousands of jobs running at all time are kept and sorted.

**File Catalogue** Another keystone of the AliEn architecture is the AliEn file system. A large database registers the location of all the files submitted and produced on the Grid. While the files are stored and duplicated on various sites on the Grid, they can be accessed by the user like a simple local file system.

Relying on these databases, handling of the jobs and data transfers is done with several optimisers. All the information about jobs and data is made available to the user via a monitoring system and web interface based on the MonALISA framework [38].

### 4.2.2 VO-box

At the site level, the workload is managed by a set of services running on one front end machine, called the VO-box. It manages the Job Agents (JAs) which run on the cluster's nodes. It is based on four services:

**Computing Element** It manages the JAs via the local resource management system (LRMS).

**Cluster Monitor** It receives information from the JAs on the sites and reports to the central database. It acts as a proxy for communication between the local services on the site (VO-box and JAs) and the central services at CERN.

**PackMan** If a job needs a package that is not already available on the site, it will install it on the shared file system.

**Monalisa** It collects information about the jobs and reports to the central monitoring system.

### 4.2.3 Job Agent

In the AliEn framework, jobs are not directly submitted to the computing nodes. The VO-box submits so-called Job Agents (JAs) which set up the environment and make sure the node is running properly before requesting actual jobs from the central database via the Cluster Monitor.

### 4.2.4 Work flow

In AliEn, the job submission process (illustrated in Fig. 4.1) is done asynchronously ("pull" model). On one hand, the user uploads files to the global file system and submits

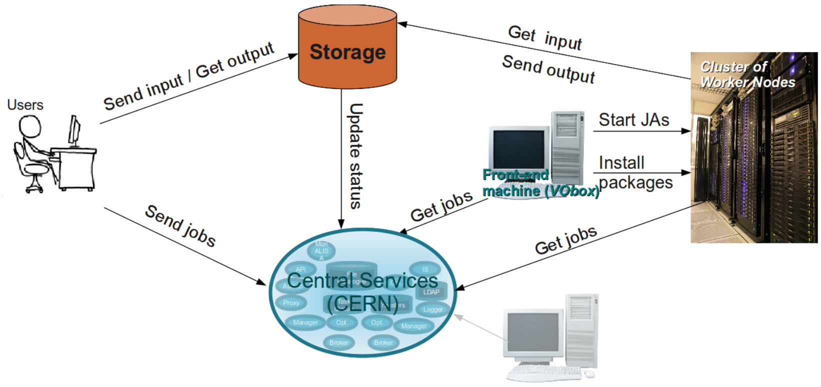



Figure 4.1: Illustration of the work flow for a user job in AliEn.

a job description to the central task queue. On the other hand, the CE in the VO-box creates JAs according to the requirements of the central task queue. If the pending jobs in the queue require new packages, they will be installed before the CE submits JAs.

The JAs connect things together by retrieving ("pulling") from the central database jobs that can run on their computing node. The JA starts the job, reports its status and eventually uploads its output to the global file system, so that it can be downloaded by the user.

## 4.3 The Nordic Grid solution: ARC

### 4.3.1 Structure

The ARC architecture is carefully planned and designed to satisfy the above mentioned needs of end-users and resource providers. The major effort is put to make the resulting Grid system stable by design. This is achieved by avoiding centralised services and identifying only three mandatory components (see Fig. 1):

**The Computing Service,** implemented as a GridFTP–Grid Manager pair of services. The Grid Manager (GM) is the "heart" of the Grid: instantiated at each computing resource's (usually a cluster of computers) front-end as a new service, it serves as a gateway to the computing resource. GM provides an interface to the local resource management system, facilitates job manipulation, data management, allows for accounting and other essential functions.

**The Information System** serves as a "nervous system" of the Grid. It is realised as a hierarchical distributed database: the information is produced and stored locally for each service (computing, storage), while the hierarchically connected Index Services maintain the list of known resources.

**The Brokering Client,** the "brain" of the Grid, which is deployed as a client part in as many instances as users need. It is enabled with powerful resource discovery and brokering capabilities, being able to distribute the workload across the Grid. It

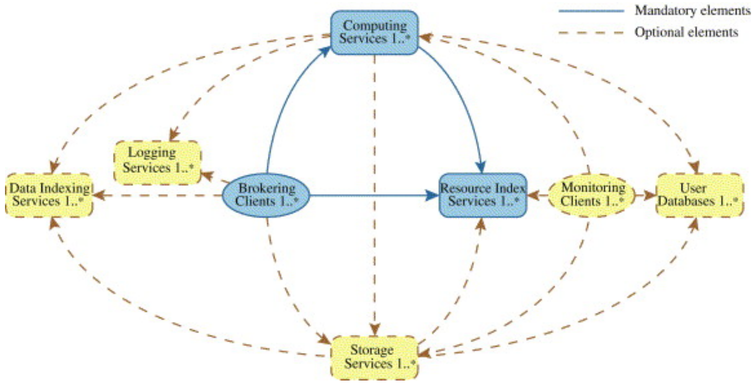



Figure 4.2: Components of the ARC architecture [37]. Arrows point from components that initiate communications towards queried servers.

also provides client functionality for all the Grid services, and yet is required to be lightweight and installable by any user in an arbitrary location in a matter of a few minutes.

In this scheme, the Grid is defined as a set of resources registering to the common information system. Grid users are those who are authorised to use at least one of the Grid resources by the means of Grid tools. Services and users must be properly certified by trusted agencies. Users interact with the Grid via their personal clients, which interpret the tasks, query the Information System whenever necessary, discover suitable resources, and forward task requests to the appropriate ones, along with the user's proxy and other necessary data. If the task consists of job execution, it is submitted to a computing resource, where the Grid Manager interprets the job description, prepares the requested environment, stages in the necessary data, and submits the job to the local resource management system (LRMS). Job states are reported in the Information System. Users can monitor the job status and manipulate the jobs with the help of the client tools, that fetch the data from the Information System and forward the necessary instructions to the Grid Manager. The Grid Manager takes care of staging out and registering the data (if requested), submitting logging and accounting information, and eventually cleaning up the space used by the job.

A non-critical component shown in Fig. 4.2 that nevertheless often comes among the top-ranked in user requirements, is the Grid Monitor. In the ARC solution, it is only a client without capabilities to manipulate resources or jobs, providing information only to the end-users. It relies entirely on the Information System, being basically a specialised user-friendly LDAP browser.

Another interesting feature of the ARC system is the RunTime Environment (RTE). An RTE advertises to the Grid users the presence of a service or a software package. When a package is installed on a site by an administrator, an RTE is associated to it. If the RTE is required by a job, a corresponding script is run before the start of the job, usually setting up the environment to use a package. Another script can be run after the job.

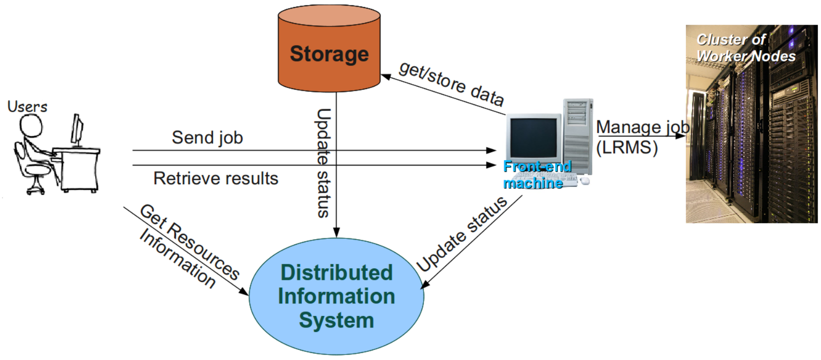



Figure 4.3: Illustration of the work flow for a user job in ARC.

Cluster-wise, ARC implements a pure front-end model: no Grid middleware has to be installed on the worker nodes. It is therefore non-intrusive and easy to handle for the site managers.

### 4.3.2 Work flow

In the ARC job submission process (illustrated in Fig. 4.3), the users submit directly their jobs to the clusters ("push" model). The ARC client gathers information about the available resources (computing and storage) from the Information System. Then, the job brokering is done by the client itself, and the job is directly submitted to the adequate cluster, along with the associated files. There, the Grid Manager will handle the job. If large datasets are used, the Grid Manager downloads or uploads the data from/to a Storage Element. The job is submitted directly ("pushed") via the LRMS. The user can monitor his job and eventually download the output by direct communication between the client and the GridFTP server on the site.

## 4.4 Interoperability

### 4.4.1 The concept

Different regions and communities came up with very different Grid solutions. However, the goal of the Grid remains to make as much resources as possible available to as many users as possible. Therefore, a concept has emerge of "interoperability". The point is to design the different Grid systems with interfaces to connect them together and allow the users of one Grid middleware to access the resources provided through another middleware. Such a system would allow both user and site administrators to choose their Grid middleware according to their own criteria, and still reach the same community.

The idea of interoperability has been discussed a lot, but can be quite difficult to implement, and even to define properly in different situations. Interoperation between AliEn and ARC has been studied in paper II. AliEn and ARC models deploy opposite architectural approaches, and as such are not easily compatible. While the AliEn model is preferred by scientists, the ARC model is preferred by the computing resource owners. Combining the two would give to scientists in ALICE access to more computing resources. In that specific case, the solution proposed is an interface on the AliEn VO-box to use ARC as a LRMS. This system would create a virtual AliEn site, which, while appearing as a single site in the AliEn monitors, would actually be constituted by distributed resources managed with the ARC middleware. This way, the Nordic sites, benefiting of the local ARC expertise, could provide resources to the ALICE experiment without requiring advanced knowledge of the AliEn middleware.

### 4.4.2 The AliEn-ARC interface

The AliEn-ARC interface is a module that had been introduced in the AliEn code [39], and that I updated and extended for more stable use. The principle is described in paper II. I will describe here the main feature of the code itself. It is mainly contained in two files of the Perl libraries of AliEn: `LQ/ARC.pm` for JA submission and `PackMan/ARC.pm` for package management.

#### Job Agent submission with ARC

The JA submission tool of the interface is quite simple, and follows the structure of the submission modules for other LRMSs. A first version of the module already existed in the AliEn code. It relies mostly on two functions:

`generateXRSL` creates the ARC job description file which will start the AliEn JA on the nodes, taking into account the requirements given by AliEn. In particular, it can require the presence of the adequate software packages.

`submit` submits as many JAs as required by the VO-box to the list of ARC sites specified in the VO-box configuration. AliEn currently only uses the local site, but the functionality to extend to several sites has been tested on a small controlled environment.

The module also include the functions required by AliEn `getStatus`, `getOutputFiles` and `kill`.

I updated the code to the newer functionalities of AliEn. I also added functions to improve the tracking of running jobs.

#### Package Management module

The PackMan module was my main effort in the interface, and had to be entirely created, and required some minor upgrades in the generic AliEn code. Its purpose is to allow the VO-box to install software packages on remote sites. The basic AliEn software already offers the functionality to install the packages locally, on a shared file system.

The module has two major functions:

`installPackage` submits a job to all the sites listed in the VO-box that will install the package on a shared file system there. The installation is done using the same script as for the normal local installation. The job is submitted by a “PackMan” user with special privileges to write on the shared file system. The job will also create an ARC RTE corresponding to the package. In case of failure, a list is kept of the non-installed packages.

`getListInstalledPackages` checks the list of packages available from the VO-box. It does so by reading the available RTEs on the sites. In the AliEn system, if a required package is not available from the VO-box, the VO-box will try to install it.

Another function is required to handle errors. At regular time intervals, the VO-box checks if packages failed to be installed. This is done by the `getListNotInstalled` function. If a package still needs to be installed, `installPackage` is called. At each installation failure, an exponential waiting time is assigned for the package. This avoids extra load from chronic errors.

The PackMan module was tested on a dummy AliEn Grid, with Central Services and VO-box installed on my desktop computer, and access to two ARC clusters in Lund and Aalborg. For operational reasons, as LHC started, ALICE could not afford to risk implemented this solution on a large scale.

# Chapter 5

# Statistical pion identification using the TPC dE/dx in ALICE

In the end of chapters I and II, I have shown that there is a lot of interest in identifying particles at $p_T$ higher than 3 GeV/c. This corresponds to the region of the relativistic rise for the TPC as described in chapter III. Here I will show the preliminary analysis we have performed using the TPC information in ALICE data to perform statistical particle identification on the relativistic rise. The first performance results for pp data of the method presented here was presented at the Epiphany conference in Krakow in January 2010, and published as proceedings (paper IV). The analysis presented here extends these results to Pb–Pb collisions, which could be used for example to calculate $R_{AA}$ for identified particles

This is the continuation of a project started by Alexandru Dobrin for his PhD thesis [40]. I adapted some details of the method and applied it to Pb–Pb data produced by ALICE at the end of 2010.

## 5.1 Method

In the relativistic rise region, the $dE/dx$ resolution of the TPC does not allow to separate clearly signals from pions, kaons and protons. However, a fit to the signal can give the yields of the different species. In this study, we fit the TPC signal in bins of transverse momentum $p_T$ with the sum of four Gaussian functions corresponding to pions, kaons, protons and electrons. The main challenge is to provide the best constraints to the fit to get a stable and accurate result.

### 5.1.1 Event selection

The very first step of the analysis is to retrieve the relevant information from the wealth of data made available on the Grid. Using the AliEn framework, we loop over the data and store the information we need.

The calibration of the detectors requires the analysis of actual data to be optimised. The calibration is therefore an iterative process, where calibrated data is used to figure out finer details required for a better calibration. In ALICE the data is generally processed twice. The first reconstructed data, using the basic calibration, is called pass1. The first iteration of reprocessed data is called pass2. The procedure can be iterated, but due to the high cost in computing resources, it is only done in special cases. The analysis presented here was performed on the early pass1 data for heavy ion runs.

The events are filtered using standard criteria to have clean data comparable to other studies within the ALICE collaboration. For each event, we store all the tracks within the TPC acceptance, with their associated momentum $p$, pseudorapidity $\eta$, TPC signal $dE/dx$ and number of clusters recorded in the TPC $N_{cl}$. When running on simulated data, we add the "true" value of the momentum and the particle species.

### 5.1.2 Calibration

Even though the TPC signal is calibrated before it is written in the reconstructed data on the Grid, we noticed that an $\eta$ dependence remained after the corrections. We assume that this variation does not depend much on the particle's transverse momentum $p_T$. We select a small region in $dE/dx$ and $p_T$ corresponding to the minimum ionising region for pions (MIP), where the TPC signal has only a very small dependence with $p_T$. The mean value of the TPC signal in this MIP region is shown as a function of $\eta$ in Fig. 5.1.

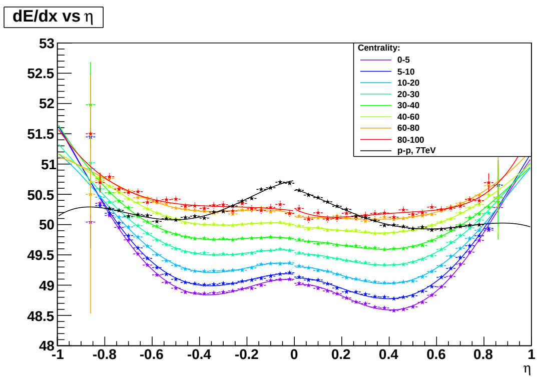



Figure 5.1: The mean value of the TPC signal in this MIP region as a function of $\eta$ for pp and different centralities in Pb–Pb. Fits to the data are used to calibrate the TPC signal for the rest of the studies.

Because of the importance of ion tails (see section 3.2.1) in the TPC with a high track density, Pb–Pb collisions, the $dE/dx$ is expected to depend slightly on the centrality. This centrality dependence is not taken into account in the global calibration of the data. In Fig. 5.1, we show the $\eta$ dependence of the TPC signal for different centralities in Pb–Pb collisions. The dependencies are only of the order of a few percent, but need to be taken into account at the level of accuracy required for PID.

Fits to the curves shown in Fig. 5.1 are used to correct the TPC signal in the rest of this analysis. The reason for the lower $dE/dx$ in central events is that the tail cancellation produces an undershoot of the baseline from which the second pulse has to grow.

### 5.1.3 Parameters constraints

The main challenge of this analysis is to provide good constraints to fit the TPC signal. A Gaussian fit to the TPC signal from one species requires a mean, a width and an amplitude. The mean of TPC signal should correspond more or less to the $\langle dE/dx \rangle$ described by the Bethe-Bloch formula shown in section 3.1.2. A small difference may arise from the fact that the TPC signal, instead of the mean value of the $dE/dx$ in the track, is taken as the truncated mean of the 60% lowest signals, which is closer to the most probable value and provides a better resolution. The width of the signal is the resolution of the TPC signal. Finally, the amplitude is found by the fit and the area gives the yield of the particle considered, and is the only value we are interested in for the final results.

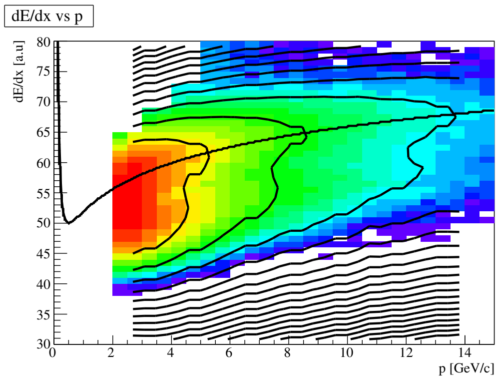



Figure 5.2: Example of 2D fit to the TPC signal as a function of $p$. The curve shown is the mean value obtained for pions. It is subtracted from the TPC signal to make $\Delta_\pi$, used for fits in bins of $p_T$.

In this study, we evaluate the $\langle dE/dx \rangle$ from fits to the data with a Bethe-Bloch like parametrisation in the following way: The distribution of TPC signals is fitted in bins of $p$ with three Gaussian functions, where the amplitudes are free parameters, the widths are fixed to the resolution, and the mean values have to follow a parametrised Bethe-Bloch function for the $\beta\gamma$ corresponding respectively to pion, proton and kaon masses. An example of such a fit is shown in Fig. 5.2. The Bethe-Bloch function is approximated with the following two-parameter function:

$$f(\beta\gamma) = \frac{p_0}{\beta^2} + p_1 \ln 1 + \beta\gamma \tag{5.1}$$

For the TPC signal resolution, we assume it only depends on the number of hits $N_{cl}$ recorded in the TPC for the track. With a naive statistical estimation, we can consider this dependence to go as an inverse square root of $N_{cl}$. However, the TPC signal and $N_{cl}$ are correlated, and the statistical argument is not completely valid. We use a simple fit in the MIP region (as shown in Fig. 5.3) to evaluate the resolution. Due to the ion tails, at high track density (low centrality), the resolution gets worse.

The parameters obtained with these methods do not seem to be accurate enough, and we are working on improving the method. In particular, we use extra information

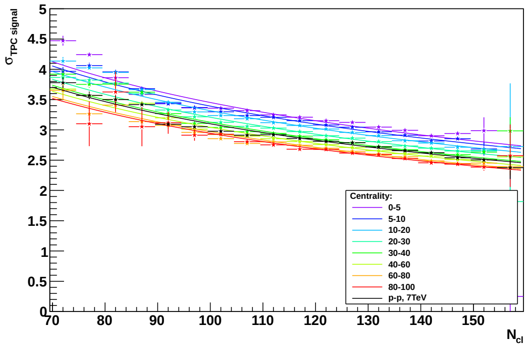



Figure 5.3: Width (resolution) of the TPC signal in the MIP region, as a function of the number of hits in the TPC.

from other detectors like TOF and V0 to test the parameters on samples of identified pions and protons.

### 5.1.4 Fitting

The parameters obtained above are $\langle dE/dx \rangle$ as a function of $p$ and $\sigma_{dE/dx}$ as a function of $N_{cl}$. We want to fit the distributions in bins of $p_T$, with all the tracks. Instead of the TPC signal, we use the value $\Delta_\pi$, where we subtract the $\langle dE/dx \rangle$ for pions obtained from the fits:

$$\Delta_\pi = dE/dx - \langle dE/dx \rangle_\pi$$

To obtain correct fitting parameters, we generate dummy tracks of pions, kaons and protons with $p$ and $N_{cl}$ corresponding to the data, and a random $\Delta_\pi$ following a Gaussian distribution with the parameters extracted. We then fit the distributions for each species in bins of $p_T$ with Gaussian curves to extract the proper parameters taking into account the smearing. For the electrons, which are on the ultrarelativitic plateau, we extrapolate the parameters from lower $p_T$.

The TPC signal in each bin of $p_T$ is finally fitted with the sum of four Gaussian functions with the mean and width fixed to the parameters explained above, and the four amplitudes as free parameters. Examples of the results are shown in Fig. 5.4. One can immediately point at a physics result visible already in the $dE/dx$ spectra. Evidently the p+K peak is much higher relative to the pion peak in heavy ion collisions than in pp collisions. The quality and statistics for the fits are not yet sufficient to judge if the excess is due to protons or kaons or both. We can just remember that a proton excess has been seen in this momentum region at RHIC.

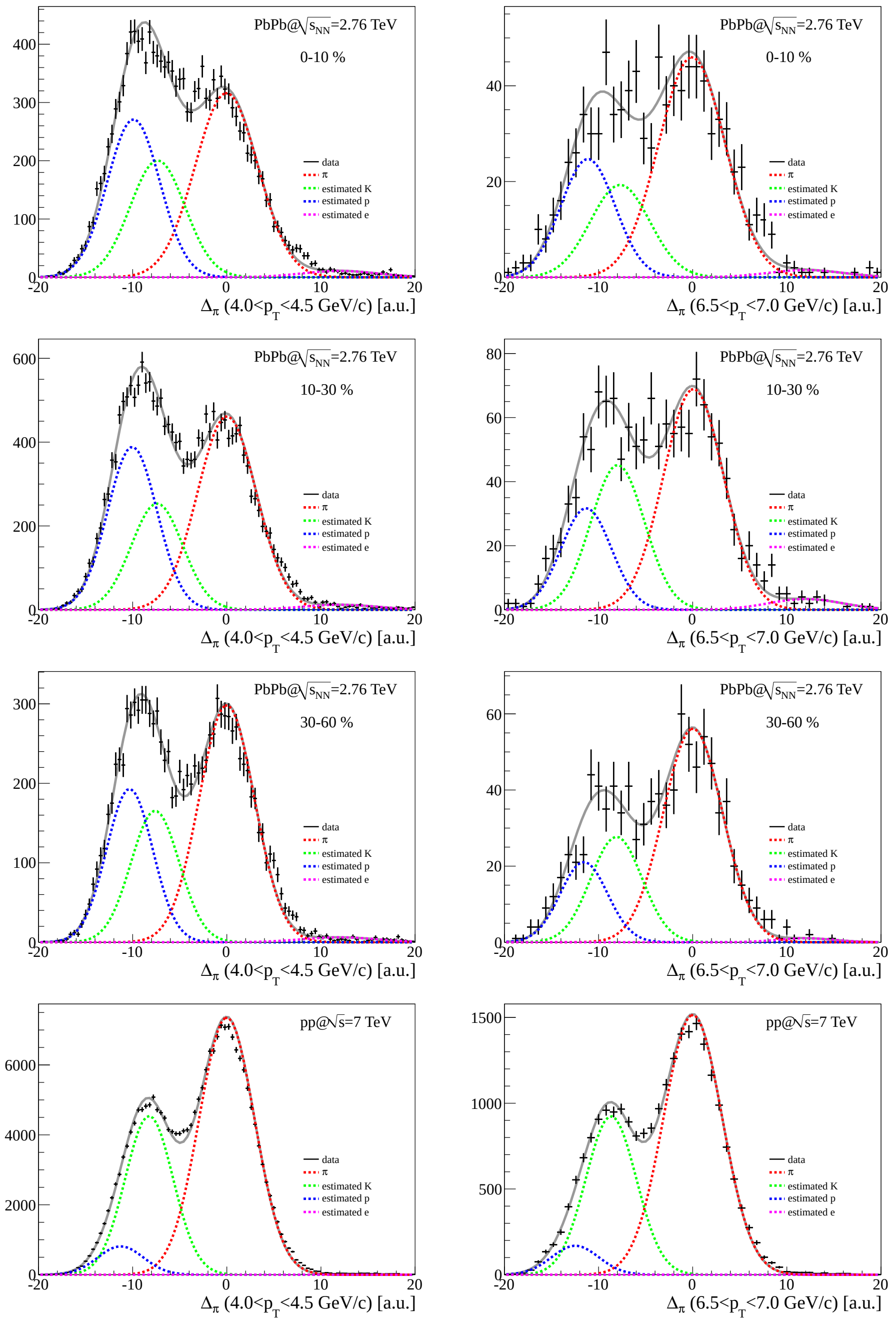



Figure 5.4: $\Delta_\pi = dE/dx - \langle dE/dx \rangle_\pi$ for two different $p_T$ bins for different centralities and for pp. The difference of the pion fraction for different centralities is visible with the naked eyes. The Gaussian fits are used to extract the yield of the different particles.

## 5.2 Results

### 5.2.1 Performance and cross checks

The first step before looking at the result we are interested in, is to check the performance and validity of the method. After many tests of the method we have found that at the present status of analysis we obtain stable results for the pion yield from the fits. We thus have to restrict the study to pion identification.

First we check the ratio between positive and negative pions. The ratio is expected to be close to unity in this $\eta$ and $p_T$ region. The measured ratio shown in Fig. 5.5 being close to one indicates that the method is consistent for positive and negative particles. The fluctuation consistent with statistical values indicate that the fits to the pions are stable. Also systematic differences may be revealed in the charge ratio but appear in Fig. 5.5 to be overwhelmed by the statistical uncertainties.

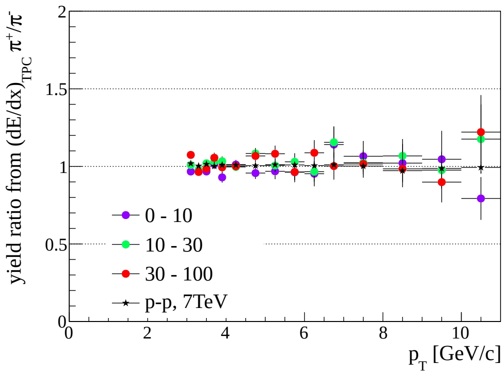



Figure 5.5: $\pi^+/\pi^-$ yield ratio. It is as expected close to unity.

From the mean and width parameters of the four Gaussians, we can calculate the so-called separation power of the TPC signal for a pair of particle types. The separation power $s_{ab}$ between two species $a$ and $b$ is defined as:

$$s_{ab} = \frac{\langle a \rangle \langle b \rangle}{\sqrt{\sigma_a \sigma_b}} \tag{5.2}$$

The larger the separation power, the easier it is to distinguish the two species. Figure 5.6 shows the separation power between pions and protons using our method. It is calculated from the parameters used for the fits. One can see that the separation power seems to reach a plateau above 6 GeV/c. If the separation power is stable, the performance of the PID method will only depend on the amount of statistics. The method can therefore be extended to higher $p_T$, when more statistics will be available.

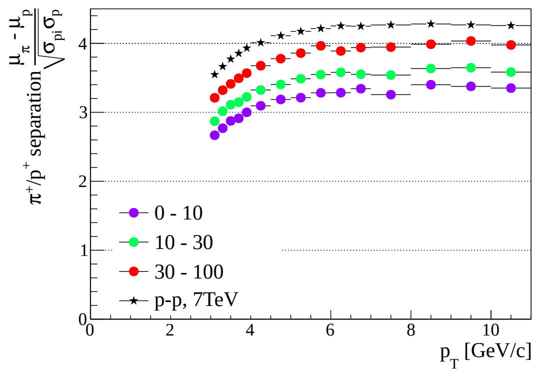



Figure 5.6: Separation power between pions and protons using our method. At $p_T \simeq 3$ GeV/c, it is difficult to distinguish pions and protons. For central Pb–Pb, the large track density in the TPC worsens the resolution and makes the separation more difficult. At high $p_T$, the separation does not change, and only statistics limits the PID performance.

### 5.2.2 Pb+Pb Results

From the fits, we obtain reliably (weakly depending on the fitting parameters) the pion fraction. This fraction is shown for Pb–Pb collision at $\sqrt{s_{NN}} = 2.76$ TeV and pp at $\sqrt{s} = 7$ TeV in Fig. 5.7. It is the direct result from the fits on the TPC signal, and is not corrected for any physics effects like secondary particles or feed down protons. However, it is already possible to see strong effects in the centrality dependence of the fractions. Evidently the pion fraction is constant with $p_T$ for pp collisions. For Heavy Ion collisions, there are relatively fewer pions at momenta 3-6 GeV/c. At larger $p_T$, the heavy ion ratio approaches that of pp collisions. The pion deficiency is clearly larger in central than in peripheral Pb–Pb collisions.

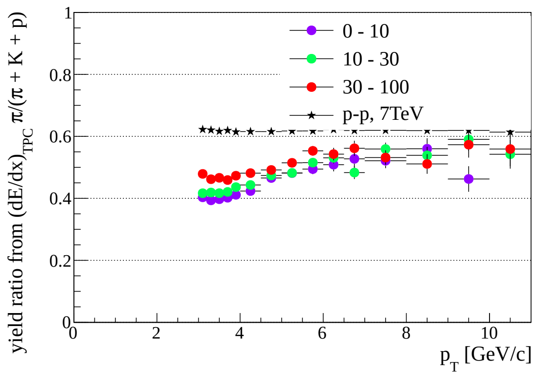



Figure 5.7: Fraction of pions in collisions for pp and different centralities for Pb–Pb.

The pion fraction, combined with the total charged hadron spectrum, can be used to obtain the $p_T$ spectra for charged pions as shown on Fig. 5.8. The $p_T$ spectrum is only corrected for global tracking efficiency with Monte Carlo using Pythia [41] and the full description of the ALICE detectors.

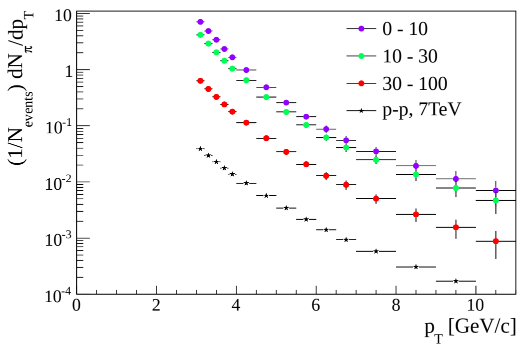



Figure 5.8: $p_T$ spectrum for identified pions in pp at $\sqrt{s} = 7$ TeV and Pb–Pb at $\sqrt{s_{NN}} = 2.76$ TeV for different centralities. The $p_T$ spectrum is only corrected for global tracking efficiency with Monte Carlo using Pythia [41]. Only statistical error are shown.

### 5.2.3 Summary and Outlook

Statistical particle identification using the TPC signal on the relativistic rise shows great promises. The rough method presented here still needs improvement and tuning, and the result should be corrected for physical backgrounds. In spite of all that, it already offers a good impression at the effects of centrality in hadron production in heavy ion collisions.

Verification of the $dE/dx$ shapes can be done by safe pions and protons from $K^0_s$ and lambda decay. Charged kaons could be identified by the kinks at the decay point. All these improvements however require large statistics to be really fruitful. Since the TOF detector has as large coverage as the TPC, it will be useful to improve the particle identification and the constraints of the statistical method in the problematic low momentum region.

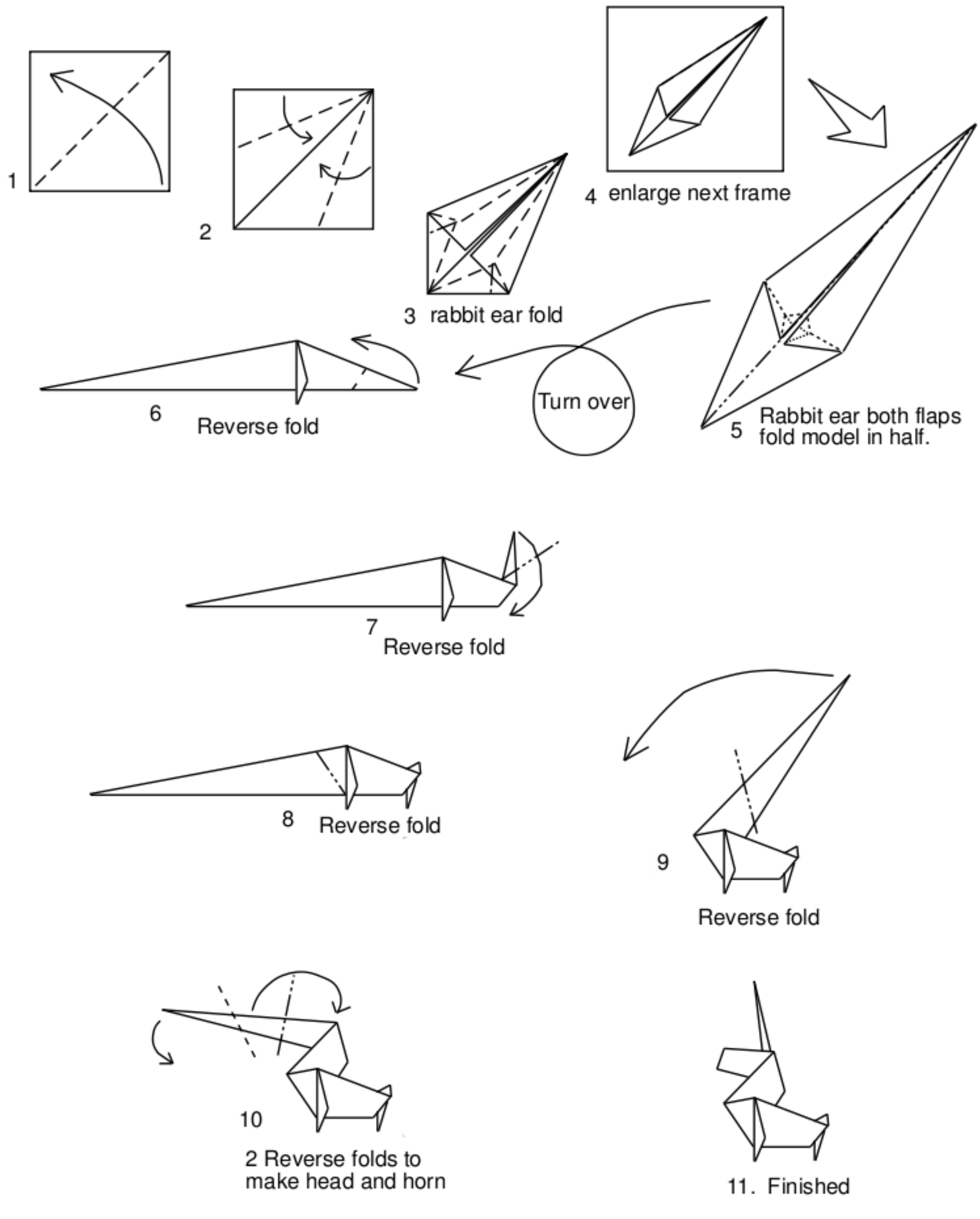
Finally, one should emphasise that the combined result of the TPC and the TOF PID may give quite good results together where one of the methods alone would be inadequate for clean identification.

# Appendix A

# Make your own Higgs Boson

If you lost track during the dissertation, and need to do something with your hands, here are instruction to produce your very own Higgs boson[1].

[1]Model "Simple Unicorn" by Perry Bailey


1
2
3 rabbit ear fold
4 enlarge next frame
Turn over
5 Rabbit ear both flaps fold model in half.
6 Reverse fold
7 Reverse fold
8 Reverse fold
9 Reverse fold
10 2 Reverse folds to make head and horn
11. Finished

# Bibliography

[1] J. Adams *et al.*, "Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR collaboration's critical assessment of the evidence from RHIC collisions," *Nucl. Phys.*, vol. A757, pp. 102–183, 2005.

[2] K. Adcox *et al.*, "Formation of dense partonic matter in relativistic nucleus nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration," *Nucl. Phys.*, vol. A757, pp. 184–283, 2005.

[3] K. Aamodt *et al.*, "Two-pion Bose-Einstein correlations in central PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *Phys. Lett.*, vol. B696, pp. 328–337, 2011.

[4] R. Snellings, "Elliptic Flow: A Brief Review," 2011.

[5] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, "Glauber modeling in high energy nuclear collisions," *Ann. Rev. Nucl. Part. Sci.*, vol. 57, pp. 205–243, 2007.

[6] K. Aamodt *et al.*, "Suppression of Charged Particle Production at Large Transverse Momentum in Central Pb–Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *Phys. Lett.*, vol. B696, pp. 30–39, 2011.

[7] B. I. Abelev *et al.*, "Energy dependence of $\pi^{\pm}$, $p$ and $\bar{p}$ transverse momentum spectrafor Au+Au collisions at $\sqrt{s_{NN}} = 62.4$ GeV and 200 GeV," *Phys. Lett.*, vol. B655, pp. 104–113, 2007.

[8] A. Adare *et al.*, "Suppression pattern of neutral pions at high transverse momentum in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV and constraints on medium transport coefficients," *Phys. Rev. Lett.*, vol. 101, p. 232301, 2008.

[9] B. I. Abelev *et al.*, "Identified baryon and meson distributions at large transverse momenta from $au + au$ collisions at $\sqrt{s_{NN}} = 200$ GeV," *Phys. Rev. Lett.*, vol. 97, p. 152301, Oct 2006.

[10] "The atlas experiment webpage." http://www.atlas.ch/photos/events-collision-heavy-ion.html.

[11] B. I. Abelev *et al.*, "Centrality dependence of charged hadron and strange hadron elliptic flow from $\sqrt{s_{NN}} = 200$ GeV au+au collisions," *Phys. Rev. C*, vol. 77, p. 054901, May 2008.

[12] A. Adare *et al.*, "Scaling properties of azimuthal anisotropy in Au + Au and Cu + Cu collisions at s(NN)**(1/2) = 200-GeV," *Phys. Rev. Lett.*, vol. 98, p. 162301, 2007.

[13] E. Avsar, C. Flensburg, Y. Hatta, J.-Y. Ollitrault, and T. Ueda, "Eccentricity and elliptic flow in proton-proton collisions from parton evolution," 2010.

[14] K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC," *JINST*, vol. 3, p. S08002, 2008.

[15] G. Aad *et al.*, "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics," *JINST*, 2009.

[16] R. Adolphi *et al.*, "The CMS experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[17] A. A. Alves *et al.*, "The LHCb Detector at the LHC," *JINST*, vol. 3, p. S08005, 2008.

[18] A. Dobrin, "Searches for New Physics with the ALICE Experiment in pp Collisions," *AIP Conf. Proc.*, vol. 1200, pp. 730–733, 2010.

[19] N. Boelaert and T. Akesson, "Dijet angular distributions at $\sqrt{s} = 7$ TeV," *Eur. Phys. J.*, vol. C66, pp. 343–357, 2010.

[20] L. Carloni and T. Sjostrand, "Visible Effects of Invisible Hidden Valley Radiation," *JHEP*, vol. 09, p. 105, 2010.

[21] K. Aamodt *et al.*, "Charged-particle multiplicity density at mid-rapidity in central Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *Phys. Rev. Lett.*, vol. 105, p. 252301, 2010.

[22] K. Aamodt *et al.*, "Alignment of the ALICE Inner Tracking System with cosmic- ray tracks," *JINST*, vol. 5, p. P03003, 2010.

[23] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," *Nucl. Instrum. Meth.*, vol. A389, pp. 81–86, 1997.

[24] W. B. Atwood *et al.*, "Performance of the ALEPH time projection chamber," *Nucl. Instrum. Meth.*, vol. A306, pp. 446–458, 1991.

[25] "Lctpc, a time projection chamber for a future linear collider." http://www.lctpc.org/e8/e57671/.

[26] K. Nakamura *et al.*, "Review of particle physics," *J. Phys.*, vol. G37, p. 075021, 2010.

[27] W. W. M. Allison and J. H. Cobb, "Relativistic Charged Particle Identification by Energy Loss," *Ann. Rev. Nucl. Part. Sci.*, vol. 30, pp. 253–298, 1980.

[28] H. Bichsel, "Approximation methods to calculate straggling functions," *Nucl. Instrum. Meth.*, vol. A566, pp. 1–4, 2006.

[29] H. Bichsel, "Thoughts about nanodosimetry," in *Theory of the Interaction of Swift Ions with Matter. Part 2* (R. Cabrera-Trujillo and J. Sabin, eds.), vol. 46, Part 2 of *Advances in Quantum Chemistry*, pp. 329 – 338, Academic Press, 2004.

[30] J. Alme *et al.*, "The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events," *Nuclear Instruments and Methods in Physics Research A*, vol. 622, pp. 316–367, Oct. 2010.

[31] A. Ishikawa *et al.*, "A GEM TPC End-Panel Pre-Prototype," 2007.

[32] T. Abe *et al.*, "The International Large Detector: Letter of Intent," 2010.

[33] A. Yamamoto and Y. Makida, "Superconducting magnet for long duration flights," Prepared for 5th Workshop on Balloon-Borne Experiment with a Superconducting Magnet Spectrometer, Tsukuba, Japan, 30 Nov - 1 Dec 1994.

[34] M. Litzkow, M. Livny, and M. Mutka, "Condor - a hunter of idle workstations," in *Proceedings of the 8th International Conference of Distributed Computing Systems*, June 1988.

[35] C. Kesselman and I. Foster, *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers, Nov. 1998.

[36] P. Saiz, L. Aphecetche, P. Buncic, R. Piskac, J. E. Revsbech, and V. Sego, "Alien–alice environment on the grid," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 502, no. 2-3, pp. 437 – 440, 2003. Proceedings of the VIII International Workshop on Advanced Computing and Analysis Techniques in Physics Research.

[37] M. Ellert, M. Grønager, A. Konstantinov, B. Kónya, J. Lindemann, I. Livenson, J. Nielsen, M. Niinimäki, O. Smirnova, and A. Wäänänen, "Advanced resource connector middleware for lightweight computational grids," *Future Generation Computer Systems*, vol. 23, no. 2, pp. 219 – 240, 2007.

[38] "Monalisa repository for alice." http://alimonitor.cern.ch/map.jsp.

[39] C. Anderlik, A. R. Gregersen, J. Kleist, A. Peters, and P. Saiz, "Alice—arc integration," *Journal of Physics: Conference Series*, vol. 119, no. 6, p. 062006, 2008.

[40] A. Dobrin, *R-hadron Searches and Charged Pion Spectra at High Transverse Momentum in Proton-Proton Collisions at the LHC using the ALICE Detector*. PhD thesis, Lund University, Lund, Sweden, October 2010.

[41] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

# Paper I

Nuclear Instruments and Methods in Physics Research A 609 (2009) 149–155

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

# The influence of detector effects on TPC performance

P. Christiansen [a,*], A. Dobrin [a], P. Gros [a,*], H.-Å. Gustafsson [a], M. Ivanov [b], M. Kowalski [c], A. Oskarsson [a], L. Österman [a], E. Stenlund [a]

[a] *Division of Experimental High-Energy Physics, Lund University, 22100 Lund, Sweden*
[b] *Gesellschaft für Schwerionenforschung mbH, Darmstadt, Germany*
[c] *Institute of Nuclear Physics, Cracow, Poland*

For the ALICE TPC Collaboration

A R T I C L E I N F O

*Article history:*
Received 3 June 2009
Received in revised form
6 August 2009
Accepted 17 August 2009
Available online 25 August 2009

*Keywords:*
ALICE
TPC
Simulation
Space point resolution
Ionization energy loss
Straggling

A B S T R A C T

TPC stand alone simulations based on different energy loss models are compared to previously reported measurements with an ALICE TPC prototype. The performance obtained from simulations is shown to depend significantly on all the included detector effects stressing that energy loss model calculations alone are inadequate for describing the performance. When all effects are included, a good quantitative description of both the specific energy loss and the space point resolution can be obtained from all models. This shows that the TPC description in the official ALICE TPC simulation is in agreement with the measured test beam data.



## 1. Introduction

The simulation of a Time Projection Chamber (TPC) is important for optimizing hit reconstruction, track reconstruction and PID algorithms and for understanding if the measured detector performance is optimal, e.g., space point and energy loss resolution. As the performance affects the PID and tracking efficiency and momentum resolution, which are used to correct physics data, it is crucial that the performance found in simulations is similar to the measured performance. In this paper we report our results on the influence of detector effects on the simulated performance of an ALICE TPC prototype.

The performance of an ALICE TPC prototype has previously been reported [1]. In that paper, the results of a Photo Absorption Ionization (PAI) calculation by Bichsel [2,3] were compared to measured charge straggling functions and truncated mean distributions. PAI calculations by Hans Bichsel have been used to improve the Particle Identification (PID) capabilities of the STAR TPC on the relativistic rise in both $p+p$ and $Au+Au$ collisions [4,5]. The test beam data with a controlled track geometry and separate PID information (from time-of-flight) provided excellent conditions for investigating to what precision the measured performance could be described by theoretical calculations. These results have also been discussed in Ref. [6]. One of the conclusions in Ref. [1] was that to reproduce the measured energy loss data one has to take into account detector effects, such as gas gain amplification variations and diffusion. In this article we have extended the study of the detector effects and the comparison now also includes the space point resolution which is important for the momentum resolution.

The starting point for the investigations was critique raised by Bichsel in Ref. [3] of the energy loss model used in the ALICE TPC simulation [7]. We compared the results obtained with the full ALICE TPC simulation to the TPC test beam data and found good agreement [8]. This was surprising since one would expect that when the energy loss models are different (see Ref. [3] and appendix of Ref. [8]) they should not both be able to reproduce the data. So we decided to investigate this puzzle more.

The ALICE TPC simulation is directly interfaced with GEANT3. To allow the underlying energy loss model to be changed and to keep all parameters under control, a stand alone simulation was developed with all the relevant detector effects from the full ALICE TPC simulation. This allowed us to use three different energy loss models (ALICE, Bichsel, and GEANT4 [9,10]) and combinations of those. The use of a full simulation, where the output has the same

* Corresponding authors.
*E-mail addresses:* peter.christiansen@hep.lu.se (P. Christiansen). philippe.gros@hep.lu.se (P. Gros).

0168-9002/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2009.08.052

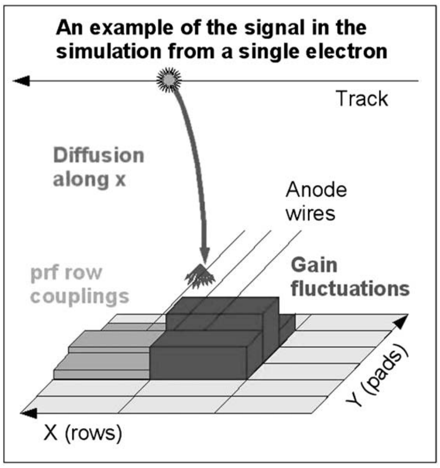



**Fig. 1.** The figure illustrates how the detector effects (steps 4–6, see text) can affect the signal from a single electron and illustrates how the correlations between signals in different rows are introduced. The figure is not to scale. The drift distance in the simulation is 90 cm and the pad plane to anode wire distance is 2 mm.

format as for recorded data (charge per pad and time bin), allowed also the comparison to the space point resolution which gave another cross check of the agreement between simulation and measurements (and the impact of detector effects).

The rest of the paper is divided as follows. First, the stand alone simulation is described in Section 2. Then the results from the simulation are compared to the test beam data in Section 3 and the consequence of excluding detector effects is shown and discussed. Finally the conclusions are presented in Section 4.

## 2. The stand-alone simulation

Here follows a detailed description of the stand alone simulation. The test beam geometry is described in Ref. [1]. In the following the direction along (perpendicular to) the anode wires is denoted $y$ ($x$), see Fig. 1.

1. The mean free path is used to randomly determine the distance to the next primary interaction. For the mean free path the calculation from Bichsel was used [3].
2. The particle is propagated to the point of the interaction and an energy loss, $E$, is randomly determined from the model distribution. The ALICE simulation and Bichsel distributions are both shown in Ref. [3].
3. The energy loss is converted into a total number of ionization electrons, $N_e$, using the expression[1]:

$$N_e = \frac{E - I}{W} + 1 \qquad (1)$$

where $I = 20.77$ eV is the first ionization potential, and $W = 35.97$ eV is the effective energy required to create an electron–ion pair.[2]

4. Each electron is individually propagated to the pad plane of the readout chamber taking into account only the diffusion: $D_T = D_L = 220\,\mu\text{m}/\sqrt{\text{cm}}$ (the tests were done without an applied magnetic field). The diffusion in the $x$ direction introduces correlations between neighboring rows [1].
5. At the wire plane each of the electrons is assigned a random gain from an exponential distribution.
6. The charge is induced on the pad plane according to a two-dimensional pad response function (see below). The time dependence is described by a gamma-4 function with a shaping time of 190 ns and the sampling time intervals are 200 ns.[3]
   The pad response in the $x$ direction also introduces correlations between neighboring rows.
7. Steps 1–6 are repeated until the particle leaves the active volume.
8. For each pad in each row the signals are then digitized for each time bin.[4]
9. In the final step, hits (charge and position) are reconstructed for each row using the same algorithm as in the ALICE software.

The detector effects in our simulations are therefore: diffusion, gas gain amplification variations, pad response, and shaping. The main difference to the previous simulation reported in Ref. [1] is that steps 6, 8, and 9 were not done. This makes a significant difference on the results.

It is important to stress that the parameters used to characterize the detector response have been taken from the ALICE TPC simulation code (see Ref. [7]) and have not been adjusted. The only exception is the absolute value for the gain. This parameter is tuned to match the charge scale between simulation and data.

The pad response function (prf) used in the simulation is factorized into two parts: $\text{prf}(\Delta x, \Delta y) = \text{prf}(\Delta x) \times \text{prf}(\Delta y)$, where $\Delta x$ ($\Delta y$) is the distance from the electron impact to the center of the pad for which the response is calculated. The induced signal on the pads has for all practical purposes the same distribution in the $x$ and $y$ direction, so the difference between the two directions enters when the induced charge distribution is integrated over the pad row geometry to form the prf.

The test beam data were all obtained with an ALICE TPC Inner Readout Chamber (IROC) with a pad size $4 \times 7.5\,\text{mm}^2$. For the response along the pads, $\text{prf}(\Delta y)$, we use a Gaussian with a width of 2 mm (half the pad width). The impact of the electron in the row direction is restricted to one of three anode (amplification) wires for each pad. The pad response function in the $x$ direction is therefore discrete. For the two wires closest to the edge the coupling to the same row is 79% while the coupling to the adjacent row is 21% (no coupling to the row on the far side). For the wire in the center, the coupling is 96% to the same row and 2% to each of the two adjacent rows. For references, see Ref. [7] Chapter 7 and Ref. [11] therein.

## 3. Results and discussion

Fig. 2 shows the comparison of simulated data to test beam data for the ALICE energy loss distribution. Here follows an explanation of the four plots:

[1] Fluctuations of the number of ionization electrons can be taken into account via the Fano factor. However, we found that this did not change the performance results even for Fano factors up to 0.3. This is supposedly because the energy loss variation from collision to collision dominates.

[2] This means that $N_e = 1$ for $E \le I + W$, $N_e = 2$ for $I + W < E \le I + 2W$, and so on. In the case of PAI models one can have energy losses, $E < I$. This always results in 1 electron in our simulation.

[3] The use of a gamma-4 function is different from the ALICE TPC simulation where a Gaussian is used. The difference is irrelevant for the results presented here.

[4] Right before the digitization a Gaussian noise with the nominal $\sigma = 1$ ADC ch is randomly added to each time bin. For the test beam data, where the gain was high, this does not affect the performance and is therefore not included in the discussion of detector effects in the following.

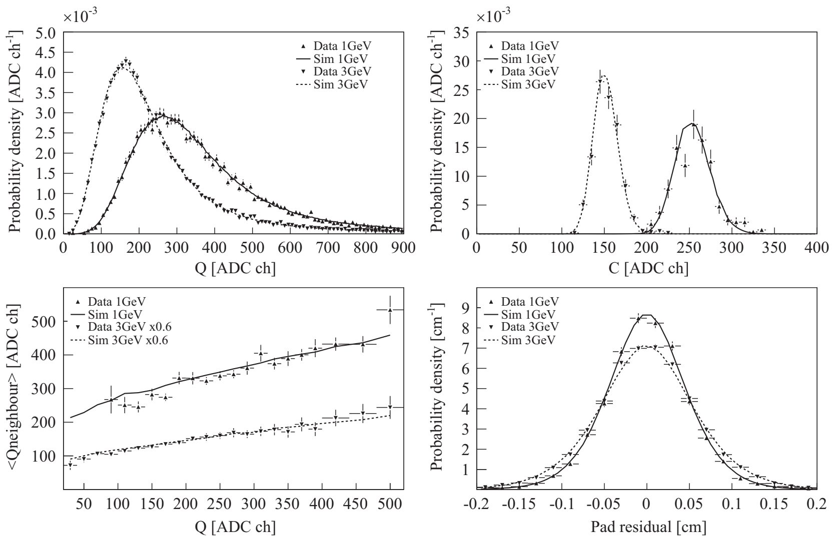



**Fig. 2.** Comparison of the stand-alone simulation using the ALICE simulation parameters (lines) to the experimental data (points) for 1 GeV/$c$ (▼) and 3 GeV/$c$ (▲) protons. For the data there are 294 (1 GeV/$c$) and 614 (3 GeV/$c$) identified tracks. For the simulations there are 40,000 tracks for each setting. The truncated mean is denoted by C. Note that in the correlation plot (bottom left), ⟨Qneighbor⟩ has been multiplied by 0.6 for the 3 GeV/$c$ protons to make the figure clearer.

- Top left: the probability density distribution for the hit charge $Q$. Also called the charge straggling function.
- Top right: the probability density distribution for the truncated mean, $C$, of a track with approximately 60 hits. The truncation is done from 0% to 60%.[5]
- Bottom left: as a function of the charge measured in one row, the mean charge in one of the neighboring rows is shown. The distribution should be flat if the charges are uncorrelated.
- Bottom right: the probability density distribution for the track fit residuals in the $x$ direction (the difference between the fitted track position and the reconstructed hit position).
  The residuals in the time direction ($z$) are similarly distributed and the simulation description is at the same level as for the residuals in $x$. Therefore they are not shown here.

Note that probability densities are per definition normalized to an integral of 1.

The absolute value of the simulated charge has been scaled for each momentum to fit the data points the best. Ideally the scale should be the same, but we find a difference of 7%. This difference could be due to variations in the gas conditions between the two runs (no correction was applied for temperature and pressure variations during the test) or to uncertainties in the absolute momentum scale. The data at the two momentum settings were accumulated in runs lasting less than 10 min (a few hours apart) and there no gain variation was observed.

For the 3 GeV/$c$ data set we made three simulations where we excluded a different detector effect in each simulation. Fig. 3 shows the results. Note that only the diffusion in the $x$ direction has been excluded; diffusion in the $y$ and $z$ directions, which is the primary source of the space point resolution, is unchanged.

Each of the three detector effects investigated affects most of the observables. Here follows the most significant changes. The gas gain variations deteriorate the space point resolution by roughly a factor $\sqrt{2}$ as shown from analytic considerations in Ref. [11] (the RMS changes from 630 to 470 μm in the simulation (third row in Fig. 3). The diffusion shifts the truncated mean distribution relative to the charge straggling function. The coupling between rows in the two-dimensional pad response function improves the pad resolution for the tracks in the test beam data (because the tracks are parallel to the read out plane and perpendicular to the anode wires, so that the pad and time position of the track in each row is similar).

Both the diffusion and the row coupling in the pad response function introduces correlations between hits in adjacent rows, but their effects are quite different. The diffusion shifts a few of the drift electrons that are created in the volume above a row to the adjacent row. This is a stochastic process where either the electron is shifted or it is not shifted and with Poisson fluctuations. The pad response function, on the other hand, smears out the charge.[6] Note that both these effects depend on the pad geometry and both are expected to be stronger for the Inner Readout Chamber (used in the test beam) than for the Outer Readout Chamber of the ALICE TPC.

Part of the aim of the work presented in this article and in Ref. [1] has been to examine to what degree the TPC performance can be described with PAI calculations.

The advantage of PAI model calculations is that they provide both the primary energy loss and the energy loss distribution and therefore provide a tool to eliminate parameters used for the PID that otherwise have to be obtained from the data itself.

We have already here used the mean free path calculated in the PAI model for the results presented above. Now, we use in step 2 of the simulation the energy loss from the PAI model and keep

[5] The truncated mean data shown in this paper are slightly different from those of Ref. [1] because a small normalization problem in the older paper was found—instead of dividing by $n$ samples, we divided by $n+1$. This was done consistently for the data and simulation.

[6] It might therefore be feasible to unfold the smearing due to the coupling between pad rows to improve the energy loss resolution.

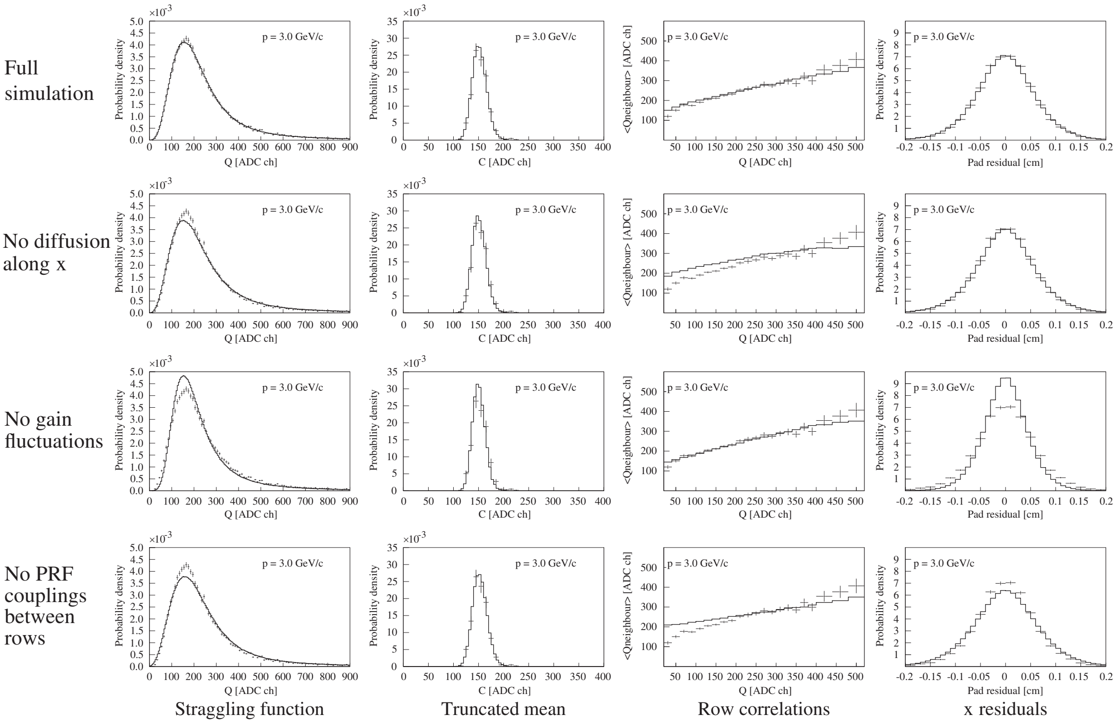



**Fig. 3.** This figure shows a comparison between data and simulation for 3 GeV/$c$ protons with all detector effects included (top) and after excluding one of the three detector effects in the simulation, see text (bottom three plots). Black points are data and solid histograms are results from the simulation (40,000 tracks).

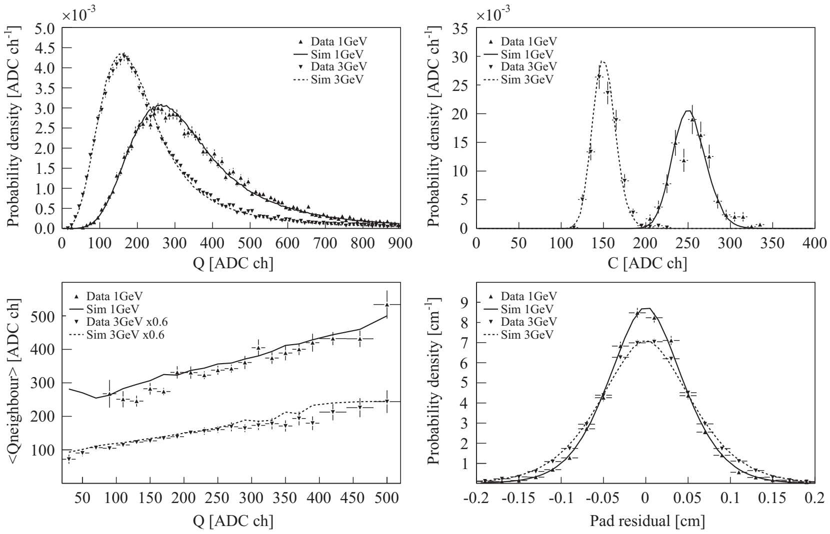



**Fig. 4.** Comparison of the stand-alone simulation using Bichsel's calculations for energy loss for Neon only (lines) to the experimental data (points) for 1 GeV/$c$ (▼) and 3 GeV/$c$ (▲) protons.

all other parameters fixed. Fig. 4 shows the results we obtain with the Bichsel calculation for pure Neon, that was used for the studies in both Refs. [1,3], with the same gas density as that of the full gas mixture (85.7% Ne, 9.5% $CO_2$, 4.8% $N_2$). The description is qualitatively good, but overshoots the peak of the straggling function and undershoots the tail. This difference with respect to the study reported in Ref. [1] is not surprising. In the first study the row coupling in the pad response function was omitted—a situation similar to the bottom row of Fig. 3—and as good agreement was found with that setup one would exactly expect the charge straggling function and the truncated mean distribution to become too narrow for the full simulation presented here.

When the three detector effects are included, the results for the two models are different and the power law energy loss function used for ALICE TPC simulations with the mean free path calculated by Bichsel, does a slightly better job of describing the data than the full Bichsel calculation does.

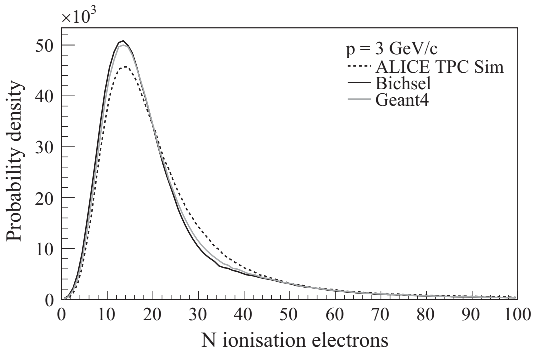



**Fig. 5.** The distribution of ionization electrons per centimeter produced by a 3 GeV/$c$ proton in the three different models: the empirical power law used in the ALICE TPC simulation and the two PAI calculations, Bichsel (pure Neon) and GEANT4 (full gas mixture). The mean free paths in all three models are similar so the main difference is in the single collision energy loss spectrum (which shows that there is almost no difference between the two PAI model calculations as also seen in Figs. 4 and 6).

The total number of ionization electrons produced by a 3 GeV/$c$ proton going through 1 cm of gas in the two models, see Fig. 5, shows a larger difference than for the full simulations. This difference seems to partly decrease due to the detector effects, but also some of the changes are absorbed in the scaling factors (gains). These factors, which are used to match the simulated charge to the data charge, are 3% higher for Bichsel than for the simple power law used in the ALICE TPC simulations.

So the solution to the puzzle mentioned in the Introduction has two parts. One is that when the same detector effects are used for both model descriptions the results are sightly different, but another is that one has the freedom to choose different scale factors and this reduces the differences.

One could ask how sensitive the PAI model results are to the fact that Bichsel's calculation was done for pure Neon. For this reason a comparison was done to the PAI model calculation by GEANT4 for the test beam gas mixture (see above) at the correct pressure and temperature. Fig. 6 shows the comparison to data when GEANT4 calculations have been used in steps 1 and 2 of the simulation. The two PAI calculations give consistent results because they have a very similar ionization spectrum as can be seen from Fig. 5.

However, a GEANT4 calculation for pure Neon was also made with the same gas density, and there the result was very different. This is because in GEANT4 only energy losses with energies above the ionization potential is allowed. This lead Bichsel to redo his calculation for the full gas mixture. Fig. 7 shows the four different model calculations for 3 GeV/$c$ protons. The agreement between the full GEANT4 calculation and the Bichsel calculation for pure Neon seems to be accidental for energy losses below the ionization potential of Neon (~21 eV). The contribution from sub-threshold excitations of the Neon electrons (E<I) in the Bichsel calculation corresponds to the ionization of $CO_2$ and $N_2$ in the GEANT4 calculation. When the calculation by Bichsel for the full gas mixture is used, Fig. 8, the agreement with the data is much worse, particular for the straggling and truncated mean distributions. That these two distributions are affected the most is not surprising, since in Fig. 7 we see that the difference in the

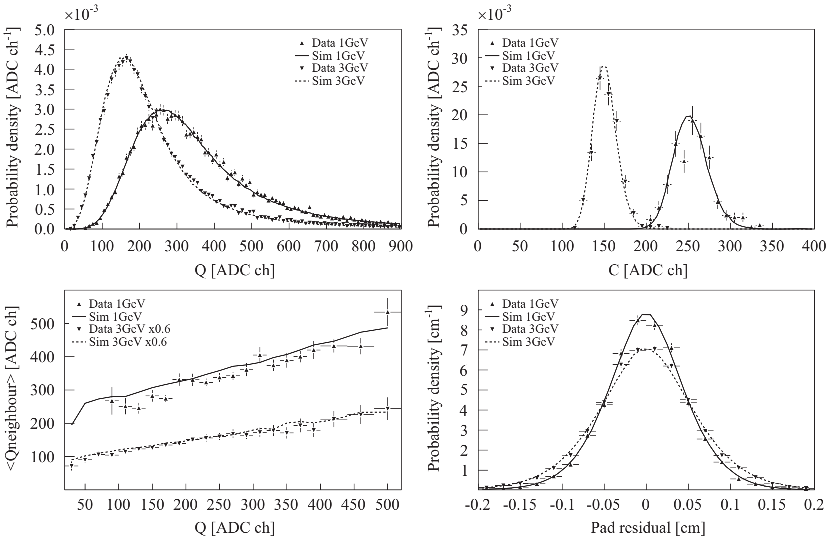



**Fig. 6.** Comparison of the stand-alone simulation using the GEANT4 simulation parameters (lines) to the experimental data (points) for 1 GeV/$c$ (▼) and 3 GeV/$c$ (▲) protons.

simulation is that for the full gas mixture we get several more collisions producing only one electron. As the charge $Q \propto N_e$, while the residuals are proportional to $1/\sqrt{N_e}$ it is clear that the charge variables are affected the most.

Our results suggest that for PAI calculations of ionization, one has to include only energy losses above the ionization threshold as is done for GEANT4. The physics picture could then be that sub-threshold excitations in general deexcite by emitting a photon which is absorbed by the rotational or vibrational bands of the quencher gas and so they only contribute to the energy loss, but not to the ionization.

The previous discussion points at what we believe is the weakest point in the simulation chain—step 3—the conversion from energy loss to ionization. To determine the number of ionization electrons from the energy loss is not straight forward especially for small energy losses. The model used in our simulation is simple and so what we actually find from our studies is that it is the ALICE energy loss function together with the simple model for calculating the number of electrons that does a slightly better job than the other models, i.e., it is the combination of energy loss and ionization that has to be right.

## 4. Conclusions

In this article we have shown how the measured charge and spatial resolution for a specific TPC test beam setup is affected by detector effects. These effects are particularly important for the ALICE TPC IROC which has short pads so that the correlations due to diffusion and charge coupling are large.

The ALICE TPC simulation (with the Bichsel cross-sections) describes the data well because it includes the important detector effects and the ionization in each primary collision is well described. The critique raised by Bichsel is therefore not correct at the level of detector performance.

The studies here indicate that the PAI model description is better when one only takes into account photo absorption cross-sections above the first ionization potential (as done in GEANT4). In this case, PAI model calculations can, at a slightly worse level, also describe the measured performance quantitatively. An open question for simulation is how to convert the energy loss into the number of ionization electrons.

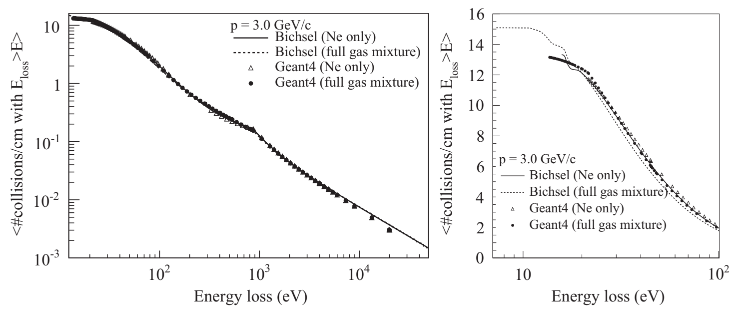



**Fig. 7.** Comparison between the PAI calculations by Bichsel (pure Neon and full gas mixture) and GEANT4 (pure Neon and full gas mixture) for 3 GeV/$c$ protons. The figures show the number of collisions per cm times the integrated probability to have an energy loss of E or larger (e.g.: a 3 GeV/$c$ proton has on the average only ~2 collisions in 1 cm of gas with an energy loss larger than 100 eV). On the left, full energy range, log scale. On the right, low energies only.

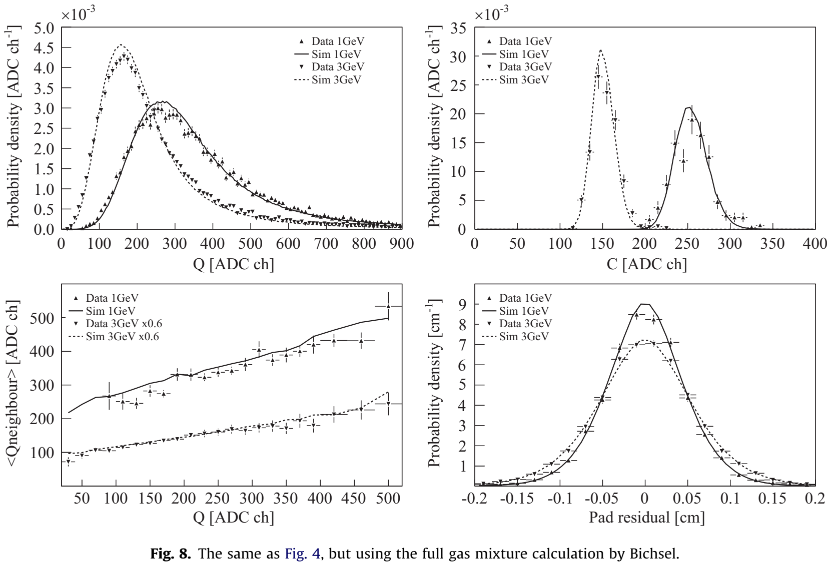



**Fig. 8.** The same as Fig. 4, but using the full gas mixture calculation by Bichsel.

Furthermore, simulations based on PAI model calculations together with models of detector effects provide a valuable tool for understanding the detector performance, and offers in principle a general approach that could be used for any gas detector.

## Acknowledgments

The authors would like to thank Hans Bichsel and Vladimir Grichine for providing the model calculations, and for valuable discussions on energy loss in gases.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.nima.2009.08.052.

## References

[1] D. Antonczyk, et al.ALICE TPC Collaboration, Nucl. Instr. and Meth. A 565 (2006) 551.
[2] H. Bichsel, Nucl. Instr. and Meth. A 562 (2006) 154.
[3] H. Bichsel, Nucl. Instr. and Meth. A 566 (2006) 1.
[4] J. Adams, et al.STAR Collaboration, Phys. Lett. B 637 (2006) 161 [arXiv:nucl-ex/0601033].
[5] B.I. Abelev, et al.STAR Collaboration, Phys. Rev. Lett. 97 (2006) 152301.
[6] P. Christiansen, ALICE TPC Collaboration, Int. J. Mod. Phys. E 16 (2007) 2457 [arXiv:physics/0703097].
[7] ALICE Collaboration, Time Projection Chamber, ALICE TDR 7, CERN/LHCC 2000-001, 2000.
[8] P. Christiansen, et al., ALICE TPC Collaboration, ALICE-INT-2008-015 (public ALICE note), 2008.
[9] S. Agostinelli, et al., Nucl. Instr. and Meth. A 506 (2003) 250.
[10] J. Allison, et al., IEEE Trans. Nucl. Sci. NS-53 (2006) 270.
[11] Y. Belikov, M. Ivanov, K. Safarik, J. Bracinik, in: Proceedings of 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, CA, 24–28 March 2003, pp. TULT011 (arXiv:physics/0306108).

# Paper II

Paper II

# Interoperating AliEn and ARC for a distributed Tier1 in the Nordic Countries

Philippe Gros, Anders Rhod Gregersen, Jonas Lindemann, Pablo Saiz, Andrey Zarochentsev

**Abstract** To reach its large computing needs, the ALICE experiment at CERN has developed its own middleware called AliEn, centralised and relying on pilot jobs. One of its strength is the automatic installation of the required packages.

The Nordic countries have offered a distributed Tier-1 centre for the CERN experiments, where the job management should be done with the NorduGrid middleware ARC.

We have developed an interoperation module to allow to unify several computing sites using ARC, and make them look like a single site from the point of view of AliEn. A prototype has been completed and tested out of production. This talk will present implementation details of the system and its performance in tests.

## 1 Introduction

ALICE [1] is a heavy ion experiment at the Large Hadron Collider (LHC) at CERN. To access the large computing power and storage required for the processing of the data created, the ALICE collaboration has developed a set of grid tools called AliEn

Philippe Gros
Lund University, Div. of Experimental High Energy Physics, Box 118, 22100 Lund, Sweden e-mail: philippe.gros@hep.lu.se

Anders Rhod Gregersen
NDGF - Nordic DataGrid Facility and Aalborg University, Aalborg, Denmark

Jonas Lindemann
LUNARC, Lund University, Lund, Sweden

Pablo Saiz
CERN - European Organisation for Nuclear Research, Geneve, Switzerland

Andrey Zarochentsev
St Petersburg State University, St Petersburg, Russia

(ALIce ENvironment [10, 3]). AliEn allows to access the resources uniformly, and manage them centrally from CERN. The sites are managed using a front-end called a VO-box, which uses the local batch system to manage the resources, without any specific configuration on the nodes themselves.

The Nordic countries are individually small, but represent together an important contributor to the LHC. The decision was taken to pool their resources to create a distributed Tier-1 for the LHC computing grid. An entity called NDGF (Nordic Data Grid Facility [9]) was created to manage the Nordic resources and acts as a single entry point for CERN. The ARC middleware [7], developed by NorduGrid [8] in the Nordic countries was chosen for the management of the resources.

The goal of this work is to create an interface between AliEn and ARC to create such a distributed site. The data management is already unified using dCache [2]. The interoperation solution presented here addresses only the job management.

In the first part, we will introduce the principle and the motivation for such an interface. Then a short description of the two middleware AliEn and ARC is given. In the fourth section, we show the mechanism in the realised interface. Finally, we describe the testing facilities and prototype performance, before concluding.

## 2 Interoperation

In order to create a distributed Tier1, the AliEn and ARC software have to be interfaced. This interoperation should allow to create a machine which would appear as a front-end to the AliEn grid, but can manage jobs over a pool of computer clusters. This task has been initiated (see [2]), but no fully functional interface is used yet.

### *2.1 Motivation*

The Grid has become a very important tool for science, but different strategies are chosen by different communities. We need then to interface the different solutions adopted to optimise the use of computing resources.

The AliEn paradigm is based on a single user (the ALICE collaboration), which receives jobs from the members of the collaboration, and submits them to the computing resources. The steering team at CERN has very much control over the whole system, but since the collaboration is not big enough, some tasks, such as the software installation, had to be automated.

ARC, on the other hand, provides a tool for many users, from different communities (different Virtual Organisations). It is very flexible, but requires involvement of the different scientific communities to provide access to useful resources.

Now when new big projects like International Linear Collider [4] or European Spallation Source, to be built in Lund [14], will start, they will also need grid solutions. If they want to access already existing resources, interoperability will be crucial.

Besides, an interface between AliEn and ARC would bring several advantages for resource management.

**Accounting** The accounting of resource usage is simplified. For the central ALICE management, only one entry exists, so the information would come from a single contact. For the Nordic community, the ALICE site would use the resources in the same way as any other Virtual Organisation, and would be managed with the same tools.

**Operation** In term of operation of the system, the ALICE grid team only has one VO-box to maintain, instead of the current nine. This would make it much easier to provide the level of service required for a Tier-1 centre.
All the sites already run ARC for other VOs like ATLAS. Besides, they benefit of the expertise of the many ARC developers in the Nordic region, while AliEn experts are mainly based at CERN.

**Flexibility** From a technical point of view, such a hierarchical system is more appealing. It should improve the scalability of the system. More clearly it greatly improves the flexibility, allowing to easily move and create new resources at a regional level, keeping only the total to the pledged value.

Of course, it also comes with some downsides. We increase the load on a single gateway machine. However, for the Nordic Tier-1, the combined size of all the federated sites should be comparable to a normal Tier-1 site, controlled by a single machine.

### *2.2 Requirements*

The main requirement of such an interface is simplicity: For ALICE, the new VO-box should not create an exception, it has to be managed with the same configuration and management tools as any other VO-box. The code must be kept as much as possible to some simple modules, non invasive to the normal system.

For NDGF, the ARC "sub-sites" should not run any special services. The site managers should need as little knowledge of AliEn as possible. The VO-specific configuration should be done through an ARC RunTime Environment (see 3.2.2)

## 3 Middleware

This project involves two pieces of middleware, with a hierarchical relation. AliEn is the master, and the very structure of the system has to be understood to create an interface. ARC is used as a client, and only the client's features are relevant, and not the actual structure of the grid.

### *3.1 AliEn*

AliEn is the grid environment developed by the ALICE collaboration at CERN. It is using some centralised services at CERN. The computer clusters are managed by a set of service running on a front-end, called the *VO-box*.

AliEn is using a “pull” model. Pilot jobs, called Job Agents (JAs) which “pull” the actual jobs from the central database.

#### 3.1.1 Job Management

The AliEn job management is done using a pull model and pilot jobs. The computer clusters are managed by a so-called VO-box, running several services to monitor and manage the pilot jobs. The three main services are the Computing Element (CE), the Cluster Monitor (CM) and the Package Manager (PackMan) (see figure 1).

The CE checks the available resources by querying the Local Resource Manager (LRMS) of the site. It also gets the list of packages available on the site from the PackMan. If there are jobs in the central queue matching these, it will send a corresponding number of Job Agents (JAs) to the node using the LRMS. If some extra packages are required, it will ask the PackMan to install them first.

The JAs (pilot jobs) check their environment on the node, and if there is a job matching in the central queue, it will pull it, download the corresponding data and start the job. When the job is finished, it uploads and registers the output. Then it can try to pull new jobs until it expires (usually after 48 hours).

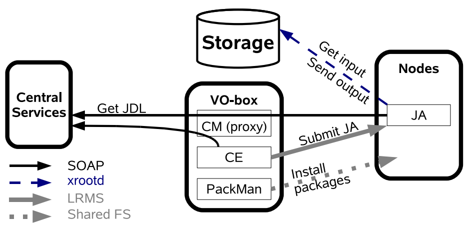



**Fig. 1** AliEn job management process

The CM handles the communications between the different services. It acts as a proxy for the communication with the outside.

#### 3.1.2 Package Management

A strength of AliEn is that it automatically installs the software required by the jobs. The PackMan [6] originally relied on two assumptions: the cluster is homogeneous (all nodes have an identical platform) and the nodes share a file system with the VO-box where the PackMan is running. There is now a prototype [11] that allows the installation of the packages locally on the nodes, thus suppressing these two requirements.

The PackMan, when requested by the CE, downloads the package (tar ball) and installs (unpacks and runs as configuration script) on the shared file system. The configuration script usually just creates a file with the relevant environment variables, to be sourced when using the package.

### *3.2 ARC*

ARC (Advanced Resource Connector [7]) has been developed originally by the Nordic community and has grown into one of the most widely deployed grid middleware. It is a highly distributed, multipurpose grid middleware.

#### 3.2.1 Job Management

In ARC, the jobs are directly submitted from the client (user) to the batch system of a cluster, following a “push” model. It is then usually possible from the client to access the running directory of the job, to see any outputs created. In some site configurations, only the standard outputs are available before the job ends.

#### 3.2.2 Package Management

In the ARC model, the packages are installed on the sites by the system administrators. Then, a RunTime Environment (RTE) script [12] is created to set the environment correctly for a job requiring the package. The RTE also advertises to the grid that the package is available on the site.

## 4 Interface with ARC

We have developed an interface to allow a VO-box to submit and manage jobs using the ARC middleware. ARC can easily be used for the submission of Job Agents on a local cluster. However, to have a single VO-box submitting to remote sites via ARC, the package management has to be adapted.

### *4.1 Job Agent Submission*

On an AliEn site, ARC can be used as a Local Resource Manager (LRMS) to manages the JAs. If the VO-box is set up for a single site (local area network and file system shared with the nodes), this works like any other AliEn site. The corresponding module is now included in AliEn and used in production on several site with high efficiency.

However, if we want the VO-box to submit the JAs to a remote site, communication with the JAs through firewalls has to be considered. Fortunately, all vital communication is done from the JA to the VO-box using SOAP. Therefore, only outbound HTTP connectivity is required, which is the most common situation.

For job monitoring, communication from the VO-box to the JA is done. With ARC, we can read the output produced by the jobs, so that the monitoring of the outputs can still be done. However, more advanced monitoring tools are still unavailable for remote JAs.

### *4.2 Package Management*

In the case of a distributed AliEn site, the PackMan cannot rely on a shared file system to make the packages available to the nodes. The packages are therefore installed on a shared file system on each sub-site, using ARC jobs with special writing privileges.

The installation jobs use an AliEn command to install the package in the exact same way as the PackMan. That way, the package installation is as reliable as on any other AliEn site and the installation script does not require maintenance in itself.

Once the package is installed, an ARC RTE is created. The RTE contains all the relevant environment variables for the package, as created by the PackMan's post-installation script. It also instantly (within the one minute refreshing time of the database server of the ARC grid manager) advertises the existence of the software on the site. The PackMan on the VO-box can therefore check the available packages on its pool of sub-sites. The RTE can then be required when submitting JAs with package requirements.

If the installation script fails, no RTE is created, and a flag is put on the VO-box to retry later.

An AliEn PackMan prototype [11] using bittorrent [13] has been developed. Since it installs packages locally on the nodes, and gets the files from the central repositories and other node through peer-to-peer protocol, it works independently from the VO-box. It could therefore in principle replace efficiently this system for a distributed site.

### *4.3 Implementation*

The AliEn native code is based on Perl. Besides a few minor modifications on the base code, the interface is contained into two Perl modules.

One handles the submission of the Job Agents using ARC. It is very similar to the other corresponding modules for other LRMS such as PBS or SGE.

The second handles the package installation. It is a completely new module, but the AliEn code already allowed the possibility to create and configure such a module.

The configuration of the VO-box in the central configuration database allows to use these modules instead of the default ones. It is also used to configure the system (by giving for instance the list of sub-sites).

## 5 Testing

An interface prototype has been tested in Lund. To avoid interfering with ALICE production, a test bed was created. A more limited version of the interface (not supporting multiple sites) is currently used for production on several Nordic sites.

This test allowed to confirm the viability of the interface, and to address some potential issues.

### *5.1 A Simple Test-bed*

To avoid interfering with the production grid at a critical time, a test bed had to be set up to test prototypes of the interface. A complete AliEn grid system was installed in Lund fig. 2), with a corresponding Virtual Organisation (VO). Its purpose was to validate the principle of the interface in a simple environment, not to test it in heavy load situation.

The test bed had three components:

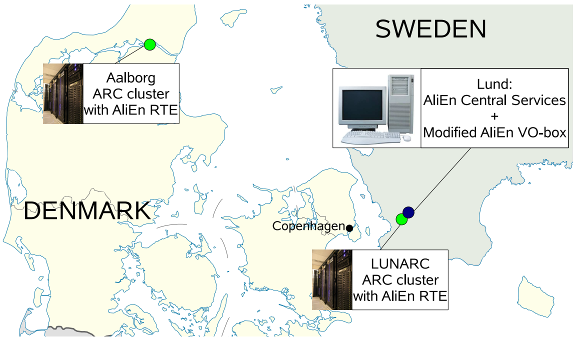



**Fig. 2** Test bed for the interface prototype. Only the computer running AliEn services in Lund is dedicated to the test.

**AliEn Central Services** The basic AliEn Central Services (CS) were installed on a machine in Lund and were associated to a test VO. They were absolutely not modified from the ones used at CERN.

**One AliEn VO-box** A single VO-box was created, on the same machine as the CS. There the interface was included (2 specific modules, and some minor modifications on some services). The site was configured accordingly in the CS database.

**Two ARC clusters running ARC** Two ARC sites participated in the exercise: LUNARC and Aalborg. These sites are used for grid production. On each site, an AliEn RTE was created. Directories were created for the installation of the packages and their associated RTE. Two grid users were added: one for running JAs, with normal user rights; the other for package installation, having writing privilege on the previous directories. A simple plugin had to be added in LUNARC to bypass a configuration that prevented writing possibility for the RTE directory.

### *5.2 Results and Observations*

Jobs requiring ROOT [5] packages were submitted to the AliEn CS. The VO-box installed the packages on the two sites. It then submitted JAs. The JAs ran the jobs successfully, producing outputs.

For the test, the ARC job manager was set to distribute the jobs randomly between the sites. Better algorithm are of course available for a large scale submission.

The test was very successful. However, it appeared that overloaded sites can have minor problems (time out) at the installation stage. This should be improved by setting high priority to the installation jobs. Checking for RTE before retrying the installation greatly reduces the impact of such situations.

Besides, the interface limited to a single sites (and not addressing the question of package management) has been used successfully on several Nordic sites (e.g. Aalborg). The efficiency and stability of the interface is comparable and maybe higher than with direct submission to the LRMS [15].

## 6 Conclusion

It is possible to create a distributed Tier-1 for AliEn. A prototype has been successfully tested in a reduced environment. A limited part of the interface is already used in production. A fully distributed site needs to be tested with high loads before it is used for ALICE production.

Since the ARC functionalities used in the interface are not very middleware-specific, it is probable that the same work could be applied to other flavors of middleware such as gLite or UNICORE.

## References


1. ALICE Technical Proposal for A Large Ion Collider Experiment at CERN LHC, CERN/LHCC/95-71, 15 December 1995
2. C Anderlik et al: ALICE - ARC intergration. J. Phys.: Conf. Ser. 119 (2008)
3. Bagnasco, S. et al.: AliEn: ALICE environment on the GRID. J. Phys.: Conf. Ser. 119 (2008)
4. J. Brau *et al.*, “International Linear Collider reference design report. 1: Executive summary. 2: Physics at the ILC. 3: Accelerator. 4: Detectors,”
5. Brun, R. Rademakers, F. Nucl. Instr. and Meth. A 389 (1997) 81; http://root.cern.ch/.
6. Buncic, P. and Peters, A. J. and Saiz, P., and Grosse-Oetringhaus, J.F.: The architecture of the AliEn system, CHEP 2004, Interlaken, Switzerland (2004)
7. Ellert, M. et al. Advanced Resource Connector middleware for lightweight computational Grids, Future Generation Computer Systems, vol 23, 2007, p. 219-240
8. M.Ellert et al.: The NorduGrid project: Using Globus toolkit for building Grid infrastructure. Nucl. Instr. and Methods A 502 (2003) 407-410
9. Fischer, L., Grønager, M., Kleist, J., Smirnova, O.: A distributed Tier-1. J. Phys.: Conf. Ser. 119 (2008)
10. Saiz, P. et al.: AliEn–ALICE environment on the GRID. Nucl. Instr. and Methods A 502 (2003) 437-440
11. Saiz, P. et al.: Modern methods of application code distributions on the Grid. CHEP09. http://indico.cern.ch/contributionDisplay.py?contribId=452&sessionId=62&confId=35523
12. ARC Runtime Environment Registry. http://gridrer.csc.fi/
13. http://www.bittorrent.com/
14. http://www.ess-scandinavia.eu
15. http://pcalimonitor.cern.ch/map.jsp

# Paper III

EUDET-Memo-2010-36

# Particle tracking and energy loss measurements with the LCTPC: A comparison to simulation models

P. Christiansen, P. Gros, L. Jönsson, M. Ljunggren, B. Lundberg, U. Mjörnmark,

A. Oskarsson, E. Stenlund, L. Österman

Lund University, Lund, Sweden

K. Fujii, R. Yonamine

KEK, IPNS, Tsukuba, Japan

20.12.2010

**Abstract**

The large prototype TPC (LCTPC), instrumented with GEM readout and front-end electronics based on the ALICE ALTRO chip was tested at the DESY T24 electron beam. Data was collected with and without an external magnetic field of 1T, and in both the standard configuration where the tracks are along the long edge (5.2 $mm$) of the pads and in a perpendicular configuration where the tracks are along the short edge (1.2 $mm$) of the pads. For electrons with momentum $p = 5$ GeV/$c$ the straggling function (charge distribution), the specific energy loss ($dE/dx$), and the spatial resolution was extracted and compared to simulations based on GEANT4 energy loss calculations.

The specific energy loss resolution was in general found to be better in the simulation than in the data, due to a narrower straggling function, raising questions about the energy loss modelling of shell effects.

The spatial resolution was found to be well described by the simulations when there was no magnetic field. For magnetic field runs the large $E \times B$ effect due to inhomogeneities in the electric field, caused by the support structure of the GEMs, is found to deteriorate the resolution from 70 $\mu m$ to 100 $\mu m$ even when an average correction is applied. This suggests that a redesign of the GEMs would significantly improve the spatial resolution.

# 1 Introduction

A possible TPC for the ILC has strict performance requirements based on the requirement for the Higgs mass to be determined from the Higgs strahlung processes. The requirement on transverse momentum resolution corresponds to a spatial resolution of better than 100 $\mu m$ at the magnetic field of 4 T.

A broad test beam programs has been carried out within the EUDET framework using a large prototype TPC (LCTPC) with different readout technologies.

In the study presented here the main focus has been on comparing the performance obtained in the test beam measurements to the performance expected from state of the art simulations based on GEANT4 energy loss calculations. The goal of this study is primarily to understand if the performance of the LCTPC is optimal or if it can be further improved, but secondarily it addresses the physics question of to what precision the performance can be simulated.

The note is organised as follows. In section 2-6 the test setup and the detector is described in detail. In section 7-11 the simulation, data analysis, and results are presented. Finally the conclusions of this study are given in section 12.

# 2 The test beam

The measurements were performed at the test beam T24 from the DESY II electron synchrotron [1]. The circulating bunches of electrons are guided towards an internal carbon fibre target to produce bremsstrahlung photons in the energy range up to 7 GeV. The so produced photons are forced to hit an external metal plate to generate electron-positron pairs. Behind the conversion target a dipole magnet causes the electrons and positrons to spread out in the horizontal plane according to their momenta. A collimator cuts out a narrow momentum slice of particles which pass through the collimator. Thus by setting the magnetic field strength, particles with a certain momentum can be selected whereas the energy spread of the electron beam is given by the opening slit of the collimator. A schematic picture of the set-up is shown in Fig. 1.

A telescope of four 3x3 $cm^2$ scintillators with a thickness of 3 $mm$, organised in overlapping pairs at a distance of $\sim$ 100 $cm$ between the pairs, provide a beam trigger [2] in a fourfold coincidence.

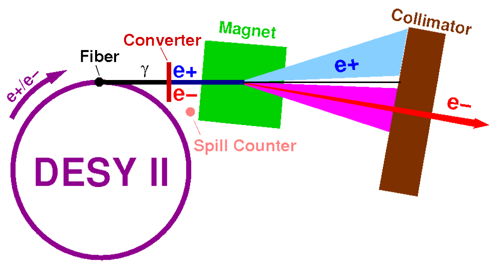



Figure 1: A schematic view of the test beam set-up

## 3 The large prototype TPC

### 3.1 The field cage

The large prototype TPC (LPTPC) [3, 4, 5, 6] consists of a cylindrical field cage and two end caps of which the anode side can be instrumented with various readout systems and the other end contains the cathode. Thus, it is essentially half a normal TPC. The inner diameter of the field cage is 730 $mm$ and is limited by the core of the superconducting magnet (PCMAG) in which it is housed. The length is limited to 600 $mm$ by the homogeneity of the magnetic field, which should not vary more than 3% over the length of the drift space. The field cage is constructed out of composite material. The cylinder is made of a 23.5 $mm$ thick aramid honeycomb wall, which is glued between two layers of glass-fibre reinforced plastic (GRP) and a polymide layer for electrical insulation. This structure provides the necessary stiffness to prevent significant deformations and presents a minimum of material ($\sim 1\%$ of $X_o$) with which the particles may interact. The field cage should be able to stand an overpressure from the drift gas of up to 10 $mbar$. The inhomogeneity of the drift field should not exceed $10^{-4}$ over the drift volume and in order to meet this requirement the electrical field is provided by a series of field strips on decreasing potential from the anode to the cathode. These are 35 $\mu m$ thick copper strips on both sides of a 75 $\mu m$ thick kapton foil, which is glued to the inside of the field cage. The copper strips on the outside are so called mirror strips which are placed to cover the gaps between the field strips, overlapping half a strip width with the inner field strips and are on intermediate potentials. This arrangement of strips provides a homogeneous field to the order which is required. The strips are connected via a resistor chain which distributes the voltage along the drift length with the last strip (counted from the anode) on the same potential as the cathode. The potential of the cathode is about -14.5 $kV$ giving a drift field of 230 $V/cm$. A schematic view of the field cage and a cross section of the cylinder wall is shown in Fig. 2.

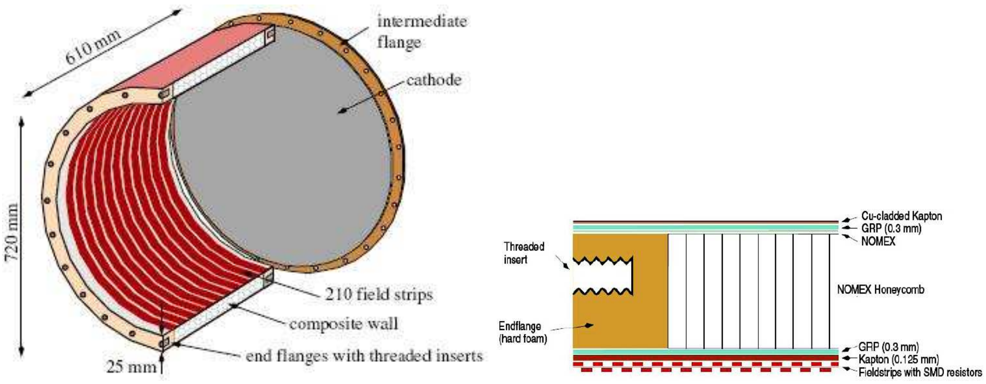



Figure 2: Left side: A schematic drawing of the field cage. Right side: A cross section of the chamber wall.

### 3.2 The end plate

The anode end plate is made from aluminium and the design is such that it should be able to accommodate seven gas amplification modules arranged in a way that corresponds to a subsection of the full scale TPC for the ILC. The accuracy for the insertion of the modules is better than 50 $\mu m$. A picture of the end plate is shown in Fig. 3.

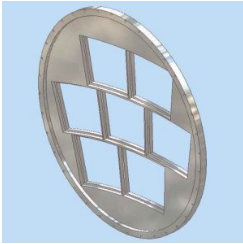


Figure 3: The end plate with the seven mounting positions for the readout chambers

## 4 The GEMs and pad plane

For this measurement, readout chambers based on gas electron multipliers (GEMs) developed by the Asian groups [7, 8] were used and 3200 pads were equipped with the ALTRO readout electronics. The GEM module has a keystone shape with the width of about 23 *cm* and the height of about 17 *cm*. The pad board has 5152 pads in total, arranged in 28 rows of about 1 *mm* wide pads separated by 0.1 *mm*. The length of the pads is 5.26 *mm*. The innermost 14 rows have 176 pads/row whereas the outermost 14 rows have 192 pads/row. Due to the shape of the module the width of the pads is increasing with its radial position so that the pad width varies between 1.15 *mm* and 1.25 *mm*. The pads of neighbouring rows are staggered with respect to each other by half a pad width. A picture of the pad plane is seen to the right of Fig. 4.

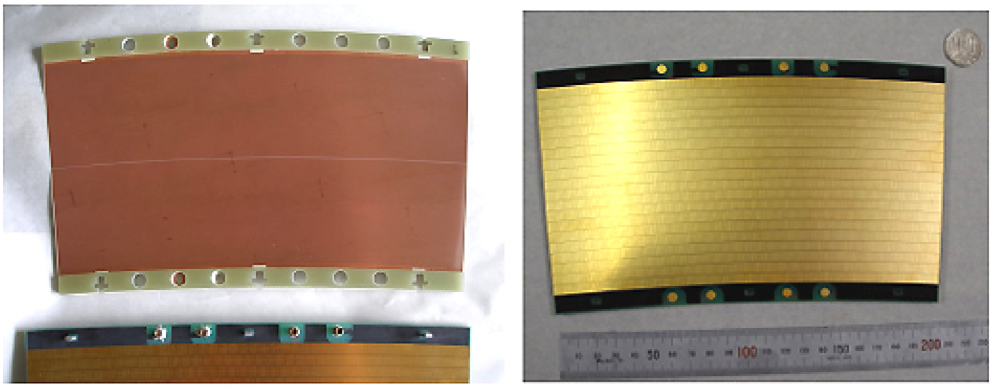


Figure 4: Left side: A GEM plane. Right side: The pad plane. The holes are for support pillars and HV supply.

The GEM modules have two copper layers of 5 $\mu m$ on each side of a 100 $\mu m$ thick liquid crystal polymer insulator. The holes are 70 $\mu m$ in diameter and the pitch between the holes is 140 $\mu m$. A picture of a GEM plane is shown to the left of Fig. 4. The GEM system included two layers of GEMs in this test. The GEM foils are supported by two thin G10 frames along the long sides of the pad board and the GEM, respectively. The short sides have no support, which minimises dead regions between the GEMs in the $\phi$ direction. The distance between the two GEM foils (the transfer gap) and the distance between the last GEM and the pad plane (the induction gap) is 4 *mm* and 2 *mm*, respectively.

Since the short sides of the GEM foil have no support, metallic pillars were glued onto the long sides of the pad board, the purpose of which is to keep the GEM foils stretched and in position. Thus, in the GEM foils and the G10 frames there are corresponding holes into which the pillars fit. An assembled module is shown in Fig. 5. A thin gating GEM was planned to prevent ions to drift back into the drift volume but was not installed at the time of the measurement. The voltage across the thick GEMs was 350 V giving a gas amplification of $10^4$ for the double GEM system.

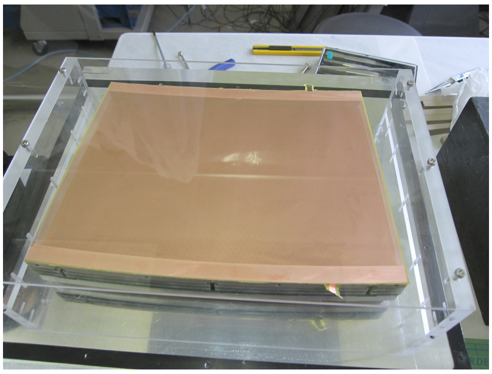


Figure 5: The assembled GEM module

The outermost GEM plane has to be on the same potential as the corresponding strip on that level in order not to cause any inhomogeneities in the drift field. For the Asian GEM module the outermost GEM should have been the gating grid on the level of the seventh strip in the field cage. However, due to the missing gating GEM this had to be accounted for in setting the potentials.

Unfortunately, since the gating grid was never introduced for this measurement, the metallic support pillars were sticking out unshielded, causing deformations of the electric field. This caused a straight track to become bent, where the magnitude of deformation depends on the distance of the track from the pillars. An example of such a reconstructed track is given in Fig. 9.

## 5 The superconducting magnet

The superconducting magnet (PCMAG) [9], which houses the LCTPC, was constructed by Japanese groups originally for the BESS-Polar balloon experiment to study cosmic anti-matter and has therefore a coil of low mass and no return yoke. A technical drawing of the magnet, with the position of the TPC indicated, is shown in Fig. 6.

The diameter of the magnet core is 85 *cm* and the length inside the magnet, with this diameter, is about 130 *cm*. The magnet can provide a field of 1.2 T but was in this measurement run at 1 T. The field is, however, not homogeneous over the full length of the magnet due to the non existing return yoke. Only in a region of $\pm30$ *cm* around the centre it is homogeneous to within 3% and this limits the length of the TPC to about 60 *cm*.

The PCMAG was at the time of these measurement sitting in a fixed position with the beam entering in the centre of the magnetic field so that the influence of the inhomogeneous regions of the field should have a small influence. The only way to measure at different drift lengths was to move the TPC inside the magnet. This is not causing any problems as long as the TPC is pushed further in to the magnet, corresponding to shorter drift lengths, but when the TPC is pulled further out the readout system is entering into the regions of inhomogeneous fields more and more. These measurements were performed at drift lengths in the range of 70 to 300 *mm*.

## 6 The readout electronics

The readout system for the LCTPC is based on the ALTRO chip [10], originally developed for the ALICE experiment at the LHC. The 16 channel ALTRO chip performs analogue to digital conversion with 10 bit precision followed by various steps of digital signal processing, including zero suppression and storage in an event buffer. The sampling can be clocked at frequencies up to 40 $MHz$ so in principle sampling at this frequency and frequencies lower by multiples of two is possible. However, at 40 $MHz$ sampling the full resolution is not maintained for the standard ALTRO chip. A limited number of ALTRO chips were modified to allow sampling at 40 $MHz$ with almost full precision [11]. Up to now the system has been operated during data taking at 20 $MHz$ only. The ALTRO chip has an event storage memory of 1 k 10-bit words, which corresponds to sampled data over a depth of 50 $\mu$s drift time at 20 $MHz$ sampling frequency. The so called T2K gas mixture (95/3/2 % Ar/$CF_4$/Isobutan) was used in the TPC, which gives a drift velocity of around 7 cm/$\mu$s at a drift field of 230 $V/cm$. This leads to a maximum drift length of 350 $cm$ that can be accommodated in the ALTRO memory i.e. much longer than the total length of the LCTPC of 60 $cm$.

In order to test recent technologies for gas amplification (GEMs and Micromegas for TPC readout) a new charge sensitive preamp-shaper has been developed. The programmable PCA16 chip [12] has, as the name indicates, 16 channels and offers different choices with respect to shaping time, gain, decay time and signal polarity. The new analogue chip required modifications to the Front End Card (FEC), compared to its original design for the ALICE TPC. These are mainly related to the programmability of the PCA16 chip which is done remotely. Data for setting the parameter values are downloaded to the board controller FPGA on the FEC via the data bus on the back plane. An 8-bit shift register delivers the digital input to set the peaking time, the gain, the polarity (common to the 8 PCA16 on a FEC) and it also provides a possibility to bypass the shaping function. An octal DAC (Digital to Analogue Converter) controls the decay time of the preamplifier. The various options will make it possible to find the optimal parameter setting for a certain gas amplification system so that the specifications for the final chip production can be quantified.

Each FEC contains 128 channels i.e. 8 PCA16 chips and 8 ALTRO chips are mounted on each

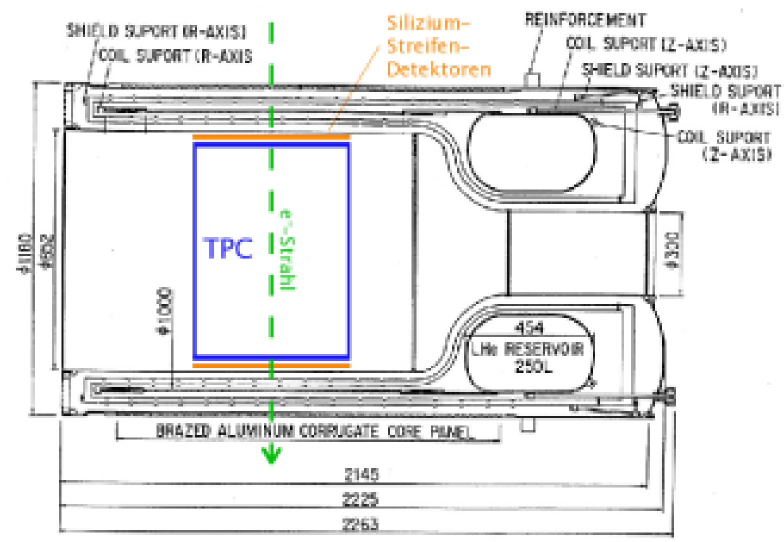



Figure 6: A technical drawing of the superconducting magnet (PCMAG) with the TPC indicated

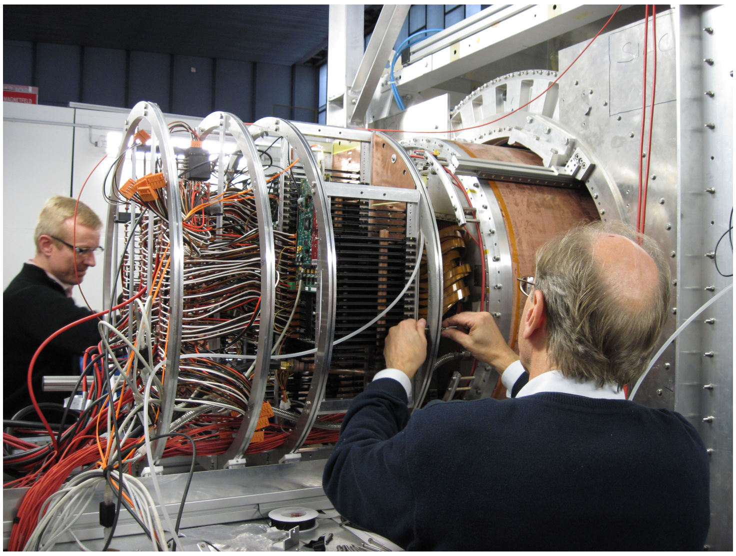


Figure 7: The front end electronics and its support structure attached to the TPC. Also seen are the optical cables for the data readout and the LV supply cables.

board. They are connected to the pad board on the TPC via flexible 30 *cm* long kapton cables. A Readout Control Unit (RCU) [13] governs the readout of the data via the back-plane to which a maximum of 32 FECs can be inserted. Data are sent via an optical cable to a Detector Read-Out Receiver Card (DRORC), which is placed in the Data Acquisition (DAQ) PC. The DAQ software uses the ALICE drivers and libraries for communication between the DRORC and the front end electronics via the optical link. At the reception of the trigger the ALTRO starts storing digitised information in the event buffer, up to a predefined number of samples. The RCU reads the ALTRO event buffer and sends the data on the optical link to the DRORC, which stores it in the memory of the readout computer. The events are stored on a local disk, a fraction of the events are sent to a monitoring program. A photograph of the test beam set-up is shown in Fig. 7.

The electronics pedestal and noise levels for all readout channels have been measured both initially as well as on a regular basis during data taking periods, mainly for pedestal subtraction and to check whether there are corrupt channels. The front end electronics has shown excellent noise performance. with a noise level of 0.5 ADC counts corresponding to an equivalent noise of 260 electrons for the longest shaping time and the lowest gain, where the noise includes random noise, coherent clock noise and long term variations on the scale of seconds. If the gain is increased to the highest value (27 mV/fC) the noise level increases to typically about 1 ADC count, which corresponds to the equivalent noise of 231 electrons at this gain.

## 7 Measurement principle

The first purpose of the test beam was to assess the performance of the LCTPC setup described above. In addition, this setup gave the opportunity to follow up on a study performed on the ALICE TPC using a detailed low level simulation of the detector [14].

The input to these simulations is the energy loss of the charged particle in the chamber gas, obtained from photo absorption ionization (PAI) models [15, 16, 17]. However, an accurate simulation also needs a very good description of the detector effects, such as diffusion, gain amplification variations, and pad response function coupling to different pads/rows. When energy loss calculations are combined with detector effects a good quantitative agreement can be obtained with data, see Fig. 8

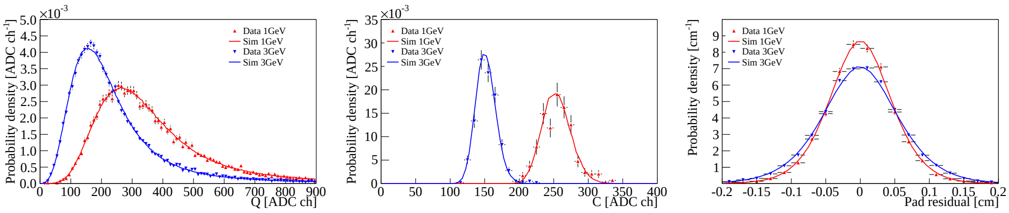



Figure 8: Comparison of simulated data to ALICE TPC test beam data (1 and 3 GeV/c protons) for the straggling function (left), truncated mean ($dE/dx$) (middle) and position resolution (right). Figures are from [14].

For these studies test beam data is of great importance because there the beam geometry is fixed so that one gather data/statistics under essentially the same conditions (diffusion, track angles etc.) and the identity of the beam particles is well known.

The PAI model calculations include atomic shell effects, but these are hard to observe in the data, because typically the pad length is much larger than the mean free path so that one collects the ionization electrons from several collisions smearing out the discrete energies. For this reason a measurement was performed with the LCTPC rotated by 90 degrees with respect to the test beam so that the short side of the pad (1.2 mm) was parallel to the beam direction whereas the long side (5.2 mm) was transverse to the beam direction. These runs take advantage of the fact that the GEM readout, unlike wire readouts, have no preference for track angles (this is of course not true for the pads but as we want to focus on the energy loss in these runs the worsened space point resolution is not a problem).

Using this configuration, we try to get insight on the energy loss process, in particular on the shell effects present in the PAI model.

This approach using low level simulation presents other advantages:

- It provides the basis for simulation that anyway has to be developed for a detector.
- It enables a cross check whether the measured performance is optimal or could be improved?
- It allows for calibration of the energy loss. The gain can be obtained from the matching of simulation and data.
- It fixes parameters for particle identification using the specific energy loss: mean, resolution, and shape of curve.

# 8 Simulation

For this study an independent TPC simulation program developed in Lund for a study of the ALICE TPC [14] has been used. The simulation has proven to give a good description of test beam data with the ALICE TPC. For the LCTPC, only the readout part was modified to describe GEMs instead of wire chambers. The processes describing the ionization of the primary particle and the diffusion in the drift volume are the same. The gas processes are:

1. The mean free path is used to randomly determine the distance to the next primary interaction.
2. The particle is propagated to the point of the interaction and an energy loss, $E$, is randomly determined from the model distribution.
3. The energy loss is converted into a total number of ionization electrons, $N_e$, using the expression:

$$N_e = \frac{E - I}{W} + 1, \tag{1}$$

   where $I = 15$ eV is the first ionization potential, and $W = 25$ eV is the effective energy required to create an electron-ion pair. These values were evaluated from the data in [18][1].

4. Each electron is individually propagated to the pad plane of the readout chamber taking into account only the diffusion. The diffusion constants used were the ones measured with the prototype during the run corresponding to this study.

The simulation of the GEM readout for each drifting electron includes the following processes:

5. At the Gem plane, the electron is assigned a random gain, following an approximate Polya function with factor 1.5.
6. The charge is transfered to the pad plane according to a two dimensional pad response function (PRF). A simple gaussian model is used. Alternative models were tested, but did not seem to induce any change in the results. The width of the PRF was determined from the data. The pad geometry used is exactly the same as used for the data reconstruction.
7. Steps 1-6 are repeated until the particle leaves the active volume.
8. Background noise is added. On top of a random gaussian noise, we add cross-talk in cables. A signal is induced in neighbouring wires in the cables from the pads to the front end card. This effect has been studied from the test beam data itself, as shown is section 10, and the amount of cross-talk was tuned to these results.
9. The data is reconstructed using the same program as for the real data.

The parameters for the energy loss in the gas are directly taken from the models, without any tuning. The parameters specific to the chamber such as the drift velocity, diffusion constant, PRF width and crosstalk were all obtained from different measurements on the test beam data. The only free parameter used to tune the simulation is the absolute value of the gain. It is hard to disentangle the amplification gain from the mean number of primary electrons, and therefore the constraints on the energy loss model are difficult to quantify.

[1]This means that $N_e = 1$ for $E \leq I + W$, $N_e = 2$ for $I + W < E \leq I + 2W$, and so on. In the case of PAI models one can have energy losses, $E < I$. This always results in 1 electron in our simulation.

# 9 Track reconstruction and track selection

## 9.1 Geometry

In the parallel configuration (standard running configuration), the pad plane is oriented with the long side of the pads parallel to the beam direction. The coordinate system used in the reconstruction defines the $z$-direction parallell to the drift direction with the origin at the pad plane, pointing towards the inner of the TPC. The $x$-direction is oriented perpendicular to the pad rows (along the long side of the pads) and the $y$-direction is accordingly parallel to the pad rows (along the short side of the pads). A sketch of the pad layout with a track signal, and the coordinate system can be found in Fig. 9.

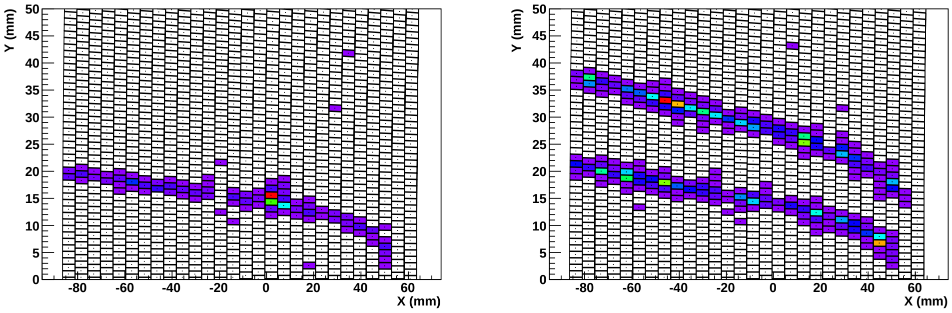



Figure 9: Display of events with one and two tracks, respectively, with the TPC in the parallel configuration. The coordinates of our reference system are also given.

## 9.2 Clustering

The first step of the *clustering* procedure is to reconstruct the pulses collected on individual pads. The charge collected on the pads is sampled at a frequency of 20 MHz and converted to digital information. Thus, the raw data contains the sampled ADC values, channel number and time information of such pulses (fig. 10). A pulse is defined by the hardware as two or more consecutive nonzero samples, where the threshold is set to 3 ADC counts. The largest ADC-value $Q_{max}$ represents the pulse height and is chosen as a measure of the charge. Time information is retrieved by a charge weighted mean of the five samples around the peak of the pulse. The coordinates of the pulse is taken as the coordinates of the pad center on which the charge was deposited.

Two charge clouds close in time and space might accidentally be included in the same pulse, for example when two tracks pass close to each other. To split them, the algorithm compares the heights of all samples included in a pulse. Starting from one end, a peak is found when the value of a sample is lower than the preceding one. A difference of a two ADC values is required to avoid false peaks resulting from noise fluctuations. All samples up to the point where the values start rising again are included in the first pulse. For the values after that, the procedure is repeated should there be more than two pulses.

The reconstructed pulses are used in an algorithm that combines the pulses in a single pad row into clusters. A search is conducted outwards from the pad having the highest charge. If a pulse is found on a neighbouring pad, within a time window around the peak value, it is added to the cluster. The

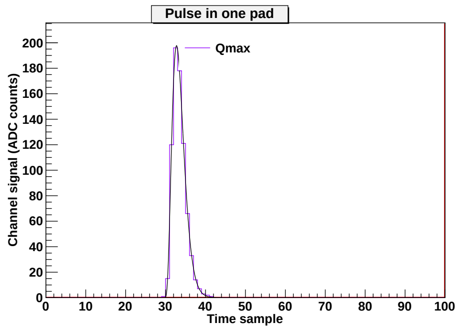



Figure 10: Typical pulse from sampling on an individual pad.

search continues until an empty pad or the end of the pad row is reached. At this point, the search algorithm stops and the pulses are flagged as used to allow the algorithm to find the remaining clusters in the row.

The $x$ and $y$ coordinates of a cluster are calculated from the coordinates of the center of the individual pads. Using a center of gravity method, the cluster coordinate $y$ for example, is calculated according to

$$y = \frac{\sum y_i Q_i}{\sum Q_i} \tag{2}$$

where $y_i$ denotes the $y$-coordinate of the pad $i$ and $Q_i$ is the charge deposited in this pad. The same method is also applied to the $x$ and $z$ coordinates, where in the case of $z$ the coordinate can be obtained from the time information of the pulse and the drift velocity of the electrons in the chamber gas. The total charge of the cluster is calculated as the sum over the individual charges of the included pulses.

Thanks to the low amount of noise present, it is not necessary to perform any cuts on the clusters. After tracking has been performed, cuts on e.g. cluster size has a very small effect.

## 9.3 Tracking

For track finding, a simple algorithm is used that performs well in the low track density and low noise environment. A cluster at the module edge is chosen as a starting point. Since the pad plane, in the parallel configuration, is oriented with the beam along the long side ($x$-direction) of the pads, the next cluster in the track should be found in the neighbouring pad row. A search window is therefore defined around the start cluster in the drift direction ($z$-direction) as well as in the $y$-direction. If a cluster is found within the $y - z$-window in the next pad row, it is added to the track. If no cluster is found, the search continues in the next row. Here, a cluster must be found or the track candidate is rejected so that a track is not allowed to contain holes larger than one row. If more than one cluster is found, the one lying closest to the previous cluster in $y$ is chosen.

The procedure is repeated with a new search window around the previous cluster until the other module edge is reached. In order to be accepted, a track needs to include at least 10 clusters and no more than two empty pad rows. The algorithm has been tested extensively and visual cross checks have confirmed that it works well.

When a track has been found, the clusters in it are flagged so that the algorithm can be run again over the remaining clusters. This allows for multiple tracks to be found. Due to conversion, multiple track events do occur and it is important to find them since they will have an effect on the energy loss distribution.

In the end of the tracking procedure, a fit is made to the coordinates of the clusters. The straigth tracks in the runs without magnetic field are fitted with a first order polynomial and the helical tracks with magnetic field are fitted with a second order polynomial.

## 9.4 Track Selection

In order to have good quality tracks, two conditons had to be fulfilled for inclusion of a track in the analysis. As mentioned, a maximum number of two missing pad rows were allowed so that the track must have clusters in all but at most two rows. In addition, the track is required to go through the middle of the module to avoid edge effects such as poorly reconstructed clusters where some of the charge ends up outside the instrumented region.

## 9.5 Perpendicular configuration

As explained in section 7, the LCTPC was also operated in a $90^o$ rotated configuration. In this configuration the beam was perpendicular to the long edge of the pads and, thus, the width of the charge distribution collected on the pads should be compared to the length ($5.2\ mm$) of the pads. Due to the fact that the width of the charge distribution is smaller than the length of the pads the space resolution is very poor and does not allow to perform a conventional tracking. The track selection is done with the following method:

First a "clean" area area is selected, far from the edges and dead regions of the module. Two "trigger" regions were defined (see fig. 11), which should contain hits to indicate the occurance of a possible track. From the signal in these two regions, a virtual track was formed. The pulses within two pads on each side of this line (in the $x$-direction) will contribute to the reconstruction of the real track.

On the signal in the corresponding region in $x - y$, we apply the same kind of clustering as in the parallel configuration. Instead of a pad row, we cluster along a "pad line" which connects each pad to one of its two neighbours in the next row.

The tracks defined have at most 52 clusters (see fig. 11).

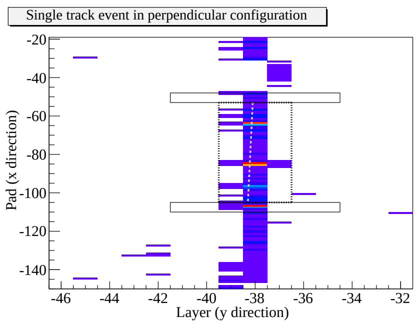



Figure 11: Display of an event with one track in the perpendicular configuration. The black rectangle represent the trigger regions. The gray dashed line is the virtual track. The fine dashed box shows the pads included in the track.

## 9.6 Event selection

For the following analysis, pure samples of 5 $GeV/c$ electron tracks with the most stable conditions possible were selected, rejecting events with the following two problems:

**Gain drops** It was observed that a few times per run, the overall gain dropped. From a fit to the average charge of consecutive events in the low gain regions, the gain drops could be identified. Using this, we clean up the data by cutting out the low gain events as shown on Fig. 12.

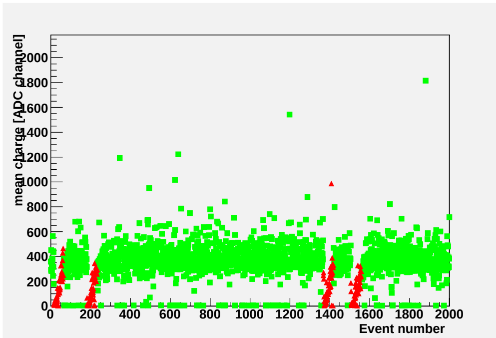



Figure 12: Average pad charge in each event for one run. Periods with gain drop (red triangles) are rejected. The events with zero charge are empty (fake trigger).

**Multiple tracks** To avoid events with multiple tracks in the perpendicular configuration, events where the ratio between the signal inside and outside a track is too high are rejected (the threshold is tuned for each run).

In both configurations, a cut on the $dE/dx$ of the track is also done to remove double tracks which are too close to be separated in space and time. These look like single track, but has a $dE/dx$ twice as large.

# 10 Cross talk

In the perpendicular run configuration, the width of the charge distribution from the track is smaller than the width of a pad row (5.2 $mm$), so that most of the time, we only get signal in one or two pad rows. This fact will be used to measure the signal induced in neighbouring channels through crosstalk in the kapton cables between the pads and the FEC.

## 10.1 Track selection

First, tracks are defined and selected as explained in section 9.5. From this track sample a refined selection is done to have only tracks that approximately follows the border between two pad rows (in the $y$-direction). Since the pad geometry is somewhat curved, and the tracks are bent in the magnetic field, a track can in practice never fulfill this requirement exactly and the poor space resolution in the perpendicular configuration does not allow a very precise reconstruction and selection of the proper tracks. However, different approximate methods of selection were tried and they all gave the same final results.

The tracks obtained are separated into two categories (see Fig. 13):

**A** tracks going between pad rows connected to different connectors

**B** tracks going between pad rows connected to the same connector

In Fig. 13, the definition of the "close" and "away" rows are given, corresponding to the two rows closest to the track, and their neighbours, respectively. Since the track is going along the border of two pad rows the charge should give signals in the pads of both "close" rows, whereas it is not expected to see any signal in the "away" rows due to the electron cloud being much smaller than the pad size tranverse to the track in this configuration. This can be verified by events from the case B where no signal should be observed on the "away" rows since these are read out by different connectors compared to the "close" rows and therefore could also not have any contribution from possible cross talk in the cables. For events fulfilling the requirements of case A the "close" rows are read out by different cables and the only explanation for an observed signal in the "away" rows must be cross talk in the kapton cable.

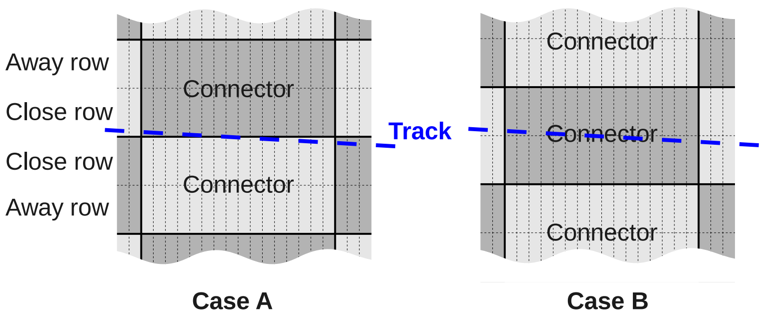



Figure 13: Schematic picture giving the definition of the "close" and "away" rows, respectively, for the two situations A and B described in the text.

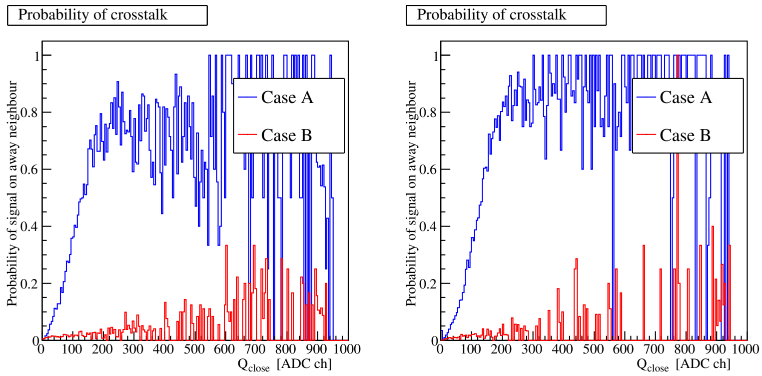



Figure 14: Probability to observe a signal in the "away" row, when a signal on the "close" row is measured. The left and right plots correspond to the pad rows on each side of the track.

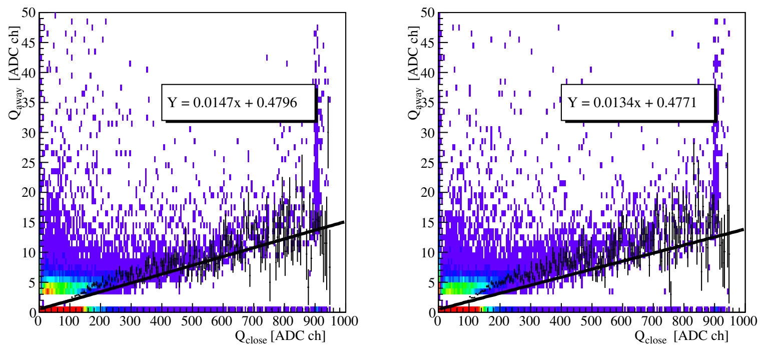



Figure 15: The charge deposited on pads in the "away" row versus that deposited in the "close" row for case A. The left and right plots show the distribution for the two neighbouring channels in the cable.

## 10.2 Results

The first goal was to see whether the results gave any indication of cross-talk. Fig. 14 shows the probability to get a signal in the "away" row, when there is one in the "close" row. It is clearly seen that in case A, if the "close" signal is large enough, there is almost always an "away" signal, while in case B, a signal almost never appears in the "away" row (as expected since these are read out by different cables). Exactly the same effect is observed for the pad rows on each side of the track, which excludes the possibility of a geometric effect from the track angle or the curved pad geometry.

The cross-talk results shown in Fig. 14 suggest a threshold effect around Q=200 ADC counts. In Fig. 15, the charge collected on a pad in the "close" row is plotted versus the charge collected on its neighbouring pad in the "away" row for the case A. Here the neighbour pad is chosen specifically to be one of the two closest cable lines in the readout cable. From this figure we see clearly the effect of zero-suppression for the low charges, for which the cross-talk is mostly below threshold. From the fit on the average correlation in Fig. 15, it can clearly be seen that the effect of the cross-talk is below 2% of the main signal. The results were confirmed by doing a test with a pulse generator.

## 11 Performance study

The data taken with the LCTPC in the two different configurations described above were analysed and compared to the results from simulations. The simulated data was created with the same format as real data. The analysis of real and simulated data were performed in exactly the same way, using the clustering and track finding as described in section 9.

### 11.1 Charge deposition

The main focus was on the charge deposition and the purpose was to confront the energy loss models with data. Since the absolute signal gain was the only free parameter of the simulation, it was important to match it to the data.

#### 11.1.1 $dE/dx$ resolution

In order to determine the gain in the simulation, the $dE/dx$ of the tracks was measured. Since we use ultra-relativistic electrons, on the plateau region of the Bethe-Bloch curve, a very clean sample for the $dE/dx$ is expected. The value of $dE/dx$ is calculated as the 60% truncated mean of the signal in the track, taking the mean of the 60% lowest signals in the track.

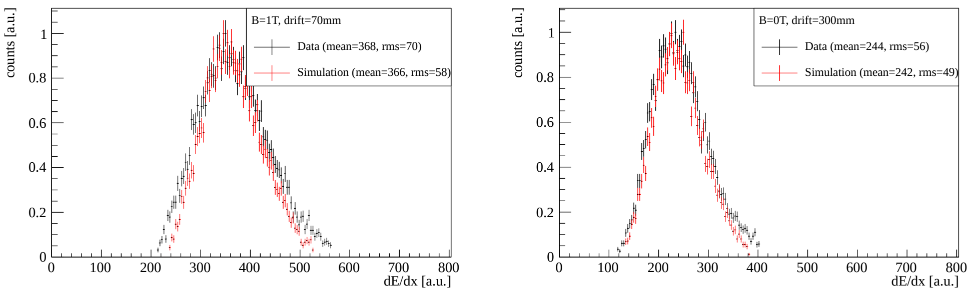



Figure 16: The $dE/dx$ spectrum, calculated from the 60% truncated mean. A comparison between simulation (in red), using Geant4's PAI model, and data (in black) is given in the case of a 1 T magnetic field (left) and no magnetic field (right).

The results are shown for the parallel configuration in Figs. 16 and 17. The gain parameter in the simulation is adjusted so that the mean value of the $dE/dx$ distribution agrees with the experimental data. It can be observed that the $dE/dx$ distribution in the simulation is somewhat narrower than that of the real data.

In Fig. 17, we see that the diffusion gets more prominent as the drift distances increase, an effect that is especially significant with no magnetic field and less significant with a 1 T magnetic field on. The correlations that the diffusion induces between pad rows (i.e. between clusters in the track) alters the charge distribution and worsens the $dE/dx$ resolution.

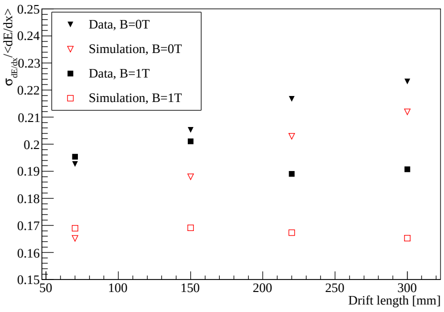



Figure 17: dE/dx resolution as a function of drift distance.

#### 11.1.2 Charge distribution

To get a better understanding of the $dE/dx$ results, the full charge distribution of the clusters in the tracks can be studied. The results are shown in Fig. 18. It can be observed that the simulation does not describe the data taken in the parallel configuration very well. The energy loss model is not precise enough to reproduce the details of the data. However, when the diffusion increases as is the case with no magnetic field and longer drift lengths, the correlations between clusters smear out the details of the model, and simulation and data start to agree.

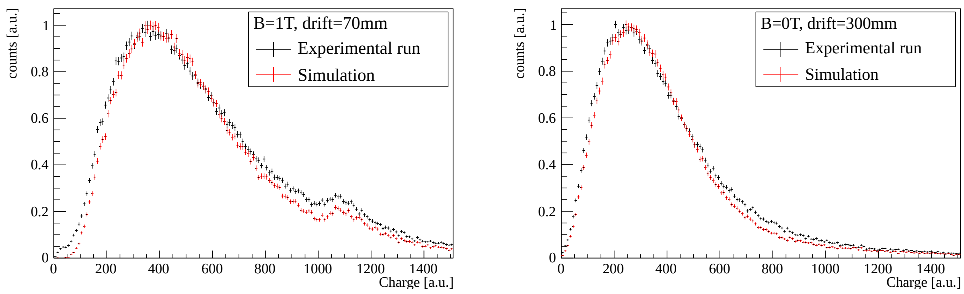



Figure 18: Charge distribution in different diffusion cases (left: minimum diffusion (short drift, B=1T), right: maximum diffusion (long drift, B=0T). With increasing diffusion, the details of the energy loss model are smeared out and the simulation gives a better description of the data. The bumps in the low diffusion case is caused by the saturation in ADC channel 1024 of the digital signal in each pad (see section 6).

#### 11.1.3 Comparison of the charge distributions for the two configuration

With the two configurations (parallel and perpendicular), we have the opportunity to test the consistency of the energy deposition comparing sampling over short track segments (1.2 $mm$) and longer ones (5.2 $mm$). Due to different levels of correlations, the charge distribution does not scale in a simple way as the sampling length is increased [16]. However, since the size of a pad is approximately 1.2 $mm$ by 5.2 $mm$, the width of four pads is 4.8 $mm$ which is approximately the same as the length of a pad. Thus, by combining signals from 4 neighbouring pads in the perpendicular configuration, the charge distribution should be comparable to that of the parallel configuration.

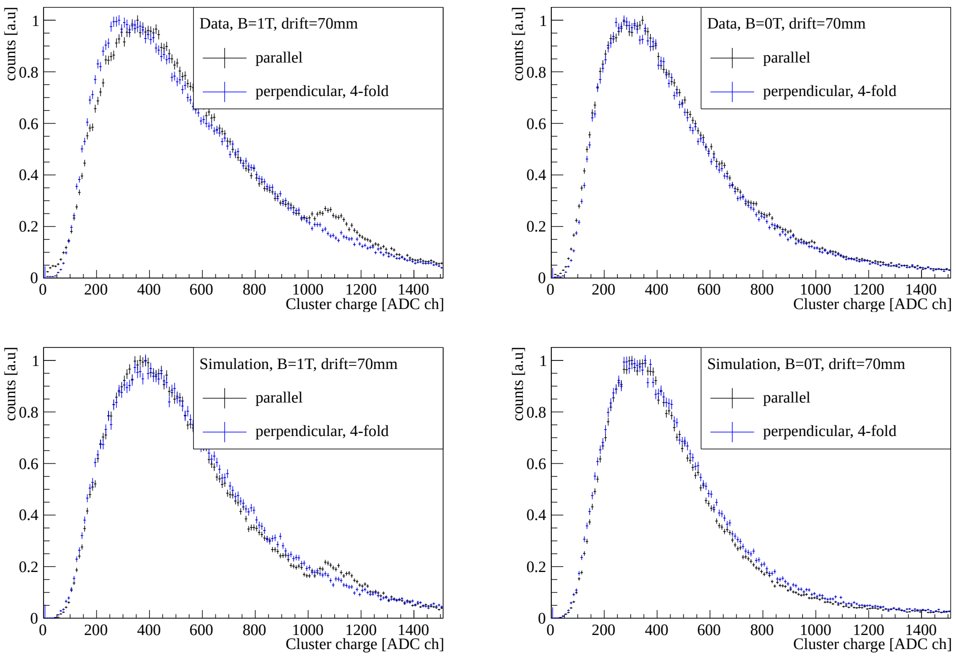



Figure 19: Comparison between the charge distribution for the parallel configuration (5.2 $mm$ long pads) and the charge collected on four consecutive pads (4x1.2 $mm$) in the perpendicular configuration. A correction factor is introduced to compensate for gain difference and possible efficiency difference. The upper plots are data and the lower are simulation with the same run characteristics.

In Fig. 19, the charge distribution from track clusters in the parallel configuration, is compared to the charge distribution of clusters grouped over four consecutive pads in a track in the perpendicular configuration. The conditions for the runs which are compared are the same i.e. the drifts length, magnetic field, electronics settings are identical. An additional scaling of the charge is done for one of the configurations to account for gain variations and possible different charge collection efficiencies.

Fig. 19 gives comparisons of charge distributions for the two different configurations with (left hand plots) and without (right hand plots) the magnetic field on, for real data (upper plots) and for simulation data (lower plots). The agreement is perfect. In spite of the differences in the charge distribution (fig. 18), it is clear that the simulation reproduces the consistency between the two configurations. This indicates that besides the energy loss process, the detector is well described by the simulation. It also means that effects observed on the charge in the perpendicular configuration should be consistent with what is observed in the parallel configuration.

### 11.1.4 Perpendicular configuration

The final step of the energy loss study was to investigate if it was possible to observe the details of the energy loss process from measurements in the perpendicular configuration with a 1 $mm$ sampling length. In Fig. 18 it was noticed that there are discrepancies between the real and the simulated data for the parallel configuration. However, in section 11.1.3, it was shown that measurements in the two configurations are consistent. Thus, the discrepancies observed in the measurements with the parallel configuration could be further investigated in the perpendicular configuration and possibly be associated with some details of the energy loss model.

In Fig. 20 the charge distribution for different steps in the simulation are shown for the lower part of the pulse height spectrum. A description of the various distributions is given below:

**Number of ionisation electrons** , represented with triangle (▼), is the number of electrons produced after ionisation, at the track level, over the longitudinal size of the pad (1.2 $mm$). For comparison, it is multiplied by the average gain.

**Number of electrons reaching the GEMs** , represented with blue squares, is the number of electrons reaching the GEM plane. The only added effect from the previous number is diffusion. For comparison, it is multiplied by the average gain.

**Reconstructed charge** , in red open circles, is the charge read out on the pads, after all effects (gain variation, noise, digitisation). It is the value to be compared directly to the data (represented with the black circles).

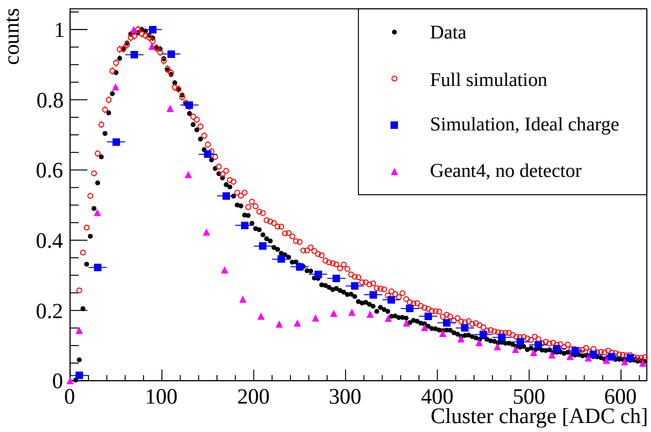



Figure 20: Charge distribution for minimum diffusion, with 1.2 $mm$ wide pads for the lower part of the pulse height spectrum. The triangles (▼) show the number of ionisation electrons produced. The blue squares shows the number of electrons reaching the GEM plane (after diffusion). The red open circles is the full simulation. The real data is represented by black full circles. We can see the bump from the second shell of Argon can be associated with the excess the simulation shows compared to data around 300 ADC.

In Fig. 20, a bump in the number of ionisation electrons (around 300 equivalent ADC channels) is observed, which is related to the L shell in Argon. Following the steps in the simulation shown in the figure, we can associate this bump to an excess in the reconstructed charge compared to data around 300 ADC channels. This suggests that the ionization from inner shells is inadequately modeled.

## 11.2 Space Resolution

Since no external reference is available, the resolution is calculated with respect to the reconstructed track. The track is obtained from a fit to the cluster coordinates and the residuals (in $y$) of the individual points are calculated using this fit. The resolution is then obtained as the width $\sigma_y$ of the residual distribution, obtained from a gaussian fit. Since the investigated point is included in the fit, this is expected to give an estimation that is better than the actual resolution. A common approach to this is to use the Geometric Mean method. According to this method, another fit is made, this time excluding the investigated point. The geometric mean of the resolutions obtained with the two methods is then calculated $\sigma = \sqrt{\sigma_{in} \cdot \sigma_{ex}}$, where $\sigma_{in}$ and $\sigma_{ex}$ is the resolution calculated with the point included and excluded, respectively, in the fitting procedure. This is considered to be a more reliable estimation of the "true" resolution.

Without magnetic field (Fig. 21), the simulation gives a perfect description of the space resolution in the data.

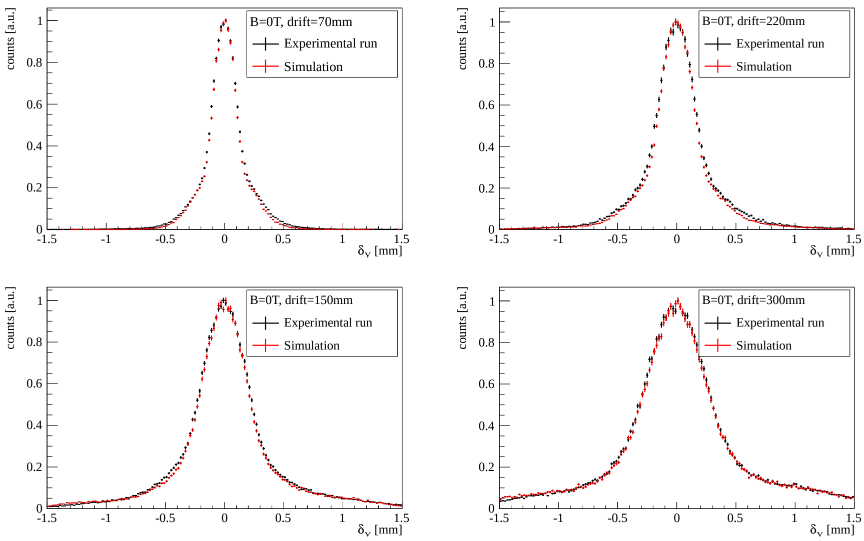



Figure 21: Space residuals for different drift distances, without magnetic field. The simulation (in red) gives a perfect description of the data.

As mentioned in section 4, there was an inhomogeneity of the electric field around the edges of the GEM plane. In the presence of magnetic field, this caused distortions of the tracks, due to the $E \times B$ effect. The tracks appeared as a consequence of this slightly deformed compared to the expected arc. This worsened considerably the resolution measurement relying on a polynomial fit to the track.

An attempt to correct for the distortion was made, using the residual distributions of the individual pad rows. The residuals are calculated assuming a helical track described by a second order polynomial in the transverse projection. Due to the know problem with the distorted electric field at the edges of the GEM the information from outermost 5 pad rows were excluded and a fit was only performed for the middle part of the track, which should be less influence by the field inhomogeneities. Fig. 22 shows the calculated residuals versus the pad rows for many events in a run. A clear S-shape can be observed. Taking the residuals averaged for each row (x-value) a distortion (y-value) can be calculated. This averaged value is used to correct the $y$-coordinates on a row-by-row basis.

To account for the distortion in the simulation, a fit with a fifth order polynomial to the distortion shape in the data was made as shown on Fig. 22. The fitted distortion was included in the simulation as follows. Since the position of the clusters in the $x$-direction cannot be measured to a better precision than the length of the pads (5.2 $mm$), the displacement (in the $y$ direction) applied to each cluster was chosen from the fit at a *random* point (in the $x$ direction) in the corresponding pad row. When the simulated data are corrected in the same way as the real data by the distortion (now *average* over

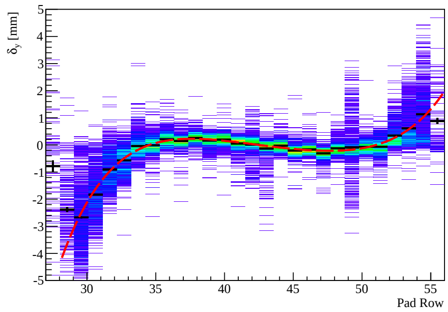



Figure 22: Space residuals $\delta_y$ versus pad row. The S-like shape is a consequence of the $E \times B$ effects associated with electric field distortions caused by the support pillars of the GEMs. This profile (black lines) is used to correct the data. The profile is fitted (dashed line) to provide a model of the distortion for the simulation.

the pad length), a spread of the residuals in each row remains due to the unknown cluster position along the pad length. Fig. 23 shows that the simulation gives a good description of the data after inclusion of the track distortion. Ideally the distortions should be done at the electron level in the simulation, but these results anyway suggest that the origin is our lack of knowledge of where along the pad the electrons dominates. Therefore this distortion cannot easily be corrected for, and, to achieve the resolution of 70 $\mu m$ predicted by the simulation, the electric field close to the GEMs has to be improved.

The space resolution ($\sigma_y$) for different drift distances are shown in Fig. 24 from which it is seen that when the drift distance, and therefore the diffusion, is increased, the resolution of the space position of the track becomes worse.

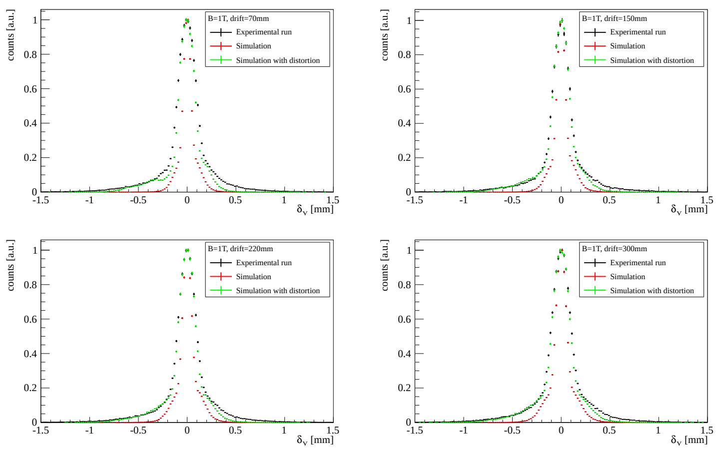



Figure 23: Space residuals ($\delta_y$) for different drift distances, with magnetic field. The $E \times B$ effects, associated to the electric field distortions caused by the support pillars of the GEMs, made the resolution worse for the data (black) than for the simulation (red). After introducing corrections to account for the distorions, the simulation (in green) gives a much better description of the data.

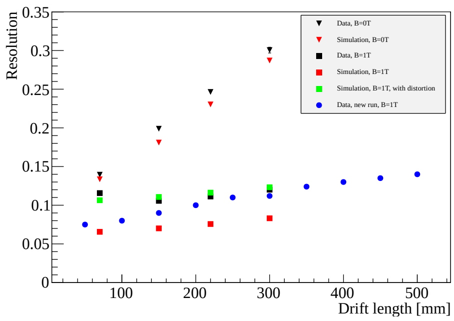



Figure 24: Space resolution for data and simulation. When the drift distance, and therefore the diffusion, is increased, the resolution of the space position of the track becomes worse. With magnetic field, $E \times B$ effects associated with electric field distortions made the resolution worse for the data (black squares) than for the simulation (red squares). After including the distortion, the simulation (green squares) matches the data.
The blue dots show for comparison the resolution measured by the Japanese group on more recent data using the same setup. The method of calculation was different, which explains the slightly different values.

### 11.3 Influence of ADC resolution

In the current version of the electronics, the analogue signal is digitised with a 10-bit precision (1024 ADC channels). This high resolution is likely to be better than needed, given that the pad width is only 1.2 $mm$. Reducing the ADC resolution to an optimal value will result in a reduction of power and cost for the electronics.

In order to test this, the digital representation of the raw signal was changed by removing the ADC bits one by one to see at which point a decrease in resolution could be observed. The regular reconstruction is then performed. The size of the ADC signal is reduced from 10 bits down to 8, 6 and 4 bits, and the performance of the LCTPC is compared for the various resolutions.

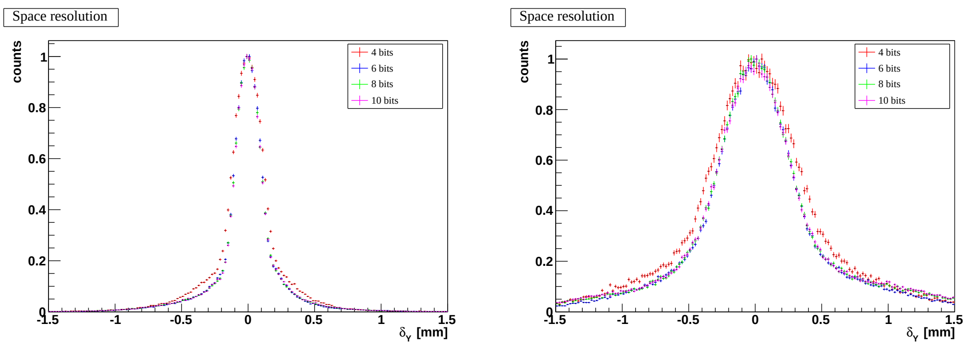



Figure 25: Examples of space residual ($\delta_y$) distributions for different values of the ADC resolution. On the left, with minimum diffusion (70 $mm$ drift and magnetic field on), and on the right with maximum diffusion (300 $mm$ drift and no magnetic field).

From Figs. 25 and 26, it is obvious that no deterioration of the space resolution is observed until

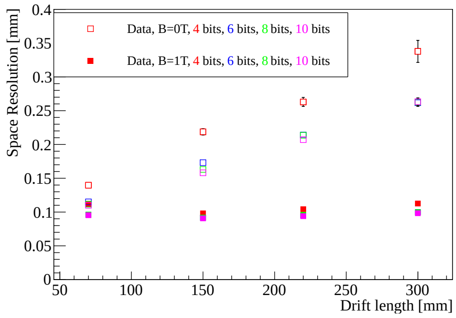



Figure 26: Space resolution as a function of drift distance for different values of the ADC resolution.

the precision has been lowered to 6 ADC bits. Below that a worse space resolution is observed, mostly because of threshold effects.

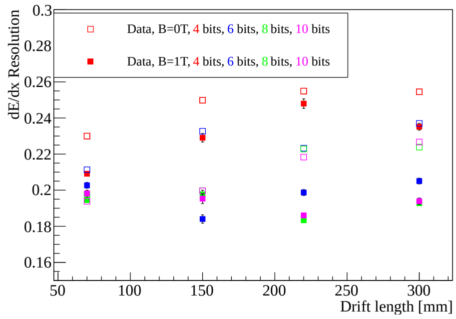



Figure 27: $dE/dx$ resolution as a function of drift distance for different values of the ADC resolution.

In Fig. 27, it is seen that, as expected, the loss in ADC resolution has more influence on the $dE/dx$ resolution. However, down to 6 bits, the loss is not excessive, and begins to be significant at 4 bits.

This shows that the ADC range can be reduced without real damage to the performance. However, this test has been done with electrons on the ultrarelativistic plateau. It should be confirmed for minimum ionising particles, where the effect should be maximal. There is in any case range for adjustment of the digital signal.

## 12 Conclusion

The large prototype TPC (LCTPC) equipped with Asian GEMs and electronics based on the ALICE ALTRO, has been successfully set up and tested thoroughly in the DESY T24 electron beam with different configurations. The data has been reconstructed, analysed and compared to simulations using software developed in Lund.

Running the LCTPC in a rotated configuration allowed to reach a high level of detail in the study of the charge distribution along tracks. The comparison to simulation allowed to probe the limits of the energy loss models, in particular the PAI model offered by GEANT4.

This unusual orientation also allowed to study cross talk in the readout cables in a real data taking situation.

In spite of imperfection of the energy loss model, the performance of the TPC has been very well described by the simulation. A loss of resolution from distortions from $E \times B$ effect related to inhomogeneities of the electric field close to the GEM was observed. The observed resolution could be described by simulation with a simple correction in reconstructed clusters to include the distortions. While it shows that there is little hope to recover space resolution with the inhomogeneous electric field, it indicates that the resolution of 70 $\mu m$ shown by the simulation should be achievable if the field is improved.

## Acknowledgements

This work is supported by the Commission of the European Communities under the 6th Framework Programme "Structuring the European Research Area", contract number RII3-026126. The project was also supported by the Swedish Natural Science Research Council. We would like to thank K. Dehmelt, T. Matsuda and H. Kuroiwah for help in setting up the equipment needed for the measurement, and H. Bichsel and V. Grichine for help with the energy loss models for the simulation.

## References

[1] T. Behnke *et al.*, EUDET-Memo-2007-11.

[2] J. Timmermans, private communication.

[3] T. Behnke *et al.*, EUDET-Memo-2006-042

[4] T. Behnke *et al.*, "A Lightweight Field Cage for a Large TPC Prototype for the ILC", JINST **5** (2010) P10011 [arXiv:1006.3220 [physics.ins-det]].

[5] K. Dehmelt on behalf of the LCTPC Collaboration, EUDET-Memo-2008-46

[6] K. Dehmelt on behalf of the LCTPC Collaboration, EUDET-Memo-2009-9

[7] A. Ishikawa *et al.*, "A GEM TPC End-Panel Pre-Prototype" *In the Proceedings of 2007 International Linear Collider Workshop (LCWS07 and ILC07), Hamburg, Germany, 30 May - 3 Jun 2007, pp TRK27* [arXiv:0710.0205 [physics.ins-det]].

[8] T. Matsuda [LC-TPC Collaboration], "Results of the first beam test of a GEM TPC large prototype", JINST **5** (2010) P01010.

[9] A. Yamamoto and Y. Makida, “Superconducting magnet for long duration flights”, Proceedings of the 5th Workshop on ’Balloon-borne experiments with a superconducting magnet spectrometer’, KEK Japan, November 30 - December 1, 1994.

[10] L. Musa *et al.*, “The ALICE TPC Front End Electronics”, Proc. of the IEEE Nuclear Science Symposium, 20 - 25 Oct 2003, Portland

[11] R. Esteve Bosch, A. Jimenez de Parga, B. Mota and L. Musa, IEEE Trans. Nucl. Sci. **50** (2003) 2460.

[12] L. Musa, “Prototype compact readout system”, EUDET-Memo-2009-31, 2009.

[13] C. Gonzales Gutrierrez *et al.*, “The Alice TPC Readout Control Unit”, *In the Proceeding of the 2005 IEEE Nuclear Science Symposium and Medical Imaging Conference, Puerto Rico, October 23-29, 2005*

[14] P. Christiansen et al., Nucl. Instr. and Meth. A 609 (2009), p. 149

[15] J. Apostolakis *et al.*, Nucl. Instrum. Meth. A 453 (3): 597-605, 2000

[16] H. Bichsel, Nucl. Instrum. Meth. A **562**, 154 (2006).

[17] H. Bichsel, Nucl. Instrum. Meth. A **566**, 1 (2006).

[18] ICRU Report 31: Average energy required to produce an ion pair. Bethesda, MD: ICRU Publications, 1979.

# Paper IV

# Charged pion identification at high $\mathbf{p_T}$ in ALICE using the TPC **dE/dx***

Philippe Gros for the ALICE collaboration

Lund University, Lund, Sweden

The ALICE TPC provides excellent charged particle tracking to study pp and Pb–Pb collisions at LHC. The TPC allows also particle identification (PID) via the measurement of the specific ionisation $dE/dx$. At high $p_T$ ($p_T > 3$ GeV/c) this is accomplished in the region of the relativistic rise of the energy loss, and provides a measurement of the yields of charged pions, kaons, and protons. Here we present the performance of such an analysis in pp collisions at 7 TeV for charged pions up to 10 GeV/c. On the relativistic rise, statistical PID can be used to evaluate the yields for pions, protons and kaons for $p_T > 3$ GeV/c. Here we show first performance results from such an analysis for charged pions up to 10 GeV/c.

## 1. Introduction

The ALICE Time Projection Chamber (TPC) is the main tracking detector in the central region ($|\eta| < 0.9$) of the ALICE heavy ion experiment at the LHC [1]. In addition to excellent tracking performance, the TPC has particle identification (PID) capabilities through the measurement of specific ionisation ($dE/dx$). This is of particular interest in the region of the relativistic rise of the Bethe-Bloch curves, i.e. at $p_T > 3$ GeV/c, where other ALICE detectors have little or no PID capabilities for charged pions, kaons, and protons. The STAR experiment at RHIC has successfully used TPC PID in the relativistic rise region in both p+p and Au+Au collisions to measure important features of heavy ion collisions in the regime of intermediate and high $p_T$ [2, 3], such as kinetic energy and valence quark scaling of the elliptic flow and the anomalous proton to pion ratio [3, 4]. These results have given rise to speculations that hadronisation at intermediate $p_T$ might be due to recombination of lower momentum "constituent quark like" degrees of freedom rather than fragmentation of high momentum partons [5]. At LHC energies recombination might result in even larger $p/\pi$ ratios in the momentum interval $p_T = 10 - 20$ GeV/c [6] because of the larger jet cross section.

Recently the ALICE collaboration has published first results from Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV on elliptic flow of unidentified charged particles [7], and the nuclear modification factor at LHC for unidentified charged hadrons [8]. In this article the method is presented which allows to extend these results to charged pions, kaons, and protons based on $dE/dx$ measurements in the relativistic rise region, and the present performance in pp collisions at 7 TeV is discussed.

* Presented at the Epiphany Conference in Krakow on January 10. 2011

## 2. ALICE TPC performance

The ALICE TPC is composed of a cylindrical gas volume, divided into two half volumes of equal size, separated by a 30 $\mu$m thick central HV electrode at -100 kV which generates a drift field of 400 V/cm. The field cage has an inner radius of about 80 cm, and an outer radius of about 280 cm, with an overall length of 500 cm in the beam direction. At both ends of the TPC, multi-wire proportional chambers (MWPC) are mounted into two end plates with 18 trapezoidal sectors each. The MWPCs have highly segmented pad readout with a total of about 560,000 pads, optimised for tracking in the high multiplicity environment of central heavy ion collisions at LHC. The front end electronics are equipped with on-board digital filters to allow baseline restoration and cancellation of the signal tails due to ion drift.

Each particle track can have up to 159 individual ionisation measurements in the TPC. The so-called TPC signal is determined as the truncated mean of the 60 % lowest hit charges associated with the track. The TPC signal determined in this way is related to the most probable energy deposit and yields a much better estimate of the $dE/dx$ than the mean value, which is influenced by large ionisation deposits in the Landau tail.

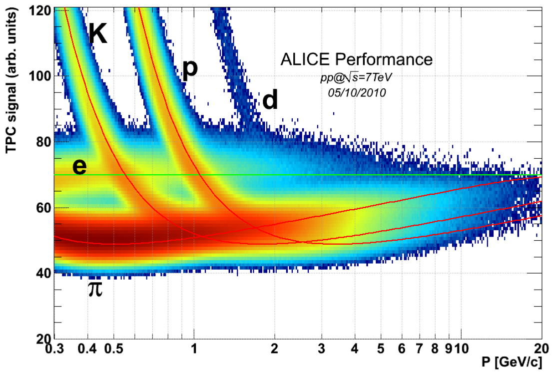



Figure 1. TPC signal as a function of momentum in pp collisions at 7 TeV.

The calibration of the TPC signal (readout chambers and electronics) has been carried out by injecting radioactive Krypton in the TPC gas and inspecting the signal amplitudes resulting from the decays [1]. The $dE/dx$ resolution ($\sigma_{dE/dx}/dE/dx$) after gain calibration was measured to be better than 5 % for full length tracks which is better than the design value of 5.5 %. The TPC signal distribution as a function of the particle momentum p in pp collisions at 7 TeV is shown in Fig. 1.

The $p_T$ resolution $\Delta(p_T)/p_T$ of the TPC was measured using cosmic muons by matching the upper and lower halves of tracks crossing the whole TPC (see fig. 2). At the present level of the calibration, the $p_T$ resolution $\Delta(p_T)/p_T$ is 7% at 10 GeV/c.

## 3. Analysis Method

In this study we use the same event and track selections as for the analysis of the unidentified charged hadron $p_T$ spectra [9]. Additionally a track selection based on the

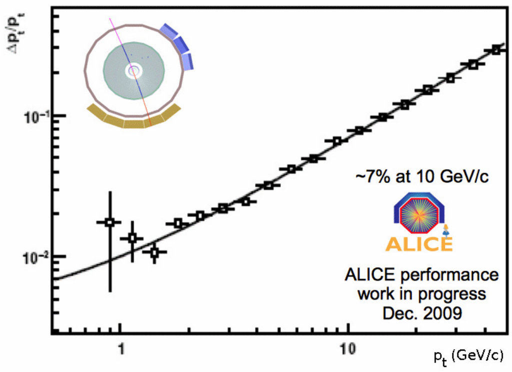



Figure 2. transverse momentum resolution as a function of transverse momentum measured by matching cosmic tracks corresponding to the same muon in the two halves of the TPC.

azimuthal angle $\phi$ is applied to remove tracks which pass through the insensitive regions between the read-out chambers, which have a smaller number of ionisation measurements and thus a worse $dE/dx$ resolution.

The challenge of using the TPC signal on the relativistic rise is thatthe TPC signal distributions of different particle species partially overlap, so that the particle yields have to be extracted statistically, i.e. from a fit to the TPC signal distribution. The TPC signal in a narrow momentum interval can be fit with a sum of four Gaussians ($\pi$, $K$, $p$ and $e$) with a total of twelve parameters: four amplitudes, four mean positions, and four widths. The main challenge of the analysis is therefore to constrain as many parameters as possible. To this end we parametrize the resolution $\sigma_{dE/dx}$ and constrain the mean values by employing the dependence on $\beta\gamma$ of the energy loss curve. In this way the mean and the width of each Gaussian can be constrained, i.e. nine of the twelve parameters can be fixed [10].

The resolution $\sigma_{dE/dx}$ is extrapolated using a clean sample of pions in the region of minimum ionisation (MIP) and depends on the number of hits in the track. The mean values are obtained from fits to the data with a Bethe-Bloch like parametrisation in the following way: The distribution of TPC signals is fitted in bins of $p$ with three Gaussian functions, where the amplitudes are free parameters, the widths are fixed to the resolution, and the mean values have to follow the parametrised Bethe-Bloch function for the $\beta\gamma$ corresponding respectively to pion, proton and kaon masses. For particle identification we subtract from the measured TPC signal the mean values for pions obtained from the fits. We call this difference $\Delta_\pi$. This first step is done using intervals in $p$ which are relevant for the parametrisation, the following steps are performed in intervals of $p_T$.

The $\Delta_\pi$ distributions are finally fitted in different bins of $p_T$ with the sum of four Gaussians where $\sigma$ and mean are constrained within 5 % of the values calculated above for the pions, kaons and protons, and extrapolated from low $p_T$ regions for electrons. Examples of such fits are shown on Fig. 3. The pion peak is centred close to $\Delta_\pi$ =0, and the width is compatible with the determined resolution function used for the fit.

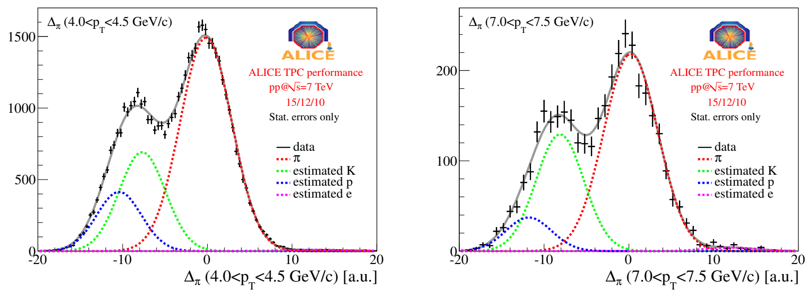



Figure 3. Examples of $dE/dx$ distributions in two $p_T$ bins. The distribution is fitted with a sum of 4 Gaussians for pions, protons, kaons and electrons. The electron contribution is very small and barely visible under the right tail of the pions distributions.

### 4. Performance of the dE/dx analysis

The fits explained above give a good estimate of the pion yield for different $p_T$ values. The smaller difference in inverse mass (and therefore $\beta\gamma$ at fixed $p$) between kaons and protons makes the separation between these two species more difficult, and a better understanding of the constraints is required to get the yield for these particles.

The fit results are first tested for consistency by checking that the ratio between positive and negative pions is close to unity as expected. The plot shown on Fig. 4 confirms this consistency.

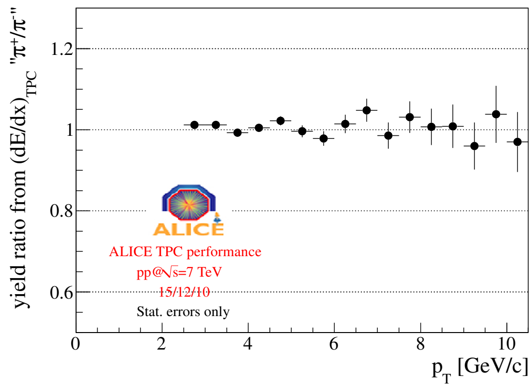



Figure 4. Ratio of the yield obtained for $\pi^+$ and $\pi^-$. The value is around unity, as expected, which confirms the consistency of the method.

The resulting fraction of pions among the charged particles is shown on Fig. 5. The ratio shown here does not include any correction for secondary particles and decays. Figure 6 shows the yields vs $p_T$ without any normalisation and corrections.

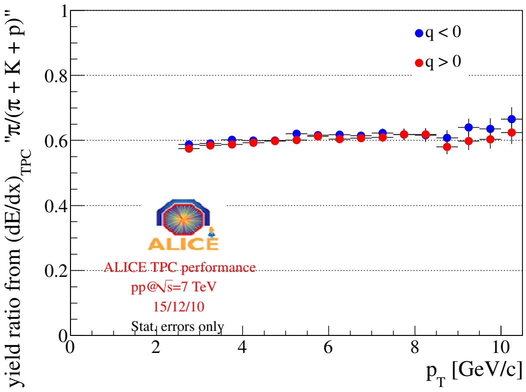



Figure 5. Estimated fraction of $\pi^+$ (resp. $\pi^-$) among the positive (resp negative) hadrons. The data represent the results from the fits to the $\Delta_\pi$ distributions without further corrections.

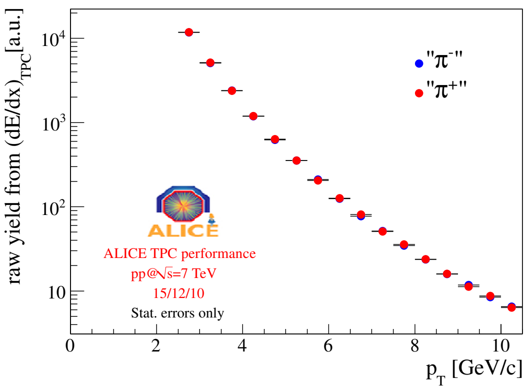



Figure 6. Uncorrected yields for $\pi^+$ and $\pi^-$ from the fits to the $\Delta_\pi$ distributions.

### 5. Conclusion

We presented the performance of the ALICE TPC for charged pion identification at high $p_T$, employing $dE/dx$ measurements in the relativistic rise region. The results demonstrate the capability to study pion production at large transverse momenta, which is of particular interest in heavy-ion collisions. Increased statistics will allow to extend the study to even higher $p_T$.

References

[1] J. Alme *et al.*, Nucl. Instrum. Meth. A **622** (2010) 316 [arXiv:1001.1950 [physics.ins-det]].
[2] J. Adams *et al.* [STAR Collaboration], Phys. Lett. B **637**, 161 (2006) [arXiv:nucl-ex/0601033].

[3] B. I. Abelev *et al.* [STAR Collaboration], Phys. Rev. Lett. **97**, 152301 (2006).

[4] A. Adare *et al.* [PHENIX Collaboration], Phys. Rev. Lett. **98**, 162301 (2007) [arXiv:nucl-ex/0608033].

[5] R. J. Fries, B. Muller, C. Nonaka and S. A. Bass, Phys. Rev. C **68** (2003) 044902 [arXiv:nucl-th/0306027].

[6] R. C. Hwa and C. B. Yang, Phys. Rev. Lett. **97**, 042301 (2006) [arXiv:nucl-th/0603053].

[7] K. Aamodt *et al.* [The ALICE Collaboration], arXiv:1011.3914 [nucl-ex].

[8] K. Aamodt *et al.* [ALICE Collaboration], Phys. Lett. B **696** (2011) 30 [arXiv:1012.1004 [nucl-ex]].

[9] K. Aamodt *et al.* [ALICE Collaboration], Phys. Lett. B **693** (2010) 53 [arXiv:1007.0719 [hep-ex]].

[10] A. Dobrin and P. Christiansen Ph.D. thesis, Lund University 2010