Charged pion spectra at high p_{T} in pp collisions at $\sqrt{s} = 900$ GeV, measured by the ALICE TPC

Lene Bryngemark

Master's Dissertation

June 2011

LUNDS UNIVERSITET

Abstract

The ALICE Time Projection Chamber (TPC) is used for tracking and measuring momentum and specific energy loss, $\mathrm{d}E/\mathrm{d}x$, of charged particles produced in hadron collisions at the LHC. This Master's dissertation describes the statistical identification of charged hadrons (pions, kaons and protons) on the relativistic rise (transverse momentum $p_{\mathrm{T}} > 3$ GeV/c) of the $\mathrm{d}E/\mathrm{d}x$ curve of the ALICE TPC. Two complementary methods are evaluated and compared. An effect of track cluster charge falling below the threshold of the read-out electronics is found to influence the species separation in $\mathrm{d}E/\mathrm{d}x$, and methods to recover the separation are suggested and discussed. Charged pion fractions and invariant charged pion yields in pp collisions at $\sqrt{s} = 900$ GeV are presented as a function of p_{T}, for pseudorapidity $|\eta| < 0.8$ and $3 \leq p_{\mathrm{T}} \leq 10$ GeV/c.

Populärvetenskaplig sammanfattning

Vid partikelkollisioner omvandlas de kolliderande partiklarnas rörelse- och massenergi till nya partiklar. Den bakomliggande fysiken kan bara förstås med en kännedom om vad som har kolliderat och vad som har producerats i kollisionen. Vid partikelacceleratorn Large Hadron Collider (LHC) vid laboratoriet CERN utanför Geneve, på gränsen mellan Schweiz och Frankrike, kollideras hadroner, vilka är sammansatta partiklar: de består av kvarkar och gluoner. Vid LHC har dels protoner (vätekärnor), dels fullständigt joniserade atomer av bly, blykärnor, kolliderats. I den höga energitäthet som uppstår vid kollisioner av tunga kärnor antas ett kvark-gluon-plasma bildas, dvs ett plasma av de partiklar hadroner består av. Detta är särskilt intressant eftersom samma tillstånd antas ha varit en fas i universums tidiga utveckling. Genom att undersöka egenskaper hos partiklar som detekteras vid sådana kollisioner kanske det går att förstå egenskaperna hos kvark-gluon-plasmat bättre. För att kunna tolka dessa egenskaper krävs också en jämförelse med proton-proton-kollisioner, där inget kvark-gluon-plasma antas kunna bildas.

Den här uppsatsen beskriver två metoder för partikelidentifiering med TPC-detektorn (Time Projection Chamber, "tidsprojektionskammare") i ALICE-experimentet vid LHC. De bygger båda på att laddade partiklar som korsar detektorn förlorar energi, och att mängden förlorad energi beror på hur fort partikeln rör sig (och på dess laddning). I en TPC kan partikelns väg genom detektorn, energiförlust och rörelsemängd (produkten av partikelns massa m och hastighet) mätas samtidigt. Eftersom både rörelsemängden och energiförlusten kan relateras till partikelns hastighet, kan partikelns massa bestämmas, vilket tillsammans med den elektriska laddningen är tillräckligt för att kunna avgöra vilken partikel det är. Fokus för analysen som beskrivs är att på detta sätt identifiera pioner, kaoner och protoner.

Rörelsemängden i sig är också intressant, eftersom olika fysikaliska processer i kollisionen antas vara olika för partiklar med olika rörelsemängd. Särskilt intressant i detta avseende är den transversella komponenten (dvs den komponent som är vinkelrät mot strålens riktning). Rörelsemängd måste bevaras till både storlek och riktning, så eftersom rörelsemängd i den här riktningen inte finns i de inkommande strålarna måste den vara resultatet av att inkommande partiklars energi fullständigt omvandlas till nya partiklar, vilka i sin tur kan röra sig i någon godtycklig riktning.

Av de två metoder som beskrivs är den ena för instabil för att ge tillförlitliga fysikresultat. Den har dock givit viktiga insikter i en systematisk skillnad för olika partikelslag vad gäller hur energiförlusten mäts i TPCn. Dessa insikter har sedermera lett till förbättringar av hur energiförlusten beräknas för varje spår. Detta, i sin tur, har lett till säkrare resultat med den andra metoden, vilken ligger till grund för de resultat som slutligen presenteras: hur många laddade pioner som har producerats vid proton-proton-kollisioner vid kollisionsenergin 900 GeV vid LHC, som funktion av transversell rörelsemängd. Detta är de första sådana resultaten för identifierade laddade pioner vid hög transversell rörelsemängd vid denna kollisionsenergi.

En jämförelse med resultat från andra analyser inom ALICE, där partikelidentifiering inte gjorts men samma data använts, visar mycket god överensstämmelse. Positivt och negativt laddade pioner är lika vanliga, vilket förväntas. En jämförelse med neutrala pioner, som detekteras på helt annorlunda sätt, stämmer också med förväntningarna, givet de systematiska osäkerheterna i de båda undersökningarna.

Acknowledgements

I would like to thank my supervisor Peter Christiansen, for continuously coming up with new ideas to take the analysis further. My predecessors in the analysis, Alexandru Dobrin and Philippe Gros – thank you for answering questions, and for your help as the analysis evolved. I also thank Peter, Anders Oskarsson and Evert Stenlund for being part of the process and commenting the results along the way, and for reading and commenting this dissertation.

Finally, I thank Bozena Wlosinska for always helping out with the administrative issues and for booking my travels, and Tuva Richert for sharing them.

Contents

Chapter 1

Introduction

Particle physics is on the verge of breaking new grounds. This is, however pretentious, not a daring claim, since reaching new energy regimes almost inevitably leads to finding new and often unexpected things. As a new particle accelerator has begun to collide particles at new world record accelerator energies, the analysis of the data begins. It aims at finding out which of all more or less plausible theories seem to hold, which can be ruled out and which should be refined – and what completely new theoretical models we might need. The flora of theories is rich; it all boils down to what is consistent with experimental data.

Experimental data in this case means the electric signals read out from particle detectors as they are traversed by particles produced in high energy hadron collisions. From these signals particle properties, such as electric charge and momentum, can be reconstructed. These properties can be combined into a particle identification. From the particle identification – giving particle constituents, where applicable – and particle kinematic properties, interactions in the collision can be deduced in the light of one model or the other. Complicated chains of processes can be reconstructed.

But most produced particles have too short lifetimes to ever be detected directly. As a matter of fact, apart from nuclear fragments, only about a handful of the resulting particle species (γ, e, μ, π, K, p (and n), with antiparticles) are both long-lived enough and interact with the detector material with large enough probability to be used for track reconstruction. These particles are then the only clue to all the processes, and hundreds of more exotic particle types, occurring in the aftermath of a particle collision. Identifying their properties as accurately as possible is therefore crucial to our understanding of the experimental outcome.

This dissertation presents and evaluates two analysis methods for particle identification at the ALICE experiment. These methods focus on identifying pions (π), protons (p) and kaons (K) via their specific energy loss ($\mathrm{d}E/\mathrm{d}x$) in the ALICE Time Projection Chamber (TPC), on the so-called relativistic rise of the $\mathrm{d}E/\mathrm{d}x$ curve. This region, sometimes called *high* p_T, will be taken to be at momentum $p > 3$ GeV/c.

The ultimate goal for the analysis is to measure the yield of identified charged particles on the relativistic rise. This can be presented as a function of transverse momentum, in so-called p_T spectra. This dissertation aims at presenting p_T spectra of charged pions reaching out to high transverse momenta. I have presented the method used to obtain these results at the *6th International Workshop on High-p_T Physics at LHC 2011*, Utrecht, the Netherlands, and it is described in proceedings from this conference (to be published).

The main questions asked in this study are:

- How does the performance of the TPC affect the $\mathrm{d}E/\mathrm{d}x$ measurement, and how can this be taken into account?

- What is the performance of the analysis methods used?
- What is the pion yield in pp collisions at $\sqrt{s}$ = 900 GeV, as a function of transverse momentum?

I will begin by describing some of the physics of interest to ALICE, with special attention to the subjects of importance to the kind of analysis I have done. Then I continue by describing the accelerator and experiment in Chapter 3. There the detectors of interest for my analysis are also described. In Chapter 4, the analysis methods are described, and the results are presented and discussed in the next chapter. Here, the performance of the two analysis methods are compared, and improvements are suggested, along with an outlook. Finally, in Chapter 6, I give a concluding summary.

Chapter 2

Physics with ALICE

This chapter gives a brief overview of some of the physics which can be investigated by ALICE (A Large Ion Collider Experiment) at the Large Hadron Collider (LHC) at CERN, outside Geneva on the border between Switzerland and France. This overview provides the theoretical context and physics motivation for the analysis described in Chapters 4 and 5. LHC and the experiment itself will be described in more detail in Chapter 3. Variables used are defined in Appendix A.

I will in the following denote proton collisions with pp, and lead ion collisions with Pb–Pb. In the most general case I will write AA for nucleus-nucleus collisions.

2.1 QCD

2.1.1 Colour

Hadrons, such as baryons (e.g. protons) or mesons (e.g. pions (or *pi mesons*)), are thought to consist of quarks and gluons. Gluons are the mediating particles of the nuclear strong force, or *colour force*, described by Quantum ChromoDynamics (QCD). The sensitivity of quarks and gluons – *partons* – to the strong force is explained by QCD through assigning them a colour quantum number, or "colour charge": conventionally (anti)red/(anti)green/(anti)blue, or linear combinations thereof. I will at times denote quarks by q and antiquarks by $\overline{\mathrm{q}}$. In strong interactions, colour can be exchanged. In the matter surrounding us, normal or *hadronic* matter, quarks are *confined*, and colours are combined in such a way that composite particles are always colourless.

A hadron consists of valence quarks, three in a baryon (qqq, or $\overline{\mathrm{qqq}}$ in an anti-baryon) and two ($\mathrm{q}\overline{\mathrm{q}}$) in a meson. These (and their binding energy) are considered to make up most of the hadron mass. Furthermore, there are *sea* quarks and gluons, virtual $\mathrm{q}\overline{\mathrm{q}}$ pairs and gluons, carrying some fraction of the hadron momentum. Confinement implies that there is no way of knocking out a quark or gluon to look at it separately at a macroscopic scale; in fact it gets more tightly bound the more energy is used in the collision. Another important concept is *asymptotic freedom*, meaning that quarks and gluons behave as if they were free when they are probed at high energies (corresponding to small scales).

2.1.2 Strings and particle jets

The fact that partons are more tightly bound the more energy is invested to separate them, makes it energetically favourable at some point of "separation" to create a new quark-antiquark pair, thus creating new hadrons, instead of separating the original partons further. A model often used to describe this is the Lund string model [1]. In this model, the quarks inside a hadron are often pictured as being connected by strings – field lines for the colour force, squeezed together by the vacuum (or, by the self-coupling between the coloured gluons). In a collision, such strings can be visualised as being stretched, and if they should break, it would be to create a $q\overline{q}$ pair so as to produce two new hadrons, and the system would still be colourless. At very high energies, such a process could be repeated many times, resulting in the production of a large number of new particles. In such a case, there could be two *jets* (collimated particle showers of high p_{T}) produced, back-to-back to conserve the original (linear) momentum. There can also be gluon jets, emerging from the $q\overline{q}$ production resulting when a gluon is radiated (by a quark or gluon). At the LHC, gluon jets are thought to dominate over quark jets.

2.2 Phenomena studied by proton collisions

One of the most often quoted motivations for the LHC and its experiments is the search for the Higgs particle(s), predicted to exist by the Standard Model (SM) of particle physics, but not yet seen at previous accelerator energies. The Higgs medium offers a mechanism by which particle mass, otherwise not emerging from the Standard Model, is generated. Colliding protons at an energy of, eventually, $\sqrt{s} = 14$ TeV is thought to be sufficiently high to cover the Higgs mass as predicted by most theoretical calculations. A simple argument for this is the electroweak symmetry breaking, giving masses to the weak Z and W bosons, indicating that the Higgs mass can't be many orders of magnitude larger than theirs.

Another key goal of the LHC is to provide the energies necessary to find "new physics", physics that goes beyond the Standard Model, which has been very successful in describing much of the particle physics we have seen up to now, but is by no means expected to be close to a final answer. One theory of what might show up is SUperSYmmetry (SUSY), in which every fundamental particle we know of has its supersymmetric (supposedly much heavier, otherwise we would already have seen it) partner ("sparticle"). If this theory holds, a wide range of new (s)particles are expected to show up if we reach high enough in energy. One motivation for SUSY is that it introduces a mechanism for the unification of all the four fundamental forces of nature: the strong, electromagnetic and weak forces and gravity.

Proton collisions at high energy (currently up to $\sqrt{s} = 7$ TeV) will provide the setting for finding which theories are more plausible than others, as well as for improving previous measurements of particle masses, branching ratios, decay modes etc. LHC will, when it reaches $\sqrt{s} = 14$ TeV, provide pp collisions at more than 7 times higher energy than the previous record holder, the Tevatron at Fermilab. Even so, the crossections for producing new phenomena are expected to be extremely small. The LHC is for this reason designed to to deliver collisions at a rate of 10^9 Hz. The two LHC experiments specialised in pp collisions, ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are designed to handle the large luminosity by using fast detectors and advanced triggering, sensitive to collisions of potential interest. A highly selective trigger (on the ppm level) is necessary to reduce the data taking rate to a tolerable level.

Another aspect of being sensitive to new pp collision physics is that one expects either very heavy particles or extremely high energy jets. ATLAS and CMS are focused on these and have traded off capability in the lower p_{T} range, by having very strong magnetic fields and no hadron identification here (which is generally very difficult anyway for p_{T} above a few GeV/c).

However, ALICE is not designed to be the experiment to discover the Higgs, nor SUSY – for example, it doesn't have a full energy coverage over a solid angle of 4π steradians. This would be necessary to be able to deduce "missing energy" i.e. energy carried away by particles not interacting strongly or electromagnetically with the detectors, thought to be a signal of much of the new physics expected at the LHC. Another shortcoming in the context of pp collisions is that ALICE uses a TPC for central tracking, and this is a slow detector which limits the raw collision rate, i.e. the rate at which rare phenomena can be produced. Efficient trigger filtering can't compensate this, so the raw collision rate is limited to about 100 kHz in pp collisions. Thus the sensitivity to very rare processes is decreased. This is, as will be discussed more in Chapter 3, motivated by the many other advantages of a TPC, and by the physics scope of ALICE: this experiment is dedicated and designed to study the physics of heavy ion collisions, where the interest lies not only in high p_{T} but also in the softest p_{T} range which contains information about the bulk, collective quark and nuclear matter properties.

2.3 Heavy ion physics

To collide heavy ions instead of single nucleons gives the possibility to reach new regimes of energy density, as well as to investigate collective behaviour of many nucleons in a high density system. In this kind of collisions the output depends quite heavily on the impact parameter b, i.e. the transverse distance between the centres of the colliding nuclei, quantifying the amount of overlap of the ions in the collision. This is often (and will from now on be) referred to as the *centrality* of the collision, given in percent, where 0% ($b = 0$) means head-on collision ("most central"), and higher percentiles correspond to more peripheral collisions (larger impact parameter). Centrality is a crucial parameter for many characteristic nucleus-nucleus collision properties. For instance, the energy density increases with centrality.

There are a few terms that are particularly important to describe heavy ion collisions. First of all, in a collision where $b \neq 0$, it must follow that only some nucleons collide, as some parts of the colliding nuclei don't overlap. The terminology used to characterise this, is to refer to the nucleons as *participants* or *spectators*, respectively. In addition, one terms each nucleon-nucleon collision a *binary collision*. In a nucleus-nucleus collision, a participant nucleon can suffer many binary collisions, while in pp collisions this number is quite naturally limited to one.

The impact parameter is of course not possible to measure as such. There are different ways of estimating the centrality. For instance, there are dedicated detectors, the two Zero Degree Calorimeters (ZDCs) placed further down the beam direction after the interaction point (those of ALICE can be seen in Fig. 3.3). These measure the number of spectators as the residual neutrons keep travelling along their path after the collision. This can also be used for determining the event plane, a parameter much used in elliptic flow measurements, see Section 2.3.3.

Another way to measure centrality is to consider different observables which relate to the centrality in a monotonous way. For example, the track multiplicity in the event increases with centrality. Picking out the 5% of the events having the largest number of tracks then gives the 5% most central events.

2.3.1 Quark-gluon plasma

It has been predicted by QCD that a new "phase" of nuclear matter, a plasma of deconfined quarks and gluons, could be reached in (ultra)relativistic heavy ion collisions [2]. In the high energy density reached in heavy ion collisions at the LHC (see Chapter 3), and already at RHIC (the Relativistic Heavy Ion Collider), the energy density (or temperature) is thought to become so high that confinement is broken in a *quark-gluon plasma*, a dense coloured medium (see Section 2.3.4). An observable effect of this medium would be that coloured particles created in a heavy ion collision suffer large energy loss as they traverse the medium on their way out. This would be particularly visible for jets, since, by conservation of momentum, single jets are never created, but one observes at least two jets. Two jets would emerge back to back from a collision. But if the jet pair is created in a hot and dense medium, they could lose energy on their way out of it. The amount of energy lost would then increase with path length through the medium (given certain assumptions on hadronisation time etc.). This implies the possibility that, in highly asymmetric cases, one jet could come out only slightly affected by the medium while the other would have lost a substantial amount (or all) of its initial energy – enough for the difference to be measurable. This phenomenon is commonly referred to as *jet quenching*. Where the energy from the quenched jet goes is not completely understood, but recent results [3] point to that it is emitted as low-momentum particles at large angles with respect to the quenched jet direction.

There are however other, albeit perhaps less spectacular, ways of characterising energy loss, at particle rather than jet level. Such a closer look might shed more light on the underlying mechanisms. Some of these will be described in the following.

2.3.2 Hadronisation: fragmentation or recombination?

Parton hadronisation by fragmentation has to some extent already been described in terms of the Lund string model (see Section 2.1.2). The picture is as follows: in a hard scattering, with large momentum transfer, an outgoing parton is produced. Each parton carries net colour and so, to conserve colour, has a colour "string" attached to it. Where this string breaks, $q\overline{q}$ pairs will be produced, leading to what will eventually be a jet. The produced particles will share the initial parton momentum and energy. This mechanism is thought to dominate for particle production at high p_T ($p_T >> 2$ GeV/c).

Another suggested mechanism in the surroundings given by QGP is *recombination* [4]. In this picture, hadrons could be formed by combining deconfined quarks (in the QGP) which are sufficiently close in phase space, rather than by string fragmentation. The resulting hadron momentum is then the sum of the parton momenta, instead of being a fraction of the momentum of a single parton. In this case there would, simply put, be an enhanced probability to form hadrons of a given (high) p_T by recombination compared to forming them via fragmentation. Note, however, that the p_T spectrum for partons is steeply falling[1], making it decreasingly probable to create high-p_T hadrons by this mechanism, simply because the number of partons close enough in phase space decreases.

The effect of increased p_T from recombination would be more pronounced for baryons (qqq) than for mesons ($q\overline{q}$), since for a given maximum parton p_T limit $p_{T,lim}$ for recombination, the baryons get a higher p_T than mesons do. After this $p_{T,lim}$, fragmentation of single high p_T partons will dominate. In particular, in the range where recombination is a more efficient production mechanism, there would be more baryons than mesons in the hadron p_T range (2* $p_{T,lim}$, 3* $p_{T,lim}$), just from summing parton momenta [4].

To see this effect of recombination, one needs identified particles in the relevant p_T range. The recombination mechanism was proposed [4] as an explanation to many of the results seen at RHIC, which will be described in Section 2.3.4. In this case a suggested limit seems to be 2 GeV/c, giving increased p/π ratios in central Au+Au collisions in the p_T region of 4-6 GeV/c.

2.3.3 Flow

The collective behaviour seen in heavy-ion collisions have been taken to justify the application of hydrodynamical models. There are strong indications that the medium behaves like a perfect liquid, meaning that it has minimal viscosity to entropy density (no energy lost to friction in the medium).

If two nuclei collide with only partial overlap, the area of the collision will have an ellipsoidal shape. In a hydrodynamical picture, this leads to a pressure gradient, as illustrated in Fig. 2.1. This would boost particles in the plane of the impact parameter (or, in a slightly more sophisticated view, in the event plane). This is often referred to as *elliptic* flow, or v_2 (referring to the decomposition of flow effects into Fourier components of the angular emission pattern of emitted particles). Effects from flow might make the interpretation of the impact of

---

[1] This is deduced from the steeply falling p_T spectra of produced particles.

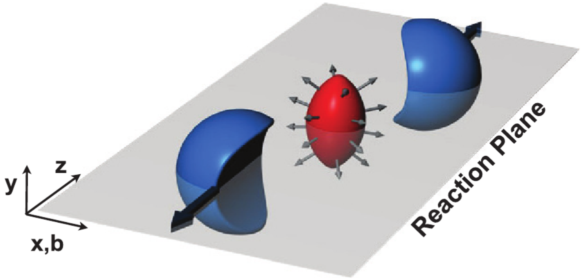


Figure 2.1: Schematic view illustrating the reaction plane, impact parameter and the ellipsoidal shape resulting from a peripheral collision [5].

the medium (energy loss) more difficult. Measurements to single out and quantify this effect are thus very important for interpreting the properties of the medium. The flow direction is strongly correlated to the reaction plane. Since the orientation of the plane can be determined event by event, these effects turn into an advantage, where e.g. the path length through the medium can be taken into account.

It has, however, recently become evident [6] that it is worthwhile to complicate this rather simple picture a little, by looking at the higher order flow components, effects that would enter from fluctuations in the distribution of colliding nucleons. Such fluctuations are illustrated in Fig. 2.2.

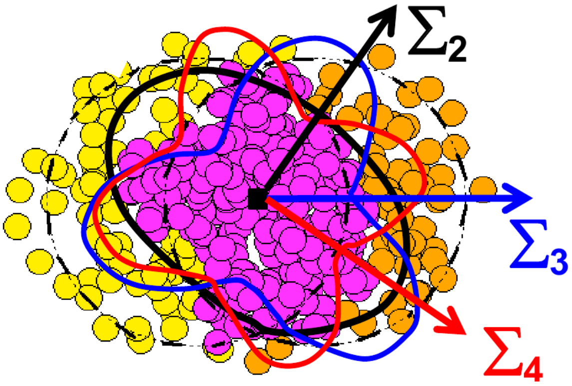


Figure 2.2: Schematic view illustrating the fluctuations in nucleon overlap giving rise to higher order flow (in directions given by vectors Σ_n) [7].

The flow of different particle species is thought to give insight on the evolution of the system, from QGP to a hadronic state, given different assumptions on the formation time of different particles.

2.3.4 Pre-LHC RHIC results

In 2001, there were several striking results [8] reported by the experiments (among them STAR and PHENIX) at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory (BNL), Long Island, USA. Here gold (Au) nuclei were collided at $\sqrt{s_{\mathrm{NN}}}$ = 130 GeV (later up to 200 GeV). One result was the suppression of hadrons at high p_{T}, in this context meaning from ~ 2 GeV$/c$ and up. This is quantifiable in the *nuclear modification factor*, R_{AA} :

$$R_{AA}(p_{\mathrm{T}}) = \frac{(1/N_{evt}^{AA})d^2N_{ch}^{AA}/d\eta dp_{\mathrm{T}}}{\langle N_{coll}\rangle(1/N_{evt}^{pp})d^2N_{ch}^{pp}/d\eta dp_{\mathrm{T}}} \tag{2.1}$$

where N_{ch} is the number of charged particles. This expression gives the ratio of the charged particles yields in AA collisions to pp collisions, in a certain pseudo-rapidity and transverse momentum range, normalised to the number of events N_{evt} and number of binary collisions $\langle N_{coll}\rangle$, so as to mimic a superposition of nucleon-nucleon collisions in the reference pp collision data at the same energy per nucleon. $\langle N_{coll}\rangle$ has to be estimated from a model, often one uses a Glauber model[2]. If no net nuclear suppression or enhancement were present, R_{AA} would be one. The results at RHIC, at $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV, showed a suppression by a factor 4-5 [9, 10].

This observation[3] can be, and generally was, interpreted as suggesting formation of a quark-gluon plasma, see section 2.3.1. At RHIC there has also been reference measurements with d+Au collisions, to make it possible to discriminate between "cold" nuclear matter effects at normal density and those of a quark-gluon plasma.

Another important result at RHIC was anomalous baryon to meson ratios in central Au+Au collisions compared to in peripheral Au+Au collisions, pp, e^+e^- or d+Au collisions, see Fig. 2.3. The ratios suggest that in a certain p_{T} range, mesons are suppressed with respect to baryons, or baryons enhanced, in central AA collisions. This has often been interpreted as an effect of recombination, as mentioned above.

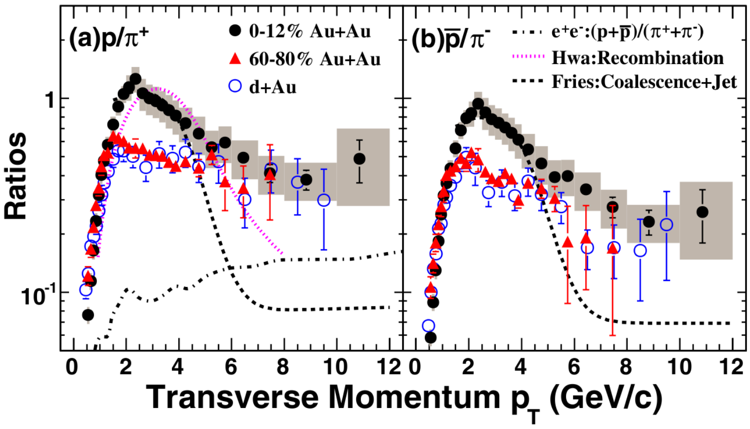


Figure 2.3: Proton-to-pion ratios as a function of transverse momentum (and centrality), measured by STAR at RHIC [9]

It is clear from the experience of RHIC that a series of measurements on different observables needs to be done for disentangling the different nuclear medium effects described above. Here, particle identification at low and intermediate p_{T} play a key role.

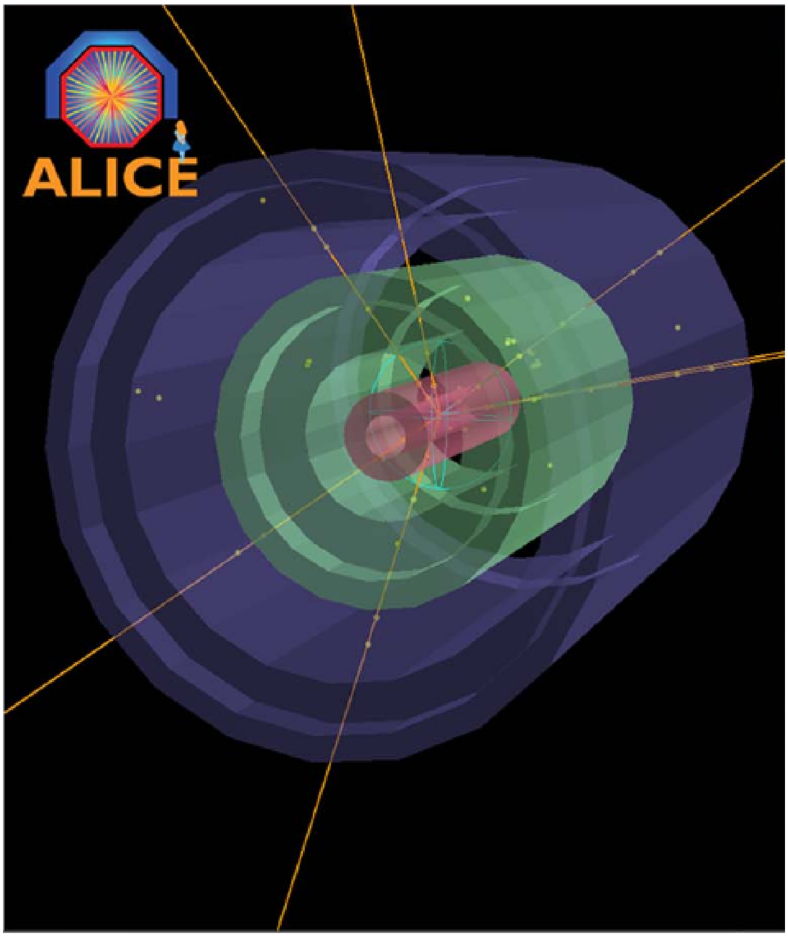
2.3.5 Heavy ions at LHC

From November 5 to December 6 2010, the LHC was colliding fully ionised ^{208}Pb nuclei. Up to 114 ion bunches, each containing about 7×10^7 lead ions, were circulated. The event rate

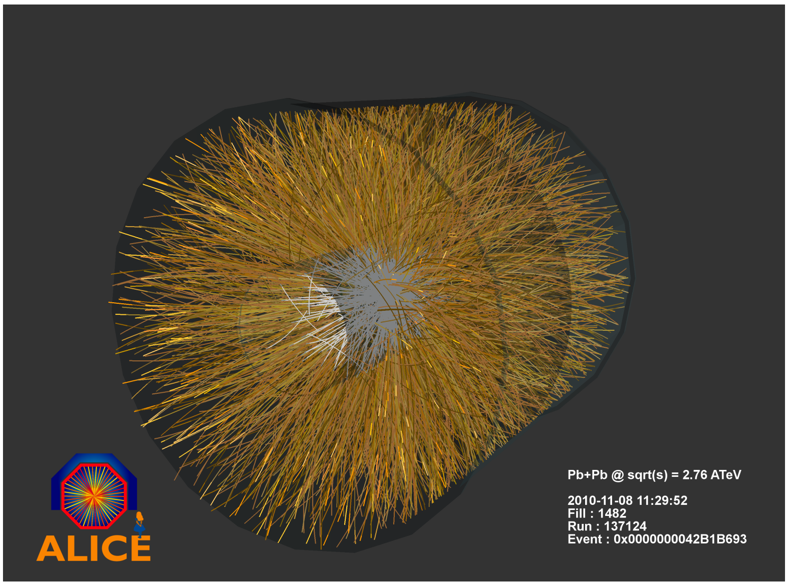
---

[2] Glauber models use a geometrical unseats, with probability distributions estimating the number of nucleons in the overlap region between two colliding nuclei.

[3] This was not the first possible observation of QGP, see for instance [11], for NA49 at the CERN SPS.



(a) pp, $\sqrt{s} = 900$ GeV, ITS



(b) Pb–Pb, $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV, ITS + TPC

Figure 2.4: Event displays from pp and Pb–Pb collisions. (a) The first pp collision candidate event in November 2009 [12]. (b) High-multiplicity event from the heavy ion run period in November 2010 [13].

in ALICE was about 100 Hz and the nucleon centre-of-mass energy was $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV [14]. While the energy per nucleon is lower than for proton collisions (currently $\sqrt{s}$ = 7 TeV for pp collisions), the total available energy can be much higher in central collisions since the number of colliding nucleons is up to 208 times higher (and N_{coll} up to 800 times larger)[4]. The result is a large number of high-energy binary collisions within the volume ov the overlapping nuclei. This large energy density is the key to reaching the new "phase" of nuclear matter, a quark-gluon plasma, which was discussed in section 2.3.1.

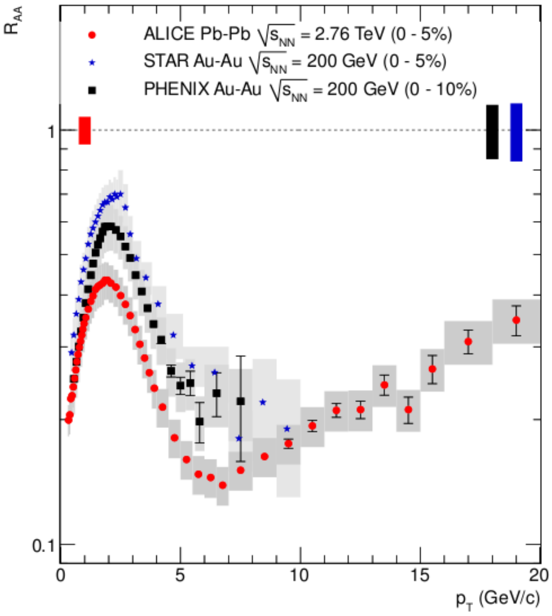


Figure 2.5: R_{AA} for charged particles as a function of transverse momentum, shown for ALICE (red circles), STAR (blue stars) and PHENIX (black squares). Systematic uncertainties in grey [14]. The rectangles around $R_{\mathrm{AA}} = 1$ show the uncertainties in the pp reference.

Heavy-ion collisions offer a high particle multiplicity environment. Fig. 2.4 shows an example of this very clearly. At such conditions, good tracking resolution (i.e., two-track separation) is crucial. This motivates the choice of a TPC for central tracking, since it provides continuous tracking in three dimensions.

One of the results obtained from this run period is the charged particle R_{AA} , seen in Fig. 2.5 for central (0-5%) events. Here the RHIC results are also shown for comparison. From $p_{\mathrm{T}} \sim 2$ GeV/c and up, an increasing suppression is visible, and it is largest at $p_{\mathrm{T}} \sim 7$ GeV/c for the ALICE curve, while it is difficult to draw any corresponding conclusions from the other two, due to large systematic and statistical uncertainties. One explanation offered for the subsequent rise in the curve is that higher p_{T} particles would be less sensitive to energy loss in the medium.

Another key point is that fragmentation is thought to dominate over recombination at higher p_{T}. This would mean that the baryon-to-meson ratios would change with p_{T}. This is not possible to tell from Fig. 2.5, since this shows R_{AA} for *unidentified* charged hadrons. Here identified hadron R_{AA} is needed.

---

[4] However, Higgs searches are not very feasible due to the large QCD background.

2.3.6 High p_{T} physics

On a very fundamental level, the high p_{T} range is interesting since, from momentum conservation, high p_{T} must stem from the creation of particles in the collision, since the incoming particles have very low p_{T}. In addition, at large momentum transfer, interactions can be calculated using perturbation theory ("pQCD"), since large momentum transfer corresponds to small distances and thus coupling strength (asymptotic freedom). This form of QCD gives a good description of pp interactions in the range where it is applicable, while low-momentum QCD is not yet very well understood theoretically.

Much high p_{T} physics will be learnt from jets at the LHC. Both CMS (Compact Muon Solenoid) and ATLAS (A Toroidal LHC ApparatuS), two of the other large LHC experiments, have good jet measuring capabilities at very high p_{T}. For instance, CMS has recently shown indications that quenched jets fragment as they would in vacuum [3].

Two especially interesting topics for ALICE in the region of the relativistic rise are to use the unique particle identification capabilities to describe the fragmentation functions (both in pp and Pb–Pb collisions) and flow for different particle species. Again, for high p_{T}, fragmentation can be calculated in pQCD.

In the hydrodynamical picture, elliptic flow gives a momentum boost to particles, with a magnitude which increases with the particle mass. Thus it should be different for, e.g., π compared to p.

2.3.7 Disentangling the effects

It should by now have become clear to the reader that there is a wide range of nuclear medium effects entering in heavy ion-collisions. To be able to understand their interplay, and to be able to rule out at least some theoretical explanations, it is very important to be able to discern the different behaviours of different produced particle species. To this end, robust particle identification methods are crucial.

In the next chapter, particularly the particle identification capabilities of the ALICE TPC will be described, along with a general description of the experiment and the LHC.

Chapter 3

LHC and ALICE

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at CERN is a particle accelerator for collider experiments. It is located in the tunnel once occupied by the Large Electron Positron collider (LEP), 100 m under ground, almost 27 km in circumference. The beam stored in the LHC has been accelerated not only here but also in many pre-accelerating steps in the large accelerator complex (see Fig. 3.1).

A large advantage of *collider* experiments compared to fixed-target experiments is that the centre of mass frame for the particle interactions coincides with the laboratory frame. This means that all the energy released in particle collisions is available for particle production (and not, as in a fixed-target experiment, partly spent on motion of the centre of mass with respect to the laboratory frame). In the case of storage ring colliders, as the LHC, another feature is that the beam is recirculated. To some extent the smaller hit probability (compared to the thicker fixed targets) is then compensated by the large number of revolutions, increasing the total probability of interaction.

The LHC is a *hadron* collider, which means the colliding particles always have a substructure: nuclei consist of nucleons, nucleons consist of partons (quarks and gluons). These constituents each carry a fraction of the overall particle energy. This means that the energy and momentum of incoming interacting particles are not exactly defined, but have a spread.The hadrons collided are protons (or hydrogen ions) and lead (Pb) ions.

The fact that the LHC is *large* means that the bending radius ρ of the accelerator is large. This reduces the amount of energy lost to synchrotron radiation, ΔE, emitted as the particle trajectories are bent. This was a major issue for the old accelerator placed in the tunnel, LEP, as it was accelerating electrons. ΔE can be related to the mass of the accelerated particle m, beam energy E (in GeV) and bending radius as

$$\Delta E \propto \frac{E^4}{\rho m^4}$$

Now since protons are $\sim$2000 times heavier than electrons, synchrotron radiation is negligible for LHC compared to at LEP (and even more so when colliding Pb ions). However the radius comes in as a limiting factor for the particle energies that can be reached, along with the magnetic field B. For the LHC, which accelerates much heavier hadrons instead of leptons in the old LEP ring, increasing B in the dipoles bending the beam is needed to reach higher energies.

The accelerated protons are produced from hydrogen gas by stripping off the electrons. A similar process, but in more steps (including intermediate acceleration and stripper foils), produces the bare lead nuclei.

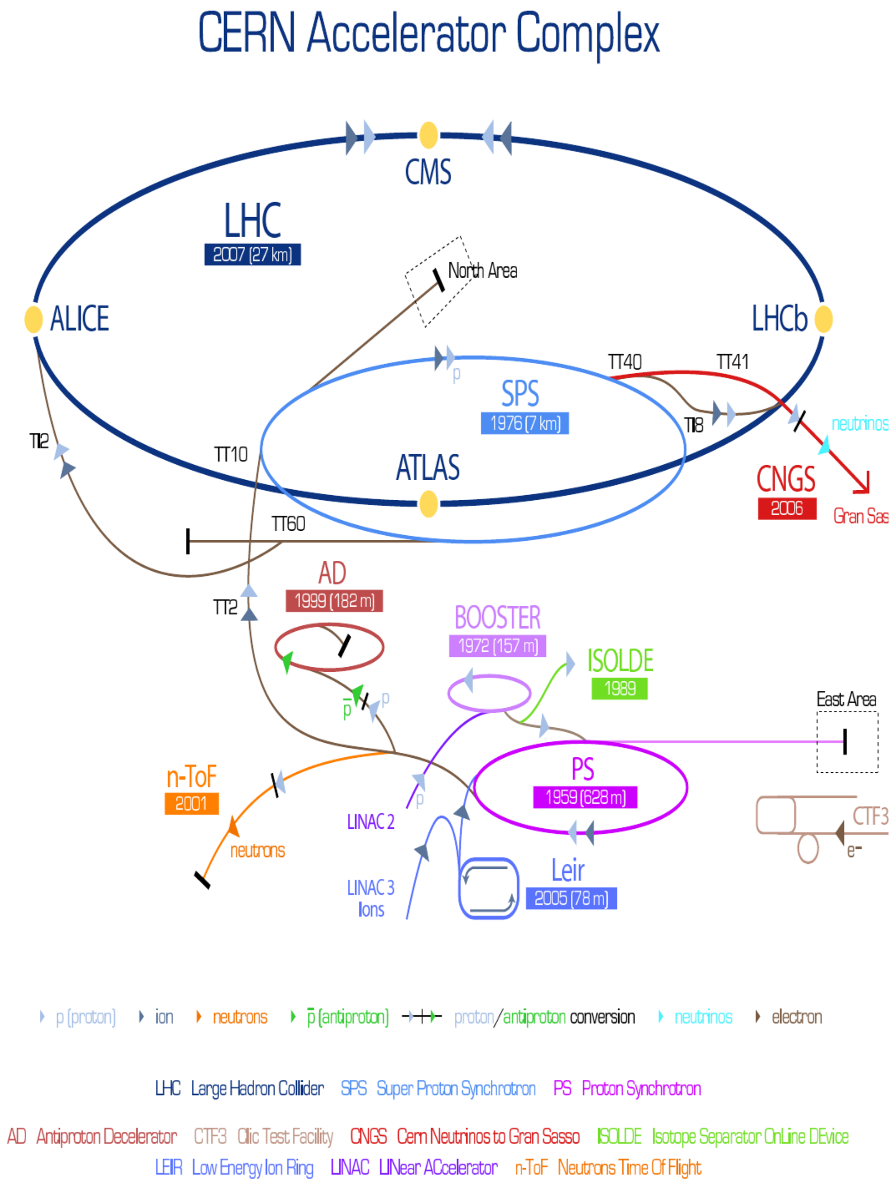


Figure 3.1: Schematic illustration of the LHC accelerator and pre-accelerator complex (not to scale) [15].

As an example, the luminosity recorded by ALICE in 2011 (up to June) is illustrated in Fig. 3.2.

3.2 ALICE

ALICE (acronym for A Large Ion Collider Experiment) [16] is one out of six experiments located at one of four interaction points along the LHC (see Fig. 3.1). It is the only dedicated heavy ion collision experiment at LHC [17].

The basic experiment layout is a central barrel placed inside a room temperature solenoid with (nominal) magnetic field strength of 0.5 T and a forward muon tracking system. It has

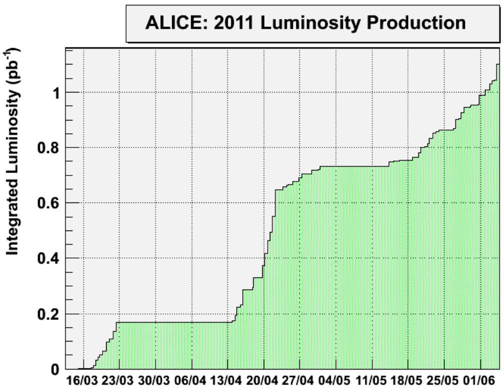


Figure 3.2: Integrated luminosity recorded by ALICE up to June 2011

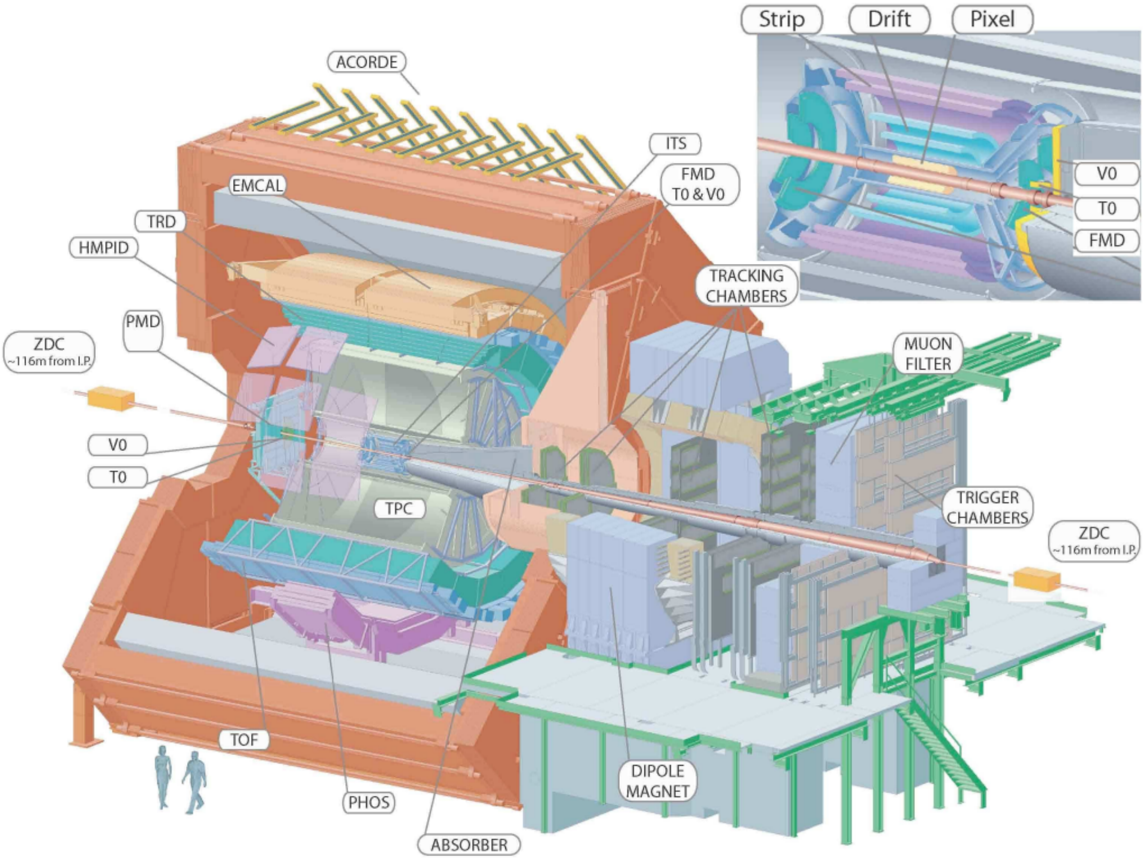


Figure 3.3: Schematic illustration of the ALICE detector with solenoid and sub-detectors [18].

18 sub-detectors[1] (see Fig. 3.3) [16]. The barrel is centred on the beam crossing point, where the collisions take place inside the 800 μm thick, evacuated Beryllium beam pipe [12]. As the main focus of this dissertation is data from the Time Projection Chamber (TPC), I will describe

---

[1] For detailed descriptions of all sub-detectors, see [16].

this detector in some more detail, and briefly mention the Inner Tracking System (ITS) which is placed inside the TPC and also used for the analysis.

At the heart of ALICE sits a cylindrical TPC. This is a gaseous detector, giving minimal influence (e.g. scattering, which influences the direction and makes back-tracking more difficult) on traversing particles before they hit detectors further out. The choice to use a TPC for tracking is motivated by the high granularity and good performance also at high particle multiplicities, as is often the case in heavy-ion collisions. It also has particle identification capabilities in a wide momentum range (more on this in Section 3.2.2) [16]. The TPC is with its drift time of $\sim 90\mu$s the slowest detector in ALICE, but at the low interaction rates in heavy ion collisions this is not a major drawback compared to the advantages.

At the centre of the TPC in turn, surrounding the beam collision point, there is the ITS. This detector (or rather, group of detectors) covers the volume where vertices from the decay of short-lived particles produced in a collision are expected. Requiring hits in the ITS can thus be used as a veto on background events in other detectors, induced by interactions in the beam halo, beam gas collisions, or by cosmic radiation [18]. Similarly, it can be used to reject tracks with a secondary vertex from weakly decaying particles (e.g. $\Lambda^0 \rightarrow \pi^-$+p, where the Λ^0 would leave no track in the ITS since it is neutral, and is generally long-lived enough to decay outside the innermost layer of the ITS). The ITS is a fast six-layer silicon detector (the inner layers being silicon pixel and drift detectors, the outer double-sided microstrip detectors), giving good resolution and hardware triggering[2] capabilities. It also has energy loss measurement capabilities and can be used for particle identification in the non-relativistic region [16].

3.2.1 The ALICE TPC

Principle of operation

A TPC is a detector with three dimensional tracking capabilities: very schematically, ionisation electrons released in the gas along the trajectory of a charged particle drift in an electric field toward a read-out plane where the position of arrival is determined in two dimensions. This is illustrated in Fig. 3.4. From the arrival time and the known drift velocity in the gas, the third spatial coordinate can be reconstructed. For the latter point, is is necessary to know at what time the original ionising particle traversed the detector. This time is given by the trigger.

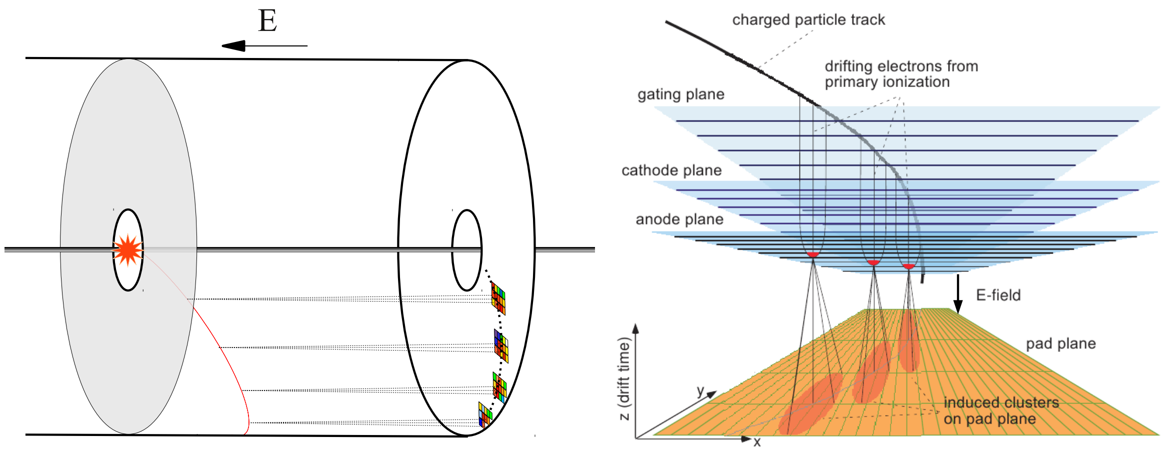
TPC specifications

Many of the ALICE TPC facts and figures are summarised in Table 3.1, and a schematic drawing of it is seen in Fig. 3.5. Here the central cathode and end plates, creating the electric field directed towards the read-out, are seen. The two end plates are segmented into 18 sectors each of Read-Out Chambers (ROCs), equipped with MultiWire Proportional Chambers (MWPCs) (also seen in Fig. 3.5). As the name implies, these read out a signal proportional to the number of incoming electrons from the traversing particle track. The signal is strongly amplified in the vicinity of the wires by a (to first approximation, in the absence of distortions) cylindrically symmetric field, proportional to $1/r$. r is here the distance to the wire acting as anode. "Below" the anode wires, there is a cathode pad plane where the induced signal is read out, as was seen in Fig. 3.4(b).

The ROCs are subdivided radially into inner and outer, respectively (IROC and OROC). The read-out pads have different sizes (see Table 3.1) to account for the radially decreasing track density [16].

---

[2] A trigger is generally speaking a signal in at least one detector which indicates that other detectors should collect data. In practise more than one detector is used, for better ability to set specific trigger conditions. For triggers used for this analysis, see Appendix B.



(a) (b)

Figure 3.4: Schematic illustration of the TPC principle of operation. (a) A charged particle produce in a collision traverses the TPC volume in a trajectory bent by the applied magnetic field (red line). It ionises the gas, and released electrons drift toward the end plate, where a signal is read out on pads. (b) Schematic illustration of the read-out principle [13].

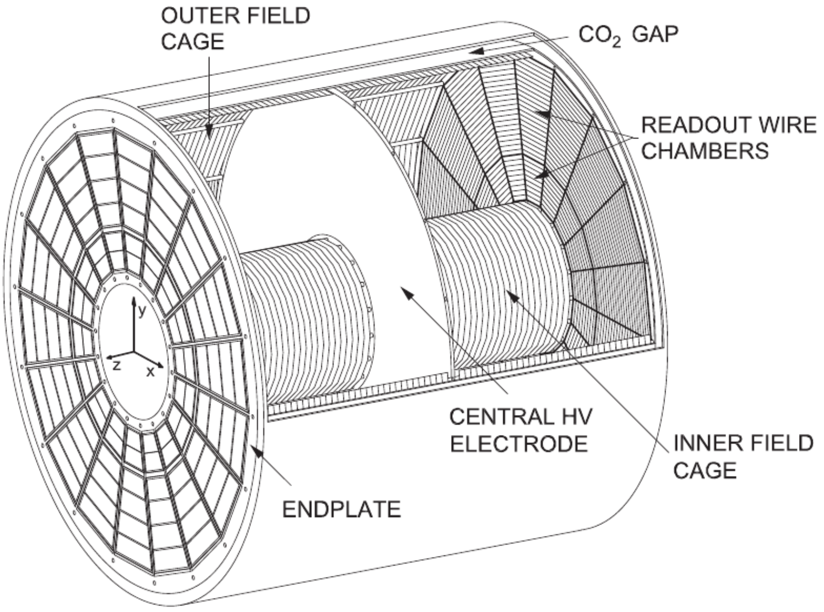


Figure 3.5: Schematic layout of the ALICE TPC [17].

3.2.2 Particle identification

There are many different approaches to particle identification. Methods such as measuring the particle's time-of-flight require dedicated detectors (as is the case in ALICE: the TOF detector). There are complementary methods which recognise particles by what interactions they are subject to, as in electromagnetic and hadronic calorimeters, and muon identifiers. In ALICE, essentially all established particle identification techniques are used [16]. This is important to cover a large momentum range, as different techniques cover slightly different sub-ranges.

Pseudo-rapidity range (full track length)	$-0.9 < \eta < 0.9$
Azimuthal coverage	360°
Active volume, radial position	$85 < r < 247$ cm
Active volume, longitudinal position (along beam axis)	$-250 < z < 250$ cm
IROC pad dimensions	4×7.5 mm^2
OROC pad dimensions	6×10 and 6×15 mm^2
Number of read-out pads	~ 560000
Drift gas mixture	Ne-CO$_2$-N$_2$ $(85.7\% - 9.5\% - 4.8\%)$ [19]
Electric drift field	400 V/cm
Maximum drift time	94μs
Cathode high voltage	(-)100 kV
Typical gas gain	$\sim 10^4$

Table 3.1: Table of TPC specifications (at the time when the data used for this analysis were taken (where applicable)) [14, 19]

The TOF, for example, can roughly speaking be used for identifying non-relativistic particles, whose velocity can be measured with enough precision.

The method used for particle identification in this analysis is to use two main physical properties of a moving particle as measured in the TPC: its momentum p and its energy loss. These measurements rely on many physical relations, as will be described below.

Momentum

Momentum is a conserved quantity, which makes it very useful for reconstructing physical processes where not every step is detected. An example would be a particle A decaying without leaving tracks within the detector, which is thus not directly seen, but where the tracks of the decay products are registered. Knowing that the momenta of these must sum up to the momentum of a decaying particle enables the reconstruction of the momentum of A.

Since mass and charge is sufficient to recognise already well-known particles, knowing for instance the momentum p and velocity v of a particle is often used for particle identification, via the relation

$$p = m\beta\gamma, \tag{3.1}$$

where m is the particle (rest) mass, $\beta = \frac{v}{c}$, c being the speed of light in vacuum, and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$.

To measure the momentum in the TPC, one uses the fact that charged particles in motion are affected by a magnetic field. As a charged particle moves in the TPC, which is surrounded by a solenoid magnet, its trajectory bends perpendicularly to both the magnetic field lines and line of motion, in one direction or the other according to the sign of the particle charge. By equating the Lorentz and centrifugal forces, the bending radius R of the particle trajectory can be expressed as

$$R = \frac{p}{qB} \tag{3.2}$$

where q is the (magnitude of the) charge. So, from the reconstructed particle track in the TPC, the particle momentum can be found. However, from the equation it is seen that for a given magnetic field, when p goes to infinity, $R \to \infty$, i.e. the track will be practically straight. Note also that it is the momentum component perpendicular to the magnetic field, i.e. the transverse momentum, that gives the bending radius.

Energy loss

As a particle traverses matter, it loses energy by different interactions. This energy loss is often denoted by $\mathrm{d}E/\mathrm{d}x$, which means change in energy E per infinitesimal unit of distance x travelled by the particle. In the TPC gas, charged particles predominantly lose energy by electromagnetic interactions with electrons in the gas.

Energy loss is described with the Bethe-Bloch function [20]

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A}\frac{z^2}{\beta^2}\left[\ln\left(\frac{2m_e\gamma^2 v^2 W_{max}}{I^2}\right) - 2\beta^2 - \delta - 2\frac{C}{2}\right], \tag{3.3}$$

where N_a is Avogadro's number, r_e the classical electron radius, m_e is the electron mass, ρ, Z and A are the density and atomic and mass number of the absorbing material respectively (in this case the TPC gas). z is the charge of the incident particle (in units of e) and I is the mean excitation potential, which is generally given by semi-empirical formulae. δ and C are density and shell corrections, respectively. The maximum energy transfer is given by (with M for the incident particle mass)

$$W_{max} = \frac{2m_e c^2(\beta\gamma)^2}{1 + 2m_e/M\sqrt{1+(\beta\gamma)^2} + (m_e/M)^2}, \tag{3.4}$$

which for $M >> m_e$, as for all hadrons, becomes $W_{max} \approx 2m_e c^2(\beta\gamma)^2$.

For low $\beta\gamma$, or momentum (see Eq. 3.1), Eq. 3.3 is dominated by the $\frac{1}{\beta^2}$ term. Here it falls steeply with increasing momentum until it reaches a minimum which is often referred to as the *MIP region* (for Minimum Ionising Particle). As momentum increases after this point, so does $\mathrm{d}E/\mathrm{d}x$, but here the logarithmic term dominates. This region is often referred to as the *relativistic rise*, as the particles are relativistic in this regime, and also because energy loss here is modified by a relativistic effect.

For a given detector gas, Eq. 3.3 is a function of $\beta\gamma$ and incident particle charge only (neglecting the small M dependence of Eq. 3.4). For the charged particles most often detected in high energy physics, the charge is $1e$. From this and Eq. 3.1, it is seen that the different particle species will follow each their own Bethe-Bloch curve when given as a function of p, displaced by an amount given by the relation between the particle masses. In particular, on the relativistic rise, the separation between particle species "1" and "2" at a given p will be nearly constant, since it is given by

$$\langle \mathrm{d}E/\mathrm{d}x\rangle_1 - \langle \mathrm{d}E/\mathrm{d}x\rangle_2 \propto \ln\left(\tfrac{p}{m_1}\right) - \ln\left(\tfrac{p}{m_2}\right) = \ln\left(\tfrac{m_2}{m_1}\right)$$

Energy loss in gas, with the mentioned possibility of large energy transfer, is often better described by an empirically determined parametrisation. This analysis uses

$$\mathrm{d}E/\mathrm{d}x = \frac{a(1+\beta^2)}{\beta^2} + b\cdot\log(1+\beta\gamma) \tag{3.5}$$

Here a and b are free fit parameters.

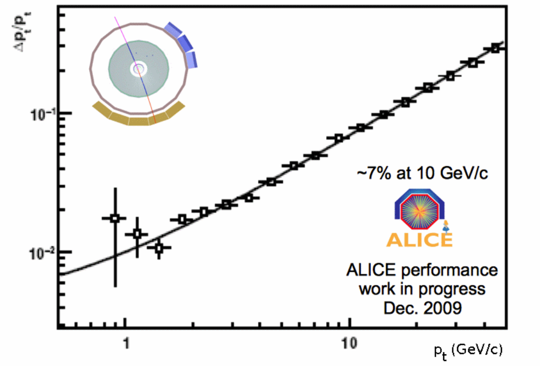
3.2.3 Position and momentum resolution of a TPC

The influence of the magnetic field on particle trajectories also holds for ionisation electrons. This is advantageous for the position resolution of the TPC, since the ionisation electrons will follow helical trajectories as they move in both the electrical and magnetic field. This reduces diffusion in the direction perpendicular to their overall motion, which enables a more accurate determination of track coordinates in these directions. However, given the gas composition and comparatively weak magnetic field, this is a small effect in the ALICE TPC.

The momentum resolution generally increases with longer track length and larger magnetic field, which increases the track curvature (as was seen in Eq. 3.2). Thus both this and the former point potentially benefit from having a strong magnetic field.

In ALICE, the choice of magnet was motivated by the possibility to use the already existing solenoid from the L3 experiment at LEP. The magnetic field was sufficient, considering the established read-out technologies (MWPCs), and the TPC gas was chosen to match the requirements set by these two other parameters, and the requirements to use non-flammable gas in the 90 m^3 volume detector.

The ALICE TPC momentum resolution as of December 2009 is seen in Fig. 3.6.



ALICE performance
work in progress
Dec. 2009

Figure 3.6: ALICE TPC momentum resolution, as determined from cosmic tracks.

3.2.4 The TPC $\mathrm{d}E/\mathrm{d}x$

The $\mathrm{d}E/\mathrm{d}x$ of a traversing particle is statistical in nature, with a mean value and large fluctuations. This results in a distribution as the $\mathrm{d}E/\mathrm{d}x$ is sampled along the track, in up to 159 points – *clusters* – in the ALICE TPC. The clusters in a track are the points on each readout pad row, along the track, where the centre-of-gravity of the deposited charge on this pad row is found (in the track reconstruction algorithms). This is schematically illustrated in Fig. 3.7.

There is quite a large probability of energy loss via few instances of large energy transfer. In a gas, being a thin medium, the $\mathrm{d}E/\mathrm{d}x$ distribution gets a corresponding high-energy tail,[3] and is generally best described by a Landau distribution, whose mean does not coincide with the most probable value. A generated example distribution is seen in Fig. 3.8. The $\mathrm{d}E/\mathrm{d}x$ is in the ALICE TPC (as is the general approach) calculated as the truncated mean, including only the lowest 60% of the distribution – this is sometimes referred to as the "TPC signal", but I will in the following simply denote it by $\mathrm{d}E/\mathrm{d}x$.

For many instances of a given particle species at a given $\beta\gamma$, the $\mathrm{d}E/\mathrm{d}x$ will in turn be a distribution around a mean value, with a certain width. This width is a measure of the $\mathrm{d}E/\mathrm{d}x$ resolution, and depends on many parameters, such as gas composition and pressure, drift velocity and even amplification. All in all, these processes are not theoretically very well understood to all details, but can be modelled on a semi-empirical basis.

The conversion from raw signal in a read-out pad into a $\mathrm{d}E/\mathrm{d}x$ value requires good knowledge of the total gain (both in gas amplification and electronics) for each pad. This is mapped

---

[3] In a thicker medium, the larger number of energy transfer collisions approach a Gaussian shape (by the Central Limit Theorem).

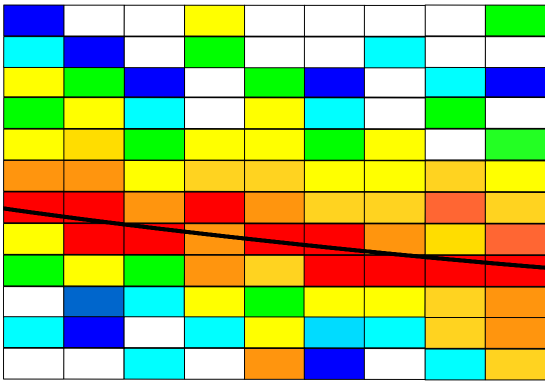


Figure 3.7: Illustration of track reconstruction from charge deposited on pads. The amount of charge is illustrated in a colour scale where red corresponds to a large charge deposit and blue to small. White means no signal is registered. The reconstructed track path is represented by the black line.

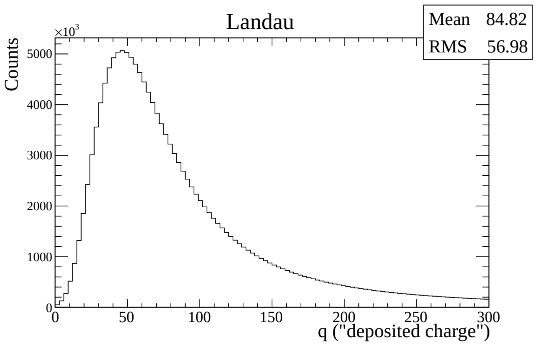


Figure 3.8: Illustration of a Landau distribution. The mean is much larger than the peak value (most probable value), here at $q \approx 85$ and $q \approx 55$, respectively.

by a reference, in the ALICE case the decay of radioactive Krypton (${}^{83}_{36}$Kr) is sed 17. Recalibration of this response is only needed after shut-downs, changes of gas etc., when the conditions have been substantially changed.

On a shorter time scale, gas parameters such as pressure and temperature vary. This will affect the drift velocity, which in turn affects e.g. the diffusion. These variations are not likely to be seen e.g. within data taken in the same run (for definition see Section 4.1).

One factor which varies from track to track, and has large impact on the $\mathrm{d}E/\mathrm{d}x$ resolution, is the drift length of the ionisation electrons. This depends on the track angle with respect to the beam axis. The influence of the drift length is complicated. On one hand, a larger drift length, as for a track perpendicular to the beam axis, leads to more diffusion, and lost electrons. On the other hand, it also implies less ionisation per cluster (the actual track length for a perpendicular track is the same as its projection on the read-out plane, while tracks at smaller angles to the beam axis have longer paths through the detector). This introduces a complicated dependence of $\mathrm{d}E/\mathrm{d}x$ on the track angle, since the $\mathrm{d}E/\mathrm{d}x$ dependence on the number of track clusters in the TPC, from now on denoted by *ncl*, is not straightforward. More details on this will be given in Section 4.5.1.

Chapter 4

Analysis: method

As described in Chapter 2, to be able to extract the properties of the medium formed in Pb–Pb collisions, all heavy ion analyses need a solid reference frame from pp collisions. In the year 2010, protons have mainly been collided at $\sqrt{s} = 0.9$ and 7 TeV. There has also been a special short reference energy pp run period in 2011, where $\sqrt{s} = 2.76$ TeV, to match the nucleon energy in the Pb–Pb collisions in the heavy ion run of 2010.

In order to extract the maximum information, observables such as elliptic flow and nuclear modification factor must be related to particle species in the interesting p_{T} range mentioned above. In other words, these measurements have to be performed for identified particles on the relativistic rise. This is one of the main motivations for doing the analysis described in the coming chapters.

The analysis methods are used to identify π, K and p on the relativistic rise. For this particular analysis, the focus with respect to physics results is on pp collision data taken with $\sqrt{s} = 900$ GeV. The main goals are to extract the charged pion fraction of particles produced, their p_{T} spectra and the cross section at this energy for charged pion production. However, other data sets are partly analysed and presented here for comparison and evaluation of the methods.

The software for the two analysis strategies presented here has been gradually developed, within the ROOT and AliROOT frameworks, primarily by Peter Christiansen and Alexandru Dobrin, mostly as part of the Ph. D. thesis work of the latter. There has later been some modifications by Philippe Gros and myself, where my major contributions are to the *ncl* method analysis software.

The analysis relies on a good understanding of the $\mathrm{d}E/\mathrm{d}x$ measurement in the TPC. In this chapter, different aspects of this are discussed, along with an outline of the analysis methods used. In addition, the different selection criteria for the data samples used are described and briefly motivated.

4.1 Run, event and track selection

The collision data consists of *events*, meaning a triggered instance of data recording, containing any number of *tracks* (including none) from particles interacting electromagnetically with the TPC gas. The events are grouped into *runs*, i.e. a sequence of events recorded during a short time period. For a run, the beam and detector conditions are considered to be constant. The runs are, for convenience, grouped into *run periods*, denoted by a number (given by the year of the data taking) and a letter, e.g. "10b". Track reconstruction etc. from the raw data is then done iteratively. Each iteration is called a *Pass* and given a number. Within each given Pass of a run period, track reconstruction algorithms are the same. The data used in the analysis are specified in Appendix B.

For the analysis, inelastic collision events are used. The event selection is done using a *minimum bias* trigger, meaning a trigger which introduces the smallest bias possible while requiring certain conditions to select inelastic events. For more details on the event and track selection criteria, see Appendix B.

The data are finally arranged using a ROOT functionality called a *tree*. In a tree, with *branches* and *leaves*, different observables of a track are still connected to the track (as opposed to the situation in, e.g., a histogram). This way, different properties can easily be correlated, for example charge, momentum, transverse momentum and energy loss. The main advantage is that it is easy to loop over data multiple times to extract and apply corrections, refine calibrations etc. on a track-by-track basis, which is necessary for this analysis.

The focus of this analysis will be the 10c Pass3 data. This Pass is likely to be the final iteration of this data set. Since it is at $\sqrt{s}$ = 900 GeV, and the LHC has since continued to reach higher energies, there will probably not be more data taken at this energy. This means that the analysis presented here is on the, most likely, final ALICE data for this energy, and it goes all the way to presenting the invariant charged pion yield results for 900 GeV.

On the other hand, statistics are limited compared to $\sqrt{s}$ = 7 TeV data (such as 10b and 10d); at lower energies, fewer tracks are produced (especially at high p_{T}). For this reason, the p_{T} reach in figures shown here will often be limited to $\sim$ 10 GeV/c, even though the methods used can go much farther (presently up to 20 GeV/c for a high-statistics 10d sample).

4.2 Quality cuts

The dE/dx analysis is focused on the relativistic rise. For this reason, in most of the analysis there is a lower momentum limit of p_{T} = 2 GeV/c, so most of the figures shown here have no points at low p_{T}. However, this is not always the case since the initial calibration steps require some data at lower p, so a fraction of them[1] are kept. This can be seen in some figures.

Since the TPC readout is segmented into sectors (see Section 3.2.1), there are some uninstrumented areas in between. Particle track sections crossing these regions will then not produce signals. This leads to fewer clusters registered for these tracks, which reduces the dE/dx and p_{T} resolution. Therefore a cut on p_{T} and track angle relative to the sector is introduced, to remove these tracks. This cut will in the following be referred to as a *ϕ cut* and is illustrated in Fig. 4.1. Here the ordinate is the angle (in radians) with respect to the sector edges, but shifted so as to get the edges in the middle. It is observed that low p_{T} tracks lose fewer clusters as they cross a sector edge, as the track bends in the magnetic field. Higher p_{T} tracks are "stiffer", and for such tracks originating near the edges, a larger proportion of the total track length will be lost since they cross the edges at smaller angles. On the other hand, the range of angles affected gets narrower with increasing p_{T}.

The effect of the ϕ cut on the distributions of $\langle ncl \rangle$ vs p_{T} is seen in Fig. 4.2. Here it is clear that the $\langle ncl \rangle$ is higher after the cut, and that the p_{T} dependence is more or less flat after p_{T} = 2 GeV/c. A lower limit cut $ncl > 70$ for tracks used in the analysis is also seen for $p_{\mathrm{T}} \geq 1$ GeV/c in Fig. 4.2(b).

4.3 Calibrations

The track angle to the beam axis, θ, is often represented by pseudo-rapidity η (see Appendix A). The pseudo-rapidity relates to the track angle in such a way that tracks orthogonal to the beam axis have $\eta = 0$ – this is often referred to as mid-rapidity, and will be so in the following.

As has been described before (Section 3.2.4), it is quite reasonable to assume that the dE/dx would be affected one way or the other by the path and drift length, and thus display a dependence

---

[1] The bulk of the data is really consisting of low-p_{T} tracks, as is realised from the shape of the p_{T} spectra. For this reason (and considering storage), a smaller fraction suffices for doing calibrations.

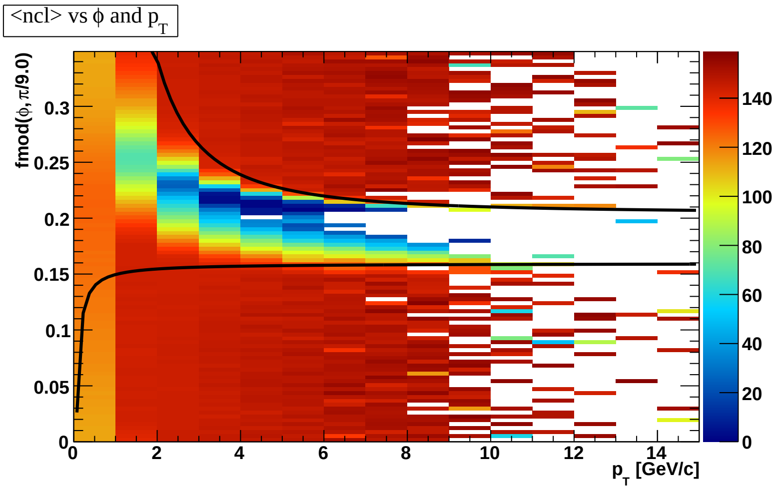
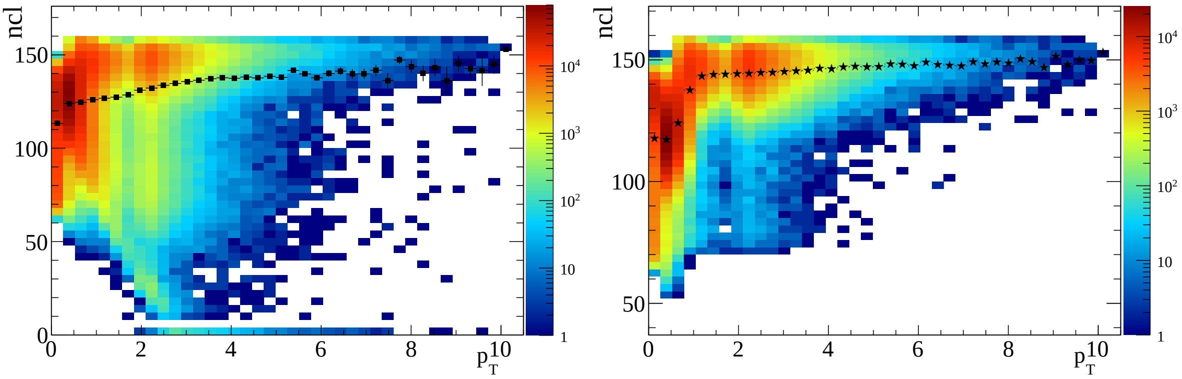


Figure 4.1: Example of the distribution of $\langle ncl \rangle$ in colour scale, with the applied ϕ cut marked by the lines (tracks in the area between the lines are rejected). Charge sign and magnetic field polarity are here compensated so all tracks seemingly bend in the same direction.



(a) No cut

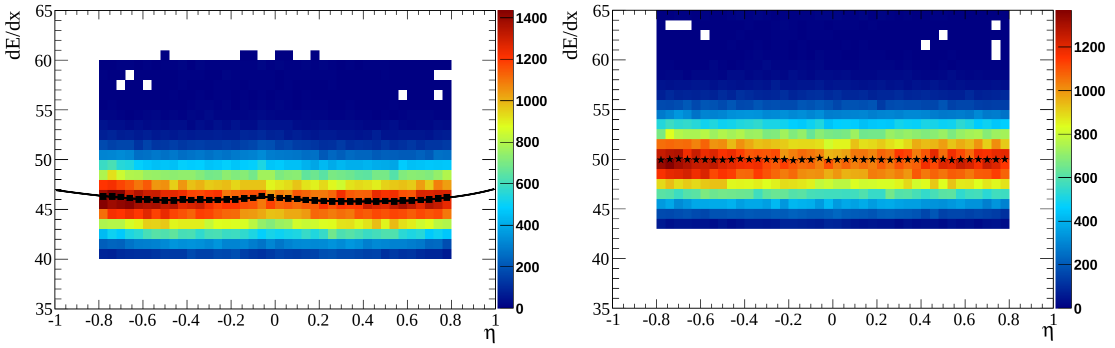
(b) With cut

Figure 4.2: Comparison of the $\langle ncl \rangle$ before ((a), mean marked with squares) and after ((b), mean marked with stars) the cut on p_T and track angle relative to sector edges, for the 900 GeV run 10c.

on η. This is also what is observed, as can be seen in Fig. 4.3, which shows the η dependence of dE/dx for MIP pions, which should all have the same dE/dx. Figure 4.3(a) shows the MIP $\langle \mathrm{d}E/\mathrm{d}x \rangle$ vs η and the calibration fit to these points. Since what has been described here clearly can be viewed as an effect entering in the TPC rather than having a more profound physics message, this dependence is corrected for in the analysis. The MIP $\langle \mathrm{d}E/\mathrm{d}x \rangle$ vs η after calibration is also shown, in Figure 4.3(b). In all these figures, the MIPs are selected with the conditions $0.4 < p < 0.6$ and $40 < \mathrm{d}E/\mathrm{d}x < 60$.

The dependence of ncl on η for the 10c MIPs is shown in Fig. 4.4.

The η calibration of the MIP dE/dx is then propagated to the dE/dx of all particles via a rescaling of dE/dx with a factor 50 divided by the calibration function value at the track η. This falls back on the MIP calibration setting $\langle \mathrm{d}E/\mathrm{d}x \rangle$ to 50 (see Fig. 4.3).



(a) Before calibration

(b) After calibration

Figure 4.3: MIP dE/dx vs η, before and after first (MIP) calibration of the 10c data. The calibration function fit is also shown.

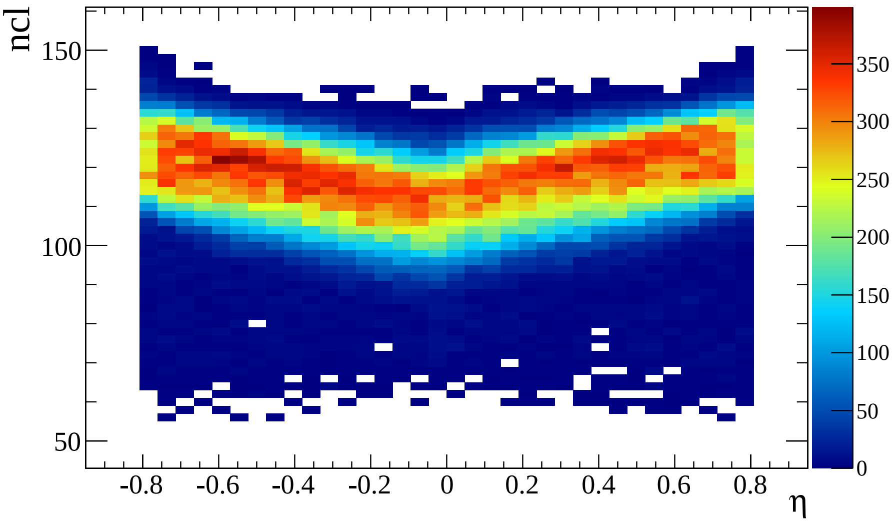


Figure 4.4: MIP ncl vs η, 10c.

Fig. 4.5 shows $\Delta_\pi\equiv$ dE/dx-$\langle$dE/d$x\rangle_\pi$ vs η in a region $|\Delta_\pi| < 5$, which (coarsely) selects π. Mean values for each η bin are shown, for two p_T ranges. Both the p_T and η dependence of dE/dx are seen. For π at low p_T, $\langle\Delta_\pi\rangle$= 0 for high pseudo-rapidity, and smaller (about 1 unit (ADC count) less) at mid-rapidity. For higher p_T, the η dependence is steeper, and $\langle\Delta_\pi\rangle$= 0.5 at the highest pseudo-rapidities. Additional cuts on higher p_T don't change this picture.

The η dependence of dE/dx is complicated since it is already corrected for, to first order, by an angle scaling in the dE/dx algorithms. The ncl η dependence is to some extent easier to understand; mid-rapidity tracks have shorter path lengths (per pad) in the TPC and thus should give fewer electrons per pad, meaning that there should be more clusters lost. The dependence is to first approximation linear, and with a p_T dependence which should be possible to parametrise. To correct for this (really, a path length correction of ncl) is an alternative to correcting dE/dx, which also has a species dependence. One possibility would be to to study the properties of a recalculated dE/dx, using an η calibrated ncl, i.e. in effect where clusters are

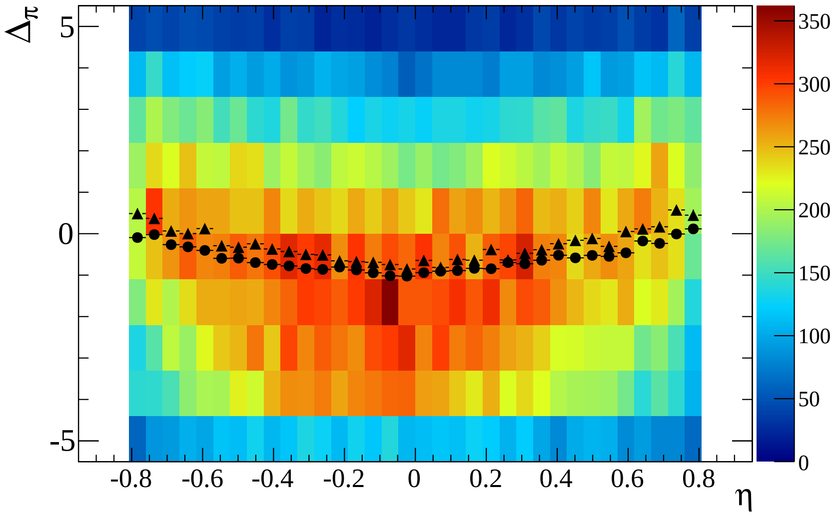


Figure 4.5: Δ_π vs η, 10c, after calibration. π means are shown with circles for $p_\mathrm{T} < 3$ and triangles for $p_\mathrm{T} > 3$.

added as a function of path length to compensate for losses. Clearly, to be done correctly, this requires some insight in the Landau distributions for the tracks, thresholds etc., and is really a deep-level study. It would however be interesting to see what correcting for the η dependence of *ncl* would do to the η dependence of $\mathrm{d}E/\mathrm{d}x$.

So, what is the interpretation of these figures in terms of implications for the analysis? Apparently, the initial calibration of $\mathrm{d}E/\mathrm{d}x$ with respect to η straightens the dependence out for the MIPs, but does not remove it in the p_T region of the analysis. In fact, for π above $p_\mathrm{T} = 3$ there is quite a strong η dependence. Most likely, this smears the $\mathrm{d}E/\mathrm{d}x$ distributions, making the separation worse at mid-rapidity.

There is an effect entering at low p_T from doing the η calibration as a function of p. In this region, p and K are not fully on the relativistic rise, where there is a simple logarithmic mass scaling giving constant separation. Here there is thus an η dependence of the separation within a given p_T bin, since higher η tracks have higher momentum than the rest at a given p_T, and the separation increases with momentum. The $\mathrm{d}E/\mathrm{d}x$ separation between, say, high-η p and π, is then inherently better than for the rest. In this region, before p are on the relativistic rise, the method might improve from also doing an η binned analysis.

A recalibration of $\mathrm{d}E/\mathrm{d}x$ vs η at $p_\mathrm{T} > 3$, from where the p_T dependence seems negligible, is an option which is more easily realised. This was attempted, but from a lack of time, not studied in detail. It remains an open question if such a recalibration should be implemented as a shift in $\mathrm{d}E/\mathrm{d}x$ with η or if it should be a scale factor as for the MIPs.

4.4 Method I: Particle identification using $\mathrm{d}E/\mathrm{d}x$

Each cluster in a track corresponds to a deposit of charge from ionisation along the particle trajectory, which is amplified and read out. As mentioned before (Section 3.2.4), the amount of charge in a cluster follows a Landau like distribution. The particle $\mathrm{d}E/\mathrm{d}x$ is calculated as the truncated mean of the lowest 60% of this distribution.

The underlying assumption of the analysis is that the amount of ionisation, or energy loss, depends only on one parameter $\beta\gamma = \frac{p}{m}$ (giving particle species and momentum) – this was shown in the Bethe-Bloch formula for energy loss given in Section 3.2.2. By looking at the dE/dx distribution in a given momentum range, one might therefore be able to distinguish and identify different particles by their characteristic energy loss. This method is often applied in the momentum region below 2 GeV/c, where many particle bands are easily discerned even just by eye, as is seen in Fig. 4.6. However (as is also visible in this figure), at higher momentum, the different particle species have a much smaller separation in dE/dx. In this range it becomes impossible to associate each track with a given particle species , just based on the track dE/dx. In this relativistic rise region of the dE/dx curves, one can instead use a statistical method, where the overall yields of different particles can be determined from fitted distributions.

The physics motivation to carry out this quite difficult task has been given in previous chapters. It should however be stressed that there are presently very few other ways of identifying charged hadrons to good precision in this p_T region. One example is to use the in-flight kink topology of weakly decaying hadrons, such as Λ and K^0, but this method suffers from limited statistics and strict (and non-trivial) selection criteria due to large background.

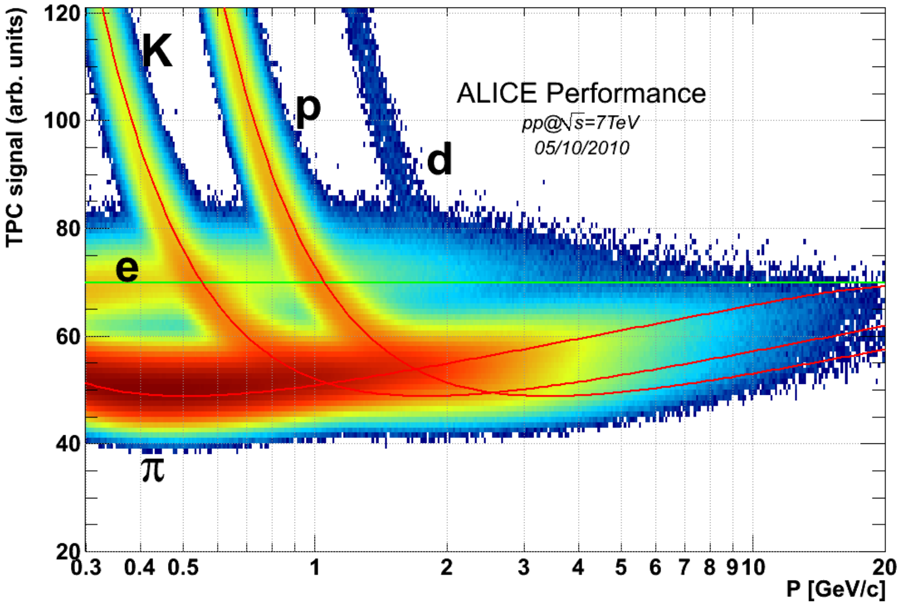


Figure 4.6: TPC signal, corresponding to arbitrary units of dE/dx, vs p.

Fig. 4.6 demonstrates the dE/dx measured in the TPC in a wide momentum range. The corresponding histogram from the analysis at hand, for the 10d data, is shown in linear scale in Fig. 4.7. Here the p region of interest for this analysis is demonstrated. In the top left corner, a cut rejecting mainly electrons is seen. It is also seen that from $p \approx 3$ GeV/c and up, protons, kaons and pions are all on the relativistic rise.

4.4.1 The Δ_π method

One of the two methods used in this analysis for particle identification on the relativistic rise is the so-called Δ_π method. The main steps are as follows:

1. The dE/dx vs p histogram is fitted with a parametrisation of the Bethe-Bloch curve.

2. From this fit, the $\langle dE/dx \rangle_\pi$ is determined. For p and K, the $\langle dE/dx \rangle$ is found from this curve via a transformation to $\beta\gamma$.

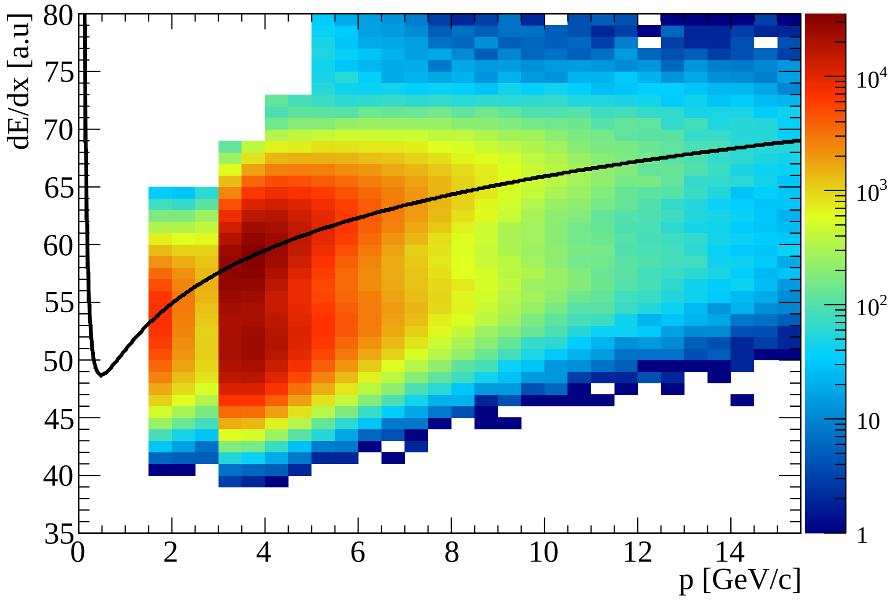


Figure 4.7: dE/dx vs p, with the π Bethe-Bloch parametrisation fit, for the 10d data.This illustrates steps 1 and 2 of the Δ_π method.

3. For each track, a Gaussian distribution around the Bethe-Bloch parametrisation value for the track p (or, $\beta\gamma$) is generated. These distributions are then used to make a mapping of the dE/dx means and sigmas from p to p_{T}, which is very difficult to obtain analytically.

4. A histogram is filled with $\Delta_\pi \equiv \mathrm{d}E/\mathrm{d}x - \langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ (for each track) vs p_{T}.

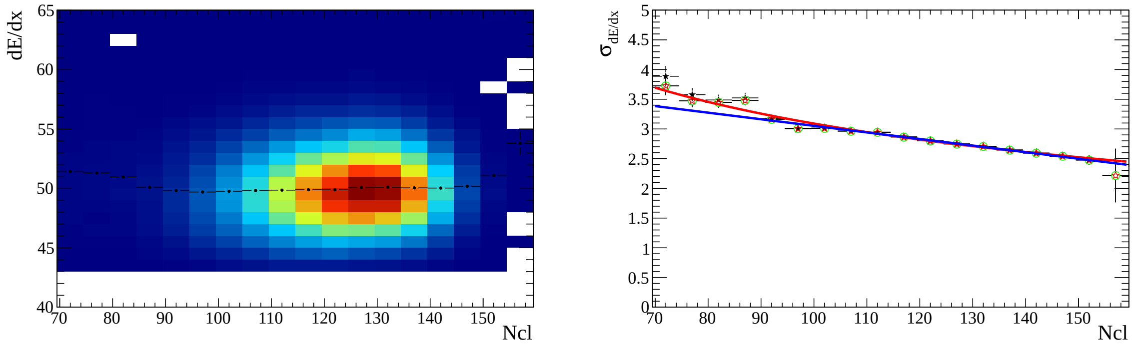
5. The dE/dx vs p_{T} histograms are fitted in bins of p_{T} with a sum of four Gaussians: one for each of p, K, π and e. Here the only free parameters are the yields; means and widths, meaning 8 out of 12 parameters, are constrained:
   - The Bethe-Bloch fit (scaled with particle masses) constrains the means of the Gaussian fits.
   - The relative widths of the Gaussians, $\sigma_{\mathrm{d}E/\mathrm{d}x}/\langle \mathrm{d}E/\mathrm{d}x \rangle$, are determined for π in the region of Minimum Ionisation (MIP), using Gaussian fits of slices in the dE/dx vs *ncl* distribution (Fig. 4.8(a)). These are then scaled with $\langle \mathrm{d}E/\mathrm{d}x \rangle$. There is also a small *ncl* dependence of $\sigma_{\mathrm{d}E/\mathrm{d}x}$, as seen in Fig. 4.8(b).
   - Electron parameters are extrapolated from the low-p_{T} region.

6. From the fits, particle yields as a function of transverse momentum can be extracted.

Positive and negative particles are treated both together, to get better statistics, and separately, to enable the study of antiparticle-to-particle ratios (expected to be close to 1 at mid-rapidity at LHC).

Fig. 4.7 illustrates the Bethe-Bloch fit curve from which the $\langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ is determined.

Fig. 4.8 illustrates how $\sigma_{\mathrm{d}E/\mathrm{d}x}$ is found from minimum ionising pions. These particles are defined as those fulfilling $0.4 < p < 0.6$, $40 < \mathrm{d}E/\mathrm{d}x < 60$ (before η calibration, cf Fig. 4.3), and should, apart from corresponding to a clean π sample (cf. Fig. 4.6), also have a very small



(a) dE/dx vs ncl

(b) σ vs ncl

Figure 4.8: (a): The MIP dE/dx vs ncl is fitted with Gaussians in bins of ncl. (b) shows the resulting fit width $\sigma_{\mathrm{d}E/\mathrm{d}x}$, along with two parametrisations. This illustrates point 3 of the Δ_π method.

momentum dependence of dE/dx. In Fig. 4.8(b), two fits to the ncl dependence of $\sigma_{\mathrm{d}E/\mathrm{d}x}$ are shown: one linear fit (blue) and one proportional to $\sqrt{\frac{159}{ncl}}$ (red). For higher ncl, the best choice of fit is ambiguous. In this analysis, the square-root fit was used, as one would expect this behaviour for counting statistics.

It might be a bit confusing that the method fits dE/dx distributions both in p and p_T. This is explained by the following: ionisation in the TPC gas is momentum dependent; this enters in the Bethe-Bloch fit. However, final physics spectra should be given in terms of $E\frac{\mathrm{d}^3N}{\mathrm{d}p^3}$. Integrating over one dimension gives $\frac{1}{2\pi p_\mathrm{T}}\frac{\mathrm{d}^2N}{\mathrm{d}y\mathrm{d}p_\mathrm{T}} \approx \frac{\mathrm{d}^2N}{\mathrm{d}\eta\mathrm{d}p_\mathrm{T}}$ for $p_\mathrm{T} >> m$ (see Appendix A). This is why the final yield is extracted as a function of p_T.

4.5 Method II: Particle identification using dE/dx and ncl

It was seen in both in simulations and in data from early 2010, that when drawing a two-dimensional distribution of dE/dx versus ncl, there was a correlation between these two observables. This is illustrated (in 2 and 3D) in Fig. 4.9. More significantly, the distribution has two clear bumps. The distribution was puzzling and the origins not known, but knowing that particles can be identified using their dE/dx, it was thought that 2D distributions like these might be used for distinguishing further between the particles. I will in the following call the use of such distributions for this purpose *the ncl method*.

4.5.1 Explanation of the ncl dependence

For a signal to be read out it must have two consecutive time samples above a certain threshold. In the TPC, the number of ionisation electrons is multiplied in avalanches at the anode wires, giving a signal which is large enough to be registered. The number of electrons in the avalanche is affected by many parameters, such as number of ionisation electrons and gas amplification. Ideally, the gas amplification gain is chosen such that small but inevitable fluctuations in these parameters do not lead to the signal falling below the threshold. In the case of the ALICE TPC, the gain is kept at moderate levels, from the experience during the detector commissioning, where a wire accidentally broke when higher gain was used. The underlying argument is that higher gain leads to more energy having to be dissipated in the system, a larger risk for discharges etc. The effect of having lower gain is however that clusters fall

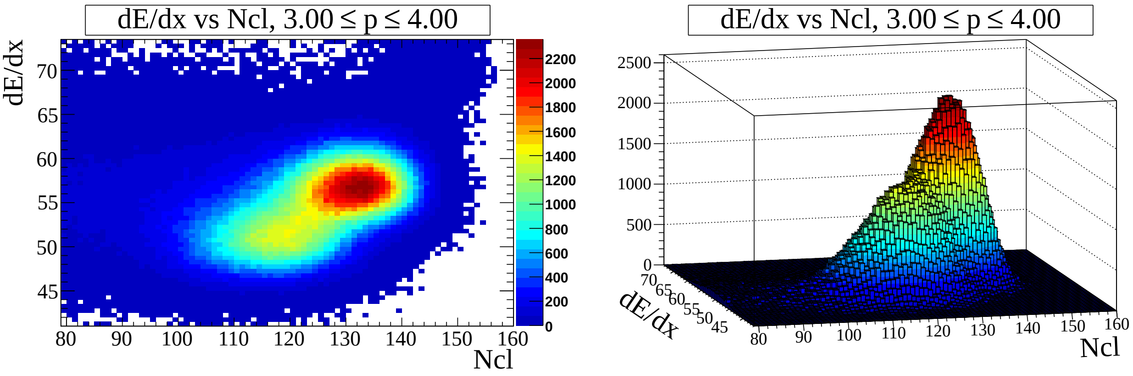


Figure 4.9: $\mathrm{d}E/\mathrm{d}x$ vs ncl, for $3 \leq p \leq 4$ GeV/c, 10b.

below threshold.

If clusters are lost, the truncated mean of the Landau distribution will be too high (since only the comparatively high charge clusters are registered), and the calculated $\mathrm{d}E/\mathrm{d}x$ will be pushed upward. For particles ionising less, more clusters will be below threshold and thus be lost. This means that the $\mathrm{d}E/\mathrm{d}x$ separation between particle species (at a given momentum) will become worse – the more ionising particles stay closer to the "true" $\mathrm{d}E/\mathrm{d}x$ value, while the less ionising particles get an increased $\mathrm{d}E/\mathrm{d}x$. It should however be noted that the effect on other observables, such as momentum resolution, is smaller (as long as the overall track length isn't notably shorter).

This dependence on lost clusters gives a complicated dependence of $\mathrm{d}E/\mathrm{d}x$ on track angle, as has been mentioned before, in Section 3.2.4. There it was explained that tracks perpendicular to the beam axis are more prone to lose clusters than those with smaller angles. This would, due to the threshold effect, mean a higher $\mathrm{d}E/\mathrm{d}x$ for these tracks. On the other hand, tracks at smaller angles which ionise more per pad (due to longer path lengths being projected onto each pad), are not as affected. The $\mathrm{d}E/\mathrm{d}x$ of these is in turn corrected (downward) for the more trivial angle dependence, by a simple scaling with the angle, which may not be sufficient. The result is an overall angle dependence of $\mathrm{d}E/\mathrm{d}x$ which is not very easy to correct for, especially since there is an "intrinsic" difference in $\mathrm{d}E/\mathrm{d}x$ which affects these corrections.

The loss of charge clusters is the result of an experimental problem (i.e. the gain is too low). Primarily, this should be solved by increased gain, but it will not be possible to entirely avoid clusters losses. As it appears today, the TPC will rather have to be operated at even lower gain. Thus the method developed here has long term general interest.

In an attempt to remedy the losses and improve the $\mathrm{d}E/\mathrm{d}x$ resolution, from the run period 10d and on, and in the iteration 10c Pass3, $\mathrm{d}E/\mathrm{d}x$ is calculated by including "one-pad-clusters". These are not included in tracking, since here a centre-of-gravity method is used for determining track cluster coordinates with better resolution than the pad size (and this requires at least two adjacent pads with signal in the cluster). For the $\mathrm{d}E/\mathrm{d}x$ this is not an issue as such, but one should be aware that a single pad firing like this might not really belong to a track, but could be just noise.

By keeping track of the number of missing clusters, the $\mathrm{d}E/\mathrm{d}x$ separation could in principle be recovered[2]. This separation is what we see in the 2D histograms as the separation between the peaks; the "true" separation S would then be recovered from that in each dimension by $S_{tot} = \sqrt{S^2_{ncl} + S^2_{\mathrm{d}E/\mathrm{d}x}}$. Given this separation, one should be aware that the $\mathrm{d}E/\mathrm{d}x$ resolution

---

[2] This has later been attempted for $\mathrm{d}E/\mathrm{d}x$ at the TPC track reconstruction level.

$\sigma_{\mathrm{d}E/\mathrm{d}x}$, given by the Gaussian fit width, isn't an unambiguous measure of the level of detail that can be resolved in the $\mathrm{d}E/\mathrm{d}x$.

Before all this was understood, it was decided to try and explore the possibilities of using the resulting distributions from the cluster loss effect as an independent method for particle identification, complementary to the one already described (in Section 4.4).

4.5.2 A closer look at the *ncl* distributions

For the two-dimensional fitting described later (see Section 4.5.3), the simultaneous fitting of two functions (one in each dimension) is used. For fitting the $\mathrm{d}E/\mathrm{d}x$ distribution, a Gaussian is used, motivated by earlier studies.

The *ncl* distribution, on the other hand, is not a priori well understood, since the number of missing clusters depends on the particle $\mathrm{d}E/\mathrm{d}x$, which depends on both charge and velocity. In a given momentum range, the number of missing clusters will thus mainly be species dependent. The underlying motivation for doing a *ncl* based analysis is that if the separation here is large enough, fitting a proper function in the *ncl* dimension would improve fits in the $\mathrm{d}E/\mathrm{d}x$ dimension.

To decide which distribution to use for the *ncl* dimension, a study on PYTHIA Monte Carlo (MC) generated tracks was done to pinpoint the shape of the *ncl* distributions. In this study, making use of the known PID, a number of functions were fitted to the *ncl* distributions of different particle species. As the number of clusters is in general much larger than the number of missing clusters, for some fit functions it made much more sense to in effect fit to the distributions of the number of missing clusters than to the actual *ncl* distributions, by shifting the mean by the maximum number of clusters achievable ($\langle ncl\rangle' = 159 - \langle ncl\rangle$).

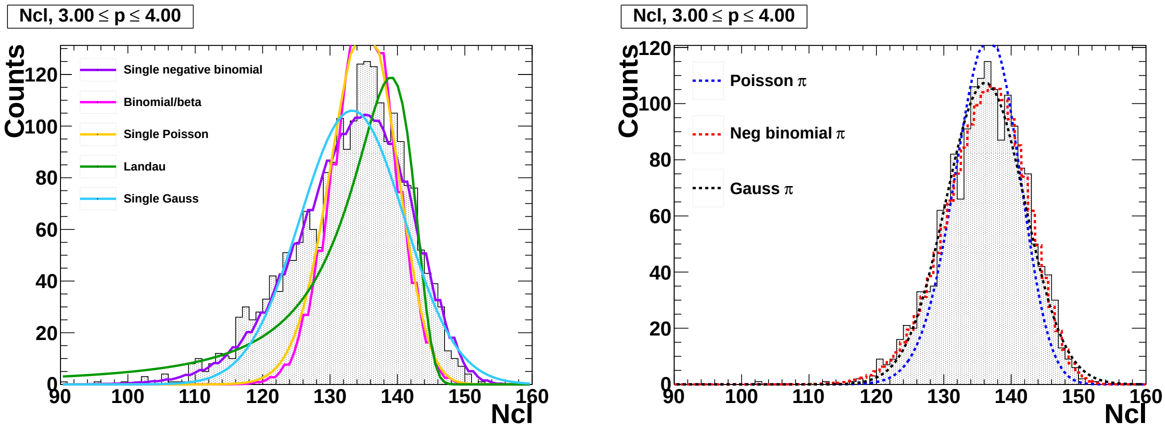
In Fig. 4.10 the fits in the MC study are seen. The *ncl* distributions are shown in grey. Input parameters to all the fits are the mean and width of the distribution at hand. In Fig. 4.10(a), a wide range of fits to the whole distribution are shown. Here it seems both a Poisson and binomial fit are too narrow, while a Landau falls off too sharply on the right edge. The Gaussian and negative binomial[3] fits are quite similar; both are too low, but the negative binomial represents both the mean and the overall (asymmetric) shape of the distribution better than the Gaussian. The study of Poisson, Gaussian and negative binomial fits will be shown in more detail. In Figs. 4.10(b) and 4.10(c), separate fits to the π and p+K peaks respectively are shown (note the very different scales on the y axis), where the PID knowledge is used to get two separate distributions.[4] Here the Poisson is again too narrow, and the Gaussian seems to be the best fit overall, especially considering the p+K distribution. This might be because the latter is actually the sum of two distributions. Finally in 4.10(d) the sum of the two former (the PID fits) are shown, along with the separate fits. Unfortunately here the Poisson fit sum isn't drawn (the "Sum Poisson" is only the π fit). Again the Gaussian and negative binomial fits seem to work quite well (and similarly).

The outcome of this study was to try both negative binomial and Gaussian fits in the *ncl* dimension of the two-dimensional fits. However, when applied to data, it turned out that the negative binomial fit wasn't very stable (for one thing, finding suitable input parameters was quite difficult). It was thus decided that the best choice was to do a two-dimensional Gaussian fit. However, some examples of negative binomial fits are shown in Section 5.2.1 to give the reader a glimpse of how well they worked.

---

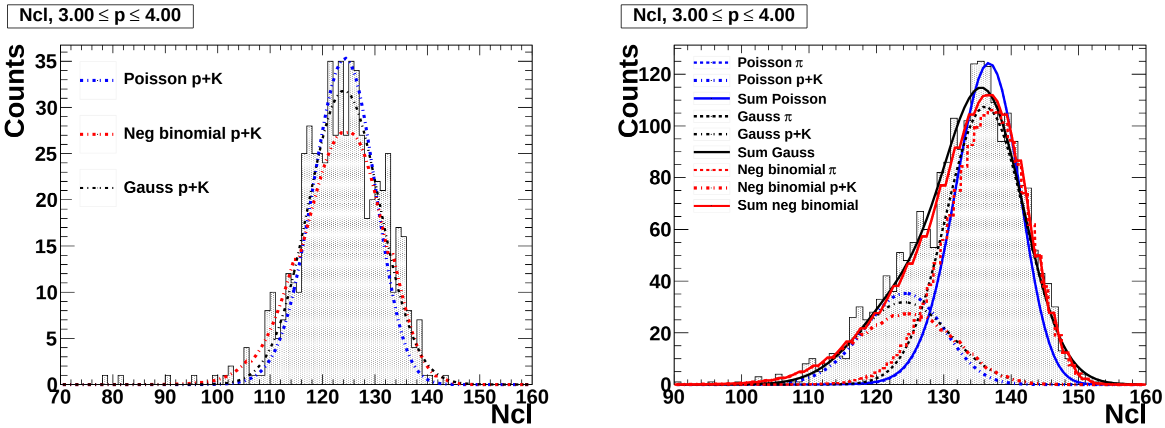
[3] This is a discrete distribution, giving the fit a ragged look

[4] The choice to study p and K together was motivated by the procedure used in the *ncl* method as a whole, see Section 4.5.3.



(a) All fits, total *ncl* distribution

(b) Gaussian, Poisson and negative binomial fits to the π peak



(c) Gaussian, Poisson and negative binomial fits to the p+K peak

(d) Gaussian, Poisson and negative binomial fits to the π and the p+K peak respectively, and the sums of the fits

Figure 4.10: Results from 900 GeV MC study of fits to the *ncl* distribution

4.5.3 The *ncl* method

This is the second of the two methods used and compared in this analysis. The basic assumptions are the same as in the Δ_π method, but the fitting is different as it uses 2D *ncl* vs $\mathrm{d}E/\mathrm{d}x$ histograms as a starting point, instead of the Bethe-Bloch fits. Steps 1 through 4 are illustrated in Figs. 4.11-4.12.

1. The 2D distributions are fitted, in momentum bins, by a two-dimensional, possibly correlated, simultaneous Gaussian fit. This fit is a log-likelihood fit, which normalises the fit to the number of entries in the histogram. First guess input parameters for this (means, widths and yields) are obtained from fitting a simpler 2D Gaussian to two different ranges, limited by an initial peak finding procedure, covering each of the two discernable peaks: one for p and K, one for π.
2. In these fits, the $\langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ is determined. A fit to $\langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ vs p gives the $\langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ curve used in the next step. Here the MIP point (see Fig. 4.12) is put in manually (since this is given by calibrations).
3. 2D histograms with *ncl* vs $\Delta_\pi \equiv \mathrm{d}E/\mathrm{d}x - \langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ (for each track) are created in transverse momentum bins.
4. These histograms are fitted like before, but now in p_T bins.
5. From the fits, particle yields as a function of transverse momentum can be extracted.

The positive particle *ncl* vs $\mathrm{d}E/\mathrm{d}x$ distributions for 10b are shown in momentum bins in Fig. 4.11(a). Here one sees that there are two "peaks"; judging from $\mathrm{d}E/\mathrm{d}x$, one corresponds to π and one to K and p, which are too close to be distinguished (in both *ncl* and $\mathrm{d}E/\mathrm{d}x$). For low momenta one also sees a patch in the top right corner corresponding to electrons.

An example of the 2D fit result in p_T is seen in Fig. 4.11(b). A major advantage of the log-likelihood fit is that it, by counting also the bins with zero content, does a reasonable job also for very low statistics (which would not be the case for a fit optimising χ^2, for example). The resulting $\langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ for each momentum bin is shown in Fig. 4.12.

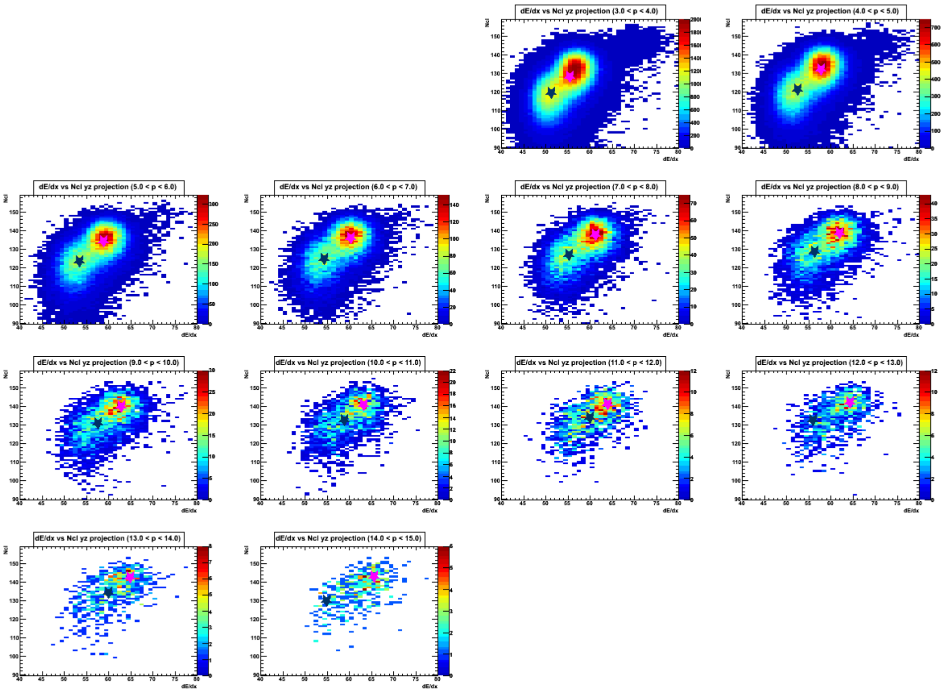
In this method, in contrast to the Δ_π method, in principle all fit parameters can remain free. This is possible due to the extra ingredient: the information in the *ncl* dimension. However, it turns out the fit stability improves when the correlation is fixed to 0.

4.6 Corrections to obtain physics spectra

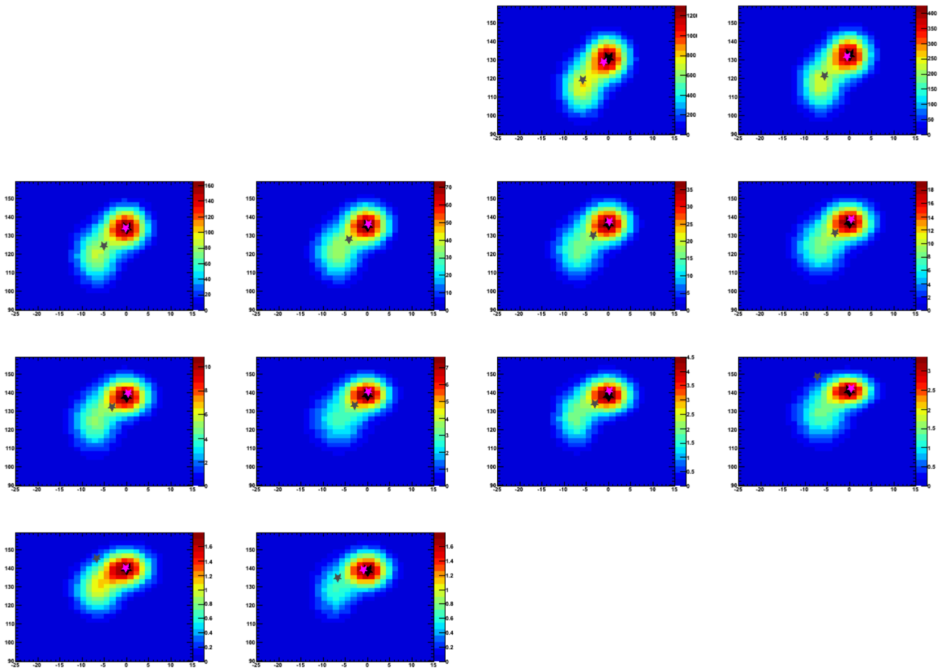
The raw yields obtained from the fits in this analysis have to be corrected. First of all, the detector acceptance has to be taken into account. The TPC, being a cylinder, has full azimuthal acceptance, but limited in the beam axis direction. The result is that tracks below a certain angle to the beam ($|\eta| > 0.8$) are not included in the analysis.

Furthermore, the track selection and reconstruction efficiency is not 100%, and neither is the trigger efficiency (for details on the trigger see Appendix B). The former, ϵ_{rec}, is estimated from a fit to the ratio of the number of generated to reconstructed tracks in MC vs p_T, as illustrated in Fig. 4.13. The latter is known from the charged particle analysis in ALICE. The trigger efficiency used in this analysis is $\epsilon_{trig} = 91.6\%$.

The fit function in Fig. 4.13 is of the form $\epsilon_{rec}(p_\mathrm{T}) = a - \frac{b}{p_\mathrm{T}}$, motivated by the assumption that the efficiency tends to some constant value a at high p_T. The efficiency shown here is after the ϕ cut has been applied. This cut reduces the efficiency especially for lower p_T.



(a) Input distributions



(b) Fit to distributions

Figure 4.11: ncl vs $\mathrm{d}E/\mathrm{d}x$ distributions for 10b, positive particles. This illustrates step 1 and 4 of the ncl method. (a): Two peaks are visible in the distribution. Stars indicate peak positions found by an initial 2D Gaussian fit. (b): ncl vs Δ_π . The additional black star indicates the final fitted pion peak from the simultaneous Gaussian fit to the two dimensions.

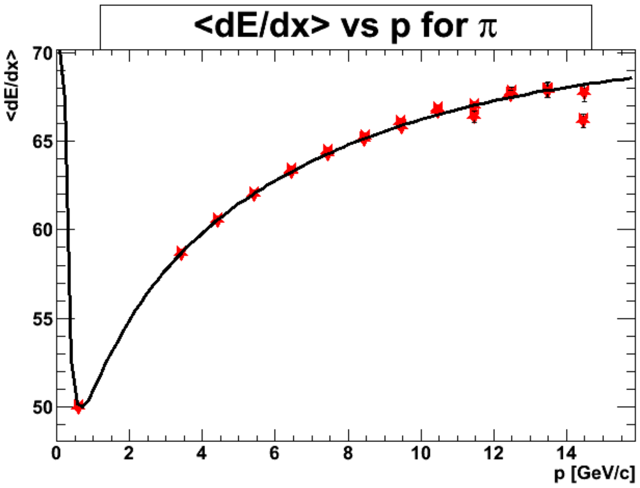


Figure 4.12: Example $\langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ curve extracted from fitting a Bethe-Bloch parametrisation to the $\langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ of each momentum bin (10d), for positive and negative particles separately, giving two points for each p. The $\langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ value subtracted in "step 3" is obtained from such a curve.

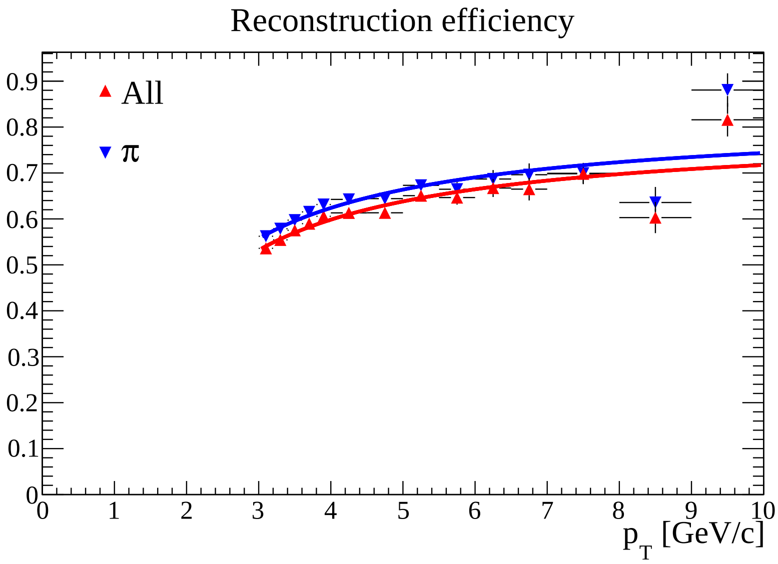


Figure 4.13: Fit to reconstruction efficiency ϵ_{rec} estimated for 10c, as the ratio of reconstructed over generated tracks vs p_T, for all charged particles (ϵ_{ch}) and for π (ϵ_π).

Finally, in all analyses, one has to take the background into account. In this analysis, the "contamination" in the pion peak consists of μ, and is known to be very small ($< 1\%$). This number would enter in a systematic error.

The formula to go from the raw pion yields of the analysis to invariant pion yields is then

$$\frac{\mathrm{d}^2 N_\pi}{\mathrm{d}y \mathrm{d}p_\mathrm{T}} = \frac{1}{N_{evt}} \times \frac{1}{\Delta y} \frac{1}{\Delta p_\mathrm{T}} \times \epsilon_{trig} \times \epsilon_{rec} \times N_\pi, \tag{4.1}$$

where N_π is the number of π and N_{evt} is the number of events. An alternative method is to use the total invariant charged particle yields as measured by ALICE, and multiply by the pion fractions found in this analysis:

$$\frac{\mathrm{d}^2 N_\pi}{\mathrm{d}y\mathrm{d}p_\mathrm{T}} = \frac{E}{p}\frac{\mathrm{d}^2 N_{ch}}{\mathrm{d}\eta\mathrm{d}p_\mathrm{T}} \times \frac{\epsilon_\pi}{\epsilon_{ch}} \times \frac{N_\pi}{N_{ch}}, \tag{4.2}$$

where the subscript *ch* is for charged particles.

Chapter 5

Analysis: results and discussion

In this chapter, the results are presented and discussed. Both methods are evaluated. First a study of the Δ_π method using Monte Carlo simulations (MC) is presented, then the performance of the *ncl* method is shown, followed by a discussion comparing the two methods. Finally the Δ_π method is used on real data, and invariant pion yields are extracted.

For all the analyses, the performance of the method is evaluated in a highly visual manner and the fit qualities are often not very easy to quantify such that they can be presented in a few result figures. For this reason, there are many figures generated for each method and data set; in the following chapter only some representative figures can be shown.

The results given here for the Δ_π method on the 10c run will be limited to $3 \leq p_\mathrm{T} \leq 10$ GeV/c. The lower bound is motivated by that π, K and p are all on the relativistic rise from p_T = 3 GeV/c on, and the upper by the little statistics for higher p_T bins, as can be seen in the fits in Appendix C.

Unfortunately, from a lack of time, the systematical errors have not been estimated. Consequently, results are shown with statistical errors only throughout the chapter.

5.1 Particle identification using the Δ_π method: MC

The real advantage of doing a Monte Carlo study is of course the benefit of knowing "the truth", i.e. knowing which particles were generated and reconstructed, and then comparing the outcome of the analysis to this. For the 10c data, two sets of MC simulations were generated, with a comparable number of events: one using PYTHIA and one using PHOJET. There are some minor differences, but I will concentrate on the PYTHIA MC production.

The disadvantage, on the other hand, is that one cannot be sure that the MC simulation really mirrors the data. A simulation always depends on a theoretical model, and an accuracy down to the level of detail seen in data can be difficult to achieve. One example of this, seen in this study, is that the dE/dx separation generally is larger in MC than in real data. One has to keep this in mind when judging on the Δ_π fit quality, before making any statements about the performance of the method for real data based on MC results.

Apart from being generated by a simulation, the MC tracks are treated exactly the same way as data (calibrations, fits etc.). The only difference is that the particle species information is available to cross check the fit results afterwards – this information is not used as input for any analysis steps.

5.1.1 Fits

Fig. 5.1 shows some fit examples for the PYTHIA MC sample, for both charge signs combined. The shaded histograms correspond to the generated input (generated particles propagated as tracks "through the detector") while the lines are the fit results using the Δ_π method. The fits

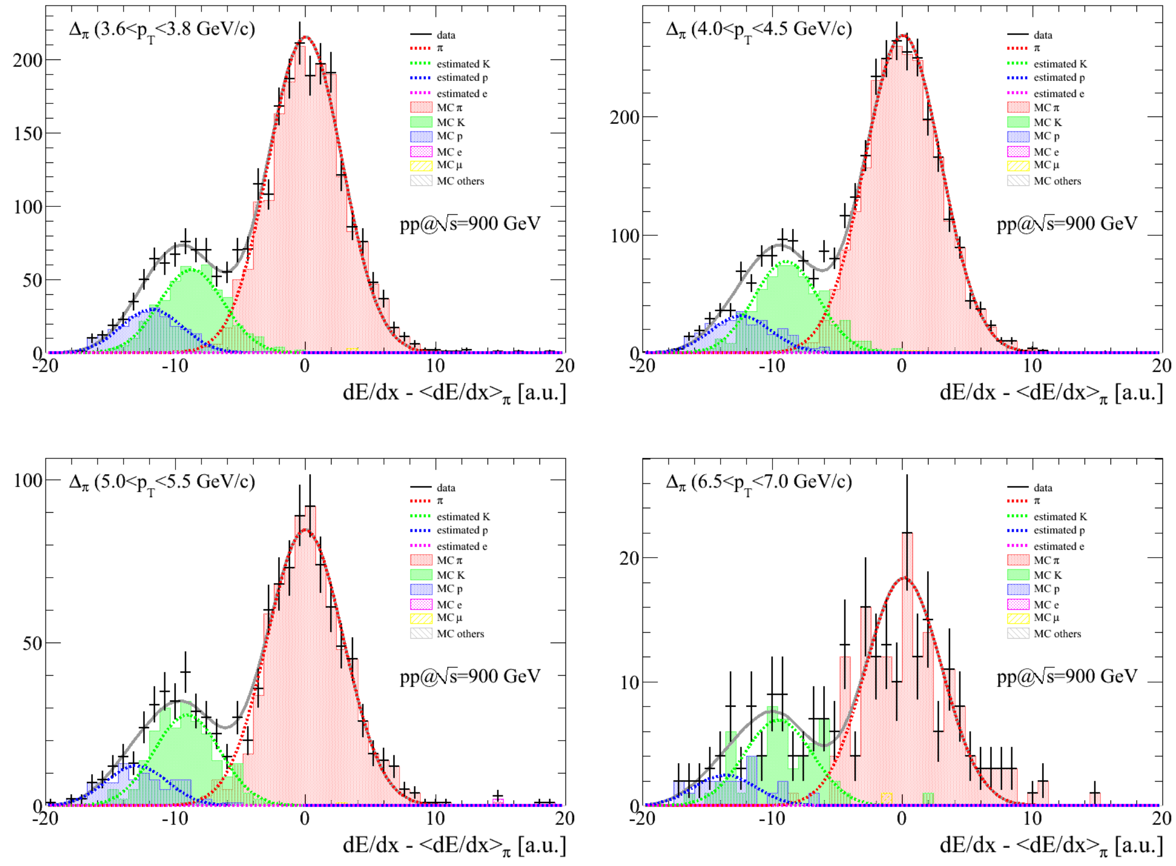


Figure 5.1: Examples of Δ_π fits for different p_T bins. The smooth lines are fits, the shaded histograms are MC truth.

show a striking agreement with the MC distributions.

5.1.2 Yield ratios

Fig. 5.2 shows the ratios from the fits and MC reconstructed truth of antiparticles to particles, for π, K and p. It is seen that the fluctuations in the ratio follows the MC truth. Remember that the error bars are statistical errors only, meaning that they do not take into account the stability of the fits. The agreement of "Fits" and "MC truth" mirrors the already known fact that the fits to the π distributions are the most stable; for π, the fits and truth agree within (correlated) statistical errors. It is known that PYTHIA generates more positive than negative particles at high p_T, due to a small contribution of valence quark[1] jets. This is also seen in the ratios, especially for protons. The disagreement between fits and truth for π, K and p could be taken as a hint on the magnitude of the systematical errors for these fits.

5.1.3 Particle fractions and p_T spectrum

Figs. 5.3-5.5 show the particle fit yield ratios of π, K and p for PYTHIA MC, compared to MC reconstructed truth. The results are given for both separate and combined charges. Note the very different scales on the ratio graph axes (bottom panels). Since the log-likelihood fits are always normalised to the number of entries in the histogram, the fractions add up to 1 for all p_T bins. The clear undershoot of the p fits in the lowest p_T bins corresponds to $\sim$ 2% too

---

[1] In pp collisions, there are 4 u ($+\frac{2}{3}e$) and 2 d ($-\frac{1}{3}e$) valence quarks)

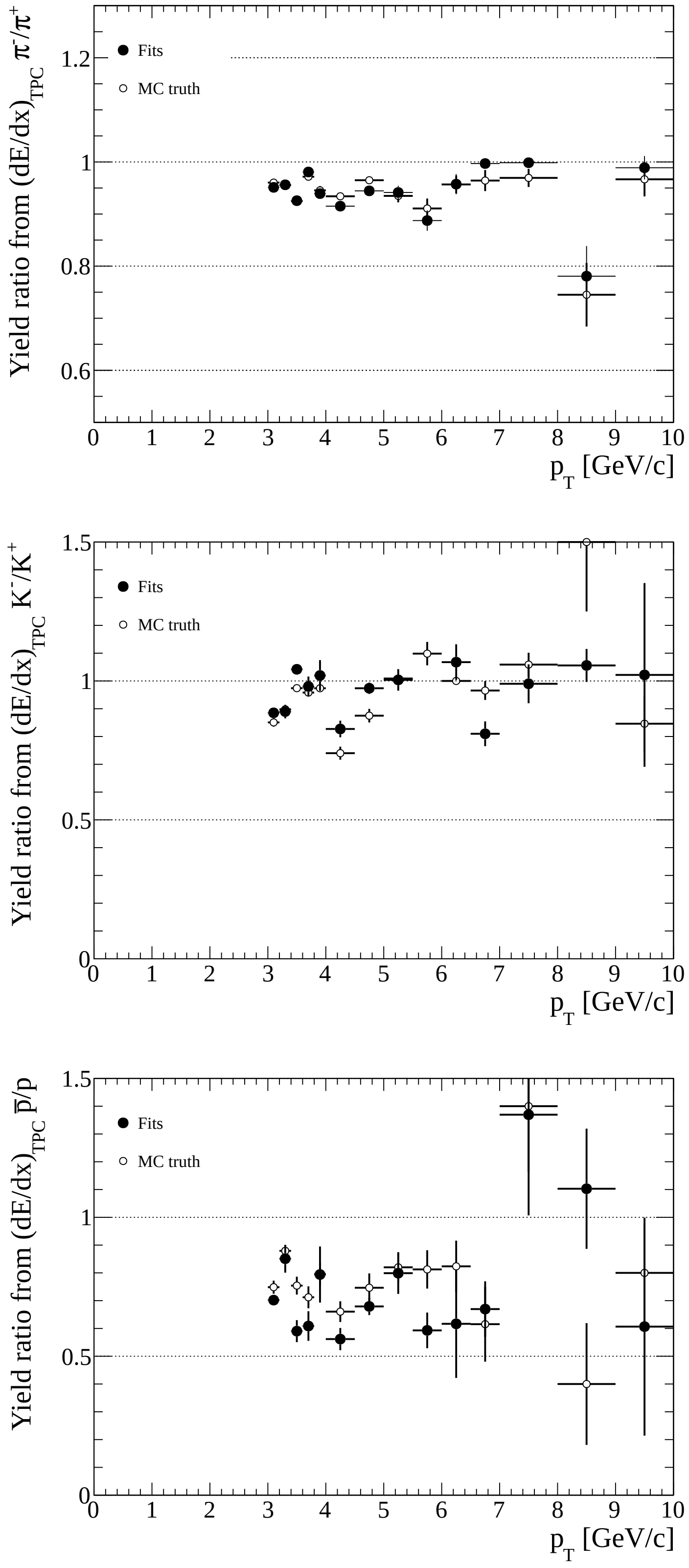


Figure 5.2: The ratio of antiparticles to particles, for π (top), K (middle) and p (bottom): open circles are the MC truth while closed circles are the fit result ratio.

many π, the rest is in the K fits. The π fits (for "All"; negative and positive particles together) are overall overestimating the yields by 2–5% over the p_T range shown. This gives another indication of systematical uncertainties to take into account.

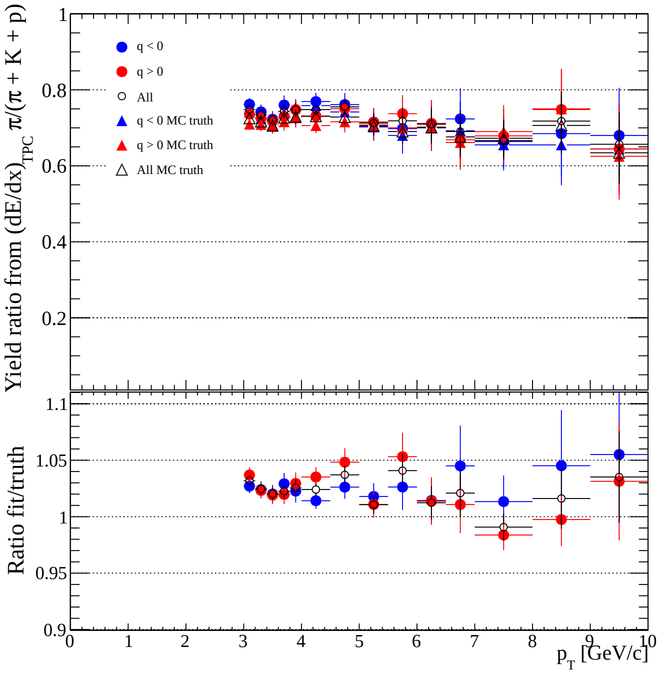


Figure 5.3: Pion fractions: comparison between the fit result (circles) and MC truth (triangles) for 900 GeV MC data.

Fig. 5.6 shows the efficiency and acceptance corrected charged particle p_T spectrum for PYTHIA MC, compared to the MC truth. Corrections are later applied to data in the same way. The ratio seen in the Fig. is thus a very good consistency check of these corrections, and as is seen, it is close to 1 all over the p_T range.

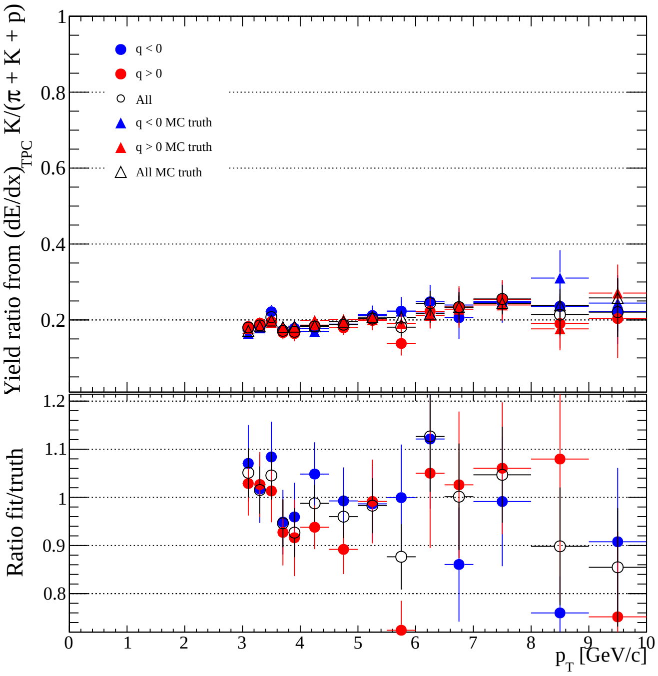


Figure 5.4: Kaon fractions: comparison between the fit result (circles) and MC truth (triangles) for 900 GeV MC data.

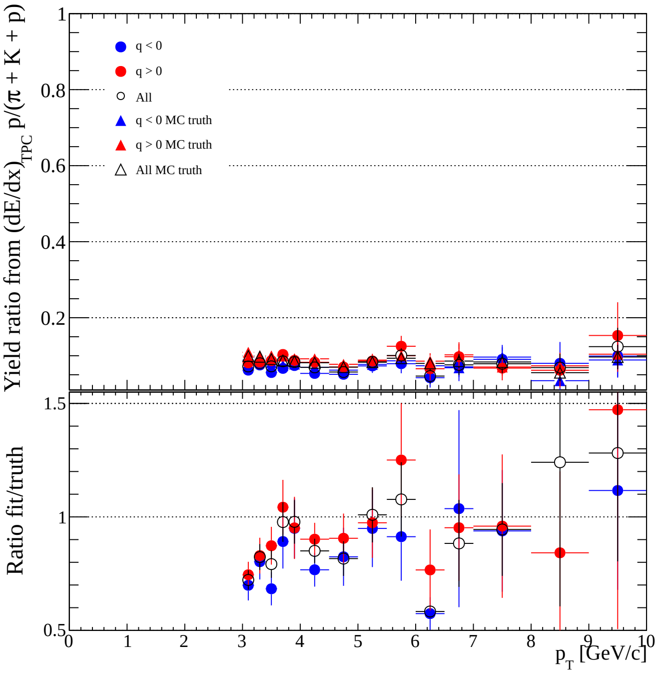


Figure 5.5: Proton fractions: comparison between the fit result (circles) and MC truth (triangles) for 900 GeV MC data.

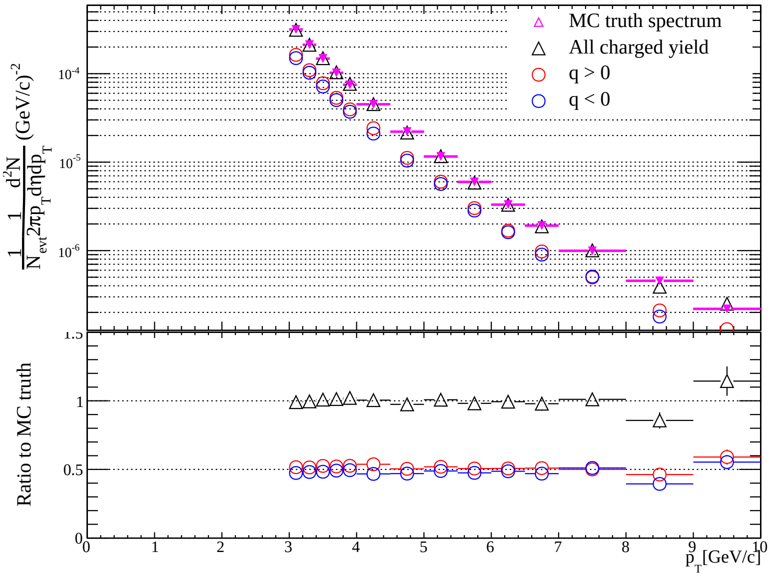


Figure 5.6: Charged particle p_{T} spectrum: raw yield as a function of p_{T}, with comparison to the MC truth.

5.2 Particle identification using the *ncl* method

With this method, it is at this stage not plausible to be able to discern protons and kaons. Here, no assumptions of Bethe-Bloch parametrisations, widths etc. are used – all this information is instead found from fits to the *ncl* vs dE/dx histograms. Separating protons and kaons would thus require that they had each their own bump in these histograms, which they do not. For this reason, they are not fitted separately, but with one 2D fit. The projections of histograms and fits on both dimensions are shown in Section 5.2.1.

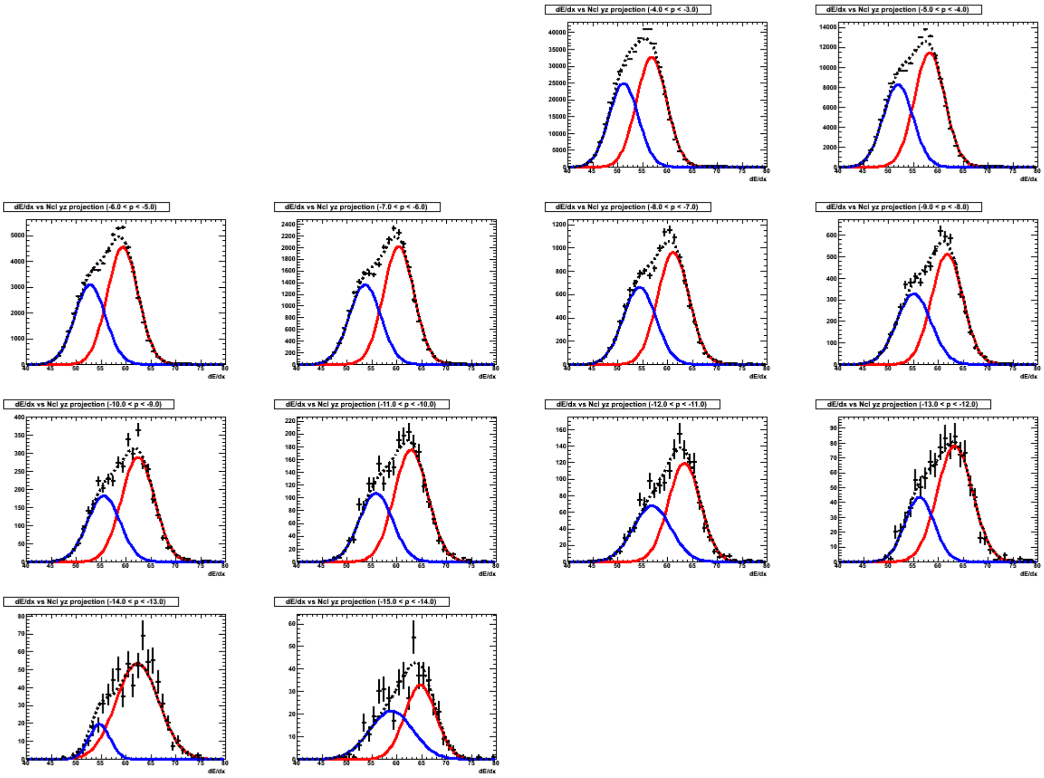
A crucial point for this analysis method is that the peak separation in *ncl* is sufficiently large, i.e. that many clusters are lost for the (less ionising from $p_{\mathrm{T}}\sim 3$ and up) protons and kaons. This is the case for the 10b run period, and for early Passes of 10c data. For 10c Pass3 and the 10d run, one-pad-clusters were included for the dE/dx calculations, increasing $\langle ncl \rangle$ and decreasing the separation in *ncl*, as explained in Section 4.5.1. For later run periods, other approaches such as increasing the gas gain (to increase the charge deposited at read-out) or recalculating dE/dx (taking missing clusters into account) were used. The outcome is that it is not feasible to be able to use the *ncl* method for 10c Pass3 and from 10d data and on, while the Δ_π method, on the other hand, performs better here.

5.2.1 Fits

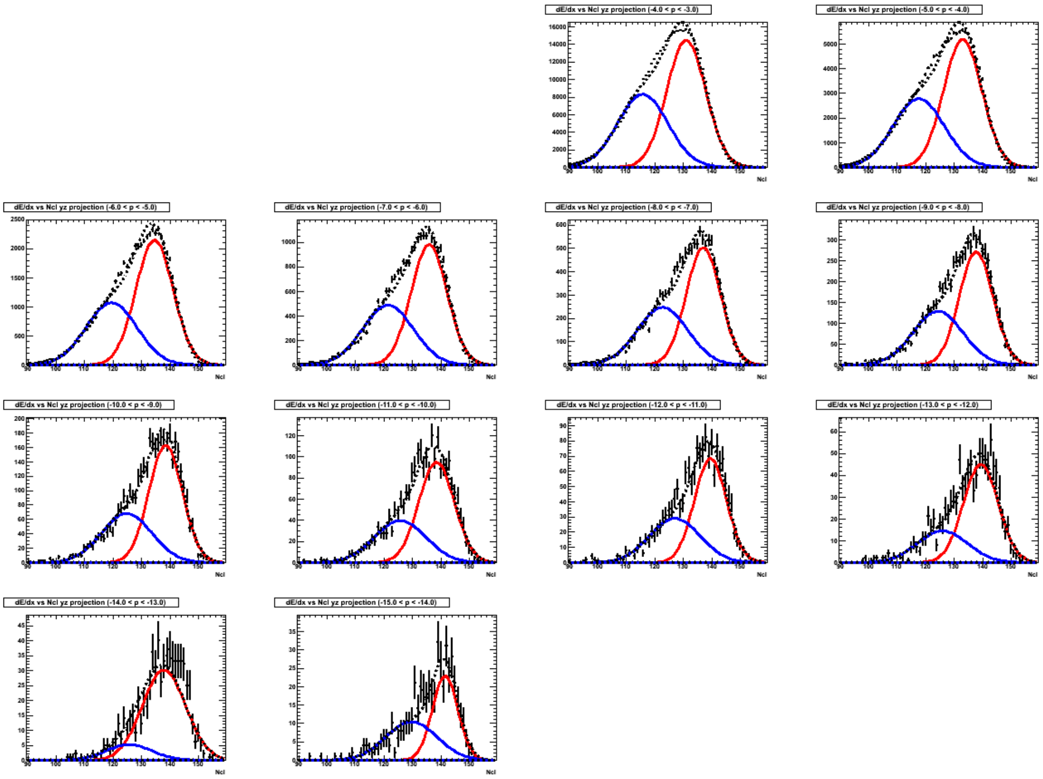
Figs. 5.7-5.8 show examples of projections of distributions and fits (red for π, blue for p+K) in momentum bins, for both dE/dx and *ncl*. They start at $p = 3$ GeV/c, which is motivated by the fact that this is were the pions have become clearly more ionising than kaons and protons.

In Fig. 5.7(a), one clearly sees the worse dE/dx separation in the 10b data (cf. the 10c fits seen in Appendix C). Overall, the outer edges of the distributions seem to be well described by Gaussians, but the sum of the two Gaussians doesn't follow the data when it comes to separating the two peaks.

The Gaussian fits to the *ncl* projections, seen in Fig. 5.7(b), do similarly well at the right edge of the π peak, but this is no longer true on the left edge of the p+K peak. This could be



(a) dE/dx distributions, Gaussian fit



(b) ncl distributions, Gaussian fit

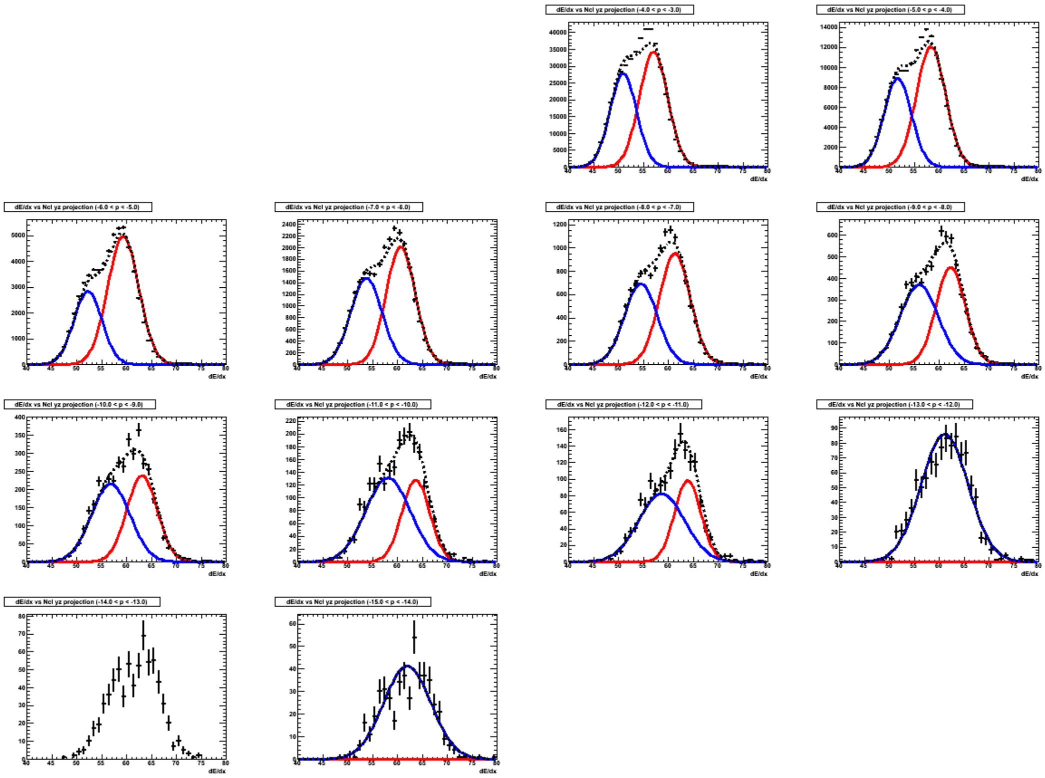
Figure 5.7: Projections of distributions in dE/dx and ncl dimension with double Gaussian fits, in momentum bins (1 GeV/c wide, 3-15 GeV/c), 10b, negative particles.

due to the asymmetry in cluster losses – for the more ionising π, not being very affected by cluster losses, *ncl* would be more truly random and a Gaussian distribution wouldn't be too surprising (just from considering the Law of Great Numbers).

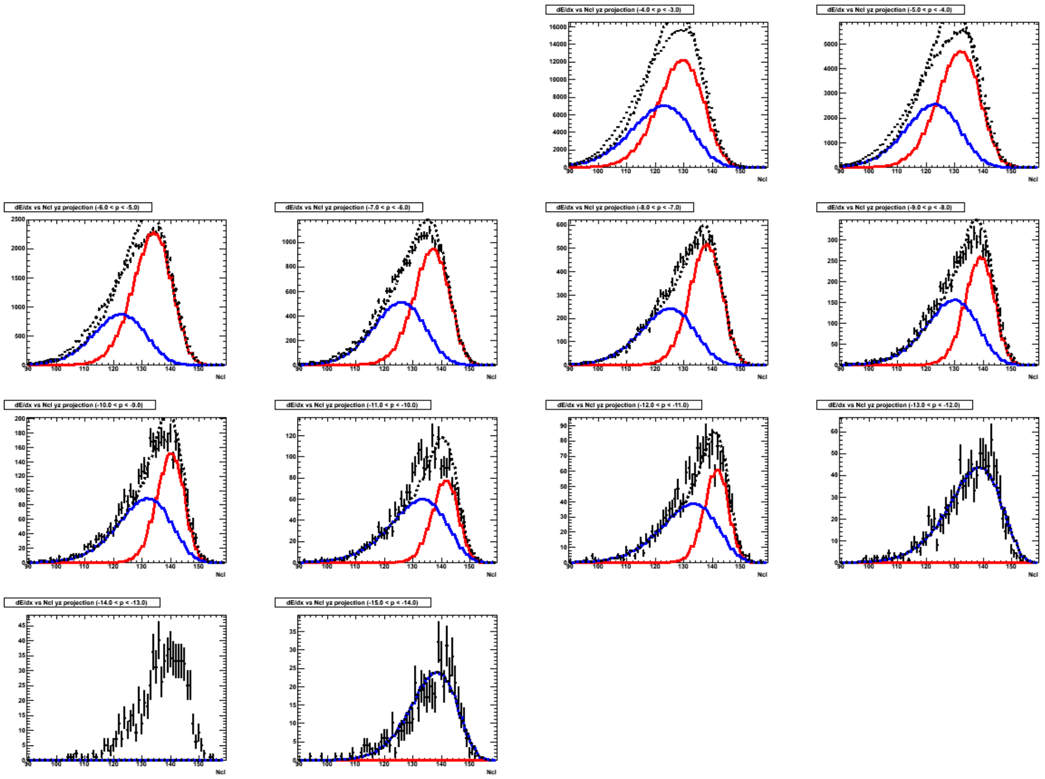
It is also seen in this figure that there are no clear peaks in the ncl distribution projections. This illustrates one of the difficulties in doing fits using this method: it is not as clear as one might have hoped what input values should be given to the fits in the *ncl* dimension. With a larger separation in this dimension too, it might have been easier to find and apply the proper function to fit to these distributions.

Figs. 5.8(a) and 5.8(b) show the same data, fitted with a Gaussian in the $\mathrm{d}E/\mathrm{d}x$ dimension and a negative binomial fit in the *ncl* dimension. One sees that using a negative binomial fit in the *ncl* dimension makes the Gaussian fits in $\mathrm{d}E/\mathrm{d}x$ less stable – in the last three momentum bins, the π peak fit doesn't converge at all. Comparing the *ncl* fits using the two approaches, one sees that the negative binomial fit to the p+K peak systematically stretches farther in under the π peak, compared to the Gaussian fit.

Overall, it seems problematic to add the p and K peaks and fit them with only one Gaussian – not very surprising after all, as the sum of their distributions is most likely not Gaussian.



(a) dE/dx distributions, Gaussian fit



(b) ncl distributions, negative binomial fit

Figure 5.8: Projections of distributions in the dE/dx and ncl dimensions, with Gaussian and negative binomial fits, respectively. Same distributions as before.

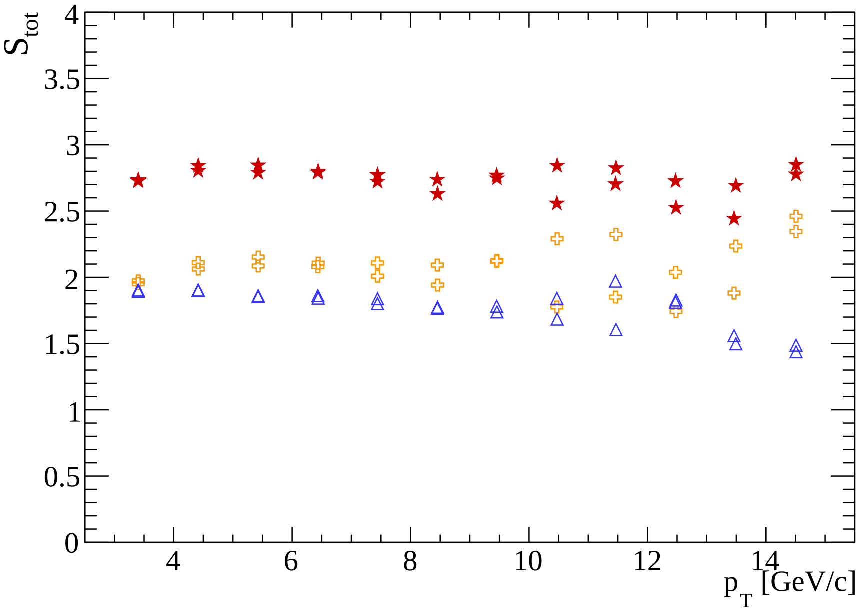
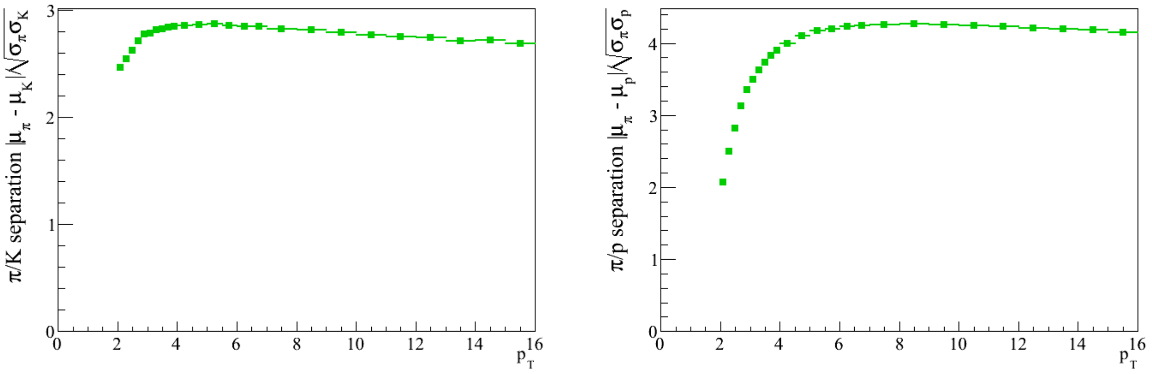


Figure 5.9: Red stars show the total separation $S_{tot} = \sqrt{S^2_{dE/dx} + S^2_{ncl}}$ for 10b, for positive and negative charge separately. Orange crosses and blue triangles show $S_{dE/dx}$ and S_{ncl}, respectively.

In Fig. 5.9, the separation $S_{tot} = \sqrt{S^2_{dE/dx} + S^2_{ncl}}$ vs p_T is shown (red stars) for the 10b data. The separation in dE/dx and ncl is also drawn, in orange crosses and blue triangles, respectively. As before, separation is defined as $S_{12} = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1 \sigma_2}}$. There is a hint of a p_T dependence (in opposite directions) for dE/dx and ncl separately, but the total separation lies stably just below 3 all the way.



(a) π to K separation vs p_T

(b) π to p separation vs p_T

Figure 5.10: Separation between particle species vs p_T for 10d ($\sqrt{s}$ = 7TeV).

Fig. 5.10 shows the separation in 10d, where $\sqrt{s}$ = 7 TeV, as for 10b, but where the dE/dx algorithm includes one-pad-clusters, reducing the separation in ncl but increasing it in dE/dx. Note that the two peaks in question in Fig. 5.9 are the π peak and the p+K peak, so the separation is not immediately comparable to any of the separations shown in Fig. 5.10, which are between π and K, p separately. A comparison of the 10b π to p+K separation to the 10d π to K separation still indicates that the increased $\mathrm{d}E/\mathrm{d}x$ separation in principle can be recovered already in 10b, from the information on the number of lost clusters. This is a strong argument that the lost clusters should be taken properly into account in the $\mathrm{d}E/\mathrm{d}x$ reconstruction algorithm.

5.2.2 Pion fractions

Figs. 5.11-5.13 show pion fractions out of the total charged hadron yield, π/h, for different data sets: 10b, 10c and 10d. The pion fraction varies a bit between the data sets and for different p_{T}. In particular, for 10c, comparatively little statistics is a problem (as both fewer events[2] and lower $\sqrt{s}$ leads to fewer high-p_{T} particles). Overall, the π fraction lies between 0.5 and 0.7.

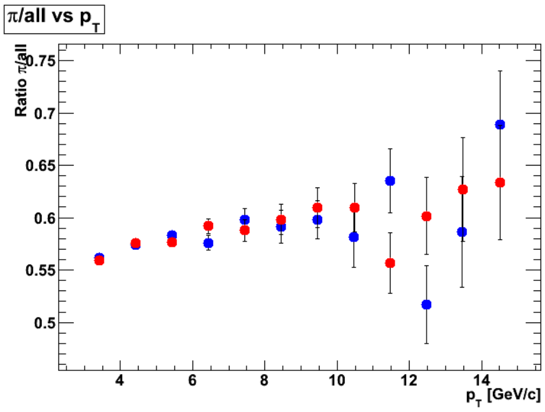


Figure 5.11: π fraction vs p_{T} for 10b. Red markers for π^+ and blue for π^-.

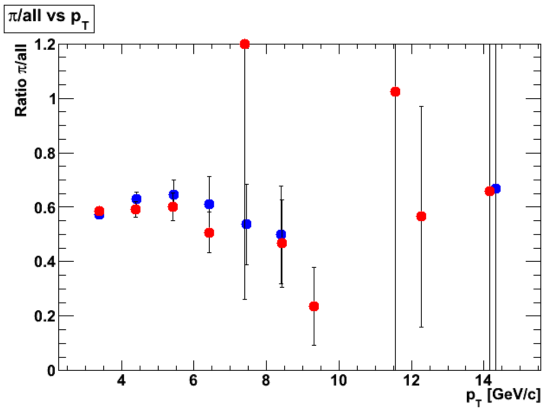


Figure 5.12: π fraction vs p_{T} for 10c. Red markers for π^+ and blue for π^-.

---

[2] For the *ncl* method analysis, the number of events available were about half of that for 10c when using the Δ_π method.

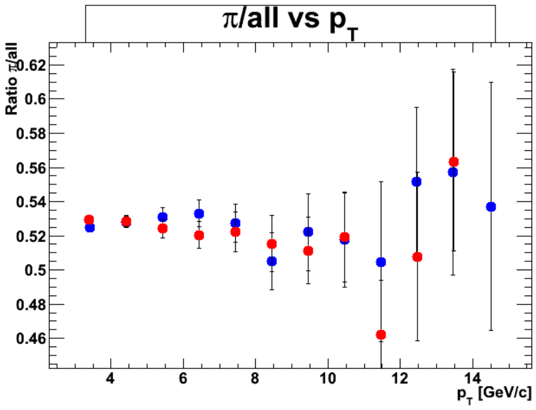


Figure 5.13: π fraction vs p_{T} for 10d. Red markers for π^+ and blue for π^-.

5.3 Comparison of the Δ_π and *ncl* method results

One of the key motivations for developing the *ncl* method was its usefulness as a complementary method to the Δ_π method already used; it hardly uses any fixed parameters, while the Δ_π method fixes or constrains as many as possible. The interesting question is then: do they give the same results?

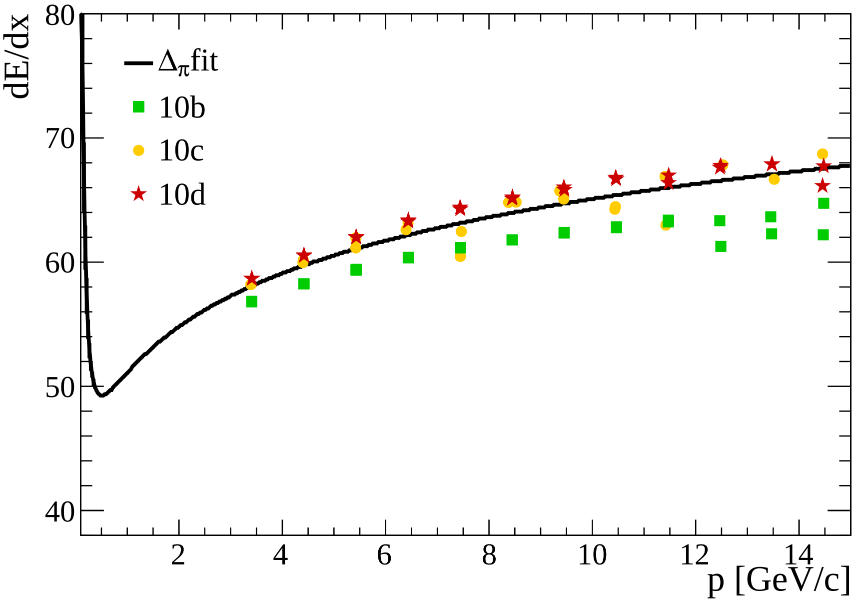


Figure 5.14: Comparison of $\langle \mathrm{d}E/\mathrm{d}x \rangle$ fit from Δ_π method for 10c and 10d to those obtained in the *ncl* method for all the different data sets.

Fig. 5.14 shows a comparison between the Bethe-Bloch parametrisation for $\langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ in the Δ_π method for 10c and 10d (the curves coincide) on one hand, and the $\langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ points (for positive and negative particles separately) found from the 2D fits in the *ncl* method on the other. It is seen that the general trend is the same for the two methods – the agreement between the 10c data is particularly striking (in these plots, the error bars have not been reproducible, but they are quite small (cf. Fig. 4.12)). The interpretation must be that the *ncl* method manages surprisingly well in nailing down $\langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ as it is given by the Bethe-Bloch parametrisation, i.e., find the pions – *without* any assumptions on energy loss. This hints at the potential in using the extra information provided by the lost clusters.

It is also clear from this figure that $\langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ is overall lower for the 10b data. This reflects the worse separation (due to cluster losses) between the π and p+K peaks in this data set, recovered to some extent in the other data, using one-pad-clusters, once the cluster loss effect was recognised. It might also point to a higher $\langle \mathrm{d}E/\mathrm{d}x\rangle_\pi$ overall in the later algorithms, which could result from adding one-pad-clusters to the $\mathrm{d}E/\mathrm{d}x$ calculation if this particular species didn't suffer much cluster losses before[3].

The main problem with the *ncl* method is that the *ncl* fits are unstable. The underlying idea was that good fits in this dimension would improve the fits in the $\mathrm{d}E/\mathrm{d}x$ dimension, and this was investigated by implementing different fit distributions in the *ncl* dimension. Unfortunately, stable fits have been very hard to obtain, even though occasions of good fits really do indicate that the desired effect on the $\mathrm{d}E/\mathrm{d}x$ fits would be realised. But then again, fits that don't converge are quite common.

One reason why the negative binomial fits are generally not performing very well for the p+K peak, even though they fit the separate particle species distributions very well (as seen in the MC fit study, Section 4.5.2), might be that the fit stretches all the way to 159. In simulations the fits go more steeply to zero already at $\sim$150. This is also the case for the Gaussian fits.

A possible way to improve this method might be to also get a $\langle ncl\rangle$ curve and subtract this from the *ncl* for each track – a Δ_π method of sorts, but in the *ncl* dimension. The advantage would be that the mean of the π *ncl* fit could be fixed to 0, which might improve the fit stability. On the other hand, this shifts the difficulty to getting a reliable curve for $\langle ncl\rangle_\pi$. This might be parametrised as a Bethe-Bloch curve, once the influence of the missing clusters on $\langle ncl\rangle$ vs p is understood. In principle, such a curve could be extracted from the *ncl* method, but then again, this requires a better fit stability, and the argument bites its tail.

Looking at the results (presented for Δ_π in the next section), the two methods give similar π-to-hadron ratios, even though the results for the *ncl* method have much larger uncertainties (the fit instability would add to the systematic errors).

It is worth pointing out that, even if fitting the distributions turned out to be difficult, the discovery of a larger separation seen in *ncl* vs $\mathrm{d}E/\mathrm{d}x$ really meant that there was room for improvements in the $\mathrm{d}E/\mathrm{d}x$ separation (as was shown in Fig. 5.9). The algorithms for calculating $\mathrm{d}E/\mathrm{d}x$ have improved in the last year due to an increasing awareness of this. Presently, it is attempted to estimate the number of missing clusters to take them into account, but it is highly non-trivial.

To really make use of the extra information given by the number of missing clusters, one would need a good handle on how many clusters could actually be found for each track. This is not the same as the global maximum *ncl*, since particles traversing uninstrumented areas, for example, would never reach this value regardless of the performance of the TPC. Only with this knowledge a proper correction of the $\mathrm{d}E/\mathrm{d}x$ can be made. An estimate of the number of clusters *below threshold* should be the next step in improving the $\mathrm{d}E/\mathrm{d}x$ algorithms.

An interesting study would be to base the $\mathrm{d}E/\mathrm{d}x$ calculations on a (Landau) fit to the charge distribution of the track, instead of calculating the truncated mean. Lost clusters at low values might then not influence the overall fit very much, making the $\mathrm{d}E/\mathrm{d}x$ less sensitive to cluster loss.

The answer to the question asked in the beginning of this section, is that the two methods

---

[3] It is clear that cluster loss increases the $\langle \mathrm{d}E/\mathrm{d}x\rangle$ for less ionising particles, so their $\langle \mathrm{d}E/\mathrm{d}x\rangle$ would decrease if this were to be properly compensated, leading to larger separation. It is less clear how the pions would be affected.

give similar results, but fits, systematics etc. are in much better control for the Δ_π method. Doubtlessly, the lost clusters contain much information, as is clear from the capabilities to determine $\langle \mathrm{d}E/\mathrm{d}x \rangle_\pi$ just from the *ncl* vs $\mathrm{d}E/\mathrm{d}x$ distributions. The best way to proceed, having investigated the effects of clusters below threshold, is most likely to make the most possible use of this information to improve the $\mathrm{d}E/\mathrm{d}x$ algorithms. Thus the TPC $\mathrm{d}E/\mathrm{d}x$ capabilities could be recovered, and the performance of the Δ_π method would improve.

With this said, the other conclusion is that the *ncl* method fits are not good enough to seem reliable for physics conclusions. Therefore, from this point on, only Δ_π results will be discussed.

5.4 Particle identification using the Δ_π method: data

5.4.1 Fits

Fig. 5.15 shows Δ_π fits in four example p_T bins (for all fits up to p_T = 16 GeV/c, see Appendix C). The smooth lines are fits to the dE/dx histogram in grey. These are the fits used for extracting particle yields.

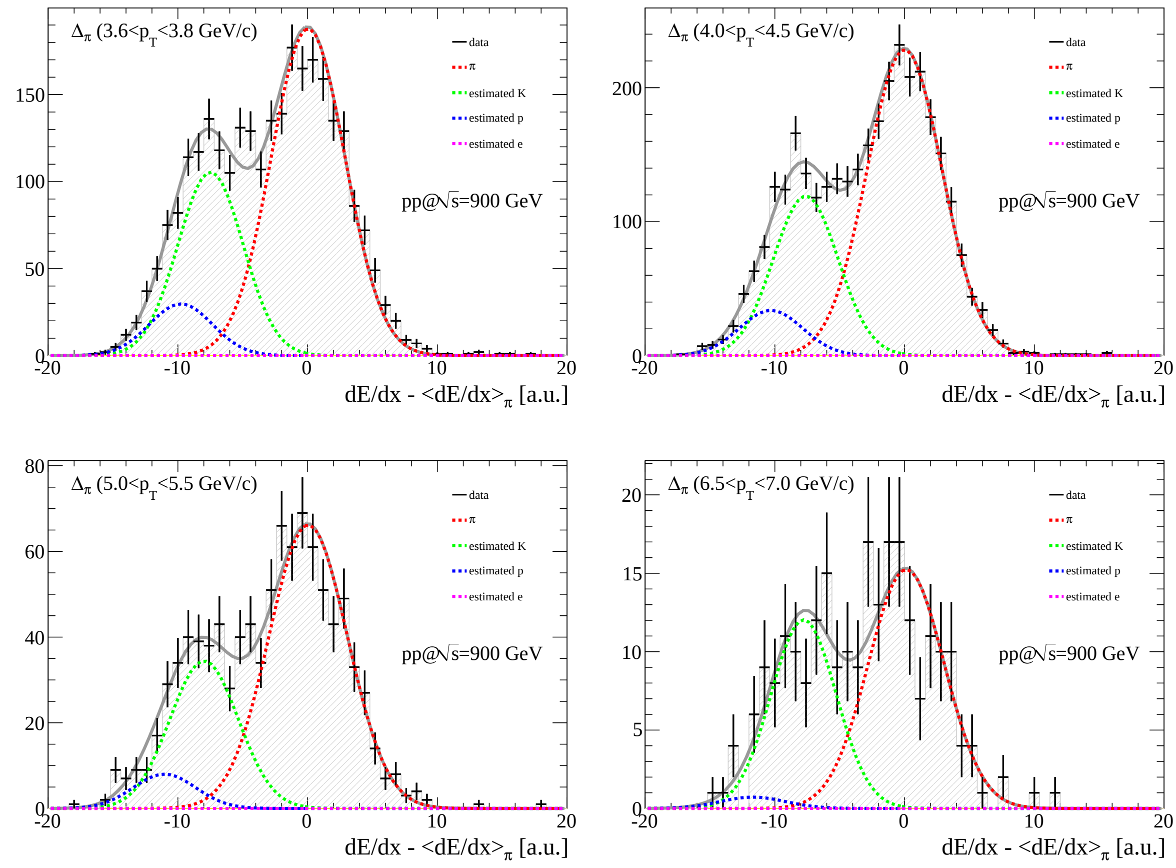
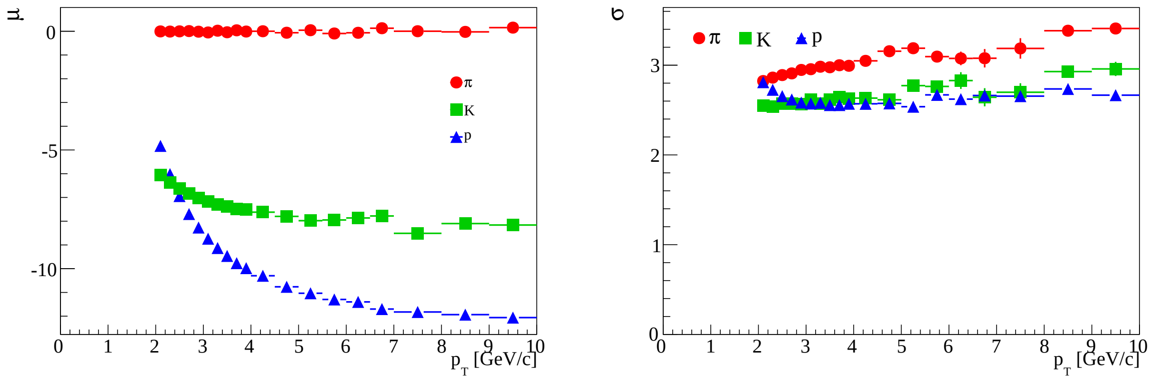


Figure 5.15: Examples of Δ_π fits.

It is seen from the examples that the fits to the π peaks are completely determined from the right-hand side. This means that the pion yields can be very well determined in this method. For K and p, on the other hand, there is more room for ambiguities. Here it is extra important to have a good Bethe-Bloch parametrisation, since this governs the means of the distributions and it is much more difficult to judge on the fit quality just from visual inspection.

Looking closer to the π peak in the examples, one sees that the fit systematically undershoots the edge on the right hand side – especially for the low p_T. This indicates that the distributions are slightly non-Gaussian here.

Fig. 5.16 shows the final Δ_π and $\sigma_{\mathrm{d}E/\mathrm{d}x}$ as a function of p_T for the fits used to determine the yields of the different particle species. Δ_π should be 0, by construction. It is seen in Fig. 5.16(a) that this holds, and that the separation between the particle species increases with p_T. This latter point is also illustrated in Fig. 5.17, which shows the separation $S_{12} = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1 \sigma_2}}$ between π and K, p respectively. Here one clearly sees that the separation tends to a constant value, as indicated by the simple mass dependence of the separation on the relativistic rise discussed in

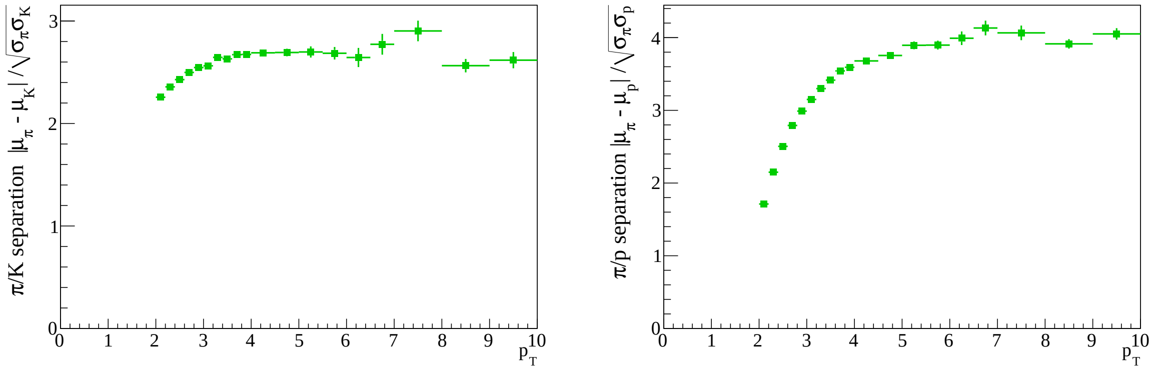


(a) Δ_π vs p_{T}

(b) $\sigma_{\mathrm{d}E/\mathrm{d}x}$ vs p_{T}

Figure 5.16: Resulting Δ_π and $\sigma_{\mathrm{d}E/\mathrm{d}x}$ from the fits vs p_{T}, for π (red), K (green) and p (blue).

Section 3.2.2. These two figures (particularly 5.16) also illustrate the difficulties to get reliable fit results in the region $p_{\mathrm{T}} < 3$ GeV/c, since here the K and p peaks overlap (protons are minimum ionising at ∼3 GeV/c).



(a) π to K separation vs p_{T}

(b) π to p separation vs p_{T}

Figure 5.17: Separation between particle species vs p_{T}.

Finally, comparing Fig. 5.16(b) and 5.17, one realises that as the resolution $\sigma_{\mathrm{d}E/\mathrm{d}x}$ (as it is often quoted when characterising detector performance) gets worse with increasing p_{T}, the separation between the species in fact gets better or tends to a constant value. This again points to $\sigma_{\mathrm{d}E/\mathrm{d}x}$ not being an unambiguous measure of the resolving power of the detector.

5.4.2 Yield ratios

Fig. 5.18 shows the ratios from the fits of antiparticles to particles, for π, K and p. There is no clear trend. The π and K fits are consistent with 1 to within ∼ 20% in the first few p_{T} bins. It is however not a priori clear what to expect. The large fluctuations, not seen in particle fractions shown later (Figs. 5.21-5.22) could indicate large fluctuations in the overall production of positive and negative particles.

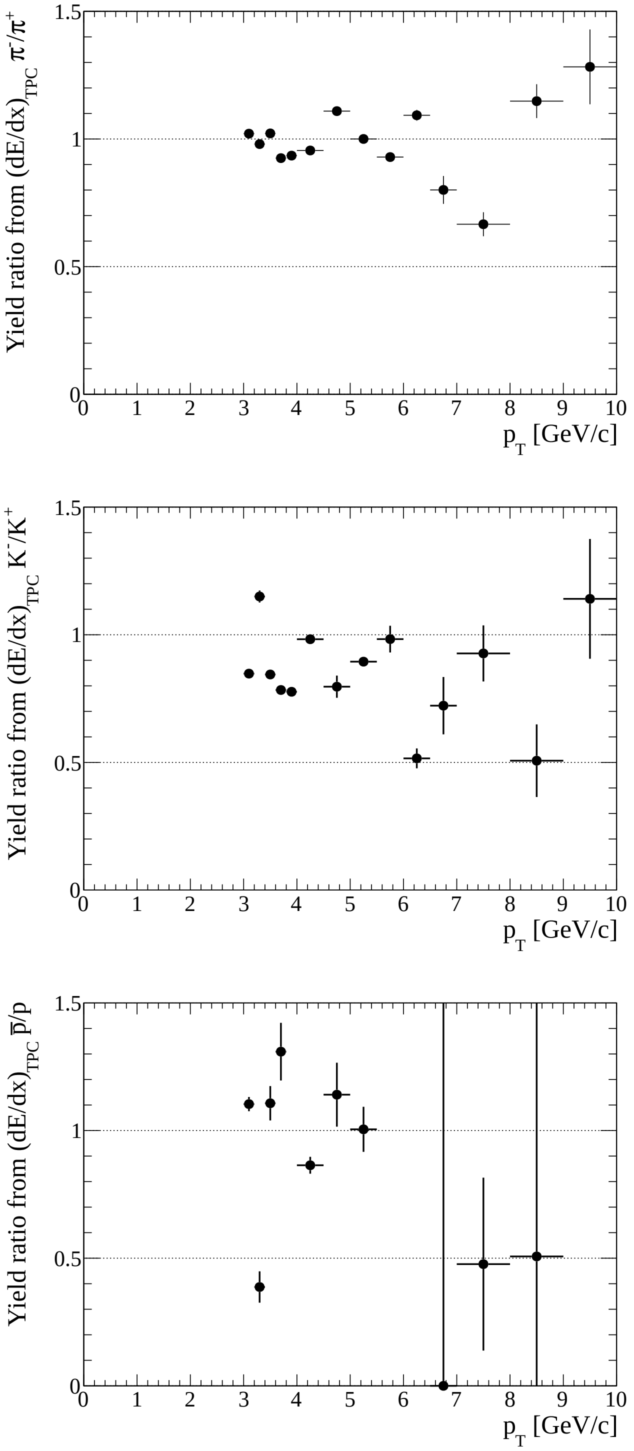


Figure 5.18: The ratio of antiparticles to particles, for π (top), K (middle) and p (bottom), for $|\eta| < 0.8$.

5.4.3 p_{T} spectra

Fig. 5.19 shows the p_{T} spectrum from the efficiency and acceptance corrected yield of the fits. The charged pions are shown as full circles. Here also the spectra from two other ALICE analyses at the same energy are drawn for comparison.

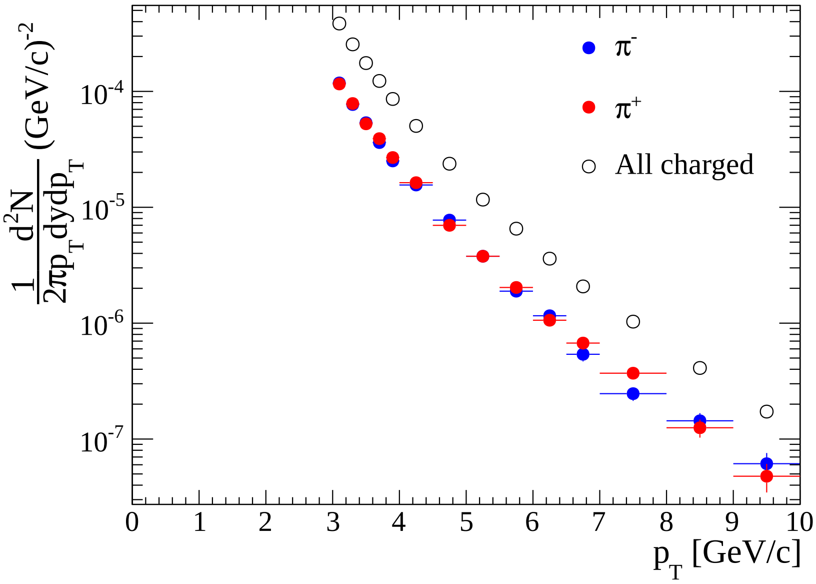


Figure 5.19: Charged particle p_{T} spectrum: efficiency corrected yield as a function of p_{T}, for $|\eta| < 0.8$.

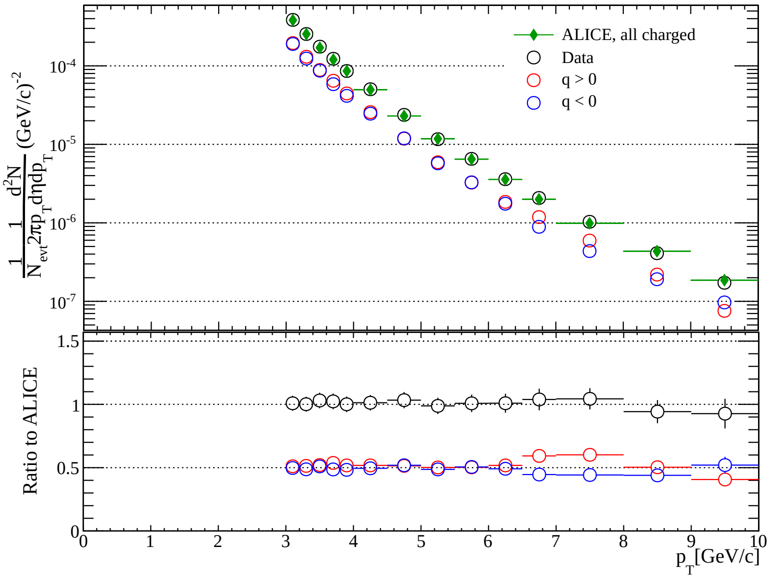


Figure 5.20: Charged particle p_{T} spectrum: invariant charged particle yield as a function of p_{T}, for $|\eta| < 0.8$, with comparison to ALICE data at 900 GeV.

Fig. 5.20 shows the charged particle yield from the analysis ("Data", open circles), and the points from the dedicated ALICE analysis of this observable ("ALICE", green diamonds), using the same data. A systematic uncertainty of 1.29% from normalisation to the number of inelastic events is not included in the points, but other systematic and statistic uncertainties are added in quadrature. The positive and negative yields are also shown separately (red and blue open circles, respectively). The bottom panel shows the ratio to the ALICE points, and it is seen that the two analyses are consistent.

Using the charged particle yield shown here, one could apply Eq. 4.2 to get the invariant pion yields. The results presented in the following will instead be based on the raw pion yields found in the present analysis, following Eq. 4.1. The very good agreement between the charged particle yields from this analysis and the ALICE results gives confidence in the applied corrections, and in deriving the invariant charged pion yields from the obtained raw yields.

5.4.4 Particle fractions

Figs. 5.21-5.22 show the particle fit yield ratios of π, K and p for 10c data. The π fractions are corrected for the slightly higher reconstruction efficiency for π compared to all charged particles (seen in Fig. 4.13), by multiplying by $\frac{\epsilon_{ch}}{\epsilon_{\pi}}$ (cf. Eq. 4.2). Looking at the white circles, representing both charges, especially the π fraction is stable over the p_{T} range.

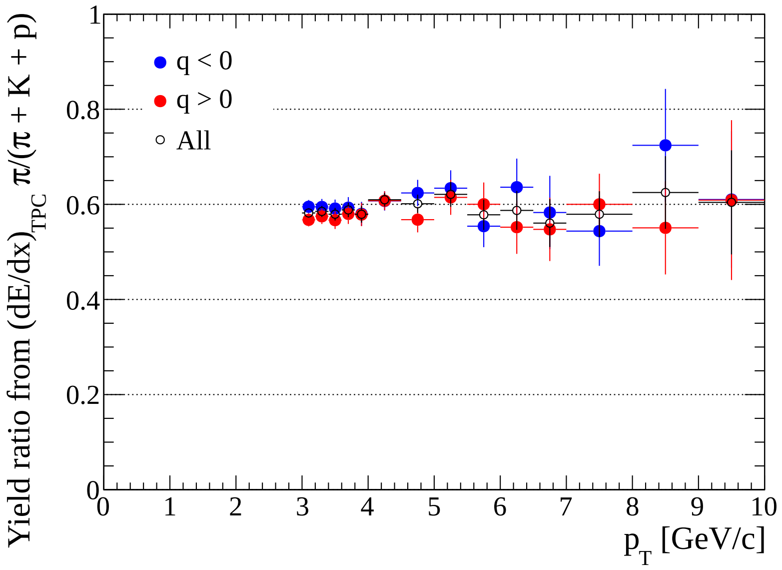


Figure 5.21: Efficiency corrected pion fraction as a function of p_{T}, for $|\eta| < 0.8$.

See Fig. 5.24 for a comparison of the uncorrected particle fractions (positive and negative together) for 10c and 10d. The main difference here is, apart from $\sqrt{s}$, the statistics; the dE/dx algorithm is the same, including one-pad-clusters.

We see in this figure that the pion fraction is $\pi/h \approx 0.6$ and more or less independent on both $\sqrt{s}$ and p_{T}. The results shown are in agreement with what was discussed by the PHENIX collaboration [10]; they quote own results and numerous observations (e.g. [21-23]) of the ratio h/π being 1.6 for p_{T} above 1.5 GeV/c, for a wide range of $\sqrt{s}$. Note that this is the inverted pion ratio compared to what is given here; it corresponds to, in our notation, $\pi/h = (1.6)^{-1} = 0.625$.

This constant ratio is of particular interest when it comes to doing relative R^{π}_{AA} , which is the ratio between the π fraction in AA and pp collisions. In this ratio, any suppression or

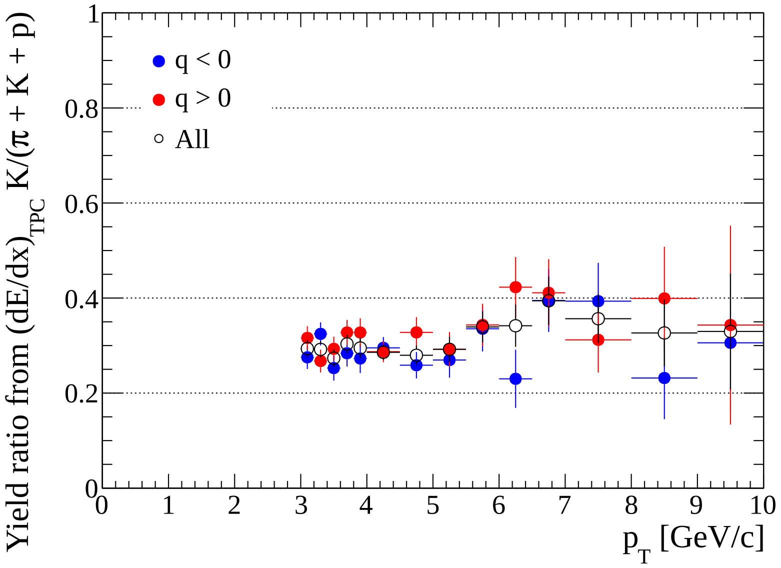


Figure 5.22: Kaon fraction as a function of p_{T}.

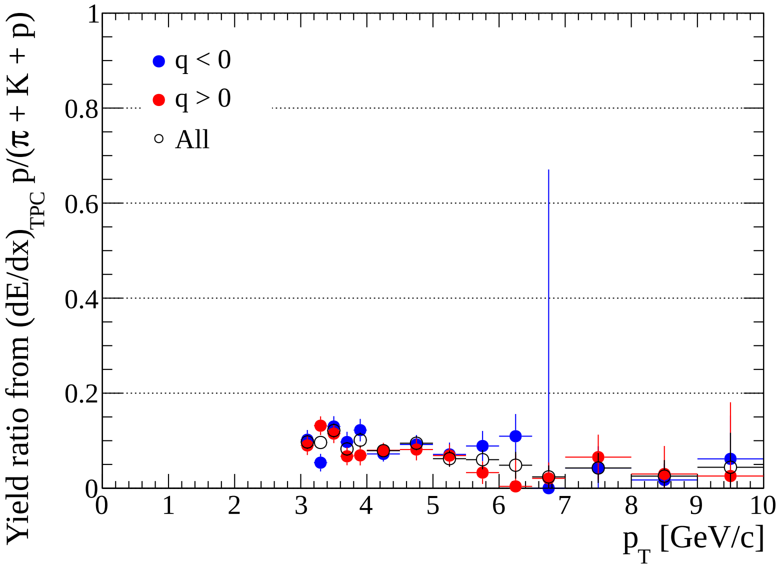


Figure 5.23: Proton fraction as a function of p_{T}. The $q < 0$ point in the $6.5 \leq p_{\mathrm{T}} \leq 7.0$ bin is from a fit that didn't converge.

enhancement in AA collisions would be directly visible. The data compared should of course be at the same nucleon energies. But, if the π fraction in pp is constant in an energy range covering the $\sqrt{s_{\mathrm{NN}}}$ of the AA collisions, a pp reference at this particular energy is not as crucial; the physics message is seen anyway.

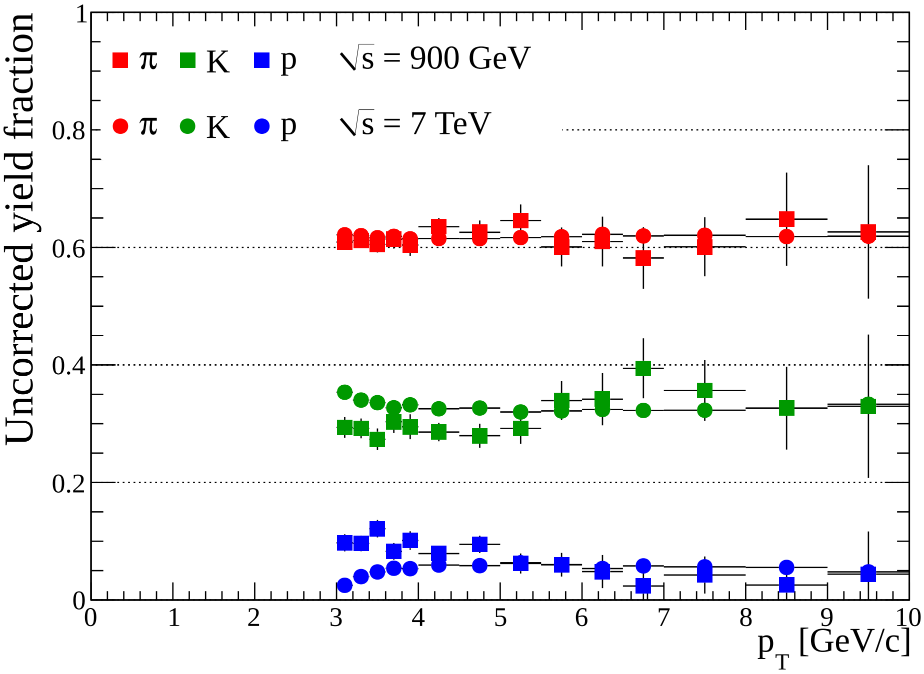


Figure 5.24: Particle fractions (uncorrected) as a function of p_T, for 10c data ($\sqrt{s}$ = 900 GeV) and 10d data ($\sqrt{s}$ = 7 TeV).

5.4.5 Pion yield

Fig. 5.25 shows the π^+ and π^- invariant yield vs p_T. It is compared to the invariant π^0 yield (blue squares), as obtained from the PHOS (PHOton Spectrometer) and conversion electron analyses in ALICE. The ratio in the bottom panel, obtained from dividing the π^0 spectrum with a power law fit to the charged pion yields as $\frac{\pi^+ + \pi^-}{2}$, is almost consistent with 1, within errors. Naively, one would expect that the yield of charged and neutral pions would be the same, from isospin symmetry. On the other hand, they might be affected differently by so-called feed-down from decaying particles. Also, the π^0 analyses make use of different detectors than this analysis does, making it hard to tell if this is really a one-to-one comparison. Finally, discrepancies might be within the (unknown) systematical uncertainties. A PYTHIA study on these yields might shed light on any physically motivated differences.

From these yields, one obtains the crossection for π^+ and π^- production by multiplying by the total inelastic crossection $\sigma_{\text{INEL}} = 50.3$ mb.

5.4.6 Final discussion

The results shown in this analysis have focused on extracting π yields, but some p and K results are also shown. To be able to extract p and K yields, a more thorough understanding of their distributions in the lower p_T bins, where they are to the left of the MIP region in the Bethe-Bloch curve, is needed. The π MIP region is well isolated in dE/dx vs p histograms, and can be fitted. From this $\sigma_{\mathrm{d}E/\mathrm{d}x}$for the rest of the Bethe-Bloch curve is extracted, and it seems to hold well for π peaks on the relativistic rise, from looking at the fits, as does the $\langle \mathrm{d}E/\mathrm{d}x \rangle$. But it is not a priori certain that this holds for the other particle species – for once, there is an *ncl* dependence of $\sigma_{\mathrm{d}E/\mathrm{d}x}$, and this is in turn species dependent, at least as long as there is a

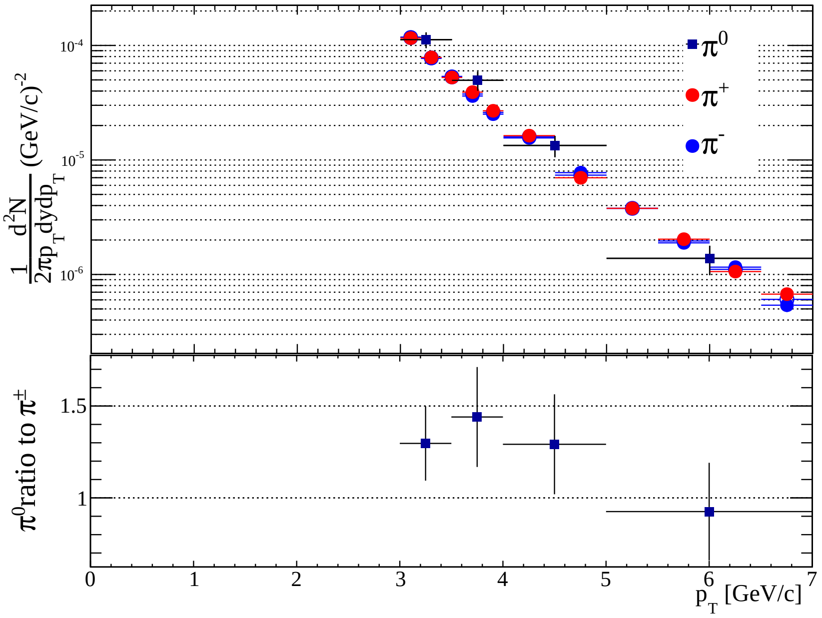


Figure 5.25: π p_T spectrum: invariant charged π yield as a function of p_T, with comparison to π^0 data from PHOS and conversion analysis at 900 GeV.

threshold effect.

One way to verify that the Bethe-Bloch parametrisation holds for $\beta\gamma$ below the MIP region would be to make use of the better species separation at low p_T. This region is presently not used in this analysis, since it would quite drastically increase the amount of data handled. However, this cross check could give valuable insights, potentially increasing the confidence in the results in the momentum region where the energy loss curves cross.

Another cross check would be to use the possibility of selecting identified particle tracks with the combined information from the TOF and the TPC. This has a limited p_T reach, but would again possibly improve the understanding of the energy loss at lower p_T.

Improvements of the Δ_π method also include getting a handle of the η dependence of *ncl* and $\mathrm{d}E/\mathrm{d}x$. It would be desirable to have greater insight in what corrections to $\mathrm{d}E/\mathrm{d}x$ are done with respect to track angle (i.e., path length), and to study what effects enter in the η dependence of *ncl*. The η dependence of the $\mathrm{d}E/\mathrm{d}x$ separation in the lower p_T bins indicates that the analysis should be binned also in η in this region.

For a cross check of the $\mathrm{d}E/\mathrm{d}x$ fit results, one could compare to the $\mathrm{d}E/\mathrm{d}x$ and $\sigma_{\mathrm{d}E/\mathrm{d}x}$ of π and p identified to be the products of Λ and K^0 decay. These are selected by, among others, cuts on invariant mass, making this a completely different analysis – but the energy loss of π and p from these decays should of course be the same.

The MC results could be cross checked by thorough comparison of results from the PYTHIA and PHOJET event generators. Finally, the systematic uncertainties should be estimated.

Among the strengths of the method, I appreciate the visual aspect. A bad fit is easily recognised and discarded – this is also a feature of the *ncl* method, which in that case made me not trust those results for physics conclusions. The robust π fits are reflected in the comparison between positive and negative π fractions, where fluctuations in the total number of charged particles cancel, and the π fraction is the same for both charge signs.

Chapter 6

Summary and conclusions

The TPC is a high-resolution instrument, which simultaneously measures track coordinates, momentum and energy loss to high precision, making the small details, such as path length dependence of the energy loss, visible. In this analysis, quite some effort has been invested in understanding the details. The influence on the energy loss measurement of clusters falling below read-out threshold has been given special attention.

This analysis shows that there is much information to gain from understanding which clusters are lost. In particular, the $\mathrm{d}E/\mathrm{d}x$ separation between different particle species is reduced by lost clusters. When the analysis started, the effects were not understood, and the algorithms for calculating $\mathrm{d}E/\mathrm{d}x$ have improved since, with an increasing awareness of the influence of cluster losses. A first step was to include one-pad-clusters in the $\mathrm{d}E/\mathrm{d}x$ calculations, and later on a correction based on calculating the number of lost clusters has been implemented. A proper correction taking into account the maximum number of clusters expected for each track would improve the $\mathrm{d}E/\mathrm{d}x$ resolution in the Δ_π method, and thus improve the performance of this method, which is one of the few ways of doing hadron identification at high p_T ($p_\mathrm{T} > 2$ GeV/c).

The MC study of the Δ_π method shows that the fits and the truth distributions coincide very well, especially for the π distributions. For the π fits, the Gaussians are determined by the right-hand side of the distributions, making these results easily verified by visual inspection. The p and K peak fits require more studies before they are entirely understood, due to the smaller relative separation of their $\mathrm{d}E/\mathrm{d}x$ distributions. In the MC study, they too are well described by the fits, however here the species separation is larger overall.

The Δ_π method has in this analysis been applied on the 10c Pass3 data, which is from 900 GeV pp collisions. Pass3 is the final iteration, including one-pad-clusters in the $\mathrm{d}E/\mathrm{d}x$ calculations, to some extent recovering the reduced species $\mathrm{d}E/\mathrm{d}x$ separation. For this data, the invariant charged pion yield in pp collisions at 900 GeV has been measured and presented in the p_T range from 3 to 10 GeV/c. The results are not in as good agreement with the π^0 invariant yield at the same energy, which may have several causes; the analyses are different, and the magnitude of the systematical uncertainties in my analysis is not known. However, the ratio is not very far from 1, which would be the expected result.

The charged particle spectrum from this analysis is in very good agreement with other ALICE results for the same data, giving confidence in the normalisation.

The pion fractions $\pi/h \sim 0.6$ are constant within errors over the given p_T range, and for both $\sqrt{s}$ = 900 GeV and $\sqrt{s}$ = 7 TeV, in agreement with results seen in previous experiments, for a wide range of $\sqrt{s}$.

Appendix A

Variables

The centre-of-mass energy $\sqrt{s}$ is given by, from s being the sum of the particle four-momenta,

$$\sqrt{s} = \sqrt{(\sum E)^2 - (\sum p)^2} \tag{A.1}$$

where E is particle energy, p particle three-momentum and the sums are over the projectile and target particles. In a collider experiment where the colliding beams consist of the same particle species at equal and opposite momenta, $\sqrt{s}$ is simply the sum of the particle energies. For AA collisions, $\sqrt{s}$ = A$\cdot\sqrt{s_{\mathrm{NN}}}$, $\sqrt{s_{\mathrm{NN}}}$ being the nucleon pair centre-of-mass energy.

The ALICE (right-handed orthogonal Cartesian) coordinate system defines the beam direction to be the z axis (positive direction towards side "A" of ALICE). The x axis is the horizontal axis orthogonal to the beam (positive direction towards the LHC centre) and y is the vertical axis (pointing up).

The energy of a particle with momentum p and mass m is , in natural units ($c = 1$),

$$E = \sqrt{p^2 + m^2}. \tag{A.2}$$

In Cartesian coordinates, the momentum vector $\mathbf{p}$ is given in three dimensions by components $\mathbf{p} = (p_x, p_y, p_z)$. The first two can be replaced by knowing the transverse momentum p_{T} and azimuthal angle ϕ (measured counter-clockwise with $\phi = 0$ at the positive x axis). The polar angle is given by θ, with $\theta = 0$ at the positive z axis.

The rapidity y is given by

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \approx \frac{1}{2}\ln\left(\frac{p + p_z}{p - p_z}\right) = \frac{1}{2}\ln\left(\frac{p(1 + cos\theta)}{p(1 - cos\theta)}\right) = -\ln\left(tan\frac{\theta}{2}\right) \equiv \eta \tag{A.3}$$

where Eq. A.2 with $p_{\mathrm{T}} >> m$ enters in the second step. The pseudo-rapidity η is a useful quantity when the mass is unknown (as the mass enters via E in the expression for y).

When going from $\frac{dN}{dy}$ to $\frac{dN}{d\eta}$, a useful relation is $\frac{dN}{dy} = \frac{d\eta}{dy}\frac{dN}{d\eta} = \frac{E}{p}\frac{dN}{d\eta}$. For π, $\frac{E}{p} \approx 1$ already at moderate momenta, making the transformation straightforward.

The particle masses used are $m_\pi = 0.13957$ GeV/c^2, $m_K = 0.49368$ GeV/c^2 and $m_p = 0.93827$ GeV/c^2.

Instantaneous luminosity L can be defined, via the integrated luminosity, as

$$N = \sigma \int L dt \tag{A.4}$$

where N is the number of interactions with a certain crossection σ.

Appendix B

Data used

B.1 Runs

Here the run periods, $\sqrt{s}$ energies, Passes, runs and number of events are listed for the data used in the analysis.

Run period: 10b, $\sqrt{s} = 7$ TeV, Pass2
Run numbers: 114931, 115186, 115193, 115310, 115315, 115318, 115322, 115325, 115328, 115335, 115338, 115345, 115393, 115401, 115414, 115514, 115521, 115890, 115892, 116102, 116112, 116287, 116288, 116401, 116402, 116403, 116561, 116562, 116571, 116574, 116609, 116611, 116643, 116644, 116645, 116684, 117048, 117050, 117052, 117053, 117054, 117059, 117060, 117063, 117065, 117077, 117082, 117086, 117092, 117099, 117109, 117112, 117116, 117118, 117120, 117220, 117222
Number of events: 27491686

Run period: 10c, $\sqrt{s} = 900$ GeV, Pass3
Run numbers: 118506, 118507, 118512, 118518, 118556, 118558, 118560, 118561, 121039, 121040
Number of events: 4962293

10c Pythia MC
Number of events: 5574400

10c Phojet MC
Number of events: 4300000

Run period: 10d, $\sqrt{s} = 7$ TeV, Pass2
125085, 125097, 125101, 125134, 125139, 125140, 125186, 125628, 125630, 125632, 125842, 125849, 125850, 125851, 125855, 126007, 126008, 126073, 126078, 126081, 126082, 126088, 126090, 126097, 126158, 126160, 126168, 126283, 126284, 126351, 126352, 126359, 122374, 126404, 126406, 126407, 126408, 126409, 126422, 126424, 126425, 126432, 126437
Number of events: 65043947

B.2 Event and track selection criteria

The trigger conditions used here are the so-called MB1 (for Minimum Bias), which require a hit in either of the V0 detectors, at least one hit in the SPD, and has a veto on beam gas interactions. In addition, a reconstructed vertex within $|z| < 10$ cm is required (because here the reconstruction efficiency is more or less constant). For tracks, the ALICE standard requirements are $|\eta| < 0.8$ (for full TPC acceptance), $ncl > 70$ and $\chi^2 < 4$ per cluster in the track fitting. There must be at least two hits in the ITS, at least one being in the SPD. Finally, primary particles (particles produced in the collision, or in subsequent strong or electromagnetic decays, or weak decays of heavy quarks) are selected by several criteria:

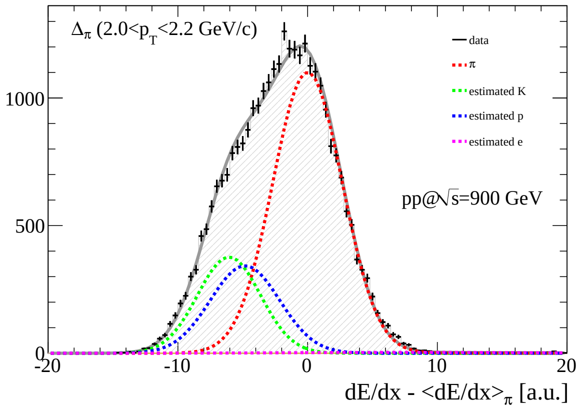
- rejecting “kink” daughters, meaning a track with a sudden change in direction
- rejecting tracks where the distance of closest approach to reconstructed vertex exceeds a certain value

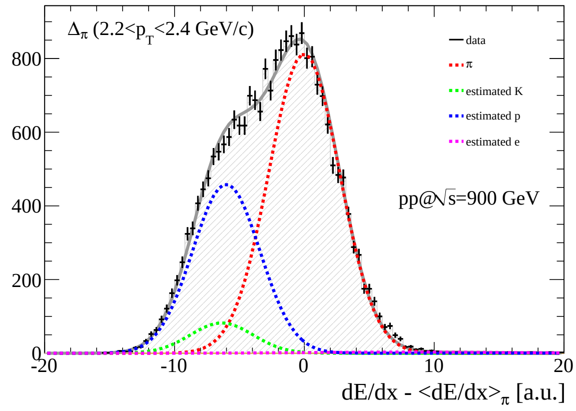
These criteria should reject weak decay products or particles emerging from photon conversions and other (including hadronic) interactions in the detector material.

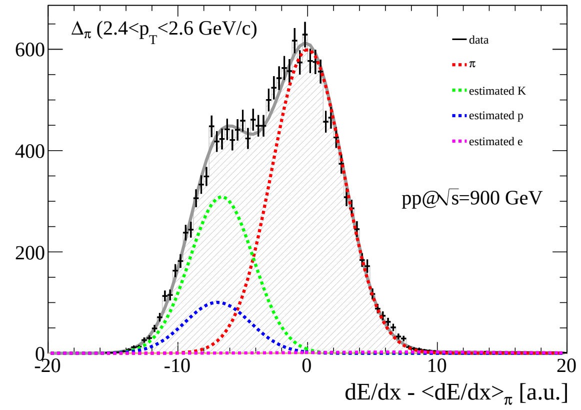
Appendix C

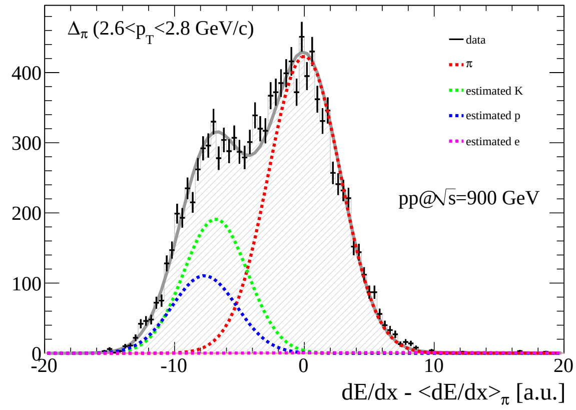
10c fits

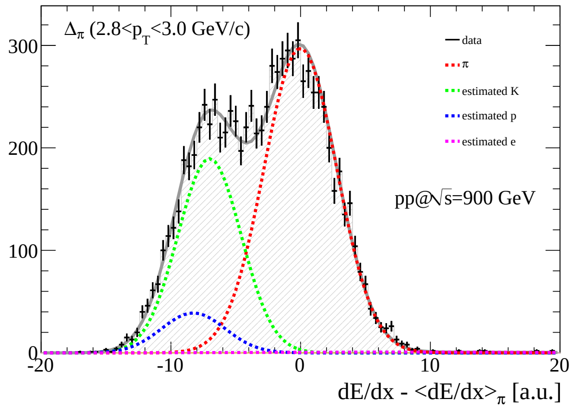
This appendix shows the fit results for the 10c data out to p_{T} = 16 GeV/c.

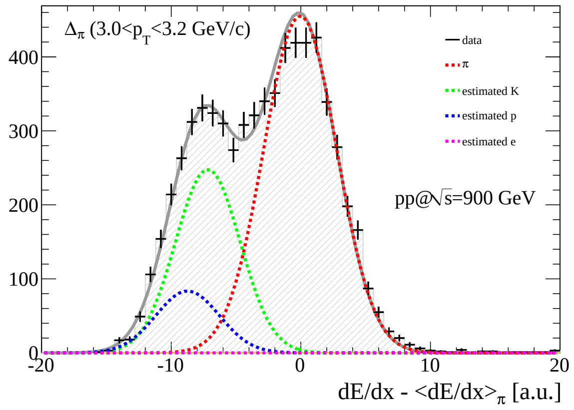


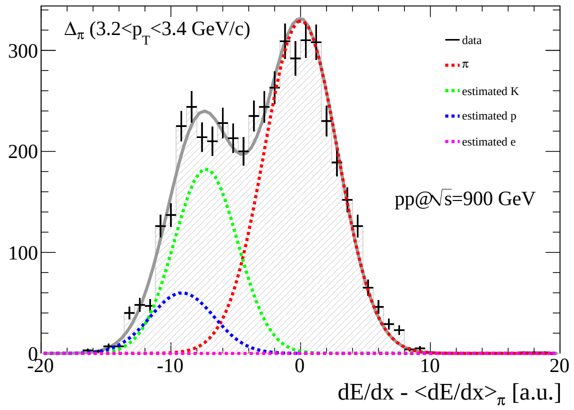


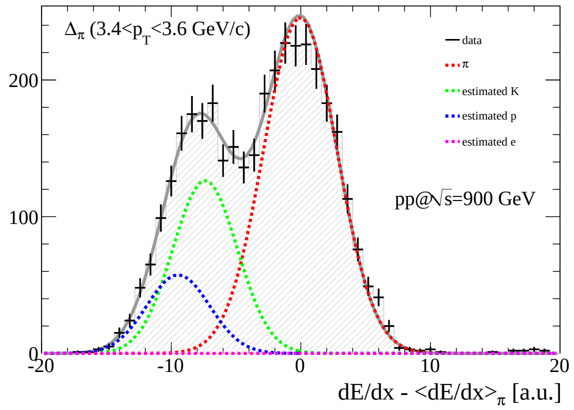


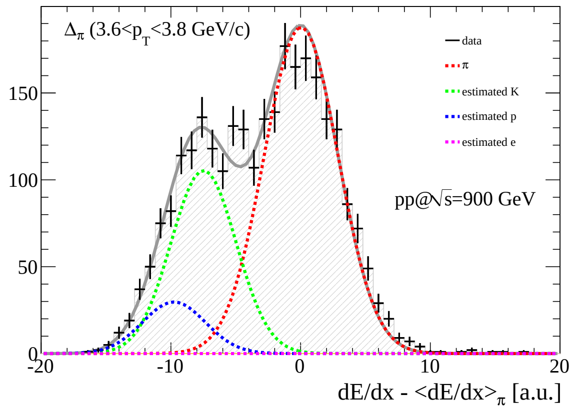


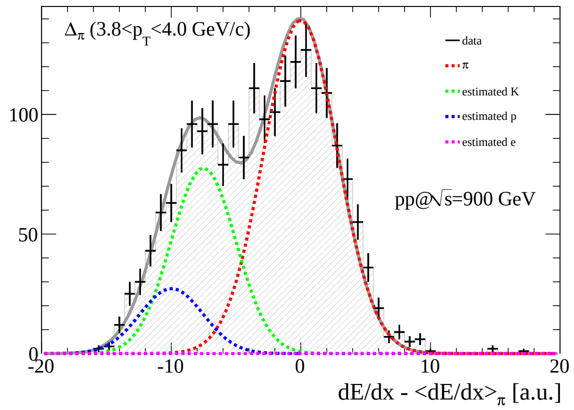

Δ$_\pi$ (2.8<p$_T$<3.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
100
200
300
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

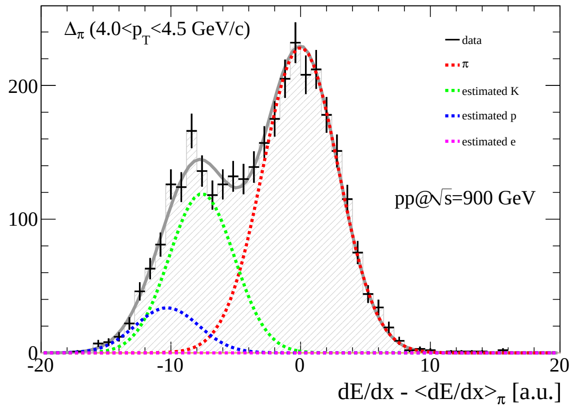

Δ$_\pi$ (3.0<p$_T$<3.2 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
100
200
300
400
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

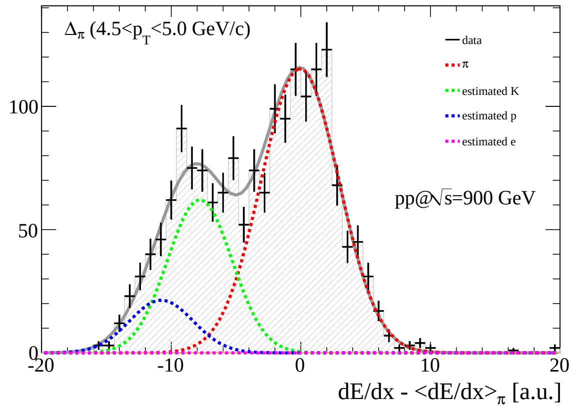

Δ$_\pi$ (3.2<p$_T$<3.4 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
100
200
300
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

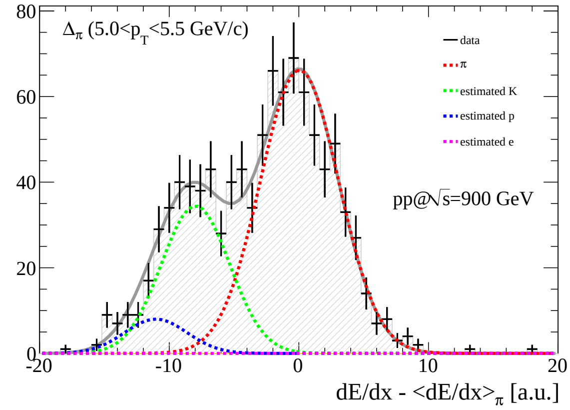

Δ$_\pi$ (3.4<p$_T$<3.6 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
100
200
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

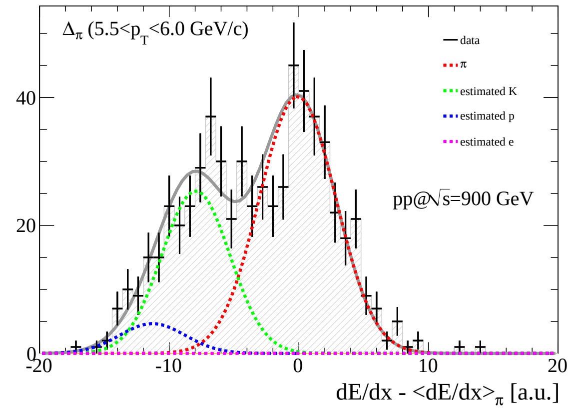

Δ$_\pi$ (3.6<p$_T$<3.8 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
50
100
150
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

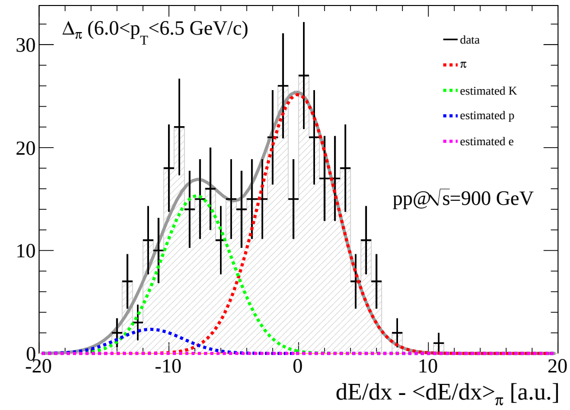

Δ$_\pi$ (3.8<p$_T$<4.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
50
100
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

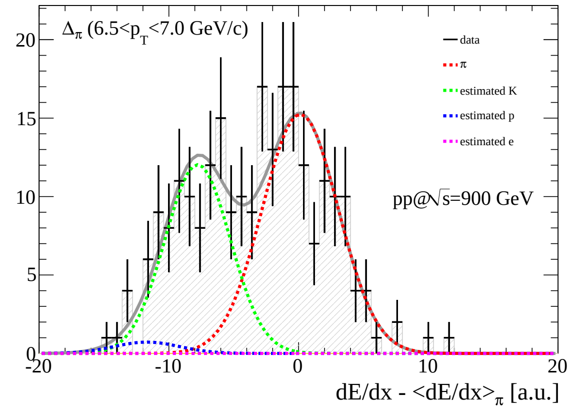

Δ_π (4.0<p_T<4.5 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
100
200
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

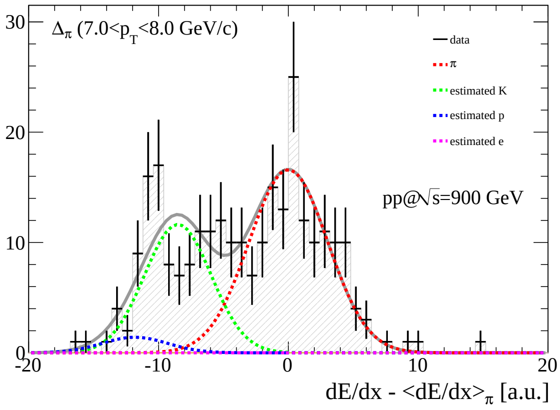

Δ_π (4.5<p_T<5.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
50
100
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

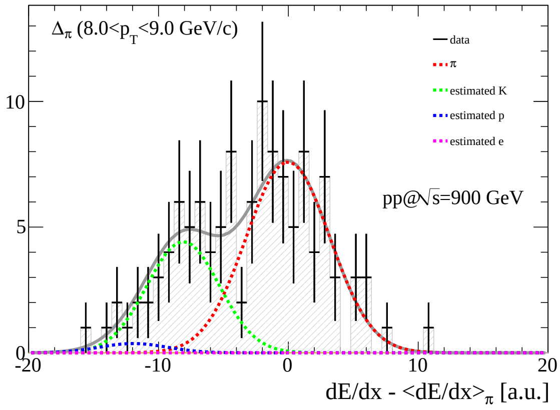

Δ_π (5.0<p_T<5.5 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
20
40
60
80
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

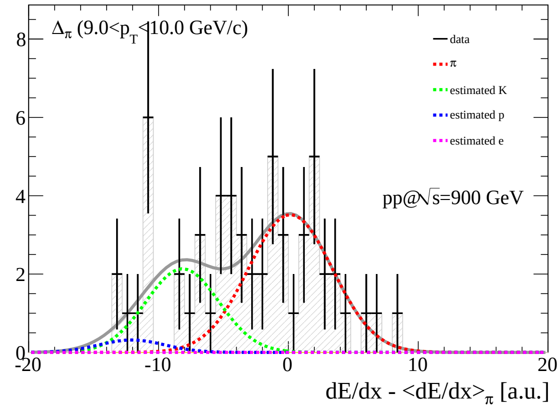

Δ_π (5.5<p_T<6.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
20
40
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

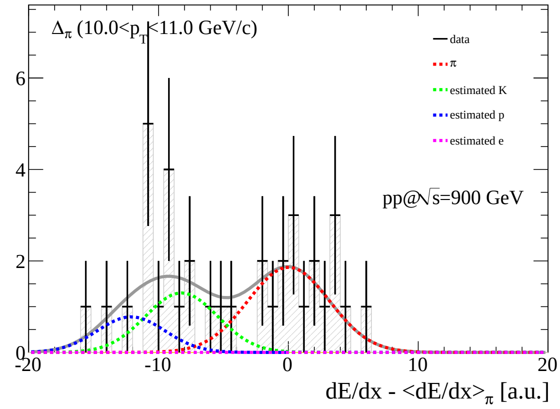

Δ_π (6.0<p_T<6.5 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
10
20
30
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

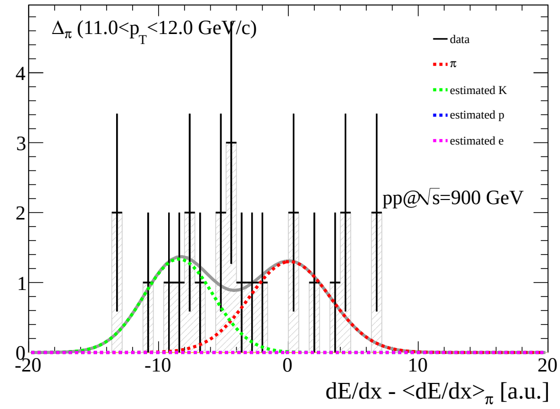

Δ_π (6.5<p_T<7.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
5
10
15
20
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

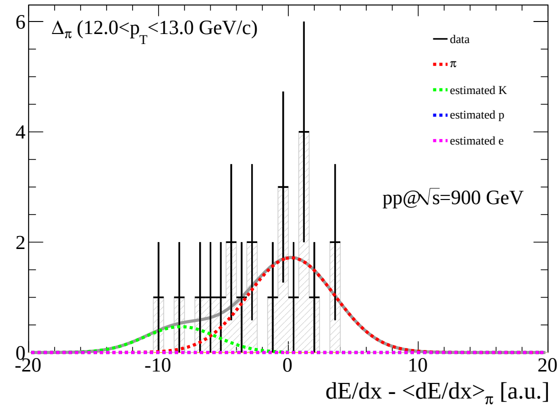

Δ_π (7.0<p_T<8.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@$\sqrt{s}$=900 GeV
30
20
10
0
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

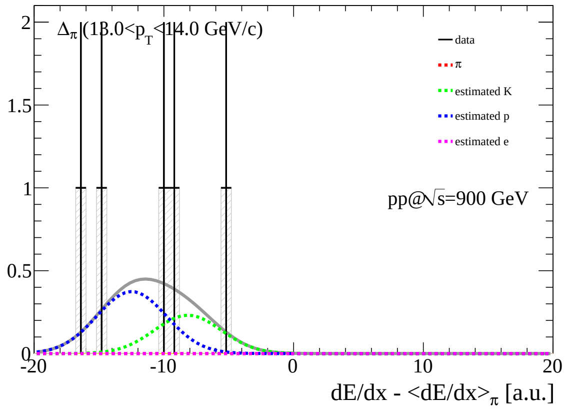

Δ_π (8.0<p_T<9.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@$\sqrt{s}$=900 GeV
10
5
0
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

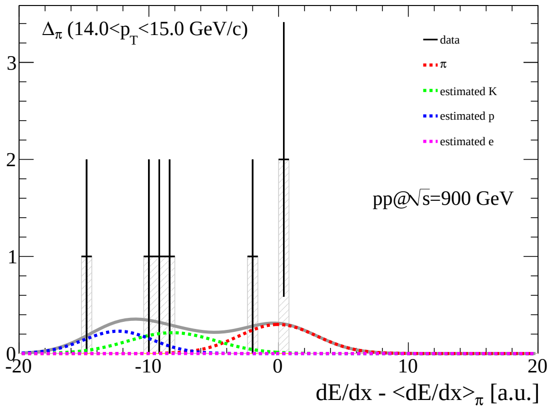

Δ_π (9.0<p_T<10.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@$\sqrt{s}$=900 GeV
8
6
4
2
0
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

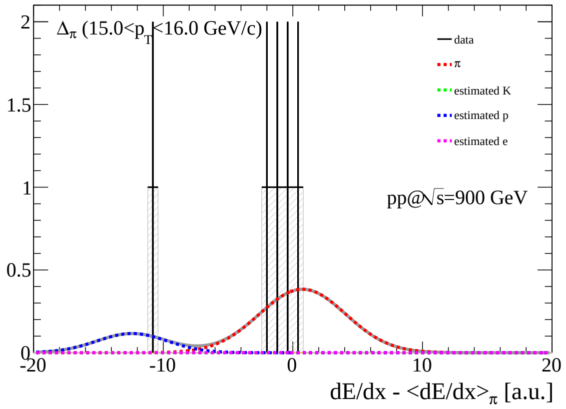

Δ_π (10.0<p_T<11.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@$\sqrt{s}$=900 GeV
6
4
2
0
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]


Δ_π (11.0<p_T<12.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@$\sqrt{s}$=900 GeV
4
3
2
1
0
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]


Δ_π (12.0<p_T<13.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@$\sqrt{s}$=900 GeV
6
4
2
0
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]


Δ_π (13.0<p_T<14.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
0.5
1
1.5
2
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]


Δ_π (14.0<p_T<15.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
1
2
3
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]


Δ_π (15.0<p_T<16.0 GeV/c)
data
π
estimated K
estimated p
estimated e
pp@√s=900 GeV
0
0.5
1
1.5
2
-20
-10
0
10
20
dE/dx - <dE/dx>$_\pi$ [a.u.]

References

1. B. Andersson, G. Gustafsson, G. Ingelmann and T. Sjöstrand. *Parton fragmentation and string dynamics.* Phys. Rept. **97** 31-145 (1983)
2. L. McLerran, Lect. Notes Phys. **221**, 1 (1985); M. Gyulassy, Z. Phys. C **38**, 1 (1988)
3. CMS Collaboration, arXiv:1102.1957v2 [nucl-ex] 10 Feb 2011
4. R. J. Fries, B. Müller, C. Nonaka, S. A. Bass, Phys. Rev. C **68**, 044902 (2003)
5. R. Snellings. *Elliptic flow: A Brief review.* New Journal of Physics 13 (2011) 055008
6. B. Alver, G. Roland, Phys. Rev. C**81**, 054905 (2010)
7. J. Jia, *ATLAS v_n results*, presentation for the ATLAS Collaboration, Quark Matter 2011, Annecy, France
8. K. Adcox *et al.* (PHENIX Collaboration), Phys. Rev. Lett. **88**, 022301 (2002), C. Adler *et al.* (STAR Collaboration), Phys. Rev. Lett. **89**, K. Adcox *et al.* (PHENIX Collaboration), Phys. Lett. B 561, **82** (2003).
9. B. I. Abelev *et al.* (STAR Collaboration), Phys. Rev. Lett. **97** (2006) 152301
10. S. S. Adler *et al.* (PHENIX Collaboration), Phys. Rev. C **69**, 034910 (2004)
11. T. Alber *et al.* (NA49 Collaboration), Phys Rev. Lett. **75**, 3814 (1995)
12. K. Aamodt *et al.* (ALICE Collaboration), Eur. Phys. J. C (2010) **65**: 111-125
    First proton–proton collisions at the LHC as observed with the ALICE detector: measurement of the charged-particle pseudorapidity density at $\sqrt{s}$ *= 900 GeV*
13. ALICE Web site,
    http://aliceinfo.cern.ch/Public/Welcome.html,
    January 8, 2011
14. K. Aamodt *et al.* (ALICE Collaboration), Phys. Lett. B **696** (2011) (arXiv:1012.1004 [nucl-ex])
15. CERN Web site,
    http://public.web.cern.ch/public,
    October 29, 2010
16. K. Aamodt *et al.* (ALICE Collaboration), JINST **3** (2008) S08002:
    The ALICE experiment at the CERN LHC
17. J. Alme *et al.* Nuclear Instruments and Methods in Physics Research A **622** (2010) 316–367 (arXiv:1001.1950 [physics.ins-det])

18. A. Dobrin, *R-hadron Searches and Charged Pion Spectra at High Transverse Momentum in Proton-Proton Collisions at the LHC using the ALICE Detector*,
Doctoral Thesis, Lund University, Lund, 2010
19. K. Aamodt *et al.* (ALICE Collaboration), Physics Letters B **693** (2010) 53–68 (arXiv:1007.0719 [hep-ex])
20. W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments. A How-to Approach*, Second Revised Edition, Springer-Verlag New York Berlin Heidelberg, 1994
21. B. Alper *et al.*, Nucl. Phys.**B100**, 237 (1975)
22. T. Alexopoulos *et al.*, Phys. Rev. D **48**, 984 (1993)
23. P. Abreu *et al.*, Eur. Phys. J. C **17**, 207 (2000); O. Klapp, Doctoral thesis, University Wuppertal, 1999