Measurement of pions, kaons and protons with the ALICE detector in pp collisions at the LHC

Meting van pions, kaons en protons met de ALICE detector in pp botsingen in de LHC

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op donderdag 3 mei 2012 des ochtends te 12.45 uur

door

Marek Chojnacki

geboren op 23 januari 1983
te Radom, Polen

Promotor:	Prof. dr. R.J.M. Snellings
	Prof. dr. R. Kamermans (†)
Co-promotor:	Dr. ir. G.J.L. Nooren
	Dr. ir. M. van Leeuwen

Dit werk maakt deel uit van het onderzoekprogramma van de Stichting voor Fundamenteel Onderzoek der Materie (FOM), die financieel wordt gesteund door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

Contents

Outline

The research reported in this thesis was performed using data collected by A Large Ion Collider Experiment (ALICE) during the first year of the Large Hadron Collider (LHC) operation. The main subject is a measurement of pion, kaon, proton and antiproton production as a function of transverse momentum (p_{t} spectra) in proton-proton collisions. The particle identification is based on the energy loss in the Inner Tracking System (ITS), a silicon detector. Chapter 1 is an introductory chapter, which presents the theoretical background of the particle production in proton-proton collisions. Chapter 2 contains the description of the ALICE detector. In Chapter 3 the calibration procedure of the Silicon Strip Detector (part of the ITS) is described. In Chapter 4 the method of the particle production measurement is described, which was used for two energies of the colliding protons ($\sqrt{s}$ =7 TeV and $\sqrt{s}$ =900 GeV). In Chapter 5 the 7 TeV results from Chapter 4 are compared and combined with results from other methods of particle identification. The combined result is compared with theoretical models, predictions of event generators and previous experimental results. In Chapter 6 the method from Chapter 4 is used to measure the spectra as a function of the event multiplicity. Chapter 7 is the concluding chapter which also contains a brief discussion of possible new measurements.

The work in this thesis was presented at two international conferences [1, 2].

Chapter 1

Theoretical background and previous experimental results

1.1 Introduction

In this chapter our theoretical understanding of some of the basic properties of proton-proton collisions is discussed, in particular the production of low energy pions, kaons, protons and antiprotons. Some models which describe heavy–ion collisions are presented because they can also be applied to proton-proton collisions. This chapter is organized as follows. Section 1.2 introduces basics properties of Quantum Chromodynamics and the proton structure. In Section 1.3 a short description of the proton-proton collision modelling is presented. In Section 1.4 some of phenomenological models used to describe the particle production in proton-proton and heavy–ion collisions are introduced. It also contains a brief discussion of the possible collective behaviour of the system produced in proton-proton collisions. Section 1.5 presents open issues for the LHC data concerning the particle productions.

1.2 Quantum Chromodynamics (QCD)

Quantum Chromodynamics (QCD) is the part of the Standard Model, the currently accepted theory of particle physics, describing the strong interaction between quarks and gluons (partons). Gluons are bosons which carry the strong interaction between quarks. There are six known types of quarks, listed in Table 1.1. Each quark has one out of three possible QCD charges. The QCD charge is labelled using colour: red (r), green (g) and blue (b). Antiquarks have an anticolour. The interaction in QCD is invariant under a SU(3) transformation in colour space (SU(3) colour symmetry). To each gluon a combination of colour and anticolour is assigned. There exists nine of such combinations. They are divided into a singlet and an octet state according to the properties of the SU(3) group symmetry. Experimental results show that gluons belong to the octet state.

Table 1.1: Classification of quarks

name	mass	electric charge in e units
down (d)	4.1-5.7 MeV/c^2	-1/3
up (u)	1.7-3.1 MeV/c^2	2/3
strange (s)	$\sim$ 100 MeV/c^2	-1/3
charm (c)	$\sim$ 1.29 GeV/c^2	2/3
bottom (b)	$\sim$ 4.2 GeV/c^2	-1/3
top (t)	$\sim$ 172.9 GeV/c^2	2/3

The strength of the interaction between quarks and gluons is determined by the strong coupling constant (α_s), which is the equivalent of the fine structure constant in Quantum Electrodynamics. As a consequence of the renormalization procedure the value of α_s depends on the amount of exchanged four-momentum between the interacting partons [3] (Equation 1.1):

$$\alpha_s(|Q^2|) = \frac{12\pi}{(11n - 2f)\ln(|Q^2|/\Lambda^2)}, \tag{1.1}$$

where $|Q^2|$ is the square of the exchanged four-momentum (energy scale), n the number of colours in QCD (equal to 3), f the number of quark flavours (equal to 6) and $\Lambda_{QCD} \sim 300$ MeV/c is a constant calculated from experimental data. A key property of QCD is that $11n - 2f > 0$. As a consequence, α_s decreases with increasing energy scale (decreasing distance). This is called asymptotic freedom.

When a cross-section in QCD is calculated based on Feynman diagrams, contributions from higher order diagrams come with higher powers of α_s. Such series can only converge if α_s is less than one. This is the case for large values of $|Q^2|$. In this regime, QCD processes can be calculated using perturbative techniques.

If $|Q^2|$ is low then α_s is large, meaning that the perturbative approach cannot be used. In such cases QCD properties can still be calculated using Lattice QCD. This regime of QCD is often called the soft regime.

An important property of QCD is colour confinement, which means that a colour charged particle cannot be isolated. As a result all hadrons (stable particles consisting quarks) are colourless. Hadrons are built from three quarks (antiquarks) with different colours or from a quark-antiquark pair in a colour singlet state. The first type of hadron is called a baryon while the second type is called a meson. Examples of hadrons can be found in Table 1.2.

In the full theory, the picture is more complicated. Quarks and antiquarks inside a hadron interact all the time producing quark-antiquark pairs and gluons. The parton content of a hadron is described using the parton density functions (PDF) $f_a(x_a, Q^2)$ which are the probabilities of finding a parton a with x_a fraction of the total longitudinal momentum of the hadron during a collision with four-momentum transfer Q^2. At any moment inside a baryon there are three more quarks than antiquarks and inside a meson the number of quarks is equal to the number of antiquarks.

Table 1.2: Properties of hadrons

name	quark content	charge in e units	mass [MeV/c^2]	cτ [m]
π^+ π^-	ud, d$\bar{u}$	+1,-1	139.569	7.8045
K^+ K^-	u$\bar{s}$, s$\bar{u}$	+1,-1	493.67	3.713
p	uud	+1	938.280	∞

1.3 Modelling of proton-proton collisions with event generators

Event generators are often used to describe proton-proton collisions. These event generators are based on perturbative QCD for the hard scattering and use phenomenological models to describe the soft processes and many-body dynamics that are present in proton-proton collisions. To perform analysis presented in this thesis two event generators were used: PYTHIA [6] and PHOJET [12, 13]. Both of them use the Lund string fragmentation model (see Section 1.3.2) for hadronization, process responsible for the particle production.

1.3.1 Classification of collisions and their mechanism

From the experimental point of view proton-proton collisions are divided into four categories: elastic (EL), single diffractive (SD), double diffractive (DD) and non-diffractive (ND) based on the rapidity[1] (y) distribution of the products. This division is illustrated in Figure 1.1 where a rough sketch of rapidity distributions of the products for each type of collisions is presented, lack of any rapidity gap in the distribution of the products is characteristic of the inelastic collisions. This division is also motivated by the collision mechanism, during an elastic collision no new particles are produced. In case of a diffractive collision at least one proton produces other particles but the protons do not exchange any quantum number (colour, flavour) while in a non-diffractive collision the protons exchange quantum numbers and produce new particles. A diffractive collision can be interpreted as an exchanges of a pomeron. The pomeron is a special case of the reggeon. This is a concept coming from the Regge theory, which was developed before QCD to explain proton-proton interactions. It assumes that the interactions can be explained as the exchange of a particle (i.e. reggeon) which fulfils the following rule [4]:

$$J = \alpha_0 + \alpha^{'} M_j^2, \qquad (1.2)$$

where J is the spin and M_J is the mass of the reggeon. In case of the pomeron $\alpha_0 = 1$. The exchange of a pomeron in QCD can be interpreted as a multi-gluon exchange.

Non-diffractive collisions are more complex. Figure 1.2 shows a schematic picture of the collision process, which can be divided into the following steps:

- initial state showers of partons inside the protons,

[1]for rapidity y and pseudorapidity η definitions see Appendix B

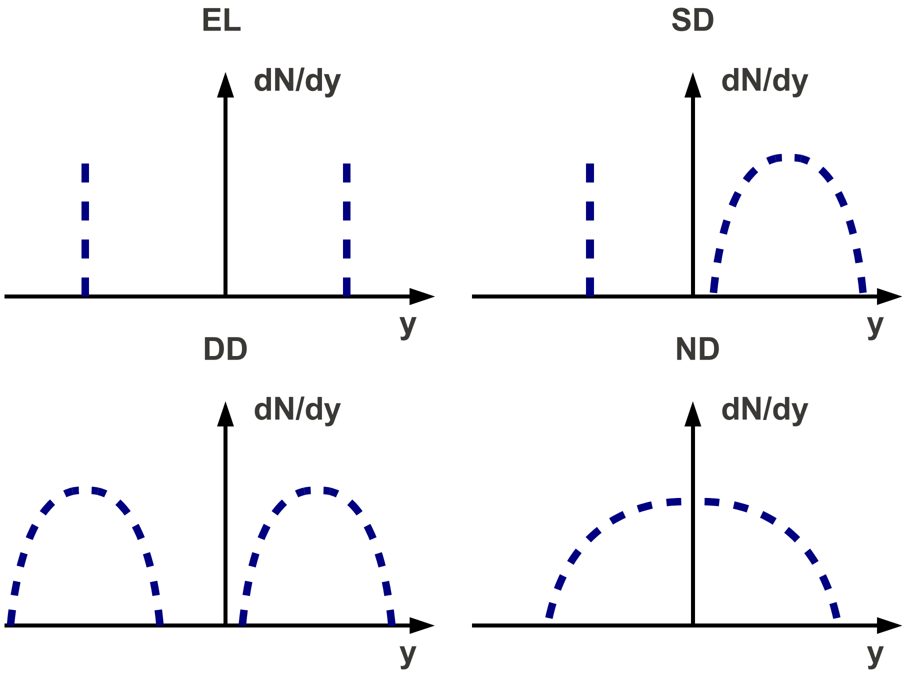


Figure 1.1: The rapidity distributions of the products for four types of proton-proton collisions: elastic (EL), single diffractive (SD), double diffractive (DD) and non-diffractive (ND)

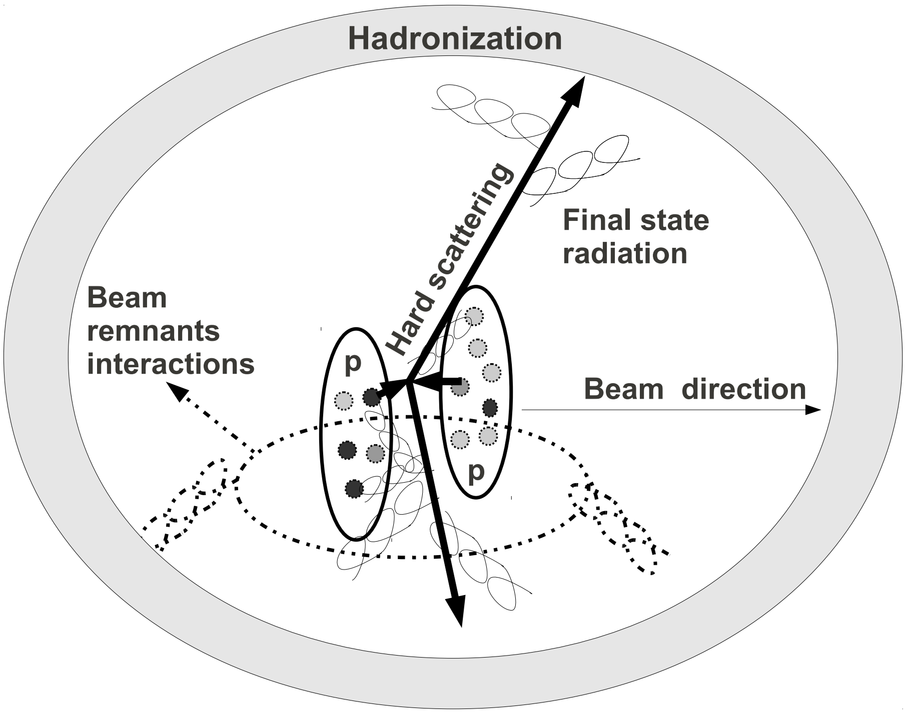


Figure 1.2: Schematic view of the non-diffractive proton-proton collision.

- hard collisions between two partons (high $|Q^2|$), involving the exchange of quantum numbers (this can happen more than once),
- radiations of partons produced in hard collisions (parton showers),
- interactions of remnants of the incoming protons (partons which do not take a part in the hard scattering),
- hadronization of all partons produced during the previous steps.

Production of pions, kaons, protons and antiprotons is part of the hadronization. It can be described using different approaches: independent hadronization, Lund string fragmentation model and a cluster model.

The production of hadrons with large transverse momentum (p_t) with respect to the proton beam axis is influenced by final state radiation of partons produced in hard scattering. This can be described by fragmentation functions [5], which give the probabilities that the high energy parton hadronizes to some hadron. Those functions are based on the perturbative calculations.

The production of hadrons with low p_t, which is a subject of the measurements in this thesis, is influenced by particle production in multiple interactions of the beam remnants and soft products of parton showers. In [6], it is stated that *"In fact, in the full event generation process, probably no other area is as poorly understood as this one"*. This is why experiments have to provide good quality measurements to constrain this part of QCD.

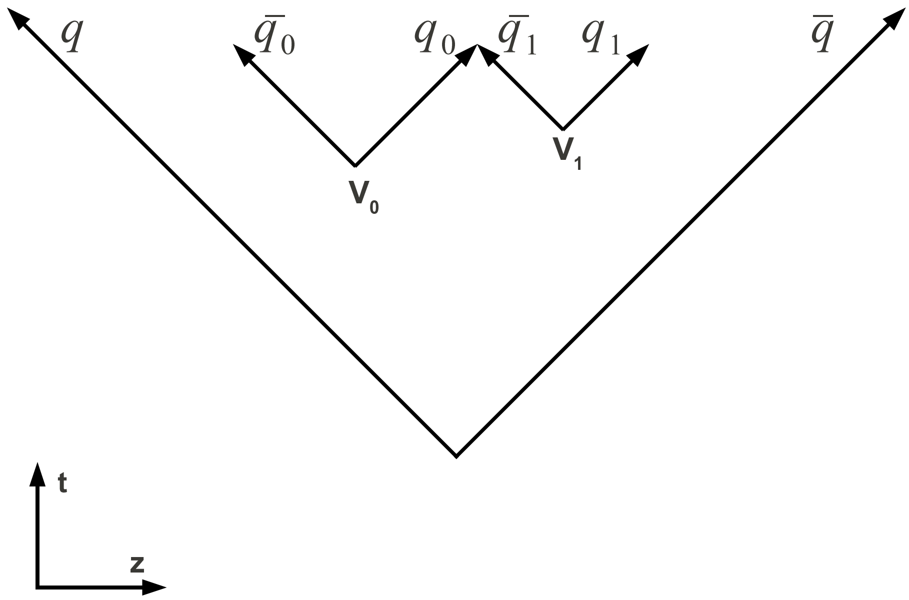


Figure 1.3: Quark and antiquark space-time trajectories in the Lund string fragmentation model. The two vertices (V_0 an V_1), where new quark-anitquark pairs are produced, are shown.

1.3.2 Lund string fragmentation model

In this section the Lund string fragmentation model is briefly discussed. For more details see [6, 7]. The main idea of the model is a creation of a colour field between a quark and an antiquark when particles move in opposite directions. This field can be approximated by a field with a linear potential (colour string).

Meson production

Quarks from the vacuum state can tunnel via the linear potential creating quark-antiquark pairs ($q_i\bar{q}_i$) at various vertices as shown in Figure 1.3. The tunnelling probability is given by $\exp(-\pi m_{\mathrm{t}}^2/\kappa)$ where $\kappa \approx 1$ GeV/fm and $m_t = \sqrt{p_{\mathrm{t}}^2 + m^2}$, where m is the mass of the pair. The generation of p_{t} (here momentum perpendicular to the axis of the initial pair) of the produced partons is done using a Gaussian distribution. The tunnelling mechanism suppresses the production of heavy quarks due to dependence of the probability on the mass (ratio of the probabilities to produce quarks is $u : d : s : c \approx 1 : 1 : 0.3 : 10^{-11}$). The pair masses are not precisely known, so that the suppression of the strange quark production is a model parameter (set to 0.3 by default in PYTHIA). A quark can join with an antiquark from the neighbouring vertex and form a meson. The created meson will pick up fraction z of the available $E + p_z$. The distribution of z is given by the Lund symmetric fragmentation function:

$$f(z) \propto \frac{1}{z} z^{a_\alpha} \left(\frac{1-z}{z} \right)^{a_\beta} \exp\left(\frac{-bm_{\mathrm{t}}^2}{z} \right) \tag{1.3}$$

where a_α, a_β, b are parameters. This function is a consequence of the assumption that there is no time ordering of the vertices [4]. The produced mesons have one out of six combinations of spin and internal orbital angular momentum. The probabilities to obtain each of these are model parameters.

Baryon production

To produce baryons in the Lund string fragmentation model, two mechanisms are used: a diquark production model and the popcorn model. The main assumption of the diquark production model is that instead of a quark (antiquark) a pair of antiquarks (quarks) in a colour triplet state is produced. This pair can connect with the neighbouring parton and form a baryon. This mechanism is based on the observation that the colour triplet states of two antiquarks (quarks) in the colour field can be considered as a quark (antiquark) [7]. The probability to create a diquark pair rather than an antiquark, the extra suppression factor for production of pairs with a strange quark and the relative contributions of the spin states of the pairs are input parameters of the diquark production model.

The popcorn model uses fluctuations of the colour field to produce quark-antiquark pairs with a different colour than the initial pair. For example if the initial pair is red-antired the standard fragmentation produces red-antired pairs, but the fluctuations can produce a green-antigreen pair. In this case both quarks (antiquarks) move towards each other (they are in a colour triplet state), producing a colour field with a blue-antiblue configuration. This string fragments as for a blue-antiblue initial pair, producing a baryon-meson -...-meson-antibaryon system. In most cases baryon-antibaryon and baryon-meson-antibaryon systems are produced (the relative contribution is a model parameter), since including the production of an additional meson introduces an extra suppression due to the increase of the transverse mass of the system. The strangeness production is also a model parameter.

Mulitparton interactions

In a full simulation of proton-proton collisions the situation is more complex because a multiparton system is produced. Strings are created by quark-gluon-..-antiquark systems. The breakup of the string is done by a breakup of each quark-gluon piece according to the standard procedure but additional kinematic conditions are introduced to ensure that the produced hadrons are on mass shell. In general the whole event is divided into groups of partons which are in a colour singlet state.

1.3.3 Event generators

In PYTHIA the total proton-proton cross-section is calculated using the Donnachie and Landshoff parametrization [9]. This cross-section is divided into EL, SD, DD and ND parts based on the Schuler and Sjostrand model [10, 11]. The number of initial hard parton-parton collisions is calculated using a Poisson distribution with mean $\sigma_{hard}(p_{\mathrm{t}}^{min})/\sigma_{ND}(s)$, where $\sigma_{hard}(p_{\mathrm{t}}^{min})$ is a hard cross-section for the parton scattering calculated with a cut-off p_{t}^{min} to avoid divergences. Partons from the hard collisions produce parton showers, which hadronizes using the Lund string fragmentation model. After the hard collisions, depending on what kind of parton took part in the hard collision (quark, antiquark or gluon), a hadron can be emitted from the proton remnant and the rest of the remnant form the multiparton system which hadronizes using the Lund string fragmentation model.

In PHOJET the total cross-section is divided into a soft and a hard part based on the momentum of the partons (around 3 GeV/c) involved in the interaction. The hard part is calculated using perturbative techniques, while the soft part is calculated using the pomeron picture. The shower evolution and hadronization is based on this same model as PYTHIA.

1.4 Phenomenological models

In this section some phenomenological models which describe the particle production in proton-proton and heavy-ion collisions are presented. Some of them only describe the total particle production (e.g. thermal model) using a simple picture based on the statistical physics. Others which have stronger assumptions like: temperature fluctuations (non-extensive thermodynamics) or a collective behaviour of the system (Blast Wave model), give a description of the shapes of the spectra.

1.4.1 Thermal model

Hadron production in proton-proton collisions can be described as a thermal production [14]. This approach assumes that the hadronization takes place from fireballs[2] according to a canonical ensemble. The quantum numbers of the fireballs are constrained by the initial conditions. In this approach the total yields are described by three parameters: temperature T, volume V and the strangeness suppression factor (γ_s), which is a phenomenological parameter introduced to reduce the phase space for the strange quark production. The volume in this model does not have a clear physical interpretation, unless the yields for the entire phase space (whole azimuthal angle and rapidity) are used. Results of the thermal model fit to the yields in proton-proton collisions at $\sqrt{s} = 200$ GeV reported by the STAR experiment are shown in Figure 1.4. The description of the yields by the thermal model is in good agreement with data but this does not imply a thermalization on an event-by-event basis, rather suggests that statistical emission is a property of the hadronization process. The obtained $T = 170$ MeV does not strongly depend on the collision energy.

1.4.2 Spectra in non-extensive thermodynamics

For a long time the transverse momentum distributions ($p_{\rm t}$ spectra) of the particles produced in heavy–ion and proton-proton collisions have been described using the function $\mathrm{d}N/\mathrm{d}p_{\rm t} \propto p_{\rm t} \exp(-m_{\rm t}/T)$, which is characteristic for an adiabatic expansion of a fireball at temperature T [16]. The increase of the energy of colliding protons has shown that this function cannot describe well the high $p_{\rm t}$ part of the particle spectra, which becomes a power-law (e.g. spectra in proton-proton collisions at $\sqrt{s} = 200$ GeV [17, 18]). The power-law part is interpreted

[2]regions of pre-hadronic strongly interacting matter

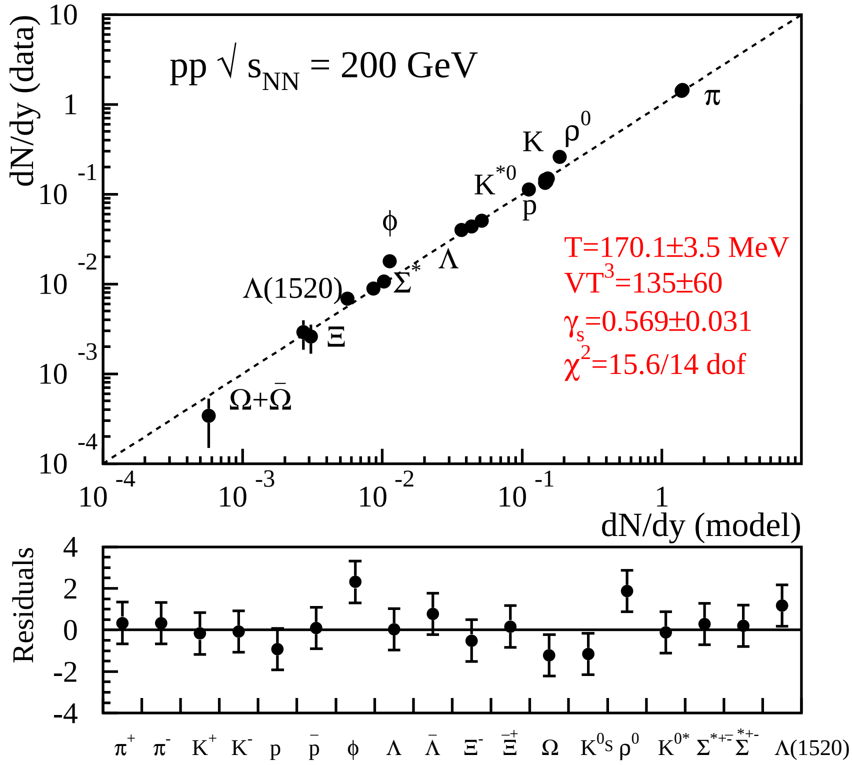


Figure 1.4: Result of the thermal model fit to the measurement of the particle productions made by the STAR experiment [14].

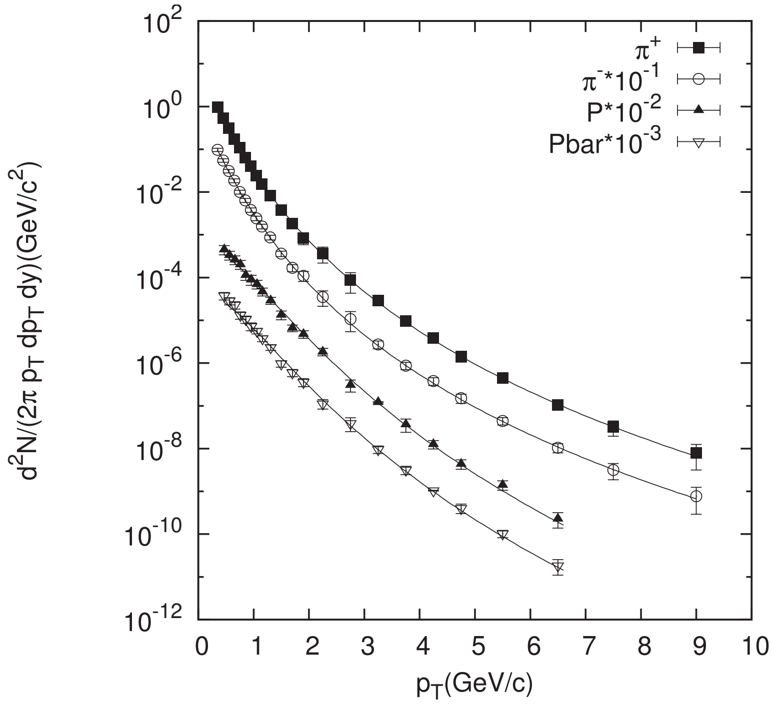


Figure 1.5: The spectra of pions, protons and antiprotons in proton-proton collisions at $\sqrt{s} = 200$ GeV fitted with functions (solid curves) based on the non-extensive approach [18].

as the region where hard processes are dominant. A function which can describe such spectra in the whole $p_{\rm t}$ range can be derived using the idea of the non-extensive thermodynamics (see Figure 1.5). Non-extensive thermodynamics is defined based on the Tsallis entropy [19] instead of the Boltzmann-Gibbs (BG) one. This entropy is defined as:

$$S_q = -\frac{1}{1-q}\left(1 - \sum_i p_i^q\right). \tag{1.4}$$

where q is the index of the non-extensivity of the system and p_i is the probability that the system is in configuration i. The Tsallis entropy is non-extensive because the entropy for sum of two independent systems A,B is not simply the sum of the system entropies:

$$S_q(A+B) = S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B). \tag{1.5}$$

If $q = 1$ the Tsallis entropy is the normal BG entropy:

$$q = 1 \quad => \quad S = -\sum_i p_i \ln p_i. \tag{1.6}$$

Based on information theory it can be shown that:

$$p_i \propto \left(1 - \frac{(1-q)}{T}E_i\right)^{1/(1-q)}, \tag{1.7}$$

where E_i is an energy of the system and T is temperature. Using the method presented in [20] it can be shown that the $p_{\rm t}$ spectra of particles are given by:

$$\frac{{\rm d}N(p_{\rm t})}{{\rm d}p_{\rm t}} \propto p_{\rm t} \int\limits_0^{p_{\rm z}^{max}} \left(1 + \frac{(q-1)}{T}\sqrt{m_{\rm t}^2 + p_{\rm z}^2}\right)^{\frac{1}{1-q}} {\rm d}p_{\rm z}. \tag{1.8}$$

Using $p_{\rm z} = m_{\rm t}\sinh(y)$ gives:

$$\frac{{\rm d}N(p_{\rm t})}{{\rm d}p_{\rm t}} \propto p_{\rm t} m_{\rm t} \int\limits_0^{y^{max}} \left(1 + \frac{q-1}{T} m_{\rm t}\sqrt{1 + \sinh^2(y)}\right)^{\frac{1}{1-q}} \cosh(y){\rm d}y. \tag{1.9}$$

Equation 1.9 simplifies to:

$$\frac{{\rm d}N(p_{\rm t})}{{\rm d}p_{\rm t}} \propto p_{\rm t} m_{\rm t}\left(1 + \frac{q-1}{T}m_{\rm t}\right)^{\frac{1}{1-q}}, \tag{1.10}$$

for $y = 0$. In the literature [17] a function (called Lévy-Tsallis) of the form:

$$\frac{{\rm d}N(p_{\rm t})}{{\rm d}p_{\rm t}} \propto p_{\rm t}\left(1 + \frac{q-1}{T}m_{\rm t}\right)^{\frac{1}{1-q}} \tag{1.11}$$

can be found. This form is equal to the 1.10 in the approximation in which thermodynamics is applied only in the transverse plane.

In [21] a possible interpretation of the parameters is given. It assumes that the system can be divided into smaller regions with temperatures which fluctuate around some common value (T_0), and those regions exchange energy. The relation between the q parameter of the p_{t} spectra and the fluctuations is:

$$q = 1 + \frac{\mathrm{Var}(T)}{\langle T \rangle^2} \tag{1.12}$$

This is an interesting interpretation, but still it is not clear how the q parameter should be interpreted in case of proton-proton collisions.

1.4.3 Collective behaviour in proton-proton collisions

An increase of the energy of proton-proton collisions provides the possibility to observe collisions with the multiplicity (total number of produced particles) comparable with a multiplicity in heavy–ion collisions at lower energies. Those events may exhibit collective behaviour of the created system. Here the collective behaviour of the system can be understood as a situation in which system is described using quantities like a density or a flow. Already some indications of the collective behaviour in proton-proton collisions are seen at $\mathrm{d}N/\mathrm{d}\eta > 6$ (multiplicity at mid-rapidity) [22] as hints of the phase transition. The evolution of the $\langle p_{\mathrm{t}} \rangle$ of identified particles as function of the multiplicity at mid-rapidity can add additional information. The shape of the spectra can provide evidence for collective behaviour if for an example it can be described using a Blast Wave model.

1.4.4 The Blast Wave model

The Blast Wave model describes the shape of the p_{t} spectra in heavy–ion collisions. It assumes a collective radial flow which modifies the thermal emission of the hadrons. The hadron emission takes place during the freeze-out described by the freeze-out hyper-surface [23]. In this model the shape of the spectra is given by the following equation:

$$\frac{\mathrm{d}N(p_{\mathrm{t}})}{\mathrm{d}p_{\mathrm{t}}} \propto p_{\mathrm{t}} \cdot \int_0^R r\mathrm{d}r m_{\mathrm{t}} I_0 \left(\frac{p_{\mathrm{t}} \cdot \sinh(\rho)}{T_{kin}} \right) K_1 \left(\frac{m_{\mathrm{t}} \cdot \cosh(\rho)}{T_{kin}} \right) \tag{1.13}$$

where $\rho = \tanh^{-1}(\beta)$, I_0 and K_1 are Bessel functions and β is the radial flow velocity. The radial velocity profile is parametrized as:

$$\beta = \beta_s (r/R)^n \tag{1.14}$$

so that $\langle \beta \rangle = 2 \cdot \beta_s/(n+2)$. The Blast Wave form derived using non-extensive thermodynamics has also been used to fit p_{t} spectra [24]:

$$\frac{\mathrm{d}N(p_{\mathrm{t}})}{\mathrm{d}p_{\mathrm{t}}} \propto p_{\mathrm{t}} m_{\mathrm{t}} \cdot \int_{-Y}^{Y} \cosh(y)\mathrm{d}y \int_{-\pi}^{\pi} \mathrm{d}\phi \int_0^R r\mathrm{d}r \left(1 + \frac{q-1}{T} \left(m_{\mathrm{t}} \cosh(y) \cosh(\rho) - p_{\mathrm{t}} \sinh(\rho) \cos(\phi) \right) \right)^{\frac{1}{1-q}} . \tag{1.15}$$

1.5 Increasing the energy of proton-proton collisions. What can we learn?

The increase of the energy of colliding protons should increase our understanding of the created system during the collisions and further test the existing models: thermal productions and non-extensive thermodynamics. It can also show the collective behaviour of the system. For sure it will deliver new constrains for the event generators.

Chapter 2

Experiment set-up and offline framework

2.1 Introduction

In this chapter the ALICE (A Large Ion Collider Experiment) detector is presented. Three sub-detectors: the Inner Tracking System (ITS), the Time Projection Chamber (TPC) and the Time-of-Flight (TOF) are described. Those sub-detectors are responsible for particle identification (PID) in ALICE, and were used to perform the analyses presented in this thesis. Descriptions of trigger detector VZERO, the ALICE Data Acquisition (DAQ) system and offline framework are also included.

2.2 ALICE detector

2.2.1 Overview

ALICE [25] is one of the four main experiments at the Large Hadron Collider (LHC) located at CERN. The ALICE detector was built to study the properties of heavy-ion collisions but it is also used to study the physics of proton-proton interactions. It is optimized for precise tracking over a broad transverse momentum region (from 100 MeV/c to 100 GeV/c) in a high multiplicity environment, good resolution of primary vertex reconstruction and particle identification over a broad momentum region. The detector set-up is shown in Figure 2.1.

ALICE consists of two main parts: a central barrel and a large muon arm with its own dipole magnet. It also has forward detectors (ZDC, PMD, FMD, T0, VZERO) which are used as trigger detectors and in multiplicity measurements. The central barrel has an acceptance of $|\eta| < 0.9$ and provides full azimuthal coverage. It consists of the following parts:

- the L3 magnet which produces a 0.5 T field parallel to the beam axis,
- the Inner Tracking System (ITS),
- the Time Projection Chamber (TPC),

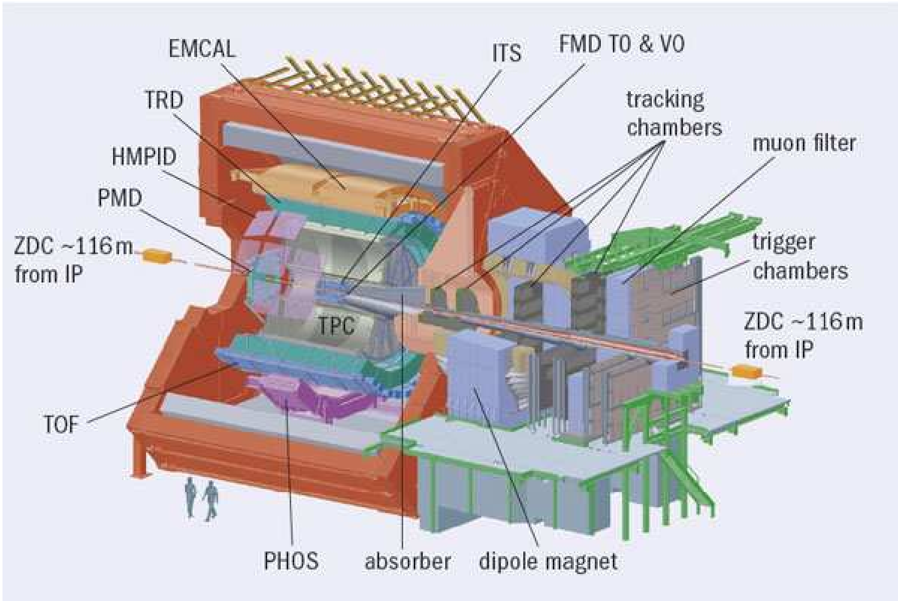


Figure 2.1: ALICE detector. All sub-detector are indicated.

- the Transition Radiation Detector (TRD),
- the Time-of-Flight detector (TOF),
- the High-Momentum Particle Identification Detector (HMPID) which is a Ring Imaging Cherenkov detector,
- two electromagnetic calorimeters PHOS and EMCAL.

The ALICE coordinate system is defined as follows: the x-axis is perpendicular to the beam direction, aligned with the local horizontal direction and points to the accelerator centre; the y-axis is perpendicular to the x-axis and to the beam direction, points upward; the z-axis is parallel to the beam direction and points in the opposite direction to the muon arm. The point of origin is at the interaction point [25].

2.2.2 ITS

The ITS detector [25, 26] shown in Figure 2.2, is the part of ALICE placed closest to the interaction point. It consists of three sub-detectors. Each sub-detector contains two concentric silicon layers. The innermost part of the ITS is the Silicon Pixel Detector (SPD), the middle part is the Silicon Drift Detector (SDD) and the outer one is the Silicon Strip Detector (SSD) (the SSD will be described in the next chapter where its calibration is also presented). The ITS provides high resolution tracking for the identification of secondary vertices of heavy flavour weak decays and improves momentum resolution at high p_{t}. The ITS also gives the possibility to reconstruct tracks produced by particles with low p_{t} which do not reach the TPC. The ITS has particle identification

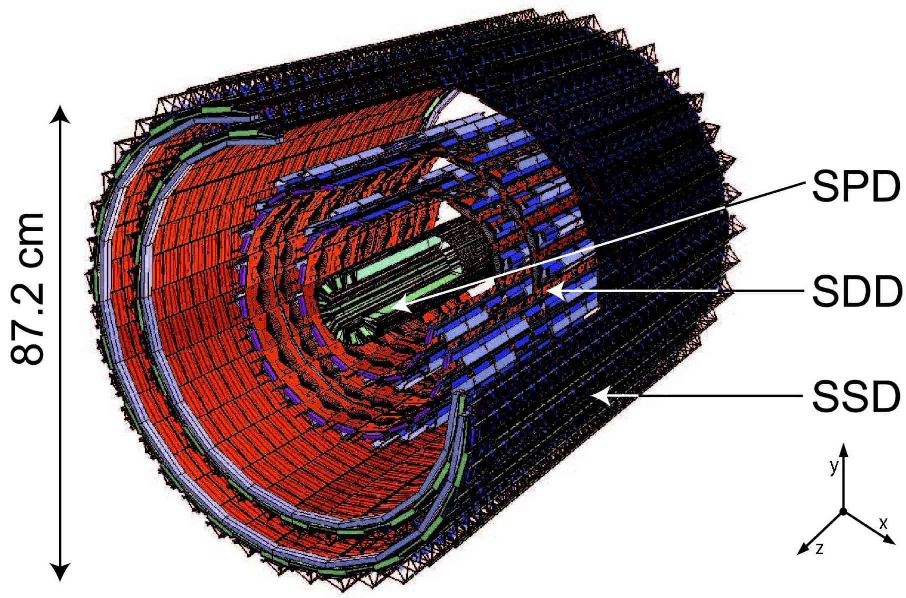


Figure 2.2: ITS detector.

capabilities based on the measurement of the particle energy loss in 4 layers (SSD and SDD). The main parameters of the ITS are summarized in Table 2.1.

Table 2.1: Main parameters of the ITS.

parameter	SPD	SDD	SSD
radius r (cm)	3.9 and 7.6	15.0 and 23.9	38.0 and 43.0
length $\pm$ z (cm)	14.1 and 14.1	22.2 and 29.7	43.1 and 48.9
Number of channels	9.8 M	133k	2.6M
spatial precision $r\phi$ (μm)	12	35	20
spatial precision z (μm)	100	25	830

2.2.3 TPC

The TPC [25, 27] is the main tracking detector of ALICE . Combined with the ITS, it provides precise momentum reconstruction, track separation, particle identification and vertex reconstruction. Figure 2.3 shows a schematic picture of the TPC. The TPC consists of an active gas volume, a central electrode, readout-chambers and two potential dividers, which create a symmetric electric field inside the gas volume. A charged particle which goes through the gas ionizes it. The produced electrons are driven by the electric field to the read-out chambers placed on the edges. The read-out chambers are multi-wire proportional chambers, where the amplification of the charge takes place. The

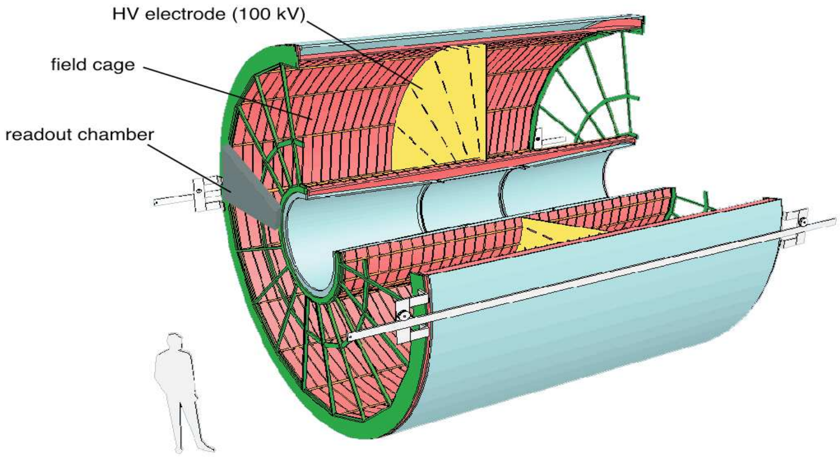


Figure 2.3: TPC detector.

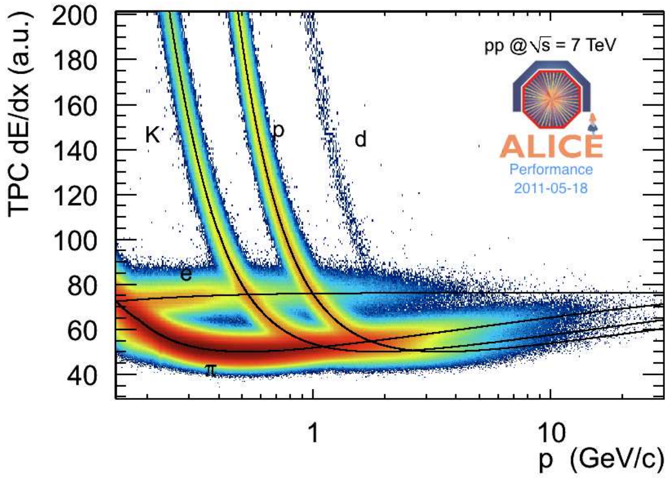


Figure 2.4: PID capabilities of the TPC. The value of the detected signal divided by the length of the track ($\mathrm{d}E/\mathrm{d}x$) as a function of the track momentum (p) is plotted.

amplified charge is read out by a grid of pads. Positions of signals on the pads (clusters) give the position of the particle in the xy plane, while the difference between the trigger time and the arrival time of the signal to the pads gives the z coordinate. The amplitude of the signal is used for the particle identification. A particle originating from the collision point can produce up to 159 clusters in the TPC. The PID capabilities of the TPC are shown in Figure 2.4, where the value of the detected signal divided by the length of the track as a function of the track momentum is plotted. The basic parameters of the TPC are summarized Table 2.2.

Table 2.2: Parameters of the TPC.

parameter	value
pseudo-rapidity coverage	$\|\eta\| < 0.9$ for full radial track length $\|\eta\| < 1.5$ for 1/3 full radial track length
Azimuthal coverage	360°
active volume radial position	$84.8 < r < 246.6$ (cm)
active volume length	2x2.5 m
segmentation in ϕ	18 sectors
segmentation in r	2 chambers per sector
total number of rows in the read-out chambers	159
maximal drift time	90 μs
Trigger rate limits in proton-proton collisions	1 kHz

2.2.4 TOF

The TOF detector [25, 28, 29] is built using Multi-Gap Resistive Plate Chambers. This technology provides a good time resolution and gives the possibility to operate at atmospheric pressure. Particle identification in the TOF is performed by measuring the time difference between the collisions time provided by the T0 detector and the time when a particle reaches the TOF. The resolution of such measurement depends on the resolutions of both times and it is around 80 ps. For proton-proton collisions it can happen that the T0 detector does not provide this information because no charged particle reaches it. In this case, the collision time is calculated using reconstructed tracks, which causes an increase of the resolution up to 120 ps. Figure 2.5 shows the particle identification capabilities of the TOF, where the measured $\beta = v/c$ as a function of track momentum divided by a track charge (p/z) is plotted.

2.2.5 VZERO

The VZERO detector [25, 30] consists of two scintillator arrays: the VZERO–A at z=340 cm and the VZERO–C at z=-90 cm. They cover $\eta \in$[2.8;5.1] (VZERO–A) and $\eta \in$[-3.7,-1.7] (VZERO–C). The VZERO is used as a trigger detector. It provides a trigger signal after detecting a particle coming for the collisions. The VZERO is also used for the centrality determination in heavy-ion collisions.

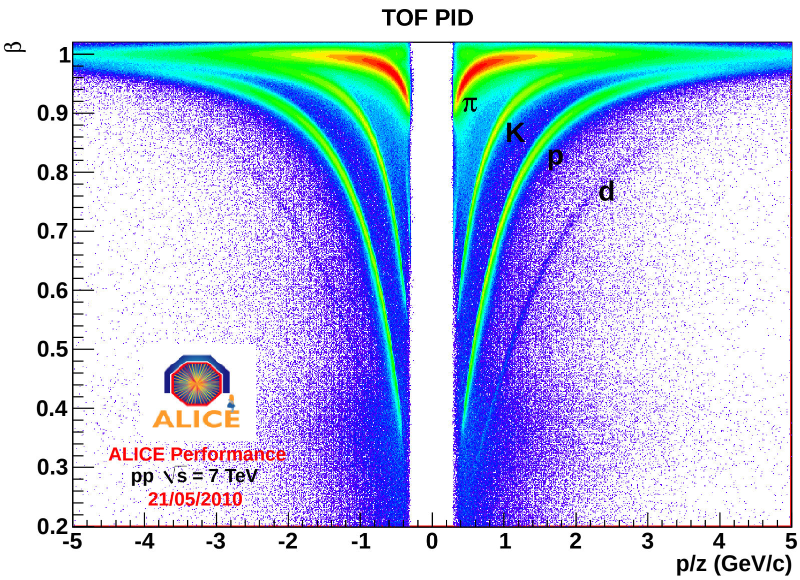


Figure 2.5: PID capabilities of the TOF. The $\beta = v/c$ as a function of a track momentum divided by a track charge (p/z) is plotted.

2.3 DAQ system

Each sub-detector of ALICE has a specially designed Front-End Read-Out (FERO) system [31], which has a standard interface of Detector Data Links (DDL). The DDL connects the FERO system to the Read-Out Receiver Card (D-RORC). Each D-RORC is connected to a Local Data Concentrator (LDC), where one LDC can have a few D-RORCs connected. When the sub-detectors receive a trigger signal they send data to the LDC. The LDC collects data from all connected D-RORCs and produces a sub-event. Sub-events are sent via the network infrastructure to the Global Data Collector (GDC). The GDC uses the trigger information stored inside the sub-event to build the whole event which is then sent to the external storage system for reconstruction.
The DAQ system is also responsible for handling the online calibration. During a calibration run, data are read out from one sub-detector. They are then analysed using a detector algorithm (calibration procedure), which can produce two kinds of output: online and offline configuration. The online configuration is stored in the part of the DAQ system called Detector Database and is loaded to the sub-detector FERO system before data taking. The offline configuration is a description of the sub-detector state, which is used during reconstruction and simulation, it is stored in the Offline Condition Data Base (OCDB).

2.4 Software tools

A special software framework (ROOT) [32] was developed to analyse data collected by high-energy experiments. It is based on the C++ computer language.

It provides a large variety of general tools to perform data analysis (e.g. packages for fitting function to data points). Those tools were used to perform all data analysis for this thesis. The software package used by the ALICE experiment (AliRoot) is based on ROOT. It consists of the code for data reconstruction, simulation of the detector performance and basic tools for the analysis of data collected by ALICE. It defines the format in which the ALICE data are stored. In the next sections the basics of AliRoot will be presented.

2.4.1 Simulation

During data analysis the performance of the detector and the reconstruction algorithms are checked using simulation. The first step of the simulation is the generation of a physics event. This is done using an event generator (e.g. PHOJET, PYTHIA) or completely artificial events designed by users. The GEANT3 [33] package is used to simulate the interactions of particles with the detector material. The results of those simulations are used to generate the response of the sub-detectors (e.g. signal in the TPC) in a format consistent with the data recorded during real data taking. Those data are reconstructed in the same way as real data recorded by the detector.

2.4.2 Reconstruction

The data reconstruction is divided into two steps: local and global. The local reconstruction is done individually for each sub-detector using only data collected by this sub-detector. An example of local reconstruction is the cluster finding in the SSD, which will be presented in the next chapter. The global reconstruction combines information from all sub-detectors (i.e. results of local reconstructions) and reconstructs the primary vertex (collision point) and the tracks.

The first step of the global reconstruction is finding the primary vertex using the SPD clusters. This is done by connecting reconstructed clusters from two layers of SPD to build tracklets. Many of the tracklets are constructed from two clusters produced by different particles. Each tracklet is propagated to the beam line. The crossing points of the tracklets originating from real particles coming from the primary vertex with the beam line are very close. Those crossing points are used to estimate the position of the SPD primary vertex.

The track reconstruction is performed using the Kalman filter approach. This approach was developed to measure dynamic system parameters or predict the system behaviour. It was adopted for reconstruction of tracks in high-energy experiments by changing the time steps to space points [34].

The track finding procedure starts with at the outer edge of the TPC where the density of clusters is small. The tracks are then propagated to the inner edge of the TPC matched with ITS clusters and propagated to the SPD. The next step is the back propagation from the SPD to the outer wall of the TPC, and further to the TRD and the TOF. The last step is a refit procedure in the TPC and the ITS, where the track parameters are updated. If the refit is successful a quality flag is set for the track. The tracks which are the result of this procedure are called global tracks. ITS clusters which are not matched with tracks in the TPC are used to form the ITS standalone tracks. During reconstruction of proton-proton collisions an independent ITS track finding procedure is run, which reconstructs ITS standalone tracks from all ITS clusters. In the end there are two samples of ITS standalone tracks, those obtained from all ITS clusters and those obtained from clusters not used by the global tracking. The reconstructed tracks (global and ITS standalone) are used to improve the reconstruction of the primary vertex. The improved vertex is called a global vertex. Each reconstructed track has a list of parameters (e.g. number of clusters in the TPC or the ITS, χ^2 per cluster), which describes its quality.

During simulation a generated particle can be connected with tracks. A track is connected with a generated particle if this particle created more than 80% of clusters which are used to reconstruct the track. If there is not such particle a track is connected with the particle that produced most of the clusters assigned to this track and the track is labelled as "fake".

2.4.3 Offline pile-up tagging procedure

An event which is a superposition of collisions is called pile-up. There are two kinds of pile-up: when collisions are from the same or from different bunch crossing. The probability to have a certain number of collisions during one bunch crossing can be expressed using the Poisson distribution. The mean value of this distribution (μ) can be calculated for each run using the number of bunch crossings without a collision. The rate of pile-up is defined as the ratio of the probability to have more than one collision to the probability to have at least one collision. This can be approximated by $\mu/2$.

A pile-up tagging algorithm is run after finding the SPD primary vertex. It checks if there is a point on the beam axis which is pointed by tracklets not used in the reconstruction of the SPD primary vertex. If such point is found the event is tagged as pile-up.

Chapter 3

Calibration of the SSD

3.1 Introduction

In the beginning of this chapter the SSD is described including software tools used in the simulation and reconstruction of the SSD data. In the second part of this chapter the procedure of the energy loss signal calibration of the SSD is presented and validated. The main goals of the SSD calibration are: provide matching between charges measured on both sides of the SSD modules, remove module to module fluctuations of the energy loss signal in the SSD and match the signal measured by the SSD with the SDD signal. That requires to define for each SSD module two calibration constants (gain factors, GF), one for each side. This approach assumes no variation of the gain factor on the chip or strip level. A linear dependence of ADC counts (strip signals) on the energy deposition in the SSD module is assumed in this procedure. Both assumptions are checked in this chapter.

3.2 The SSD

3.2.1 Principle of operation

A charged particle which passes through a silicon wafer produces electron-hole pairs. The number of pairs is proportional to the energy loss of the particle. An electric field is applied to the wafer, so electrons and holes are driven along the field lines to opposite sides of the wafer, where they are collected by strips. The strips are implants of n-type to collect electrons (N-side) and p-type to collect holes (P-side), as shown in Figure 3.1. The charge collected by the N-side strips is the same as the charge collected by the P-side strips. The N-side strips are not parallel to the P-side strips, so that the point where a particle crossed the wafer can be reconstructed from the crossing point of the strips which collected the charges.

3.2.2 The SSD set-up

The SSD sensor is a 300 μm thick silicon wafer with 768 parallel strips on each side. The active area of each sensor is 73x40 mm^2. It is surrounded by a bias

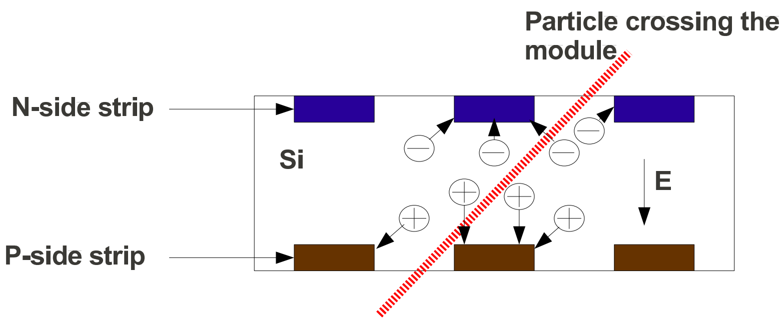


Figure 3.1: SSD operation principle (see Section 3.2.1)

ring, which creates an electric field inside the sensor. The angle between the N-side and P-side strips is 35 mrad, so that one N-side strip crosses 15 P-side strips and the other way around (strips on the edges of the sensor are shorter and they have fewer crossing points). Each side of the sensor is connected to a hybrid circuit with 6 HAL25 chips. The sensor and the chips create the SSD module. The HAL25 chip is connected to 128 consecutive strips from one side. It collects, amplifies, shapes and holds strip signals, which are proportional to the collected charge. The configuration of the chips is done using the JTAG interface. Part of the configuration is the adjustment of the shaping time between 1.4 μs to 2.2 μs. This time should match the time between a collision and read-out of the strip signals and it can be adjusted using the HOLD delay settings.

The modules are mounted on carbon fibre ladder structures. A layer 5 ladder has 22 modules while a layer 6 ladder has 25 modules. In total there are 34 ladders in layer 5 and 38 in layer 6, this gives 1698 modules in the whole SSD. Endcaps are mounted on each side of a ladder and they are connected with the HAL25 chips. They provide a read-out and control of the HAL25 chips. The number of chips connected to each endcap depends on the layer and side of the SSD.

Each endcap is connected to a read-out card (AD modules) using cables of approximately 40 m length. The AD modules digitize the strip signals. The read-out cards are placed in crates, which have 20 slots. Nine consecutive read-out cards are connected to a LINK card which has the DDL interface. Nine AD modules and one LINK card form one unit of the SSD FERO. The LINK card provides a connection to the DAQ system, the ALICE trigger system and the ALICE Detector control system. The SSD FERO performs the pedestal correction and the zero suppression (it is described in the next section) using values set before data taking. It also calculates online common mode (the part of noise common to all strips connected to the same HAL25 chip) and corrects strips signals for it, used values of the common mode are sent to the LDCs.

3.2.3 SSD operation

Before each physics run (recording of collisions) the SSD requires a calibration run, during which all strips are read out 500 times without zero suppression. The SSD detector algorithm calculates the pedestal, noise and common mode for each strip using the collected data. Based on those values the dynamic map of the dead channels in the SSD is updated in the OCDB. The updated dead channel map, noise and pedestal values are stored in the Detector Database. They are sent to the LINK cards before the physics run. Only if the value of a strip signal after pedestal and common mode correction is above threshold (by default three times the noise value measured during the calibration run) it is sent to the LDC during a physics run. The noise values from the calibration runs are also sent to the OCDB and are used during the reconstruction of the SSD data and simulation of the SSD.

3.2.4 SSD reconstruction

The goal of the SSD local reconstruction is to provide a list of points where charged particles went through the SSD modules (recpoints). As a first step, the signal is converted from ADC counts to keV. First the signal on each side is multiplied by a gain factor, different for the P-side and N-side strips (determination of the gain factors will be described in the next sections of this chapter). This corrects for fluctuations of the calibration factor at the module level and the difference between signal from N-side and P-side strips. Then the signals are multiplied by an overall calibration factor (0.72 keV/ADC). This factor was defined based on the beam test of the SSD modules. The next part of the local reconstruction procedure looks for groups of consecutive strips from one side of one module with the signal above noise threshold, those groups are called digits. Digits from both sides are connected to form recpoints. The connection is made by checking if N-side strips forming one digit cross with the P-side strips forming a digit on other side. The coordinates of a recpoint are calculated using the weighted mean of the coordinates of the strips, where the signal takes the role of weight. The charge of a recpoint is the mean value of the signal of the N-side and P-side strips. In the case when a digit from one side (e.g. N-side) can be connected with two digits from the other side (e.g. P-side), the charge of the recpoint is calculated using information from one side (e.g. P-side). This situation is rare in proton-proton collisions, due to the low density of particles. The same approach of calculating the cluster charge is also used when one side is labelled as dead. Besides the charge and the coordinates, the recpoint also stores its size (number of strips on each side) and relative difference between N-strip and P-strip signals. The whole SSD reconstruction is coded in AliRoot as the AliITSClusterFinderV2 class.

The list of the recpoints is sent to the track finding procedure. Normally the recpoint properties are not used by the track propagation procedure. The track finding procedure looks for recpoints, which can be assigned to the track, using only the predicted track shape. When the high density of the recpoints makes the propagation complicated the recpoint properties are used (e.g. size, charge). This situation is extremely rare in proton-proton collisions. Recpoints which are assigned to the track are transformed into track points. The track point stores less information than the recpoints (e.g. size is not stored), but it gives the possibility to calculate the track inclination inside the SSD module. This is important for the energy loss signal calibration of the SSD.

3.2.5 SSD simulation

In the simulation the value of the energy loss in the SSD module is calculated using GEANT3 . It is scaled to the strip signals using the conversion factor. The strip signals are then multiplied by the gain factors to introduce fluctuations of the signal at the module level while noise is added to each strip signal. Noise is calculated using information from the OCDB. The simulation assumes that the strip noise is described by a Gaussian distribution.

3.3 Calibration method

3.3.1 Charge matching calibration

Charge matching is the requirement that for each SSD cluster (recpoints or track points) the charge measured by the N-side strips is equal to the charge measured by the P-side strips. It is a consequence of the principles of the SSD operation presented in Section 3.2.1. To check the performance of the charge matching, the charge ratio (CR) is used:

$$CR = \frac{Q_N - Q_P}{Q_N + Q_P}, \tag{3.1}$$

where

Q_N is the charge measured by the N-side strips,

Q_P is the charge measured by the P-side strips,

$CR \in [-1; 1]$;

For a linear dependence of strip signals on the energy deposition the value of the CR should be zero and does not depend on the value of the measured charge. The charge matching part of the SSD calibration procedure can be performed based on the charges measured without correcting for the track inclination inside the module (recpoints charge), since this part of the calibration should not depend on the inclination. It should fulfil two requirements: the new mean value of the CR (calculated after the calibration) should be zero and the sum of the Q_N and Q_P should not change. In order to obtain the new gain factors, the old values of the gain factors (i.e. used in reconstruction of the calibration data sample) should be multiplied by:

$$\begin{aligned} C_P &= \frac{1}{1 - \langle CR \rangle}, \\ C_N &= \frac{1}{1 + \langle CR \rangle}, \end{aligned} \tag{3.2}$$

where $\langle CR \rangle$ is the average CR in a calibration data sample. Those corrections should be applied at the level on which $\langle CR \rangle$ was calculated (e.g. whole detector, module).

3.3.2 Calibration of the energy loss signal

For the absolute calibration (i.e. the part of the calibration responsible for reducing fluctuations and matching the signals between the SDD and SSD) the correction for the track inclination inside the SSD module is needed. This can only be calculated using track points. Knowledge of the track inclination gives the possibility to define the energy loss signal as the measured charge (energy loss) normalized to the module thickness (300μm). After applying the absolute calibration at the module level the precisely defined way of the measuring the energy loss signal for a module (reference value) should give value equal to a nominal value for all modules. In order to be able to combine the energy loss signals in the SDD and the SSD, common reference and nominal values have to be defined for both detectors.

The distribution of the energy loss in silicon for a charged particle can be described by the Landau distribution [35]. In reality, the noise induced by the experimental set-up (e.g. electronics) results into a distribution that can be better described by the convolution of Landau and the Gaussian distributions. (see Figure 3.3, where the distribution measured by the SSD is plotted.)

The Most Probable Value (MPV) of the energy loss in the material is used in the calibration. The dependence of the MPV in silicon on $\beta\gamma = p/m$ is shown in Figure 3.2, where functions for four track lengths in silicon are plotted. (The influence of this dependence on the calibration procedure will be discussed later.) When $\beta\gamma$ is larger than 2 the MPV for silicon does not depend strongly on the particle momentum, as can be seen in Figure 3.2. Based on that property it was decided by the SSD and SDD groups to use as the reference value of the energy loss signal for the SSD and the SDD, the MPV of the Landau distribution, obtained from the fit of the convolution of the Landau and Gaussian distributions to the energy loss signal distribution for the tracks reconstructed with momentum larger than 1.2 GeV/c. The groups also agreed that the nominal value (mpv_{NV}) is 84 keV/300μm, corresponding to the MPV in 300 μm of silicon for particle with $\beta\gamma > 10$. (It has to be mentioned that the numerical value of the nominal value is not important as long as it is the same for the SSD and the SDD because both detectors do not perform a precise measurement of the energy loss but use this measurement to do particle identification.) So if in a calibration data sample the reference value for some module is equal to mpv_i, the old gain factors for this module should be multiplied by mpv_{NV}/mpv_i to obtain the new gain factors. In summary the final values of the gain factors for each side of the SSD module are given by:

$$\begin{aligned} GF_{Nnew} &= GF_{Nold} \cdot C_N \cdot \frac{mpv_{NV}}{mpv_i}, \\ GF_{Pnew} &= GF_{Pold} \cdot C_P \cdot \frac{mpv_{NV}}{mpv_i}, \end{aligned} \tag{3.3}$$

where GF_{old} stands for the old gain factors used in the reconstruction of a calibration data sample and GF_{new} are the new values of the gain factors which will be used in the subsequent reconstructions.

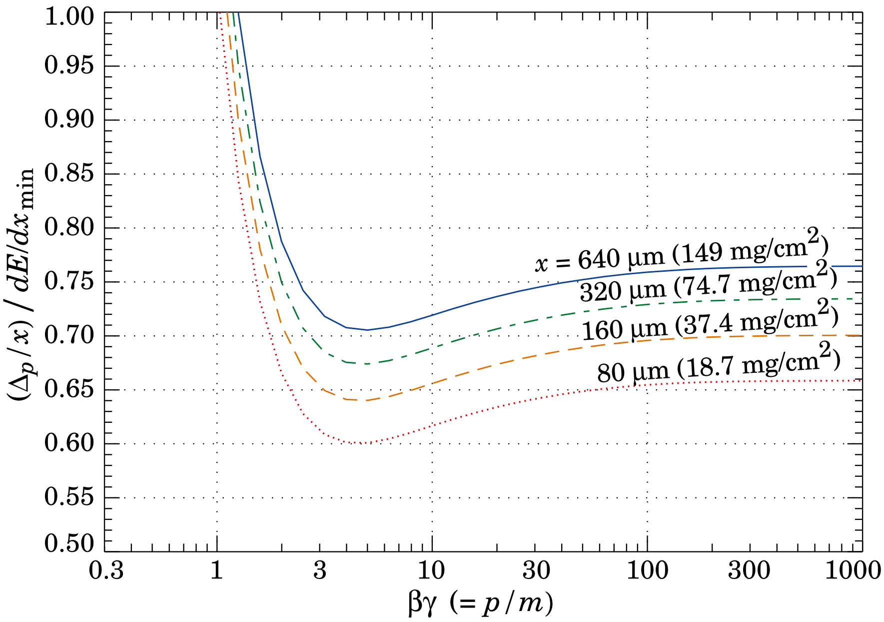


Figure 3.2: MPV of energy loss in silicon divided by mean energy loss of a minimum ionizing particle (388 eV/μm) as a function of a particle momentum (p) [38].

3.4 Calibration

3.4.1 First calibration using the test set-up

The first calibration of the SSD was performed using the test set-up in the laboratory using cosmic radiation (Utrecht calibration). The test set-up consists of spare SSD ladders connected to a copy of the ALICE DAQ and a trigger system. The data collected by this set-up were reconstructed using AliRoot (only the local SSD reconstruction) with all gain factors equal to one. Using reconstructed data the charge matching calibration was performed. The new gain factors were: 0.86 for the P-side strips and 1.16 for the N-side strips. Those values were stored in the OCDB as the gain factors for all the SSD modules. The absolute calibration could not be done because the track inclination could not be calculated from the collected data.

3.4.2 Calibration based on data from the 2009 run

In December 2009 the ALICE experiment recorded the first proton-proton collisions at $\sqrt{s} = 900$ GeV. A data sample containing around $3 \cdot 10^5$ events was reconstructed using information from the entire ALICE detector. This data sample was not large enough to perform the SSD calibration at the module level but was large enough to check the charge matching for each module and detect modules with effects larger than 10% in the absolute calibration.

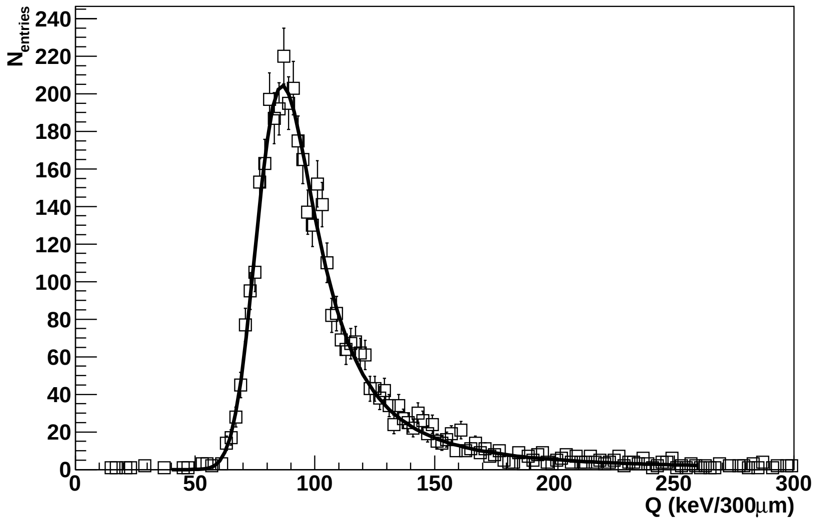


Figure 3.3: Example of a fit of the convolution of the Landau and the Gaussian distribution to the measured distribution of the energy loss signal for one SSD module.

Track points from tracks with $|\eta| < 0.9$, refits in the TPC and the ITS were used to calculate for each module the CR distribution. Most of the CR distributions were strongly peaked around zero which proves that the charge matching was already well controlled using the Utrecht calibration.

Due to the small number of recorded events the requirement of the momentum value ($p > 1.2$ GeV/c) was released to perform a check of the energy loss signal calibration. That made it possible to perform fits of the convolution of the Landau and the Gaussian distributions to the track point energy loss signal distributions of single modules, obtained using the same tracks as for the CR distribution. Figure 3.3 shows an example of such a fit. Out of all the SSD modules 1505 modules had enough data for the fits. Out of them, 8 showed problems with gain, 5 had a lower gain for one side and 3 had problems with gains on both sides. The lower gains for the 8 modules are caused by a hardware problem. Their gains were updated in the OCDB using Equations 3.3 with $mpv_{NV} = 86.24$ keV/300μm (mean value of the MPV for all modules in this data sample). The default value $mpv_{NV} = 84$ keV/300μm was not used because the momentum requirement was released and this can lead to an increase of the calculated MPV. These gain factors were used in the final reconstruction of the 2009 data. Results from this data sample are presented in the next chapter.

3.4.3 Calibration based on data from the 2010 run

In 2010 the LHC started to produce proton-proton collisions at $\sqrt{s} = 7$ TeV. The amount of data recorded in the first few weeks was sufficient to perform a calibration of the SSD at the module level as defined earlier in this chapter. The procedure was successful (the distribution of the energy loss had enough entries that the fit converged) for 1522 modules. This number of modules depends on the on the number bad channels and is consistent with the number of active modules in that period. For each module, the $\langle CR \rangle$ and the MPV were used to obtain new gain factors using Equations 3.3. From all modules under study (including those for which the fit did not converge) 8 showed problems: 7 of them have recorded a smaller number of track points compared to the other modules and one module showed two peaks in the charge distribution. All of those modules had non-standard noise. Based on these studies those modules were masked as dead. During this procedure the same track quality requirements were used as for the 2009 calibration.

To check the effect of the new calibration, the same procedure was repeated for a new sample of proton-proton collisions reconstructed using new calibration. The results of a comparison of MPVs for both data samples are shown in Figure 3.4. The spread in the MPVs distribution decreased from 3.5 keV/(300μm) to 0.78 keV/(300 μm), which proves that the new gain factors reduced the noise.

Figure 3.4 also shows that the gain is stable in time because the data samples were taken in different periods. The 2010 calibration was used in the 7 TeV data analysis presented in the next chapter.

3.4.4 MPV as a function of the track length

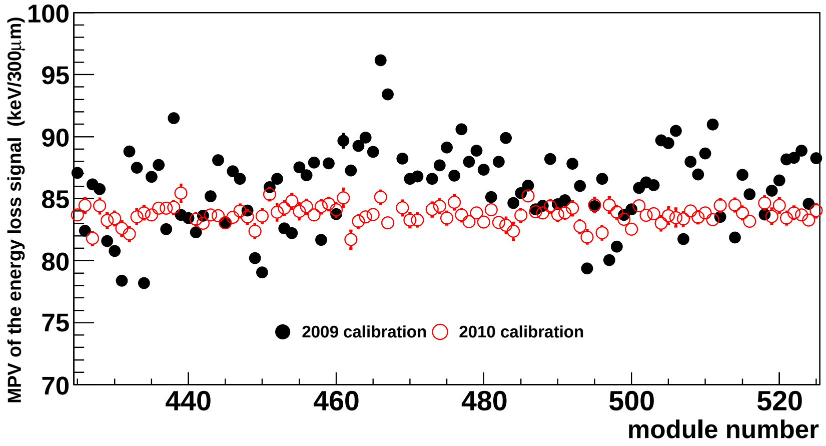
Figure 3.2 shows that the MPV of the energy loss of a particle in silicon depends on the track length (d). According to [38] the MPV is proportional to $d(a+b\ln d)$ where a and b are parameters which depend on $\beta\gamma$.

The collisions in ALICE take place in the centre of the detector (z=0). As a consequence, particles going through the SSD modules in the centre of the detector on average have a shorter length in silicon that those passing modules on the edges of the SSD. By normalizing the measured charge to the module thickness the linear part of this dependence is corrected but not the logarithmic part. This can be seen as the flat shape of the MPV distribution for the 2009 calibration in Figure 3.4(b). The gain factors which were calculated also correct for the logarithmic part. As a result, they consist of two parts: a random one which corrects the gain fluctuations at the module level and a deterministic one which corrects on average for the logarithmic part of the track length effect. Due to that the gain factors should not be used in a generation of the gain variations at the module level in a simulation of the SSD. From the gain factors the random part should be extracted and used for this generation, while the whole gain should be used in the reconstruction. One other solution is to correct for both parts of the dependence of the MPV on the track length during the normalization of the charge. None of these solutions have been implemented yet.

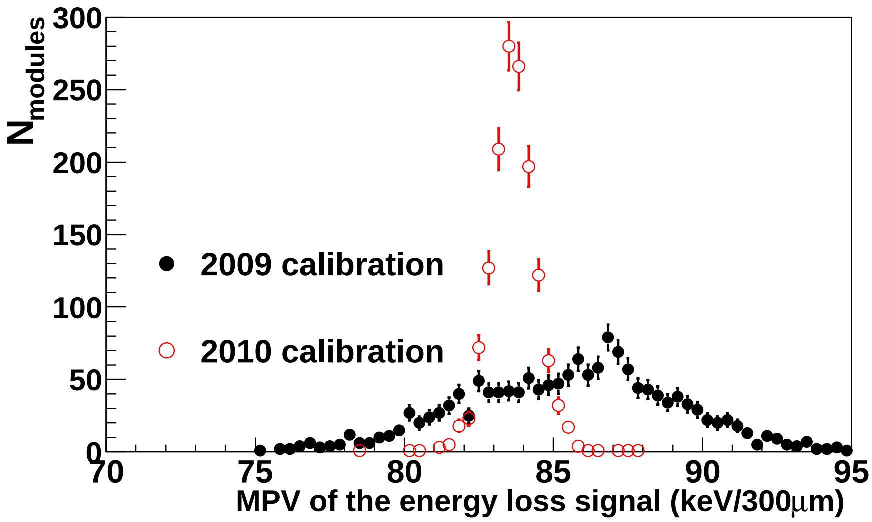
3.5 Performance studies

3.5.1 Calibration at the level of a single HAL25 chip

To check the calibration at the HAL25 chip level the energy loss signal distributions at this level have to be calculated. To do that the track point energy loss



(a) MPV of the energy loss signal in the SSD for each module.



(b) Distribution of the MPVs of the energy loss signal in the SSD (one entry = one module)

Figure 3.4: The effects of the calibration made using the 2010 data sample (see Section 3.4.3). The spread of the MPV distribution decreased from 3.5 keV/(300μm) to 0.78 keV/(300 μm). The flat shape of the MPV distribution for the 2009 calibration is related to the dependence of the MPV on the track length (see Section 3.4.4).

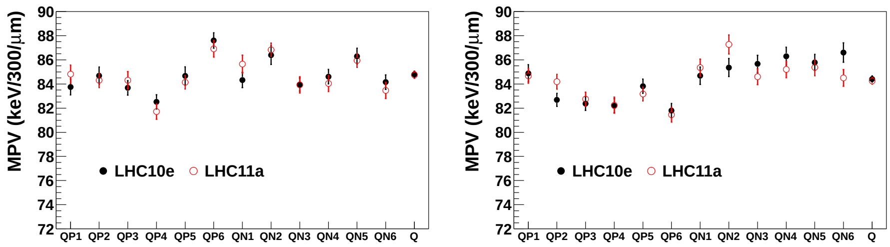


Figure 3.5: Two examples of the MPVs of single chips. Measurements for 2 different data taking periods are plotted. The first 6 bins on the X-axis show gains for the P-sides chips and the next 6 for the N-side chips. The last bin shows the MPV for entire module.

signal has to be divided into an N and a P part (charge measured by N(P) strips normalized to module thickness) using the values of CR. The coordinates of a track point are used to assign the N and P parts to chips. In principle the MPV obtained from fitting the convolution of the Gauss and the Landau distribution to a distribution of the energy loss signal for each HAL25 chip should be equal to 84 keV/(300μm). Figure 3.5 shows the MPVs for all the chips for two modules using two different data samples (periods): LHC10e (proton-proton collisions recorded in summer 2010), LHC11a (proton-proton collisions recorded in March 2011). Those plots raise two questions:

- Is the difference between MPVs measured for one chip in two periods a statistical or systematic effect?
- Is the difference between MPVs of the different chips in one module in one period a statistical or systematic effect?

To check whether the difference between two independent measurements of the same value is a statistical or systematic effect, the following variable can be calculated:

$$x_{diff} = \frac{x_1 - x_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \tag{3.4}$$

where

x_i is the value of a measurement i;

σ_i is the statistical uncertainty of x_i;

If the distribution of x_{diff} can be described by a Gaussian distribution with mean 0 and standard deviation 1, that means that the differences have a statistical origin.

To answer the first question the MPVs calculated for two periods were compared by calculating mpv_{diff}. The distribution of mpv_{diff} is shown in Figure 3.6 and was fitted using a Gaussian distribution. The fitted Gaussian distribution has a mean of 0.058 and the standard deviation equal to 1.13. The values obtained from the fit are close to the values expected for statistical fluctuations.

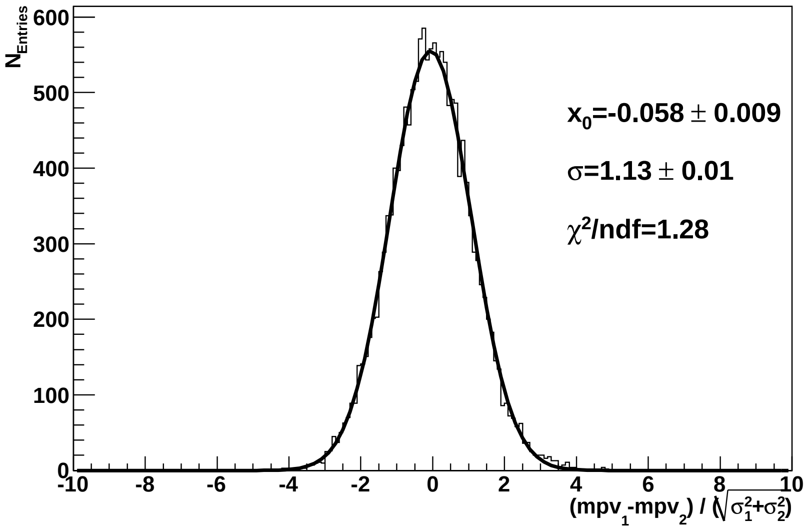


Figure 3.6: Distribution of the difference between the chip MPVs measured in two data samples (LHC10e and LHC11a) normalized to the statistical uncertainty (Equation 3.4).

This result suggests that the difference between the MPV values measured for two periods is mainly caused by statistical fluctuations. This shows that the gain at the chip level is stable in time.

To answer the second question the MPVs calculated for the chips from one side of one module were compared by calculating mpv_{diff}. The distribution of the mpv_{diff} was calculated for each data sample. Those distributions are shown in Figure 3.7. They cannot be described by a Gaussian distribution with standard deviation equal 1.This means that the chip-to-chip fluctuations of the gain are not a statistical effect.

3.5.2 Charge matching vs. charge

Figure 3.8 shows the correlation Q_N vs. Q_P for recpoints (see 3.2.4), the line corresponds to the perfect matching (Q_N =Q_P). A dependence of the charge matching on the charge value is observed. For charges larger than 400 keV, the distribution deviates from the perfect matching, a structure is seen in this part of the distribution. To investigate this structure closer, the Q_N vs. Q_P distributions were calculated but only using recpoints with two (2x2) or three (3x3) strips on both sides. Those distributions are shown in Figure 3.9. In case of the 3x3 recpoints the part of the distribution which deviates from the perfect matching is wider than in case of the 2x2 recpoints. This suggests that the mentioned structure on Figure 3.8 is a result of the recpoint size. There are few 2x2 recpoints with $Q_N > 1300$ keV suggesting a saturation of the N-side strip at the level of 650 keV. Those distributions prove that the responses of the N-side and P-side strips in ADC for the same charge are different, especially

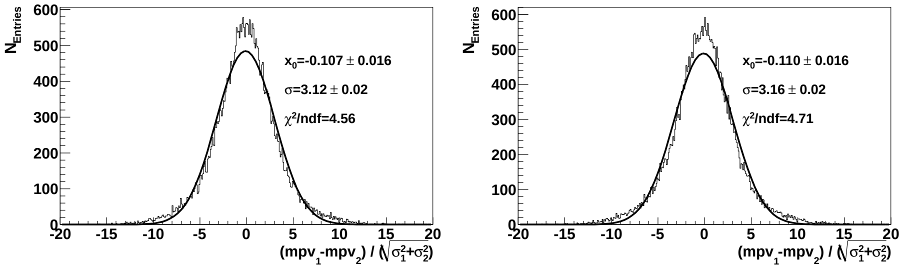


Figure 3.7: Distribution of the difference between of the chip MPVs, normalized to the statistical uncertainty (Equation 3.4), for chips from the same side of one module, calculated using all available modules. Distributions for two periods of data taking are plotted LHC11a (left), LHC10e (right).

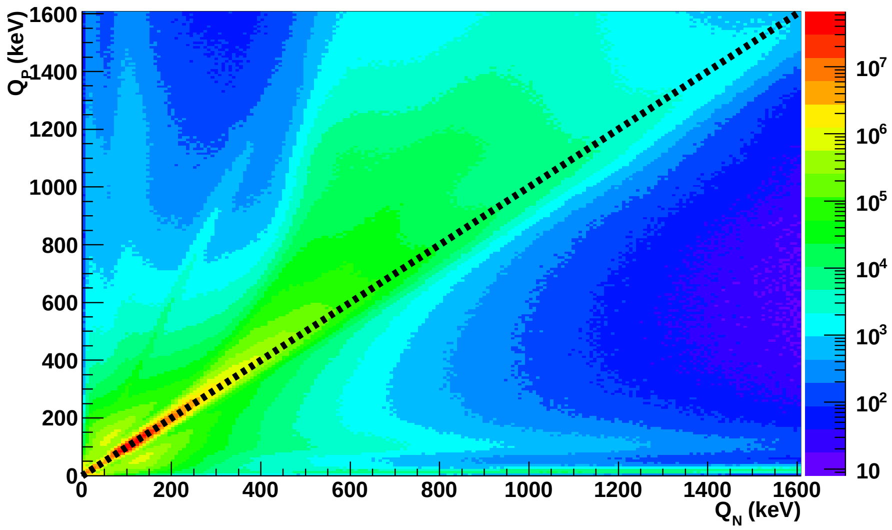


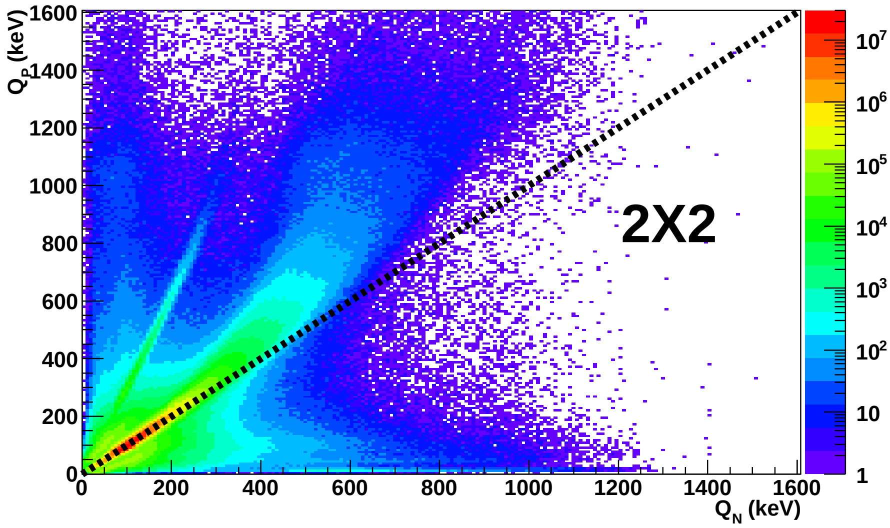
Figure 3.8: The correlation plot Q_P vs. Q_N for recpoints (see Section 3.2.4) in data. The line shows perfect matching $Q_N = Q_P$. The distribution for higher charges deviates from the perfect matching this is studied in Section 3.5.2. The structure for $2Q_N = Q_P$ is caused by the module with wrong gains. That was corrected in a next pass of the reconstruction. Other structures with a small charge on one side are related with noisy recpoints or shared recpoints.

for large charges. This difference can be caused by different behaviour of the HAL25 chips for positive and negative signals.

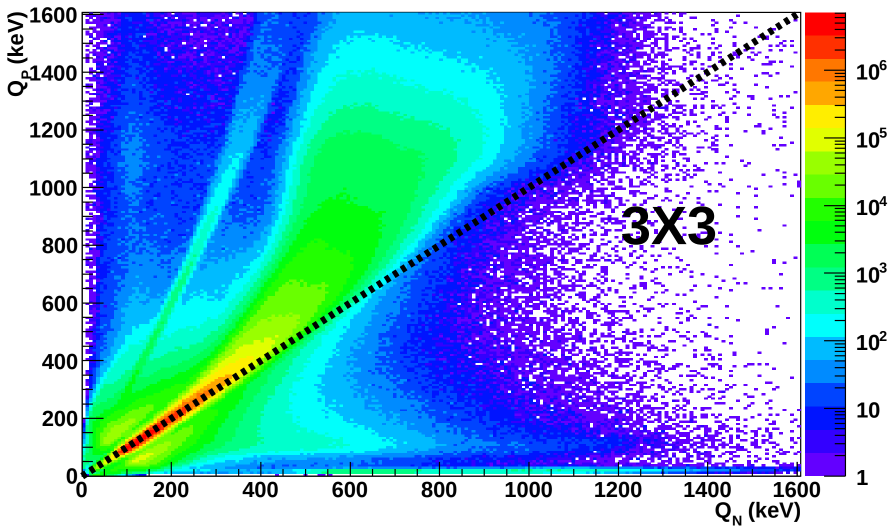
Figures 3.8 and 3.9 also illustrate special types of the SSD recpoints. The structures with a high charge on one side and small on other correspond to the noisy recpoints (signal on one of the sides is noise) or recpoints which do not have the corrected charge assigned because signal on one of the side is shared by two recpoints (see Section 3.2.4). The structure for $2Q_N = Q_P$ is caused by the module with wrong gains. That was corrected in a next pass of the reconstruction.

3.6 Conclusions

The energy loss signal in the SSD was calibrated at the module level with an accuracy of 1%. The applied calibration is stable in time. The calibration can be improved by defining gain factors for individual HAL25 chips. Saturation of the measured charge is observed as a lack of the charge matching for high charges. This effect can be caused by the HAL25 chips.



(a) recpoints with two strips on both sides



(b) recpoints with three strips on both sides

Figure 3.9: The correlation plot Q_P vs. Q_N for two sizes of recpoints (see Section 3.2.4) in data. The line shows perfect matching $Q_N = Q_P$. The structure for $2Q_N = Q_P$ is caused by the module with wrong gains. That was corrected in a next pass of the reconstruction. Other structures with a small charge on one side are related with noisy recpoints or shared recpoints.

Chapter 4

Measurement of the π, K, p transverse momentum spectra in proton-proton collisions

4.1 Introduction

In this chapter the measurement of pion, kaon proton and antiproton p_t spectra using the energy loss signal in the ITS for global tracks (see Section 2.4.2) is presented. Results for two proton-proton collision energies $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV are discussed. The measurement was performed for primary particles, those directly produced during the collision including products of decays, but not products of weak decays of strange particles.

4.2 Data sample and event selection

4.2.1 Data sample

For the analysis of proton-proton collisions at $\sqrt{s} = 900$ GeV the final reconstruction of the good quality runs from the LHC09d period (runs from December 2009) was used. The corresponding simulation is based on the PYTHIA 109 tune (D6T) [36]. The data sample has around $3 \cdot 10^5$ of good events and the number of simulated events is around $3 \cdot 10^6$.

For the analysis of proton-proton collisions at $\sqrt{s} = 7$ TeV the second reconstruction pass of runs from LHC10b and LHC10c periods (the collisions recorded in spring 2010) was used. The corresponding simulations are based on the PYTHIA 320 tune (Perugia0) [37] and PHOJET. The selected runs have all 3 main PID detectors (ITS, TPC, TOF) read out and a low probability of pile-up events ($\mu < 0.1$, see Section 2.4.3). They were marked as good runs in the ALICE run conditions table [39]. All the runs had the same L3 magnet configuration. This data sample had more than $5 \cdot 10^7$ good events and a comparable number of simulated events for each generator. In the 7 TeV analysis the PHOJET simulation was used as the default for corrections.

4.2.2 Event selection

ALICE physics selection

The online minimum-bias (MB) trigger selection is based on the following conditions:

- Two or more hits in the SPD are detected.
- At least one hit in the SPD and a signal in at least one of the VZERO arrays are detected.
- Signals in both VZERO arrays are detected.

During the offline data analysis the trigger conditions are recomputed using reconstructed data with the additional rejection of beam-gas events. The beam-gas event rejection is based on the arrival times of the VZERO signals (if the collision takes place around $z = 0$ they are 11 ns and 3 ns) and correlation between the number of tracklets and number of hits in the SPD. The corrected correlation was defined using simulation. This offline procedure is called the ALICE physics selection.

Selected events

Only events which pass the ALICE physics selection and have a global or an SPD vertex (see Section 2.4.2) reconstructed within $|z| < 10$ cm were considered. Those are common ALICE conditions for selecting minimum-bias events. The selection on the z vertex position provides a good acceptance in the $|y| < 0.5$ region.

4.3 Measurement of the p_{t} spectra

In each selected event, only tracks fulfilling certain track quality requirements (track cuts) are used. For each track the rapidity is calculated using three different mass hypotheses: pion, kaon and proton. If a track has $|y| < 0.5$ for one of the mass hypotheses a histogram is filled with the difference between the logarithm of the measured energy loss signal in the ITS and the expected value calculated using the parametrization ($\ln[\mathrm{d}E/\mathrm{d}x]_{meas} - \ln[\mathrm{d}E/\mathrm{d}x]_{calc}$). This is done in p_{t} bins while the parametrization is a function of total momentum. This procedure gives a series of six histograms (three types of particles, two charges) in each p_{t} bin. Each histogram has one, two or three maxima depending on the momentum. The area below the peak centred at zero is mostly the raw yield of particles from the mass hypothesis. The raw yield is determined by fitting the histograms around zero with an analytical function. The raw yield is then corrected for the reconstruction inefficiency, contamination and other effects:

$$\frac{\mathrm{d}N}{\mathrm{d}p_{\mathrm{t}}}||_{|y|<0.5} = \frac{1}{\Delta_{p_{\mathrm{t}}}} \frac{(1-\omega)C_{cross-section}}{\varepsilon\varepsilon_{Event}\delta_{TPCITS}} n_{fit}, \tag{4.1}$$

where

n_{fit} is the raw yield (number of tracks) extracted from the fit,

ε is the efficiency of the tracking and the PID procedure, including acceptance,

ω is the contamination,

δ_{TPCITS} is the correction for the TPC-ITS matching,

ε_{Event} is the efficiency of the event selection,

$C_{cross-section}$ is the correction for incorrect simulation of the interaction with the detector material,

$\Delta_{p_{\mathrm{t}}}$ is the p_{t} bin size.

The statistical uncertainty of the spectra ($\Delta_{\frac{\mathrm{d}N}{\mathrm{d}p_{\mathrm{t}}}}$) is given by:

$$\Delta_{\frac{\mathrm{d}N}{\mathrm{d}p_{\mathrm{t}}}} = \frac{\mathrm{d}N}{\mathrm{d}p_{\mathrm{t}}} \left(\left(\frac{\Delta_{n_{fit}}}{n_{fit}} \right)^2 + \left(\frac{\Delta_{\omega}}{1-\omega} \right)^2 + \left(\frac{\Delta_{\varepsilon}}{\varepsilon} \right)^2 + \left(\frac{\Delta_{\varepsilon_{Event}}}{\varepsilon_{Event}} \right)^2 \right)^{0.5} . \quad (4.2)$$

In the next paragraphs those quantities as well as the whole procedure are described in detail.

4.3.1 Definition of the energy loss signal in the ITS

The first step of the analysis was to combine the measurements of the energy loss signals in the SSD and the SDD layers into the energy loss signal in the ITS. The signal has to have a good resolution (ratio of the spread of measured values to the mean measurement) and should not introduce too many requirements on a track because not all of the global tracks are reconstructed with 4 points in the SSD and the SDD. This is mainly caused by dead zones in the detectors. Those two conditions are contradictory, a better resolution requires more energy loss samples.

Different scenarios of combining signals from the layers were investigated, using the energy loss signal in the SSD and the SDD, corresponding to the measured response for each layer. It was observed that in scenarios, which use the highest measured signal the resolution is the largest. This is caused by the influence of the tail of the Landau distribution for values higher that its mean. In the end as a compromise the energy loss signal in the ITS ($(\mathrm{d}E/\mathrm{d}x)_{meas}$) for global tracks was defined as follows:

- for tracks with 4 points in the SDD and the SSD as the mean of the two lowest energy loss signals in the ITS layers,

- for tracks with 3 points in the SDD and the SSD as the weighted mean of the two lowest energy loss signals in the ITS layers, with weights 1 and 0.5 for the lowest and second lowest energy respectively.

In case of tracks with less than 3 points in the SDD and the SSD the signal was not defined. In this approach the resolution is around 12%.

Table 4.1: The list of the standard ALICE track cuts.

	Common cuts for both analyses	
successful refit procedure (see Section 2.4.2) in:		
the TPC	yes	
the ITS	yes	
rejection of kinks, tracks which are a reconstruction of the charged particle decaying to the charged and neutral particles	yes	
requirements on the ITS reconstruction	At least one reconstructed cluster in the SPD layers	
Maximal value of χ^2 per cluster in the TPC	4	
	Specific cuts for each analysis	
	$\sqrt{s}$=900 GeV	$\sqrt{s}$=7 TeV
Minimal number of the TPC clusters, where 159 is the maximum number	80	70
Maximal values of the distance of closest approach of the track to the primary vertex (dca):		
dca_{xy} [μm]	$7 \cdot \sigma_{dca}$ $\sigma_{dca} = 50 + 60/p_{\mathrm{t}}^{1.01}$ for p_{t} in GeV/c	$7 \cdot \sigma_{dca}$ $\sigma_{dca} = 26 + 50/p_{\mathrm{t}}^{1.01}$ for p_{t} in GeV/c
dca_z [μm]	$5 \cdot \sigma_{dca}$ $\sigma_{dca} = 146 + 70/p_{\mathrm{t}}^{1.115}$ for p_{t} in GeV/c	$2 \cdot 10^4$

4.3.2 Track selection

In this analysis the standard ALICE track cuts for selecting primary tracks were used. Those cuts are summarized in Table 4.1. The performance of the cuts

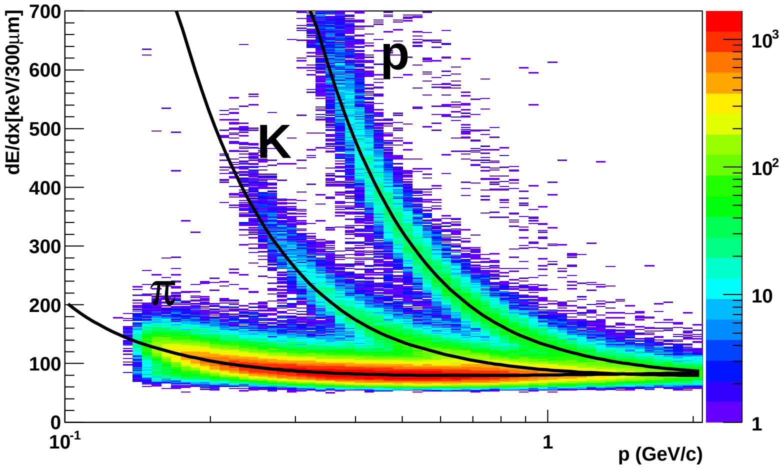


Figure 4.1: Energy loss signal in the ITS as function of the track momentum at the primary vertex in data (proton-proton collisions at $\sqrt{s}$ = 900 GeV) compared to parametrization (Equation 4.3). Three lines indicate bands for each particle type.

will be studied in Section 4.4.3.

In the table dca stands for the distance of closest approach of the track to the primary vertex. In addition to the standard ALICE track cuts, the requirement of 3 or 4 points in the SSD and the SDD with energy loss signal higher than 50 keV/300μm was introduced (ITSPID cut). This was needed to calculate the energy loss signal in the ITS. The requirement for the value of the signal in the ITS layers was introduced to reject tracks to which a noisy track point in the SDD or the SSD was assigned during the tracking. Noisy track points have a small energy loss signal.

4.3.3 Determination of the dE/dx parametrization

The energy loss signal in the ITS for tracks fulfilling the cuts was calculated as a function of the track momentum at the primary vertex. This is shown in Figure 4.1. Three main bands for pions, kaons and protons are seen. The energy loss signal in the ITS for the momentum higher than 200 MeV/c does not distinguish pions from electrons and muons. In momentum bins where a particle band is well separated, the peak position for that band was extracted. The dependence of the peak position on $\beta\gamma$ was parametrized using the following equation:

$$(\mathrm{d}E/\mathrm{d}x)_{calc} = \begin{cases} a_0 \frac{a_1 + 2\ln(\gamma) - \beta^2}{\beta^2} \cdot ((\beta\gamma - a_3)^2 + a_4) & \text{if } \beta\gamma < a_2 \\ a_0 \frac{a_1 + 2\ln(\gamma) - \beta^2}{\beta^2} \cdot ((a_2 - a_3)^2 + a_4) & \text{if } \beta\gamma \geq a_2 \end{cases} \tag{4.3}$$

where

$\beta\gamma = p/m$,

p is the momentum at the primary vertex,

a_i are free parameters.

This equation is based on the parametrization from [40], with an additional polynomial describing the low $\beta\gamma$ part. The lines in Figure 4.1 are calculated using this parametrization. All parameters were separately calculated for data and simulations. They are also different for both data samples, since a better calibration of the energy loss signal in the SSD was used (the 2010 calibration from Chapter 3) during the reconstruction of the 7 TeV data.

4.3.4 Extraction of the raw yields in p_{t} bins

The raw yields were extracted by fitting an analytical function (minimalization of χ^2) to the histograms of $x = \ln[\mathrm{d}E/\mathrm{d}x]_{meas} - \ln[\mathrm{d}E/\mathrm{d}x]_{calc}$ in p_{t} bins. This function was defined as follows (the justification of this choice will be presented later):

$$\begin{aligned} f(x) &= s(x) + b(x), \qquad (4.4) \\ s(x) &= \begin{cases} \frac{C}{N\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-x_0)^2}{2\sigma^2}\right) & \text{if } x < x_0 + d\sigma \\ \frac{C}{N\sqrt{2\pi}\sigma} \exp(a_s(x_0 + d\sigma - x) - 0.5d^2) & \text{if } x \geq x_0 + d\sigma \end{cases} \\ b(x) &= A \cdot \exp(a_b(x_{min} - x)), \end{aligned}$$

where N is a normalization factor given by:

$$N = 0.5 \cdot (1 + \mathrm{Erf}(d/\sqrt{2})) + \exp(-0.5d^2) \cdot \frac{1}{\sqrt{2\pi}\sigma a_s}. \qquad (4.5)$$

The derivative of $f(x)$ is continuous if:

$$a_s = \frac{d}{\sigma} \qquad (4.6)$$

which was a requirement during the fitting procedure. The function $f(x)$ is a sum of two functions: signal ($s(x)$) and background ($b(x)$). The signal part describes the raw yield of particles from the mass hypothesis and it has two parts: a gaussian and an exponential. The background function describes the yield of particles not from the mass hypotheses around $x = 0$. In the case of the pion mass hypothesis there is no need to use the background function. The x_{min} is the lower limit of the x range where the fit is made. For the 7 TeV data sample the fit ranges in each p_{t} bin were defined based on the positions of histogram extrema and the expected values of σ. Due to a low number of events in the 900 GeV data sample the fit ranges were tuned individually for each p_{t} bin. In each p_{t} bin the same range was used to calculate the raw yield of positive and negative particles of the same type. The fit parameters were:

C is the raw yield (to get the number of tracks n_{fit} for Equation 4.1, it should be divided by the x bin size),

x_0 is the mean of the gaussian part of the signal function,

σ is the standard deviation of the gaussian part of the signal function,

d is the start of the exponential part of the signal function in standard deviation units,

A is the scale factor of the background,

a_b is the parameter of the background.

Fits for the lowest and highest p_{t} bin used for all types of positive particles in the 7 TeV data sample are shown in Figure 4.2. Solid lines indicate the $f(x)$ functions in the ranges where the fits were performed. Dashed lines are the signal and background functions. The lower p_{t} limit of the spectrum calculation for each particle was defined by requiring the efficiency (see Section 4.3.5) to be larger than 0.25, the upper limit by a reasonably small value of the systematic uncertainty presented in Section 4.4.1.

The performance of the fit function was studied in simulations. The whole procedure of calculating the raw yields was applied to simulation results, including the fit ranges definition. Figure 4.3 shows an example of the proton yield calculation, including contributions from all sources of protons. Using yields from simulation, the accuracy of the fit was defined as:

$$\Delta_{PID} = \frac{n_{fit} - n_{NGP} - n_{TP}}{n_{fit}}, \tag{4.7}$$

where

n_{fit} (see Equation 4.1),

n_{TP} is the number of reconstructed tracks corresponding to primary particles from the mass hypothesis (e.g. "primary p" in Figure 4.3),

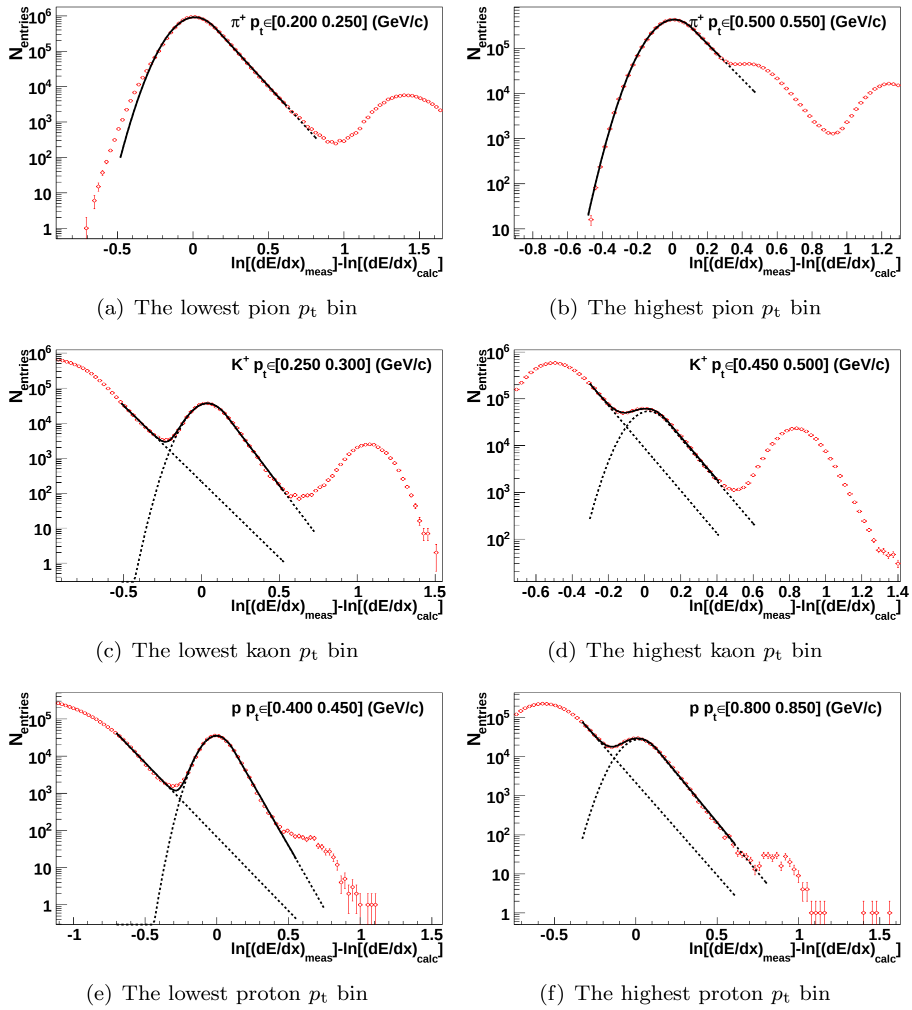
n_{NGP} is the number of reconstructed fake tracks and tracks of secondary particles[1] created by particles from the mass hypothesis, including muons and electrons in case of the pion yield (e.g sum of "secondary p" and "fakes primary p" in Figure 4.3).

The term n_{NGP} is calculated inside the signal region, while n_{TP} in the whole x range. The signal region is defined based on Equation 4.4 as follows: $[x_0-5\sigma, \infty)$. It was introduced to take into account that some tracks created by secondary particles have x smaller than $x_0-5\sigma$ and they do not contribute to the calculated raw yield.

Deviations from zero of Δ_{PID} are caused by:

- Other sources of contamination in the signal region (e.g. tracks with incorrectly assigned charge, fakes from other particles), this effect increases Δ_{PID}.
- A tail of the primary particle distribution outside the signal region, such as seen in Figure 4.3 where the distribution of tracks created by primary protons is not described by the signal function for $x < -0.4$, this effect decreases Δ_{PID}.

[1]Secondary particles are not primary (e.g. products of weak decay of strange hadrons and particles produced by interactions with the detector material).



(a) The lowest pion $p_{\rm t}$ bin

(b) The highest pion $p_{\rm t}$ bin

(c) The lowest kaon $p_{\rm t}$ bin

(d) The highest kaon $p_{\rm t}$ bin

(e) The lowest proton $p_{\rm t}$ bin

(f) The highest proton $p_{\rm t}$ bin

Figure 4.2: Illustration of the calculation of raw yields in the lowest and the highest $p_{\rm t}$ bin in the 7 TeV data sample for positive particles of all types. The points are the measurement and solid lines indicate the $f(x)$ function used to determine the yields. Dashed lines are the signal and background functions (see Equation 4.4).

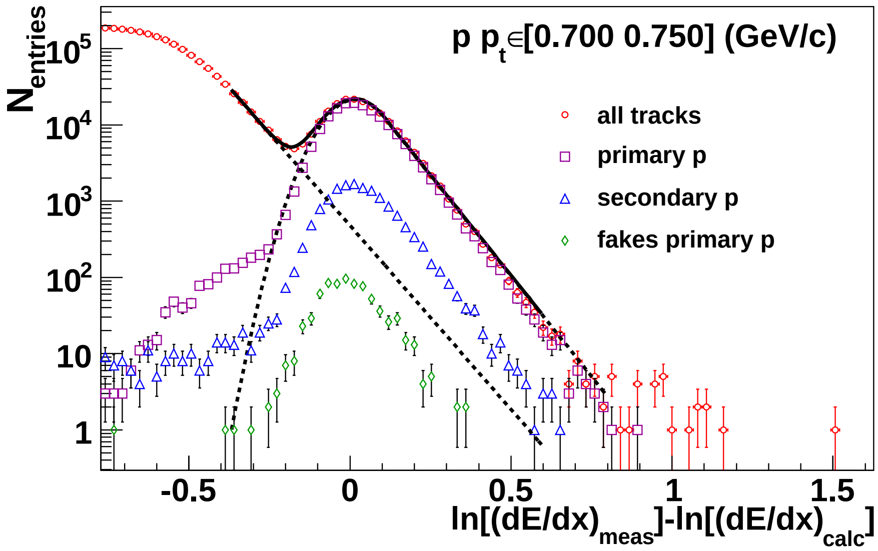


Figure 4.3: Determination of the proton raw yield in one p_{t} bin in the simulation for 7 TeV analysis. Contributions from all sources of protons are shown. For the definition of the fake see Section 2.4.2.

- The range where the fit was made, definition of the range can decrease or increase Δ_{PID}.

The final value of Δ_{PID} is a sum of those effects. In the analysis of the 900 GeV data sample Δ_{PID} was assigned as part of the systematic uncertainty of the whole measurement. In the 7 TeV data sample those effects were studied in detail. The results of those studies are part of the discussion about the systematic uncertainties (see Section 4.4.1).

Δ_{PID} was also calculated for other fitting functions (e.g. sum of three Gaussian functions). The function given by Equation 4.4 had the smallest absolute values of Δ_{PID}, which was the main reason why it was used to extract raw yields.

4.3.5 Tracking and PID efficiency

The tracking and PID efficiency is defined as follows:

$$\varepsilon = \frac{n_{fit} - n_{cont}}{n_{MC}}, \tag{4.8}$$

where

n_{fit} (see Equation 4.1),

n_{cont} is the number of tracks created by secondary particles from the mass hypothesis inside the signal region, including muons and electrons in case of pions,

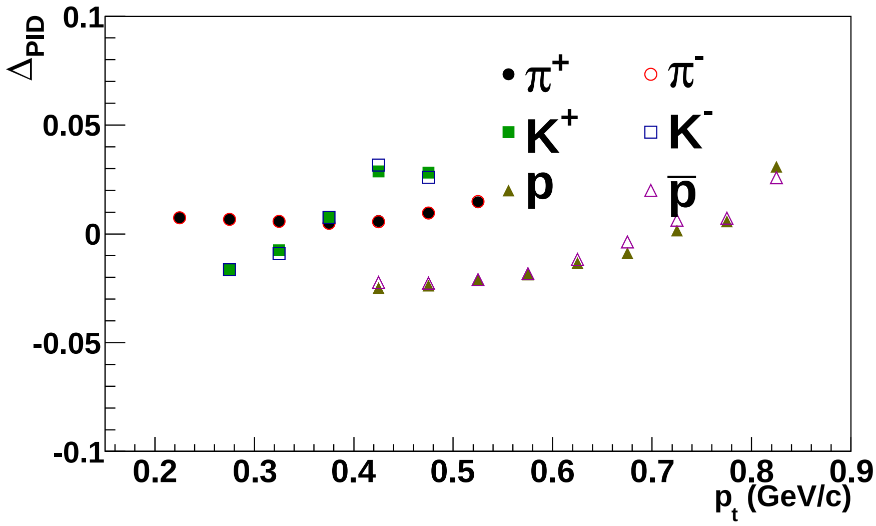


Figure 4.4: Δ_{PID} (Equation 4.7) for all particle types in the 7 TeV analysis.

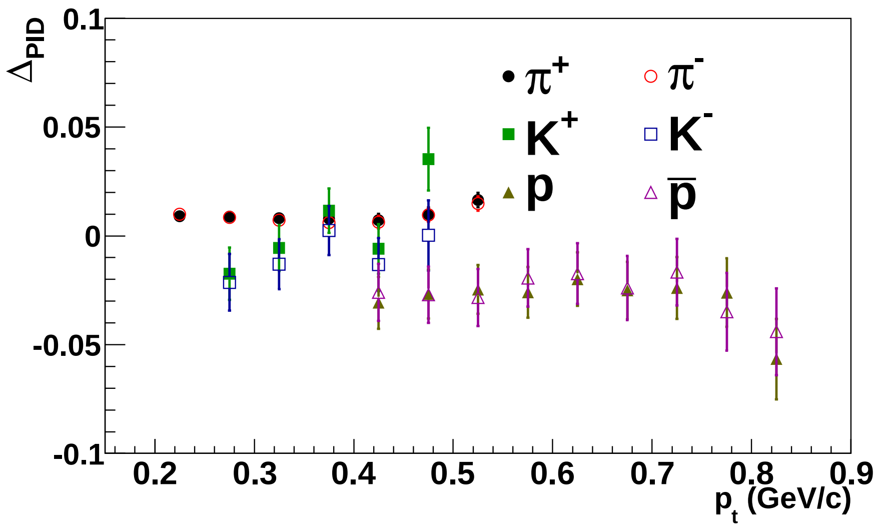


Figure 4.5: Δ_{PID} (Equation 4.7) for all particle types in the 900 GeV analysis.

n_{MC} is the number of the generated primary particles from the mass hypothesis in events which pass all the event cuts (physics selection and the z vertex cut).

This definition of efficiency takes into account the tracking as well as the PID efficiency and the acceptance. It also directly corrects for all effects which contribute to the Δ_{PID} if they are the same in data and simulation. The efficiencies

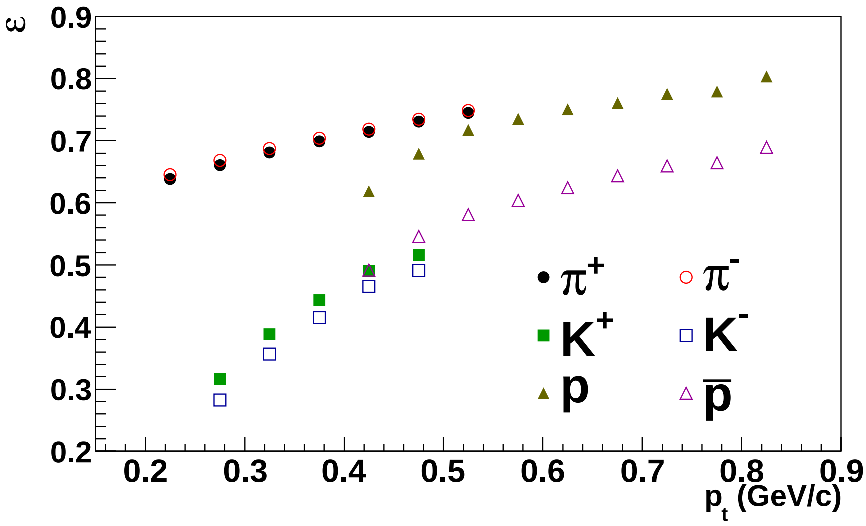


Figure 4.6: The efficiencies (Equation 4.8) for all particle types in the 7 TeV analysis.

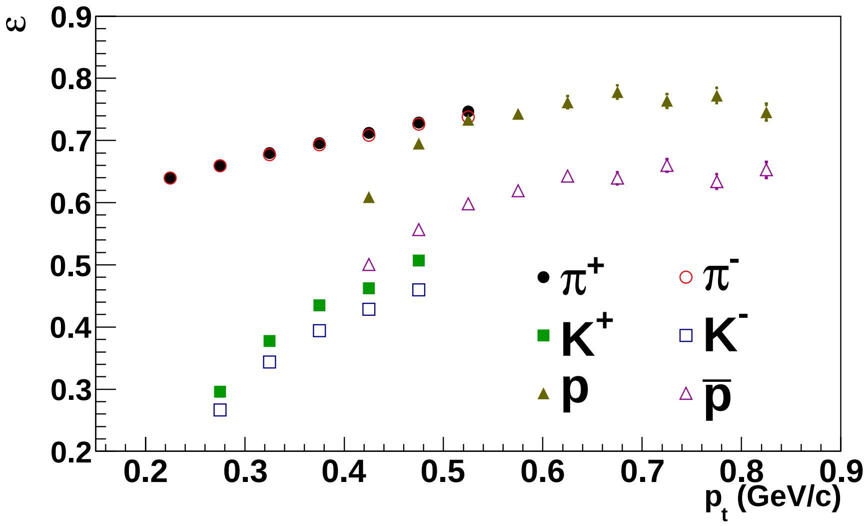


Figure 4.7: The efficiencies (Equation 4.8) for all particle types in the 900 GeV analysis.

are shown in Figures 4.6, 4.7, where plotted uncertainties are calculated using a binomial approximation. The difference between the efficiencies for positive and negative particles, especially for protons and kaons, is caused by the difference between the cross-sections for interactions with the detector material of positive and negative particles (see Section 4.3.8). The kaon efficiencies are lower than for other particles because $c\tau$ of the kaon is the smallest.

4.3.6 Contamination

The contamination is defined as follows:

$$\omega_i = \frac{n_i}{n_{fit}} \tag{4.9}$$

where

n_{fit} (see Equation 4.1),

n_i is the number of tracks inside the signal region coming from a source i of contamination.

The ω in Equation 4.1 is a sum of ω_i over all relevant sources. The contamination for all the particles includes the effect of secondary particles, but for the kaon spectra this is negligible. The contamination for the kaon, proton and antiproton spectra does not include misidentified particles because they are subtracted from the raw yield during the fitting procedure. The contamination for the pion spectra includes the muon, electron and positron contributions because they are indistinguishable from pions based on the energy loss signal in the ITS.

Protons and antiprotons

For protons (antiprotons), the only sources of contamination are secondary protons (antiprotons) from weak decays of strange particles, mainly Λ ($\overline{\Lambda}$) and Σ^+ ($\overline{\Sigma^+}$) particles, and protons produced by interactions of particles with the detector material. To calculate those contaminations, a data driven method was developed, consisting of the following steps (for each p_{t} bin).

- The dca$_{xy}$ cut (see Table 4.1) is released and dca$_{xy}$ distributions for tracks with $-0.2 < x < 0.5$ for the proton and antiproton mass hypotheses are calculated. This method of selection also selects a small fraction of pions and kaons to the distributions.
- Based on simulations three dca$_{xy}$ template distributions of protons and antiprotons are calculated: for primary particles, particles coming from the weak decays (WD) and products of interactions with detector material. Those distributions are shown in Figures 4.8 (a,b).
- The data distribution is fitted (in [-3 cm;3 cm] range) with the template distributions (in case of antiprotons only two distributions: the primary and the products of weak decays are used) to calculate the relative contributions of each source in the data distributions. The fit is made using the method described in [41] (TFractionFitter class in ROOT). Examples of fits are shown in Figures 4.8 (c,d).
- The fit result is used to calculate the contamination inside the dca$_{xy}$ cut.
- The contamination is corrected for the amount of pions and kaons that contribute to the data distribution. This is done based on the fit used to calculate the raw yield.

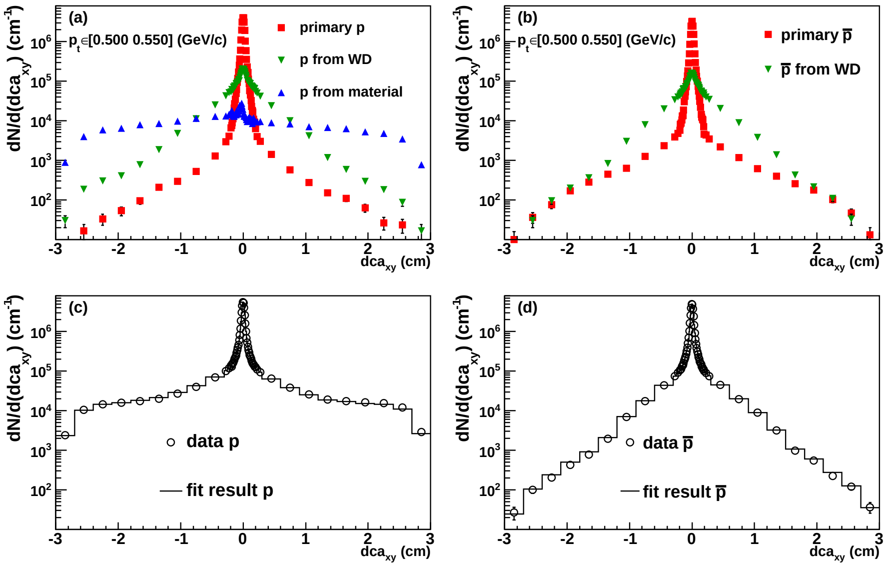


Figure 4.8: Illustration of the method of calculating proton and antiproton contaminations by fitting the dca$_{xy}$ distributions. The top row shows the template distributions of the dca$_{xy}$ distributions for protons (left) and antiprotons (right). The bottom row shows results of the fits to the data distribution the using method in [41].

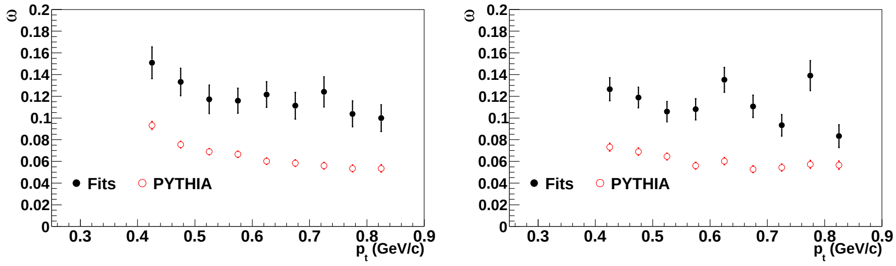
This procedure is based on two assumptions. The contamination does not depend on the energy loss signal in the ITS so the contamination calculated for $-0.2 < x < 0.5$ is the same as for the raw yield. The main background for protons (antiprotons) in Equation 4.4 are primary kaons and pions and they have the same dca$_{xy}$ distributions as primary protons (antiprotons). This assumption is used in the last step of the contamination corrections.

The results obtained with this method were compared with the results from pure simulation, see Figures 4.9, 4.10. For both energies the event generators underestimate the contamination by factor around 2. The proton contamination is larger than the antiproton contamination due to the protons coming from interactions with detector material. The statistical uncertainties obtained from the fits using the dca$_{xy}$ distributions depend on the statistical uncertainties in the data and simulation. This is why the uncertainties of the contamination from the dca$_{xy}$ fits are larger than when using the model directly.

Pions

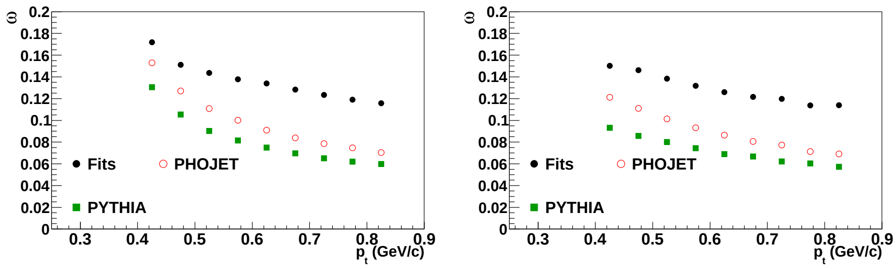
The main sources of the pion contamination and their contributions are presented in Figure 4.11. This result is based on the PHOJET simulation for the 7 TeV analysis. These studies show that significant sources of pion contamination beside the products of weak decays of strange particles are:

- primary electrons (positrons), mainly π^0 and η decays products,
- secondary electrons (positrons) from γ conversions inside the detector ma-



(a) p contamination (b) $\overline{p}$ contamination

Figure 4.9: Comparison of the proton and antiproton contaminations (Equation 4.9) calculated using the fit method to the contamination from PYTHIA (D6T) simulation for 900 GeV data.



(a) p contamination (b) $\overline{p}$ contamination

Figure 4.10: Comparison of the proton and antiproton contaminations (Equation 4.9) calculated using the fit method to the contaminations from PYTHIA (Perugia0) and PHOJET simulations for 7 TeV data.

terial,

- muons from Drell-Yan processes and charged pion decays.

The method developed for the calculation of the proton contamination could not be applied to calculate the pion contamination. The main problem is that in the data samples the "ExB effect"[2] is present. The tracking procedure did not correct for this effect during the reconstruction, which caused an asymmetry of the dca_{xy} distributions. This asymmetry was also not simulated. This leads to an incorrect estimation of the secondary electron and positron contributions in the data distributions by the fit method because the dca_{xy} distributions of the electrons and positrons coming from γ conversions are asymmetric. This asymmetry is caused by the propagation of secondary electrons (e.g. created in the beam pipe) to the primary vertex. In addition it is impossible to distinguish primary pions from primary electrons, positrons and muons based on the dca_{xy} distributions. The difference between those templates is smaller than the sensitivity of the fit method. Still the fit method was used to check the systematic uncertainties of the secondary pion contaminations in the 7 TeV data sample. This will be described in Section 4.4.2.

It was decided in the 7 TeV analysis to use the contaminations from the PHOJET simulation to correct the pion yields. There is no significant difference between the contamination obtained using PYTHIA and PHOJET(see Figure 4.12). This way of defining the contamination was supported by the observation that beside the pions coming from weak decays the other parts of the contamination are not expected to be strongly model dependent. Primary electron and muon production should be proportional to the pion production and the uncertainties of the secondary electron production are mainly caused by the uncertainty in the description of the detector material.

For the 900 GeV data sample, the K^0 and Λ yields in the simulation were scaled to match the measured results from [47]. A weight depending on the mother particle p_{t} was assigned to each pion coming from the decay of K^0 and Λ. Using this approach the part of the contamination caused by pions coming from weak decays was calculated. It was found that it was 1.2 times higher than the contamination in the simulation without the weights. The contaminations obtained with the weights were used to correct the pion yields because they are more accurate. The difference between contaminations with and without the weights was added to the total systematic uncertainties. The electron and muon contribution to the contamination was taken directly from the simulation.

4.3.7 TPC-ITS matching correction

To verify that the matching of the tracks between the TPC and the ITS is the same in data and simulation the δ_{TPCITS} variable (Equation 4.1) was introduced. It is defined as follows:

$$\delta_{TPCITS} = \frac{\varepsilon^{data}_{TPCITS}}{\varepsilon^{MC}_{TPCITS}}, \tag{4.10}$$

[2]A shift of the cluster position in the SSD and the SDD due to the Lorentz force that acts on the electrons and holes as they drift through the sensors.

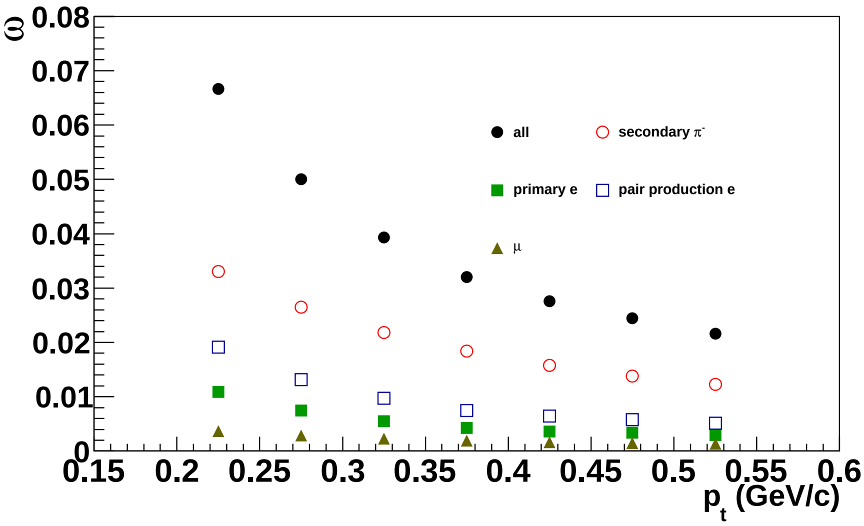


Figure 4.11: All sources of π^- contamination (Equation 4.9) in the 7 TeV data analysis. Results are based on the PHOJET simulation.

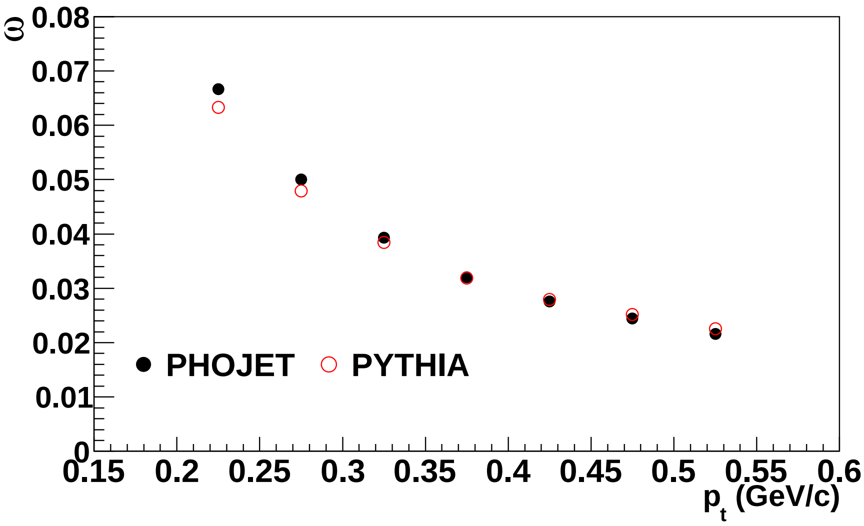


Figure 4.12: Comparison π^- contamination (Equation 4.9) from PHOJET and PYTHIA (Perugia0) simulations for $\sqrt{s}$=7 TeV proton-proton collisions.

where ε_{TPCITS} is the ITS-TPC matching efficiency defined as follows:

$$\varepsilon_{TPCITS} = \frac{n_{ITSPID}}{n_{Standard}}, \tag{4.11}$$

where

n_{ITSPID} is the number of tracks used in the analysis, fulfilling the standard ALICE track cuts and the ITSPID requirement (see Section 4.3.2),

$n_{Standard}$ is the number of tracks fulfilling the standard ALICE track cuts (see Table 4.1).

The term δ_{TPCITS} was calculated for all three mass hypotheses. It was found that it is on the level of 0.99 for all three mass hypotheses. The deviation of δ_{TPCITS} from unity can be caused by the incorrect simulation of the influence of dead zones in the ITS on the reconstruction. The influence of the particle charge on δ_{TPCITS} was also studied. The observed difference between the values for positive and negative particles is below 0.005. It was decided to use δ_{TPCITS} as a correction in the spectra calculation.

To check the influence of the particle type on δ_{TPCITS}, samples of pions, kaons and protons (antiprotons) based on the TPC PID (see Figure 2.4) were used. The TPC PID was used because the energy loss signal in the ITS cannot be calculated for all tracks fulfilling the standard ALICE track cuts. In the 7 TeV data sample it was found that δ_{TPCITS} for kaons and protons (antiprotons) calculated using the TPC PID is 0.005 higher than calculated using all tracks fulfilling the rapidity cut. That difference is included in the systematic uncertainties. In case of the 900 GeV sample the effect is lower than the statistical uncertainties of the measurements.

4.3.8 Interactions with the detector material

GEANT3 describes incorrectly the interaction of antiprotons with the detector material, since both elastic and inelastic cross-sections are wrongly parametrized. The description of the interaction of protons is also not accurate. This is discussed in [42]. A special correction for the inelastic part was developed using the FLUKA transport code [43], which describes those interactions better. The correction factor ($C_{cross-section}$ in Equation 4.1) is from 0.86 to 0.92 for the antiproton spectrum in the used p_{t} range. The obtained correction for protons is constant and equal to 0.99. This correction was applied in both analyses.

The effect of the elastic cross-section on the efficiency can be observed in simulation using the following function:

$$\rho = \frac{\varepsilon^{-}_{TPC-stand}}{\varepsilon^{+}_{TPC-stand}} \tag{4.12}$$

where $\varepsilon_{TPC-stand}$ is the ratio of the number of tracks created by primary particles of the certain type reconstructed by the TPC to the number of tracks created by primary particles of the same type fulfilling the standard ALICE track cuts. The indexes + and − define the charge of the particle type. The standard ALICE track cuts require a reconstruction of a cluster in the ITS. So $\varepsilon^{\pm}_{TPC-stand}$ depends on the fraction of dead zones in the ITS, the tracking

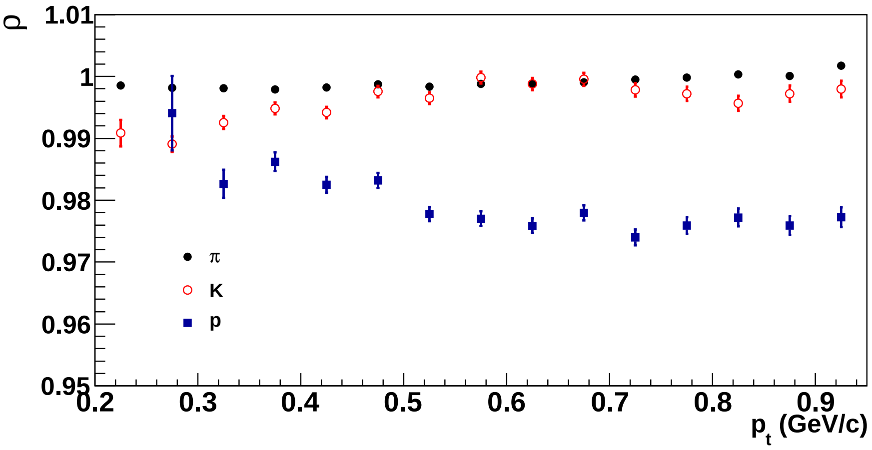


Figure 4.13: The ρ function (see Equation 4.12) for all types of particles in the 7 TeV analysis.

algorithm efficiency and the elastic cross-sections for interactions with the detector material. The dependence on the cross-section is caused by particles which had many elastic interactions and the TPC tracks for those particles cannot be matched with ITS clusters because clusters are too far from the prediction of the Kalman filter. The ratio of $\varepsilon^{-}_{TPC-stand}$ to $\varepsilon^{+}_{TPC-stand}$ depends only on the elastic part because all other effects should be the same for positive and negative particles. In the end the ρ value can be approximated as:

$$\rho \approx \frac{1 - A\sigma^{-}_{EL}}{1 - A\sigma^{+}_{EL}} \tag{4.13}$$

where σ_{EL} stands for the elastic cross-section and A is a scaling factor. A value of ρ larger than one means that $\sigma^{+}_{EL} > \sigma^{-}_{EL}$.

Figure 4.13 shows ρ for all types of particles. The ratios of elastic cross-section for interactions with carbon for negative particles to the same cross-section for positive particles in GEANT3 were also calculated and are shown in Figure 4.14. The rise of $\sigma^{\overline{p}}_{EL}/\sigma^{p}_{EL}$ around $0.45 < p < 0.6$ GeV/c corresponds to the drop of ρ around $0.45 < p_{\mathrm{t}} < 0.7$ GeV/c.

In [43] it was reported that in GEANT3 $\sigma^{\overline{p}}_{EL}$ (for carbon) for momentum around 0.5 GeV/c is overestimated by a factor around 2 while the σ^{p}_{EL} is in good agreement with available experimental data and it is twice smaller than the measurement of $\sigma^{\overline{p}}_{EL}$. This means that ρ for protons and antiprotons in simulation is underestimated, which leads to overestimating the antiproton spectrum. To correct for that an additional correction of 0.985 was applied to the antiproton spectrum in the 7 TeV analysis. An additional 1% systematic uncertainty was assigned to the antiproton spectrum due to that correction.

For kaons and pions the GEANT3 cross-sections were also compared to existing data [44, 45, 46] (see Figures 4.15, 4.16). It can be seen that the K^{-} cross-section in GEANT3 rises for low momentum and the pion cross-sections are underestimated by GEANT3. To check the influence of the inelastic interac-

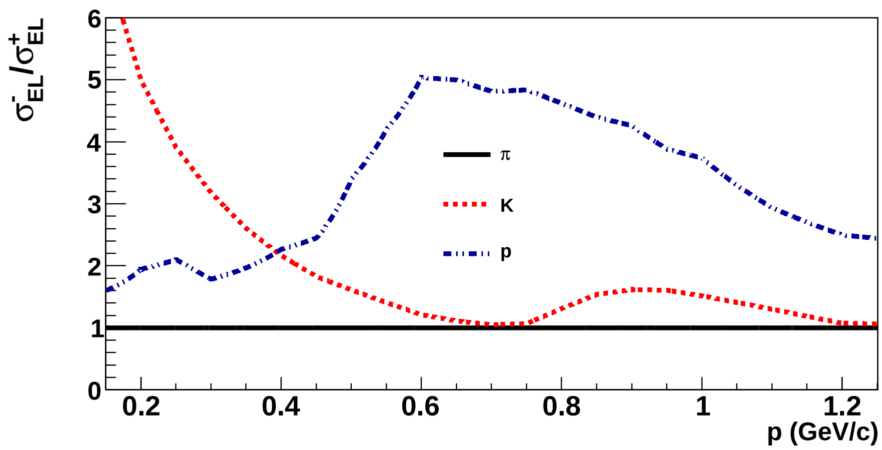
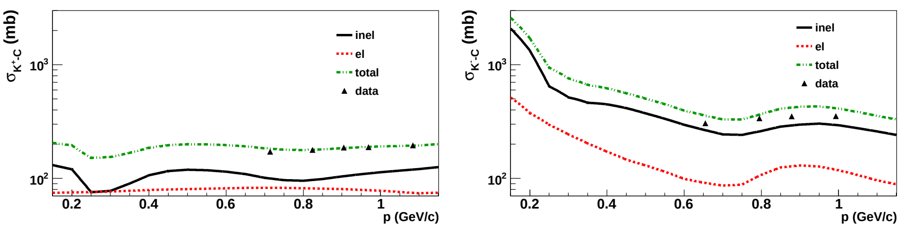
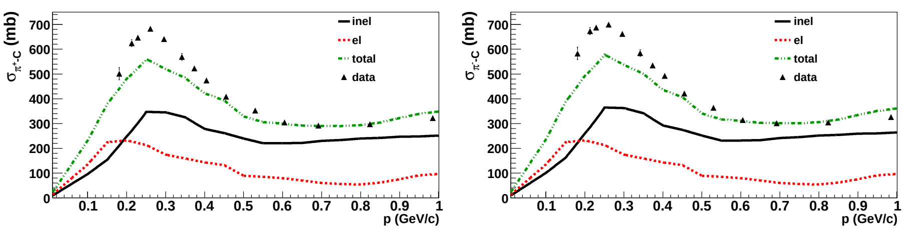


Figure 4.14: Ratio of the GEANT3 elastic cross-sections for interactions with carbon for negative particle to the same cross-section for positive particle.



(a) K^+ (b) K^-

Figure 4.15: Comparison of kaon cross-sections for interactions with carbon between the GEANT3 parametrization and experimental data [45, 46]. The GEANT3 cross-sections (labelled as *total*) are divided to the elastic (labelled as *el*) and inelastic (labelled as *inel*) parts.



(a) π^+ (b) π^-

Figure 4.16: Comparison of pion cross-sections for interactions with carbon between the GEANT3 parametrization and experimental data [44]. The GEANT3 cross-sections (labelled as *total*) are divided to the elastic (labelled as *el*) and inelastic (labelled as *inel*) parts.

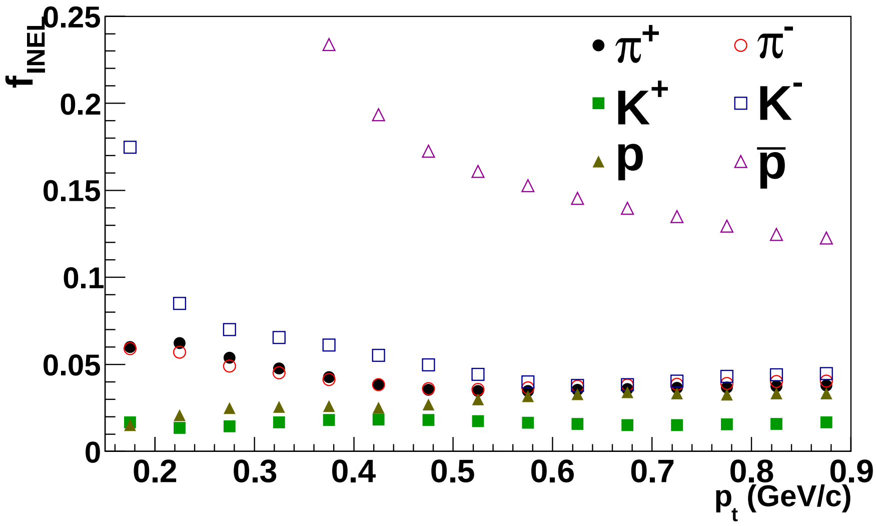


Figure 4.17: Fraction of the primary particles lost due to inelastic interactions (f_{INLE}) in radius $R < 80$ cm. Calculations were done using simulations for the 7 TeV analysis.

tions on the spectra, the fractions of particles suffering an inelastic interaction in radius $R < 80$ cm (f_{INEL}) were calculated. The values of f_{INEL} for all types of particles are shown in Figure 4.17. It has to be mentioned that those fractions can be underestimated because in the simulation information about inelastic interaction is not stored if products of the interaction cannot be detected by the ALICE detector. The difference between the values of f_{INEL} for positive and negative particles is the main cause of the difference between the tracking efficiencies (see Figure 4.6). The $p_{\rm t}$ dependence of f_{INEL} values can be explained by the momentum dependence of the inelastic cross-sections.

For the 7 TeV analysis it was decided to use $0.5 f_{INEL}^{K^-}$ as uncertainty of the inelastic negative kaon cross-section. This propagates as $0.5 f_{INEL}^{K^-}/(1 - f_{INEL}^{K^-})$ to systematic uncertainties of the final negative kaon spectrum. Due to the underestimation of the pions cross-sections it was decided to assign $0.2 f_{INEL}^{\pi}/(1 - f_{INEL}^{\pi})$ as the systematic uncertainty to the pion spectra. The details of the influence of the cross-sections on the spectra were not noticed in the 900 GeV analysis due to the large statistical uncertainties. So only a 2% correction to the lowest $p_{\rm t}$ bins of negative kaon spectrum was applied, based on the comparison of the cross-section between GEANT3 and FLUKA.

4.3.9 Event efficiency

The event efficiency (ε_{Event} in Equation 4.1) is defined as follows:

$$\varepsilon_{Event} = \frac{n_{MC}}{n_{MCall}} \tag{4.14}$$

where

n_{MC} (see Equation 4.8),

n_{MCall} is the number of MC particles of a given type from all generated events.

It corrects for the loss of particles due to the trigger inefficiency and the vertex cuts. This correction can depend on the event generator which was used to calculate it. In both data samples it was found that this correction is constant in the used p_t ranges. In addition, the values for the PYTHIA and PHOJET simulations in the 7 TeV analysis are the same. The calculated values 0.99 for the 7 TeV data sample and 0.97 for the 900 GeV data sample are compatible with the loss of events due to the z vertex cut.

4.3.10 Pile-up correction

For the 7 TeV data the influence of pile-up was investigated. The average value of the μ parameter (see Section 2.4.3) for selected runs is 0.06 which means 0.03 probability for pile-up. Pile-up was not an issue for the 900 GeV data sample due to the low luminosity of the proton beams.

The procedure of calculating the 7 TeV spectra was performed with and without rejection of the pile-up events based on the relevant ALICE rules. It was found that the pile-up tagging excludes 1% of the events passing the ALICE physics selection, which is in agreement with expected efficiency of the pile-up tagging method. It was also found that the rejection of pile-up decreased all the spectra by around 1%. It was decided to use the pile-up tagging in final analysis and add a 2% uncertainty to the normalization procedure uncertainty.

4.4 Evaluation of systematic uncertainties

The remaining sources of systematic uncertainties are: the fitting procedure used in the calculation of the raw yields, the contamination correction and the detector material budget. The stability of the final spectra on track cut variation was also checked.

4.4.1 Fitting procedure

One source of systematic uncertainties is the difference between the x distribution in data and in simulations. This can be seen in Figure 4.18 where those distributions are compared for one p_t bin for the kaon mass hypothesis. There are two main differences between these distributions: the data distribution has a larger spread and smaller separation between the peaks than the simulated distribution. This means that a change of the range where the fit is made in data and simulation can have a different influence on the calculated yields. The fraction of primary tracks lying outside the signal region can also be different in data and simulation.

The determination of the fit ranges (see Section 4.3.4) was different in both data samples, so different methods were used to evaluate the systematic uncertainties. In case of the 900 GeV analysis the difference between the number of events in data and simulation is large, meaning that the change of the ranges in simulation can be sensitive to effects which are irrelevant for data. This is not the case for 7 TeV analysis.

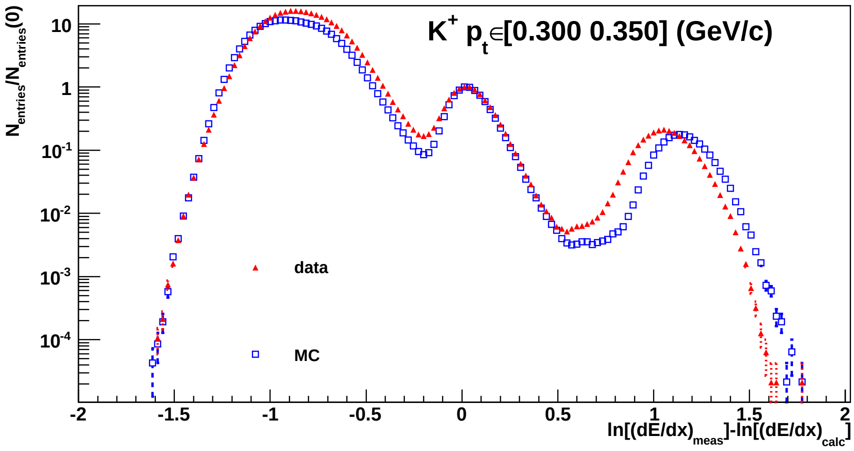


Figure 4.18: Comparison of the energy loss signal in the ITS between data and simulation (MC) for the 7 TeV data sample.

In case of the 900 GeV data sample, only the influence of the fit range on the raw yield was studied. This was done by calculating the relative change of the raw yield due to a change of the fit range (Δ_{range}), where a negative value of Δ_{range} means that the yield from the changed range is larger than the standard one. The ranges were changed by δx =0.03, which corresponds to excluding (including) 1 or 2 bins from (in) the ranges. The results of those calculations are shown in Figure 4.19.

The calculation of the statistical uncertainty of Δ_{range} was not trivial because the raw yields in the standard and the changed ranges are strongly correlated. This correlation has to be taken into account to avoid overestimating the statistical uncertainty of Δ_{range}, which makes it impossible to evaluate the systematic change of the raw yield due to the change of the range. To evaluate those uncertainties a special procedure was developed. From each histogram a new histogram was made, where new data points were generated from the Gaussian or Poisson distributions with the spread corresponding to the statistical uncertainty of each point. Using the new histogram, new raw yields and Δ_{range} were calculated for all x ranges under study. For each histogram this procedure was repeated 500 times, to obtain the Δ_{range} distributions. The standard deviations of those distributions were assigned as the statistical uncertainties of the Δ_{range}. It was concluded that in most of the p_{t} bins the effect connected to the change of the range is much smaller that the statistical uncertainty of the calculated raw yield.
Those studies show a few important properties of the fits:

- For all particle types the change of the raw yield due to the change of the fit range increases with p_{t}, it is caused by the increase of the background from other particles.
- The change is the lowest for the pion yield. It is only significant when the upper limit of the range is changed because that causes a change of the

fitted exponential part of the signal function.

- For kaon and proton yields the changes are similar. They are the largest when the lower limit of the range is changed because that causes a change of the fitted background function.

In the 900 GeV analysis besides the results of those calculations the Δ_{PID} (Equation 4.7) was also included as a systematic uncertainty of the fitting procedure.

In case of the 7 TeV analysis the fit ranges were changed in data and simulation at the same time. They were changed by half of the pion σ (Equation 4.4), 0.055 for data and 0.05 for simulation. The entire procedure of calculating the corrected spectra was performed with the changed ranges. The new spectra were compared to standard ones and the relative difference between them (Δ_{range2}) was assigned as the systematic uncertainty. Values of Δ_{range2} for positive particles are shown in Figure 4.20. Due to a correlation between the two spectra (for the standard and changed range) the statistical uncertainties of Δ_{range2} are overestimated but they are still smaller than the observed effects. Those calculations show a similar influence of the fit range on the corrected spectra as on the raw yields in the 900 GeV sample.

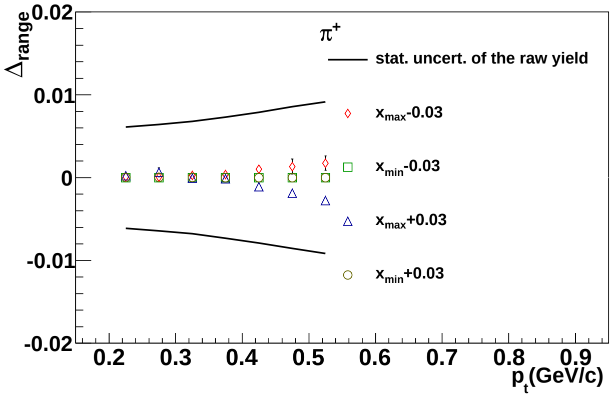
The basic check for the primary particles lying outside the signal region (see Section 4.3.4) is Δ_{PID} for low p_{t} bins where the background contribution is small. To cross-check that value on data all histograms of x were made using pure samples of pions, kaons and protons, selected using the TPC PID. That gives the possibility to reduce the background, especially in low p_{t} bins. The function $f(x)$ (Equation 4.4) but without the background part was fitted to those histograms. The relative difference between the total number of entries in each histogram and the yield obtained from the fit was calculated (Δ_{PID2}). The same procedure was applied to the simulation but in addition Δ_{PID} was calculated in the standard way. All three values (Δ_{PID2} for data and simulation and Δ_{PID} for the pure samples) were compared with the standard Δ_{PID} for low p_{t} bins of kaons and protons and all p_{t} bins of pions. It was found that the difference between them is below 0.01. Based on those studies it was decided to assign 1% systematic uncertainty for all types of particles in all p_{t} bins in the 7 TeV analysis.

4.4.2 Contamination correction

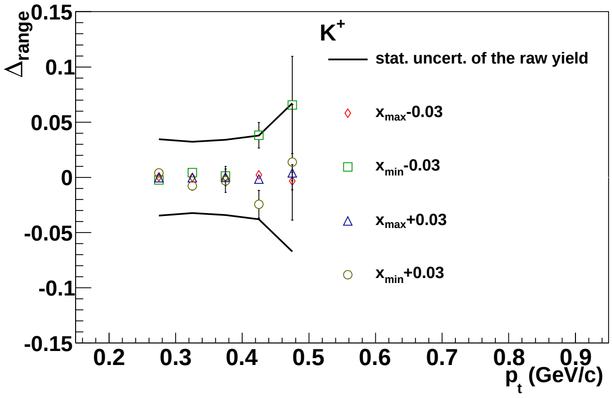
Influence of the Σ^+ ($\overline{\Sigma^+}$) to Λ ($\overline{\Lambda}$) ratio on the proton (antiproton) contaminations

The Σ^+ ($\overline{\Sigma^+}$) and Λ ($\overline{\Lambda}$) particles produce a proton (antiproton) during their weak decays. Due to different lifetimes of those particles ($c\tau = 2.4$ cm for Σ^+ and $c\tau = 7.9$ cm for Λ) the dca$_{xy}$ distributions of the produced protons are different. The final shape of the dca$_{xy}$ distribution of all protons (antiprotons) coming from weak decays is a function of the Σ^+($\overline{\Sigma^+}$) to Λ ($\overline{\Lambda}$) ratio in the model used to calculate the templates.

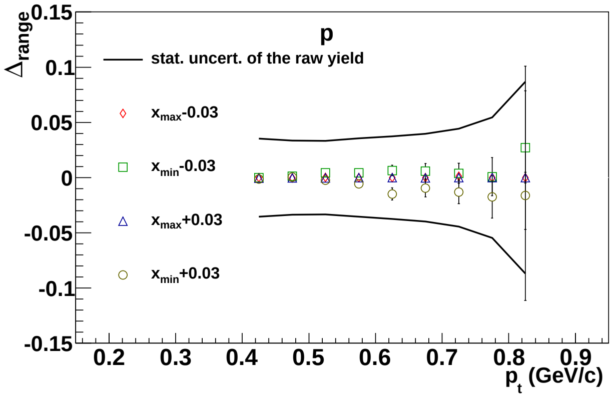
For the 900 GeV data sample, the collected number of events was not large enough to observe the influence of the ratio on the calculated proton (antiproton) contamination. In case of the 7 TeV data sample an attempt to fit two template distributions for protons (antiprotons) coming from the weak decays was made.



(a) pion yield

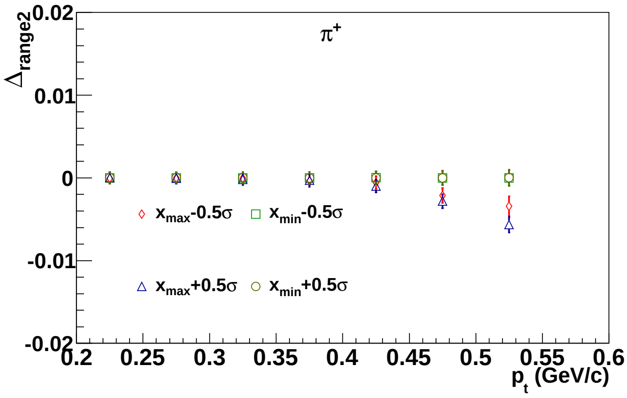


(b) kaon yield

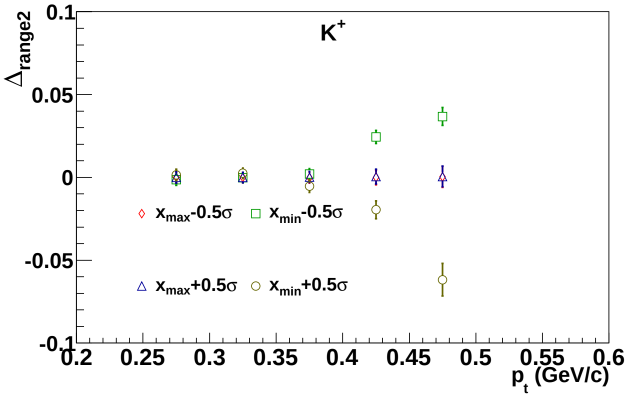


(c) proton yield

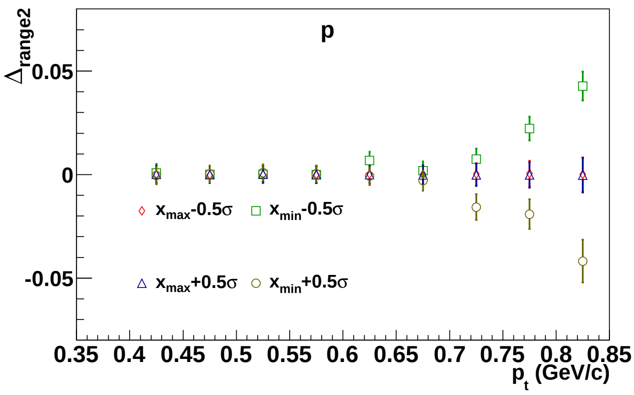
Figure 4.19: The relative changes of the positive particle raw yields (C in Equation 4.4) using different ranges during fitting function $f(x)$ (Equation 4.4) in the 900 GeV analysis.



(a) pion spectrum



(b) kaon spectrum



(c) proton spectrum

Figure 4.20: The relative changes of the positive particle spectra using different ranges during fitting function $f(x)$ (Equation 4.4) in the 7 TeV analysis.

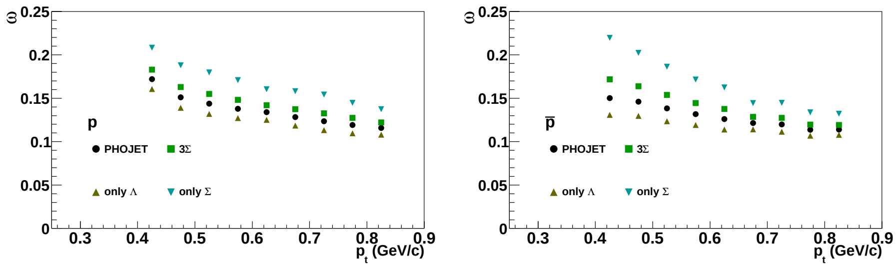


Figure 4.21: Test of the influence of different shapes of the weak decay templates (different ratios of Σ^+ ($\overline{\Sigma^+}$) to Λ ($\overline{\Lambda}$) see Section 4.4.2) on calculated proton (right) antiproton (left) contamination (Equation 4.9).

This attempt was not successful because the main difference between those distributions is in the region where the fits are constrained by the primary particles ($|\mathrm{dca}_{xy}| < 0.2$ cm).
To evaluate the systematic uncertainty caused by the Σ^+ ($\overline{\Sigma^+}$) to Λ ($\overline{\Lambda}$) ratio, fits using different templates for protons (antiprotons) coming from weak decays were performed. The following scenarios of the weak decay template were tested:

- the Σ^+ ($\overline{\Sigma^+}$) to Λ ($\overline{\Lambda}$) ratio as it was simulated by an event generator, which is the standard approach (labelled as PHOJET in Figure 4.21),
- only Σ^+ ($\overline{\Sigma^+}$) or Λ ($\overline{\Lambda}$) template (labelled as *only* Σ and *only* Λ in Figure 4.21),
- scaling the Σ^+ ($\overline{\Sigma^+}$) contribution by a factor (3 in Figure 4.21) in the event generator template (labelled as 3Λ in Figure 4.21).

The results based on the PHOJET simulation are shown in Figure 4.21. The main observation from those studies is that the calculated contamination increases when the contribution of the protons (antiprotons) coming from Σ^+ ($\overline{\Sigma^+}$) in the weak decay template increases. This effect is a consequence of the difference between the shapes of the distributions of protons (antiprotons) coming for Σ^+ ($\overline{\Sigma^+}$) and Λ ($\overline{\Lambda}$). Increasing the Σ^+ ($\overline{\Sigma^+}$) contribution in the weak decay template increases the relative yield in $|\mathrm{dca}_{xy}| < 0.2$ cm for this template, but the contribution of the weak decay part which is the result of the fit is mainly constrained by the $|\mathrm{dca}_{xy}| > 0.2$ cm region.

It was decided to use as systematic uncertainty the difference between the contaminations obtained with the PHOJET template and the pure Λ ($\overline{\Lambda}$) template. This corresponds to an increase of the contamination caused by increasing the Σ^+ ($\overline{\Sigma^+}$) contribution by a factor of 3. The scenario with a pure Σ^+ ($\overline{\Sigma^+}$) template is unrealistic because ALICE has detected Λ ($\overline{\Lambda}$) particles in 7 TeV proton-proton collisions. The final uncertainties on the spectra decrease with p_{t} for protons from 1.4% to 1.0% and for antiprotons from 2.4% to 0.7%.

Systematic uncertainties of proton contamination

During the reconstruction and simulation the “ExB” effect was not properly taken to an account (see Section 4.3.6). This is one of the reasons why the

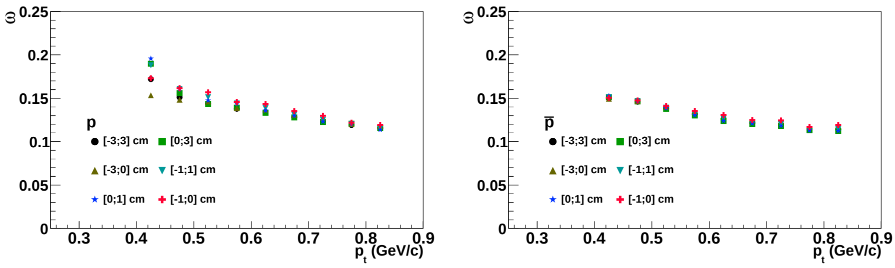


Figure 4.22: Comparison of the proton and antiproton contamination (Equation 4.9) for 7 TeV data obtained using different fit ranges in method from Section 4.3.6.

measured dca$_{xy}$ distribution is not perfectly reproduced by the simulation. This could affect the calculation of the proton contamination based on the fit method. To check that effect in the 7 TeV sample, the proton (antiproton) contamination was calculated using different dca$_{xy}$ ranges during the fit procedure. The results of those calculations are shown in Figure 4.22. Only for low p_{t} protons a significant difference was observed. It was decided to assign a systematic uncertainty to the proton spectrum in the 7 TeV data sample due to the "ExB" effect and other possible deviations of the simulated dca$_{xy}$ distribution from the measured one. It decreases with p_{t} from 3% to 0.5% . In the 900 GeV analysis this check could not be performed due to limited statistics.

Scaling the Λ and K^0 yields

In the 900 GeV analysis, the Λ ($\overline{\Lambda}$) and K^0 yields in simulation were scaled to match the measured values [47]. The contaminations using the scaled yield were calculated. For protons (antiprotons) the contamination is consistent with the contamination obtained using dca$_{xy}$ fits, even though the Σ^+ ($\overline{\Sigma^+}$) contribution was not scaled.

Pion contamination uncertainties

The uncertainties for the 900 GeV analysis are described in Section 4.3.6. For the 7 TeV analysis the dca$_{xy}$ fit method was used to evaluate uncertainties. It was used only in positive or negative dca$_{xy}$ range to reduce the influence of the asymmetry of the dca$_{xy}$ distribution. In p_{t} bins where the fits converged the calculated parts of contaminations caused by the weak decay products do not differ more than 0.01 from the calculated using simulations. Based on those studies, a 1% systematic uncertainty was assigned to the pion spectra in the 7 TeV analysis. Systematic uncertainties related to the muon and electron (positron) part of the contaminations were not assigned because muon and electron production is small and proportional to the pion production (see Section 4.3.6).

4.4.3 Track cut variations

The influence of the track cuts (see 4.1) on the final results was checked. The main goal was to explore uncertainties in the efficiency caused by imperfect representation of data by simulation. The calculations of spectra were run with several different sets of track cuts and the results were compared to the standard ones.

For the 900 GeV sample the effect of the following track cut variations on the corrected spectra was checked:

- changing the minimum number of clusters in the TPC from 80 to 65 and 90,
- including kink daughters to check the influence of the kink tagging procedure on the kaon and pion spectra,
- changing the maximum dca_{xy} from $7{\cdot}\sigma_{dca}$ to $5{\cdot}\sigma_{dca}$ and $9{\cdot}\sigma_{dca}$ (see Table 4.1).

In the 900 GeV analysis those variations change the spectra by less than statistical uncertainties. It is therefore not necessary to include the effect in the total systematic uncertainties.

In case of the 7 TeV analysis the influence of the following cut variations was checked :

- changing the minimum number of clusters in the TPC from 70 to 60 and 80,
- including the kink daughters,
- changing the maximum dca_{xy} from $7{\cdot}\sigma_{dca}$ to $5{\cdot}\sigma_{dca}$ and $9{\cdot}\sigma_{dca}$ (see Table 4.1),
- changing the maximum χ^2 per cluster in TPC from 4.0 to 3.5 and 4.5,
- removing the requirement Q>50 keV/(300μm) for the energy loss signal in the SSD and the SDD.

Only the pion spectra in the highest p_{t} bin showed an effect on the level of 1%, caused by the number of TPC cluster cut. As a result of this calculation a systematic uncertainty of 0.5% was assigned for all spectra in the entire p_{t} range.

4.4.4 Material budget

In order to estimate the effect of the uncertainty in the detector material on the spectra for the 900 GeV data sample, a simulation with increased detector material budget by 20% was performed. The efficiencies were calculated using that sample and compared to standard ones. The differences between the efficiencies were scaled to 7% uncertainty in the ALICE material budget to calculate the final uncertainties for measured spectra. The uncertainties are p_{t} independent.

The same studies were also performed for the 7 TeV data sample. In this case, the simulations with an increase and a decrease of the material budget by 10% were performed, and a better estimate of the uncertainty of the material budget was also known (-6.2%;3.4%). As a final value of the uncertainties the changes of the efficiencies caused by the lower limit (-6.2%) of the material uncertainty were used. For the kaon and pion spectra the number of simulated events was large enough to make those uncertainties p_{t} dependent. The uncertainties are summarized in Table 4.2. In case of the 7 TeV pion and kaon spectra the values for the lowest and highest p_{t} bin are given.

Table 4.2: Uncertainties due to the uncertainty of the material budget.

particle	uncertainty [%]	
	900 GeV	7 TeV
π^+	0.5	0.8-0.1
π^-	0.5	0.8-0.1
K^+	0.7	1.5-0.1
K^-	1.1	2.5-0.8
p	0.6	0.7
$\overline{\mathrm{p}}$	1.5	1.7

4.4.5 Total uncertainty

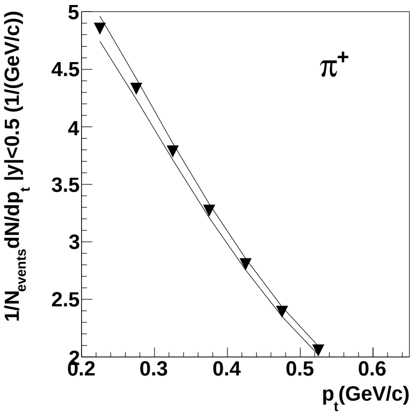
Tables 4.3, 4.4 summarize all parts of systemic uncertainties. If an uncertainty depends on p_{t}, the values for the first and last bin are given. The final uncertainties were calculated as the sum of all contributions in quadrature.

Table 4.3: Total systematic uncertainties of the spectra in the 900 GeV analysis.

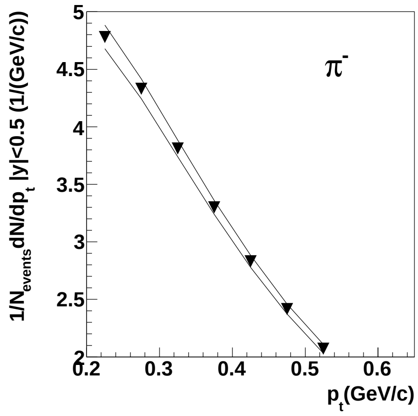
uncertainty	π^+ [%]	π^- [%]	K^+ [%]	K^- [%]	p [%]	$\overline{\mathrm{p}}$ [%]
function shape	0.9-1.6	1-1.5	1.7-3.5	2.1-0.2	3.8-5.7	2.6-4.4
fit range	<stat. uncert.		0.0-4		0-4	
contamination	0.4-0.2		0		<stat. uncert.	
material budget	0.5		0.7	1.1	0.6	1.5
sum	1.1-1.7	1.2-1.5	1.9-5.4	2.4-4.2	3.1-6.4	3.0-6.1

4.5 Final spectra and uncertainties

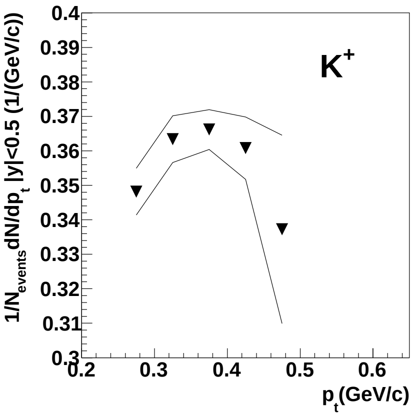
All spectra for both data samples are presented in Figures 4.23 and 4.24 where lines indicate the systematic uncertainties. The 7 TeV spectra are normalized to the number of events passing the ALICE physics selection which were not tagged as pile-up and the 900 GeV spectra to the number of inelastic events, using the approach from [48, 49]. In case of the 900 GeV spectra the statistical uncertainties dominate while for the 7 TeV spectra the systematic uncertainties are larger than the statistical because the 7 TeV analysis was done on the bigger data sample.



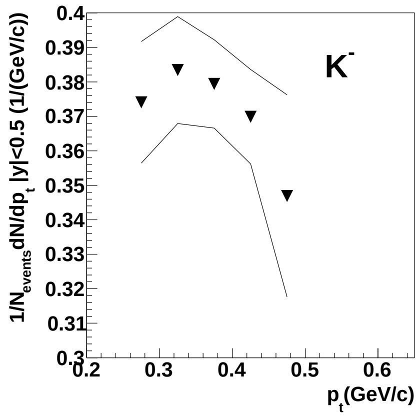
(a) π^+ spectrum



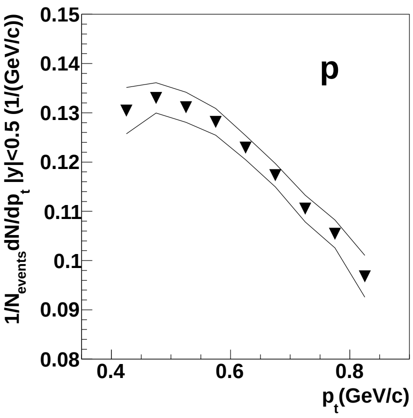
(b) π^- spectrum



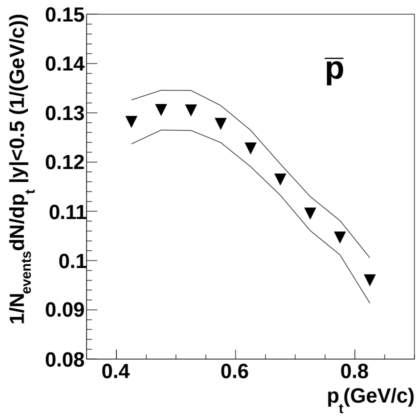
(c) K^+ spectrum



(d) K^- spectrum

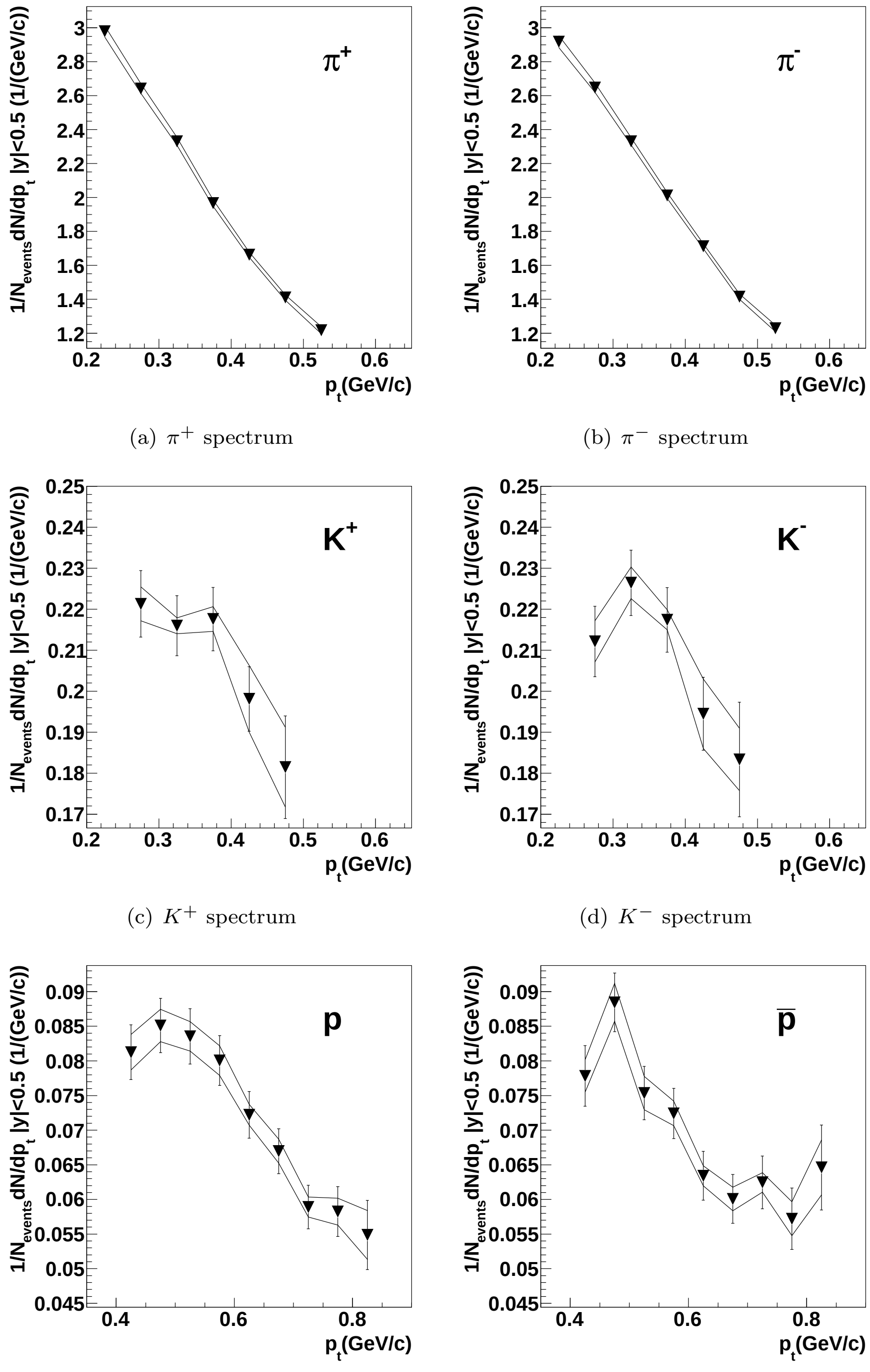


(e) p spectrum



(f) $\overline{p}$ spectrum

Figure 4.23: Spectra in proton-proton collisions at $\sqrt{s}$=7 TeV. Normalized to the total number of events used in the analysis. Lines show the systematic uncertainties.



(a) π^+ spectrum

(b) π^- spectrum

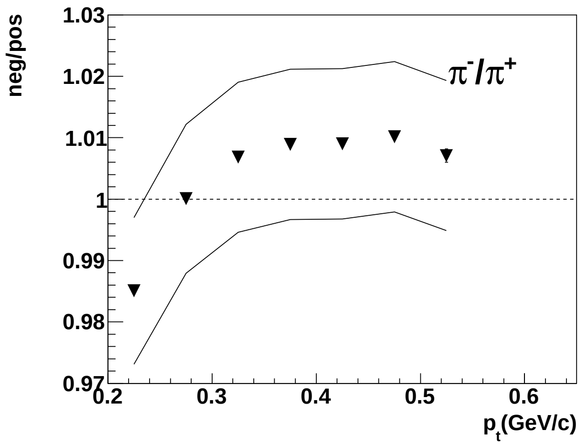
(c) K^+ spectrum

(d) K^- spectrum

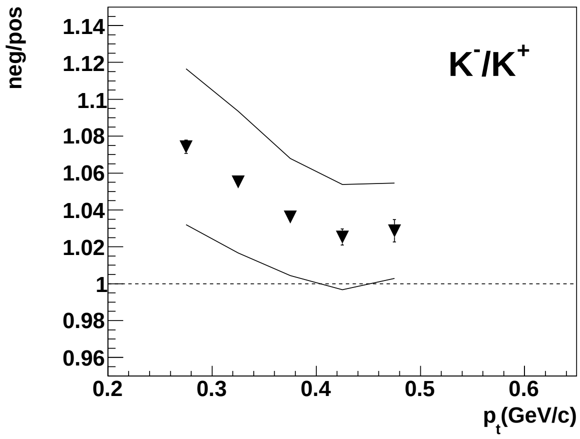
(e) p spectrum

(f) $\overline{p}$ spectrum

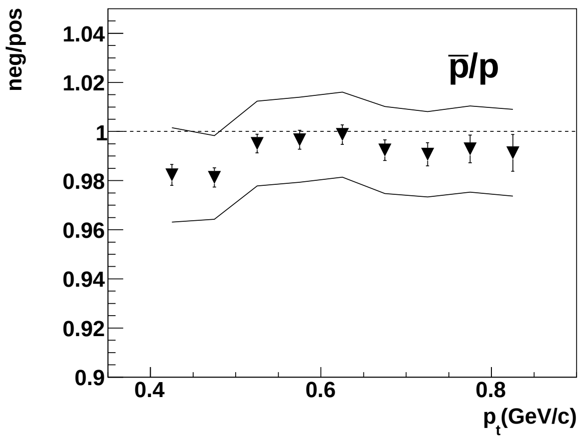
Figure 4.24: Spectra in proton-proton collisions at $\sqrt{s}$=900 GeV. Normalized to the total number of the inelastic collisions. Lines show the systematic uncertainties



(a) π^- over π^+ ratio

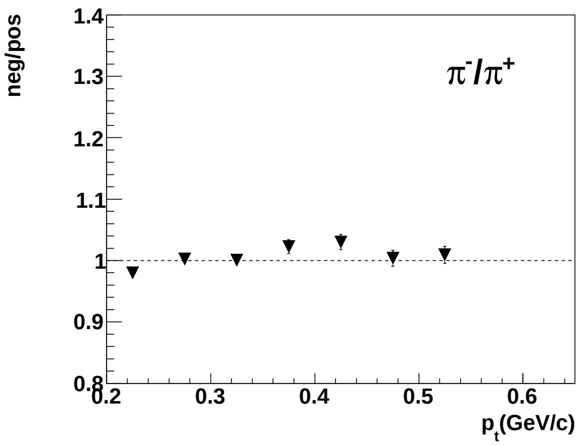


(b) K^- to K^+ ratio

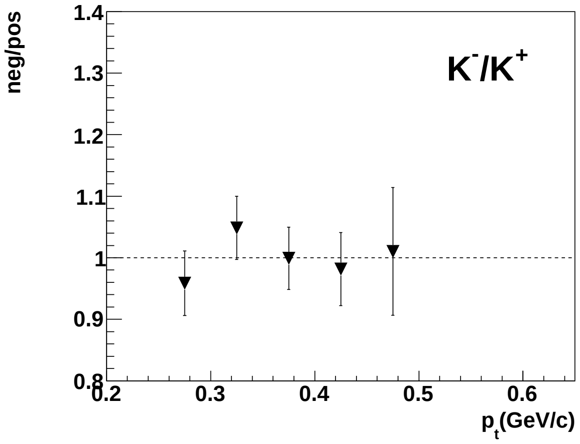


(c) $\overline{p}$ to p ratio

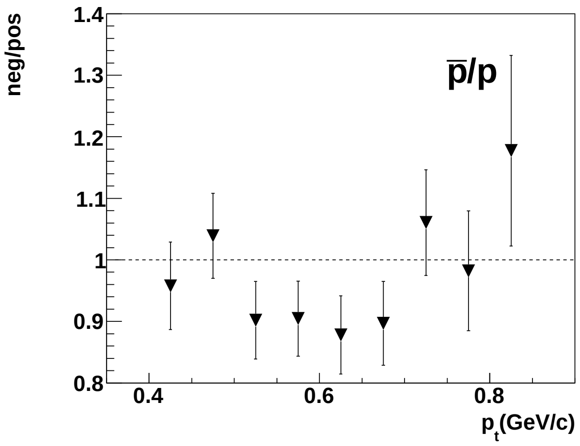
Figure 4.25: Ratio of the negative to positive particle spectra in proton-proton collisions at $\sqrt{s} = 7$ TeV. Lines show the systematic uncertainties.



(a) π^- to π^+ ratio



(b) K^- to K^+ ratio



(c) $\overline{p}$ to p ratio

Figure 4.26: Ratio of the negative to positive particle spectra in proton-proton collisions at $\sqrt{s} = 900$ GeV.

Table 4.4: Total systematic uncertainties of the spectra in the 7 TeV analysis.

uncertainty	π^+ [%]	π^- [%]	K^+ [%]	K^- [%]	p [%]	$\overline{p}$ [%]
function shape	1		1		1	
fit range	0-0.7		0.2-8		0-4	
contamination	1		0		3.3-1.1	2.4-0.7
ITS-TPC matching	0.5		0.5		0.5	
cut variations	0.5		0.5		0.5	
material budget	0.8-0.1		1.5-0.1	2.5-0.8	0.7	1.7
interactions with detector material	1.4-0.7	1.2-0.7	0	3.8-2.3	0	1.4
sum	2.3-1.8	2.1-1.8	1.9-8.1	4.7-8.5	3.6-4.4	3.5-4.8

4.6 Negative to positive ratios

The ratios of a spectrum for negative particles to a spectrum for positive particles were calculated for all types of particles. Not all parts of the systematic uncertainties of the spectra are relevant for the ratio uncertainties, some of them cancel out, because they are the same for positive and negative particles. Systematic uncertainties of the ratios were calculated for the 7 TeV spectra. The parts of the spectra uncertainties connected with the material budget, the track cut variations, the contamination correction and interactions with the detector material were used. Only the uncertainties coming from interactions with the detector material were propagated directly from the spectra measurement, all others were calculated specially for the ratios to avoid overestimating them. For 900 GeV ratios the systematic uncertainties were not calculated because the statical uncertainties are dominant. Ratios are shown in Figures 4.25(lines indicate the systematic uncertainties), 4.26. The antiproton to proton ratio for the 7 TeV data sample is consistent with the result reported in [42]. For the pion and kaon 7 TeV ratios, a deviation from unity is seen. For the 900 GeV ratios due to the statistical uncertainties this cannot be concluded. The deviation is probably caused by the incorrect description of interactions of particles with the detector material. This effect should be studied in more detail because it is the main source of systematic uncertainties for the pion and kaon ratios. Redoing the 7 TeV analysis for a different magnetic field configuration could also improve the understanding of the systematic uncertainties on the percent level.

4.7 Conclusions

The developed method measures spectra with an accuracy of few percent, but it still can be improved by better evaluation of few corrections. The contamination correction can be an example, it is the main source of the systematic uncertainties of the proton and antiproton spectra. Those uncertainties could be reduced if in the reconstruction and simulation a proper treatment of the "ExB" effect would be used, and if the Σ^+ ($\overline{\Sigma^+}$) to Λ ($\overline{\Lambda}$) ratio would be better known. The ratio of the negative kaon spectrum to positive kaon spectrum indicated problem with the description of the interactions of particles with the detector material, which should be further studied.

Chapter 5

Combined Spectra

5.1 Introduction

In this chapter the 7 TeV spectra calculated in the previous chapter (ITSTPC) are compared and combined with preliminary spectra (as shown at the Quark Matter Conference in May 2011), measured using different methods and PID detectors. This enlarges the p_t coverage of the measurement. The mean p_t ($\langle p_t \rangle$) and integrated yields (dN/dy) are calculated for each particle type using the combined spectra. Those results are compared with previously published results and with predictions of event generators. A corresponding analysis for the 900 GeV data is presented in [51], including spectra from the previous chapter.

The 7 TeV results which are presented in this chapter, are an update of the results which were shown on the Quark Matter Conference, but they are not official results of the ALICE collaboration. The paper which will contain the official results is under preparations.

5.2 Other methods of measurement

Beside the ITSTPC method the ALICE experiment uses other methods to measure p_t spectra of pions, kaons protons and antiprotons. Each of the methods contributes to partially overlapping but also complementary p_t ranges. Those ranges are determined by the PID capabilities of the detectors which are used for particle identification. All methods use the same rapidity selection $|y| < 0.5$.

5.2.1 ITS standalone (ITSsa)

A measurement is done using standalone ITS tracks obtained using all ITS clusters (see Section 2.4.2). The calculation of the raw yield is performed by selecting particles with the measured energy loss signal in the ITS close to the expected one for each particle species. This method gives the possibility to measure spectra up to p_t of a few hundred of MeV/c. Calculations of the pion, proton and antiproton contaminations are done in the same way as for the ITSTPC method.

5.2.2 TPC-TOF combined measurement (TPCTOF)

The TPC-TOF combined measurement is done using global tracks. The raw yield is calculated by selecting particles with the measured energy loss signal in the TPC (see Figure 2.4) and the measured time of flight in the TOF (see Figure 2.5) for higher p_t close to the expected one for each particle type. The pion, proton and antiproton contaminations are calculated using the idea of the dca_{xy} fit method.

5.2.3 TOF

The TOF measurement is done using global tracks. The calculation of the raw yield is performed by fitting an analytical function to the time of flight distribution in a given p_t bin. For this method all contaminations are calculated based on simulations. It is therefore only used for larger p_t, where the contamination is low.

5.3 Combined spectra in proton-proton collisions at $\sqrt{s} = 7$ TeV

5.3.1 Combining spectra

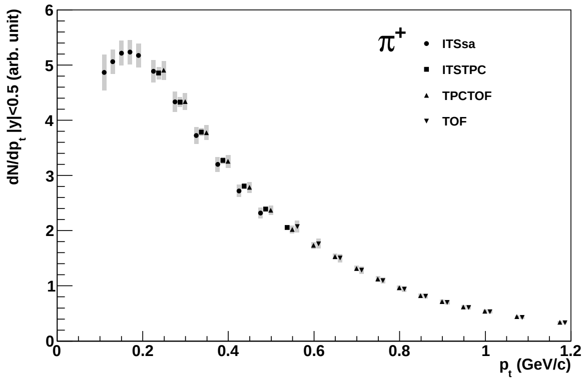
Spectra measured by the 4 methods were compared and combined to obtain the final spectra. The comparison of the spectra is shown in Figure 5.1. The results agree on the level of systematic uncertainty of each method. Table 5.1 lists the ranges where spectra from each method were used.

Table 5.1: The p_t ranges where spectra from each method were used during combining them.

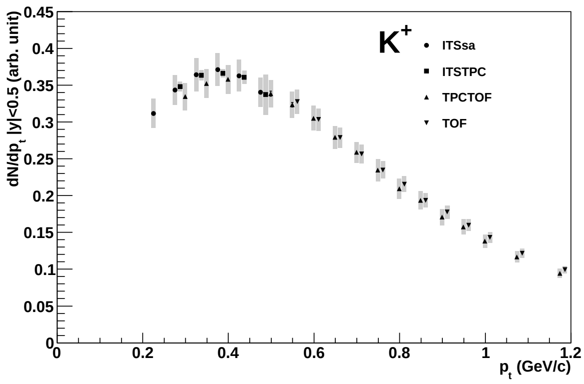
	ITSsa	ITSTPC	TPCTOF	TOF
ranges for π (GeV/c)	[0.1-0.5]	[0.2-0.55]	[0.2-1.3]	[0.5-1.6]
ranges for K (GeV/c)	[0.2-0.5]	[0.25-0.5]	[0.25-1.3]	[0.5-1.6]
ranges for p (GeV/c)	[0.3-0.55]	[0.4-0.85]	[0.45-1.7]	[0.9-2.5]

Spectra were combined using the weighted average, where weights were 1 over the square of the total uncertainty (statistical and systematic added in quadrature). The uncertainties of the combined spectra are the weighted averages of the total uncertainties of the spectra. They are used in all fitting procedures described in this chapter, so the uncertainties which will be quoted as fit results have both parts systematic and statistical. Due to the preliminary character of the spectra many uncertainties are overestimated. Several parts of the systematic uncertainties of the ITSTPC, TPCTOF and TOF methods are common. For each method total uncertainties are also point-to-point correlated. Those aspects lead to an underestimation of the χ^2/ndf in all fit procedures.

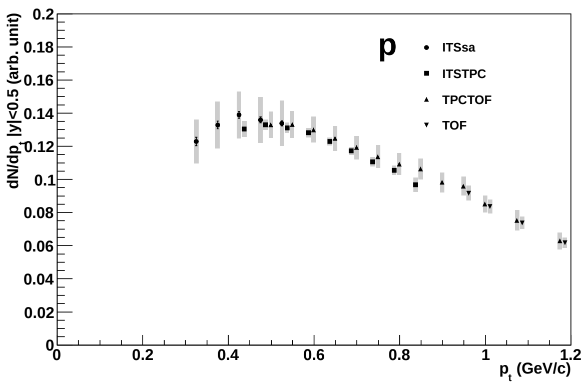
In the final analysis the procedure of combining the spectra should be changed. All analyses should provide the systematic uncertainties, which are specific for each method (e.g. PID procedure). Those uncertainties should be used when



(a) pion spectra



(b) kaon spectra



(c) proton spectra

Figure 5.1: Spectra for positive particles measured by the 4 methods in proton-proton collisions at $\sqrt{s} = 7$ TeV. A p_{t} range where all methods can be applied is shown. Points are shifted from bin centres for a presentation purpose. The grey bands indicate systematic uncertainties.

combining the spectra. The rest of the uncertainties which are common for several analyses (e.g. uncertainties related to the tracking for ITSTPC, TPCTOF and TOF methods) should be added to the uncertainties of the combined spectra.

After combining the spectra, they were normalized to the number of inelastic collisions. This was done using the trigger efficiency which is 0.86±0.07 [50]. This uncertainty includes the 2% pile-up uncertainty described in Section 4.3.10.

5.3.2 Lévy-Tsallis fits

The combined spectra were fitted using the Lévy-Tsallis function, the same form as in [51]:

$$\frac{\mathrm{d}^2N}{\mathrm{d}y\mathrm{d}p_\mathrm{t}} = p_\mathrm{t}\frac{\mathrm{d}N}{\mathrm{d}y}\frac{(n-1)(n-2)}{nC(nC+m_0(n-2))}\left(1+\frac{m_\mathrm{t}-m_0}{nC}\right)^{-n}, \tag{5.1}$$

where $\frac{\mathrm{d}N}{\mathrm{d}y}$ (integrated yield) n and C are fit parameters, m_0 is the particle mass. This equation is the same as Equation 1.11 when $n = 1/(q-1)$ and $nC = m_0 + nT$. The fits are shown in Figure 5.2. The parameters obtained from the fits and χ^2/ndf values are summarized in the following table:

Table 5.2: Results of the fit of Lévy-Tsallis function (Equation 5.1) to the combined spectra.

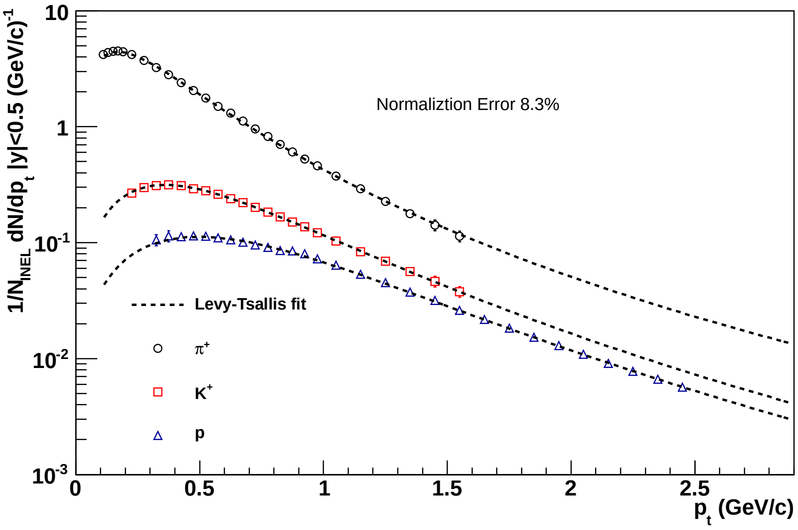
particle	dN/dy for $\|y\| < 0.5$	C [GeV/c]	n	χ^2/ndf
π^+	2.26 ± 0.02	0.116 ± 0.003	5.6 ± 0.2	4.63/24
π^-	2.23 ± 0.02	0.122 ± 0.003	5.9 ± 0.2	3.9/24
K^+	0.277 ± 0.004	0.195 ± 0.008	6.5 ± 0.8	0.43/19
K^-	0.274 ± 0.005	0.194 ± 0.010	6.9 ± 1.0	2.14/19
p	0.125 ± 0.001	0.213 ± 0.007	6.3 ± 0.4	6.38/26
$\overline{\mathrm{p}}$	0.124 ± 0.001	0.211 ± 0.007	6.3 ± 0.5	3.32/26

The values of the χ^2/ndf are small as expected.
Other functions were also fitted to the spectra:

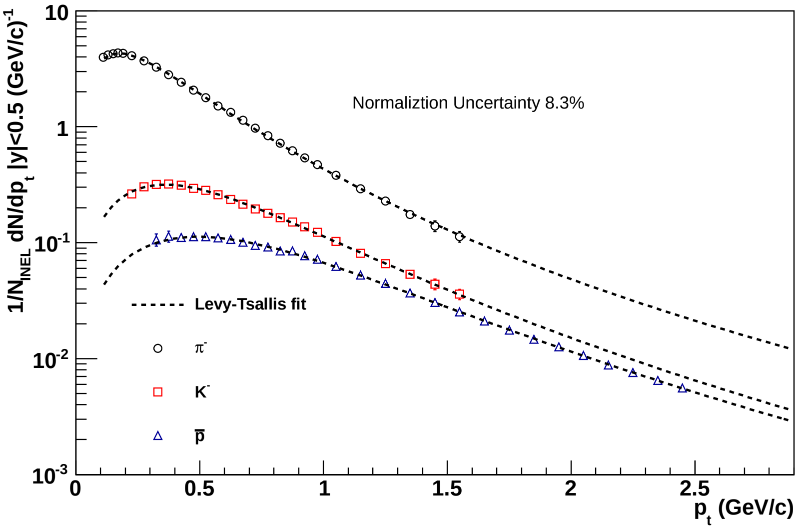
$$\frac{\mathrm{d}^2N}{\mathrm{d}y\mathrm{d}p_\mathrm{t}} = p_\mathrm{t}\frac{\mathrm{d}N}{\mathrm{d}y}\frac{\exp(-p_\mathrm{t}/T)}{T^2}, \tag{5.2}$$

$$\frac{\mathrm{d}^2N}{\mathrm{d}y\mathrm{d}p_\mathrm{t}} = p_\mathrm{t}\frac{\mathrm{d}N}{\mathrm{d}y}\frac{1}{T\exp(-m_0/T)(m_0+T)}\exp(-m_\mathrm{t}/T), \tag{5.3}$$

Those functions were used to describe the measured spectra at lower energies. Figure 5.3 shows the ratios of data points to the fitted functions (only for positive particles) which were calculated to compare how well each of the functions describe the spectra. The Lévy-Tsallis function gives the best description of the spectra in the used p_t ranges; all the experimental points agree with the fitted functions within 5%. The two other functions cannot describe the spectra.

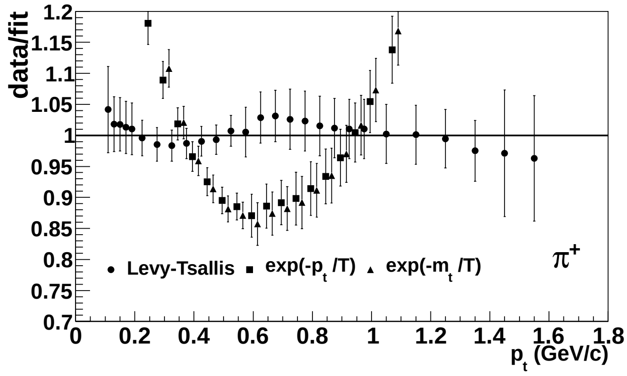


(a) Positive particles

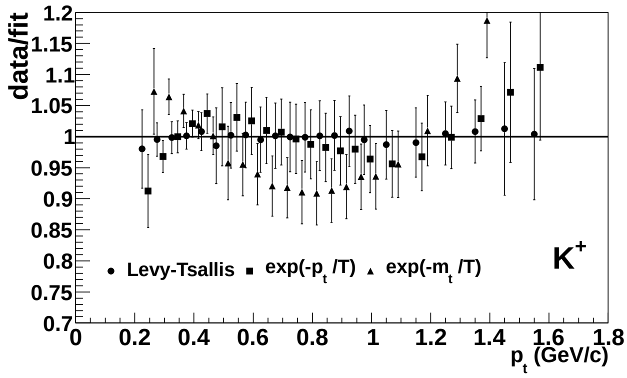


(b) Negative particles

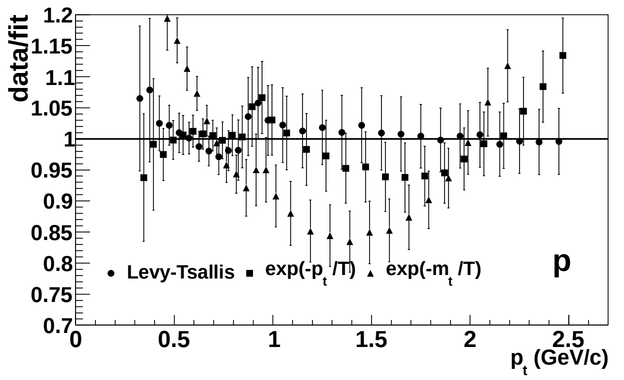
Figure 5.2: Combined p_t spectra fitted with the Lévy-Tsallis function (Equation 5.1).



(a) for π^+ spectrum



(b) for K^+ spectrum



(c) for p spectrum

Figure 5.3: Ratios of data points to different fitted functions (Equations 5.1,5.2 and 5.3) for positive particles. Points are shifted from bin centres for a presentation purpose.

5.4 Calculation of $\langle p_{\mathrm{t}} \rangle$ and integrated yields

The parameter $\mathrm{d}N/\mathrm{d}y$ from the fitted Lévy-Tsallis function was taken as the integrated yield. The results of the Lévy-Tsallis fit were also used to evaluate $\langle p_{\mathrm{t}} \rangle$. As a cross-check the yields and $\langle p_{\mathrm{t}} \rangle$ were calculated using the data points in the p_{t} ranges covered by the measurement and the fitted Lévy-Tsallis functions outside them. The differences between the values calculated using those two methods were below 1%.

The yields in the measured p_{t} ranges are between 80% and 90% of $\mathrm{d}N/\mathrm{d}y$ of the fitted Lévy-Tsallis functions. To estimate the systematic uncertainties caused by the extrapolation of the spectra to the p_{t} regions not covered by the measurement other ways of calculating the integrated yield and $\langle p_{\mathrm{t}} \rangle$ were also checked. The results of the fits of the functions given by Equations 5.2, 5.3 were used to calculate $\mathrm{d}N/\mathrm{d}y$ and $\langle p_{\mathrm{t}} \rangle$. Two ways of calculating the yields and the $\langle p_{\mathrm{t}} \rangle$ were used: only a fit result or the combination of a fit result and the data points. $\mathrm{d}N/\mathrm{d}y$ and $\langle p_{\mathrm{t}} \rangle$ were also calculated using the shapes of the spectra from few event generators to extrapolate the spectra. The spectra from event generators were rescaled to match the measured spectra in the lowest and the highest p_{t} bin. The rescaling was done independently for the lowest and the highest p_{t} bin. This gives the possibility to reduce the effect of the normalization uncertainty during the extrapolation. The results of the different extrapolations are compared to the standard values of $\mathrm{d}N/\mathrm{d}y$ and $\langle p_{\mathrm{t}} \rangle$ in Figure 5.4. Based on these studies a 5% uncertainty was assigned as the extrapolation uncertainty, indicated by the lines in Figure 5.4. The values obtained from the functions given by Equations 5.2, 5.3 are outside the 5% band, but as was shown in Figure 5.3 those functions do not describe the data accurately, especially for high p_{t} where they underestimate data.

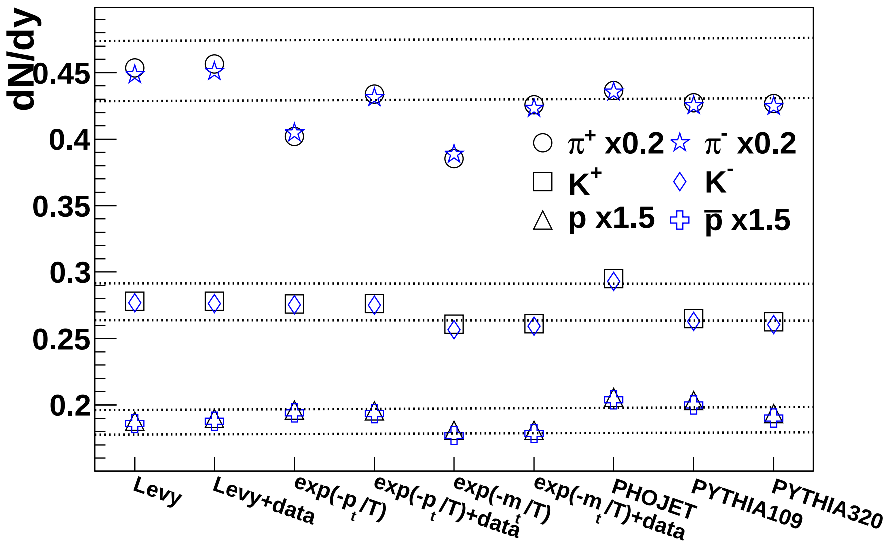
The final uncertainty of the integrated yield is the quadratic sum of the 5% extrapolation uncertainty, the uncertainty which results from the fit procedure and the 8.3% normalization uncertainty. The final value of the $\langle p_{\mathrm{t}} \rangle$ is the average value of positive and negative particle $\langle p_{\mathrm{t}} \rangle$. The $\langle p_{\mathrm{t}} \rangle$ uncertainty is the quadratic sum of the 5% extrapolation uncertainty and the uncertainty which results from the fit procedure. All values are summarized in Table 5.3.

Table 5.3: The values of the integrated yields ($\mathrm{d}N/\mathrm{d}y$) and the $\langle p_{\mathrm{t}} \rangle$.

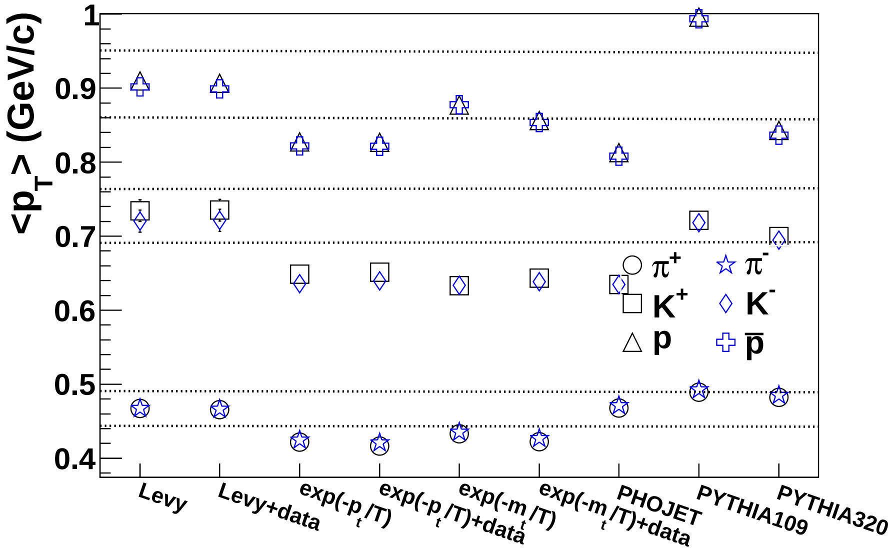
particles	$\mathrm{d}N/\mathrm{d}y$ for $\lvert y \rvert < 0.5$ positive	$\mathrm{d}N/\mathrm{d}y$ for $\lvert y \rvert < 0.5$ negative	$\langle p_{\mathrm{t}} \rangle$ [GeV/c]
π^+ , π^-	2.26 ± 0.22	2.23 ± 0.22	0.46 ± 0.03
K^+ , K^-	0.28 ± 0.03	0.27 ± 0.03	0.73 ± 0.04
p , $\overline{\mathrm{p}}$	0.125 ± 0.012	0.124 ± 0.012	0.90 ± 0.05

5.5 Comparison with data at lower energies

Figure 5.5 shows a comparison of the calculated $\langle p_{\mathrm{t}} \rangle$ in proton-proton collisions at $\sqrt{s}$ =7 TeV with previous results ([51] and references therein). The $\langle p_{\mathrm{t}} \rangle$ increases with the mass of the particle as was already seen by previous experiments at lower energies. This increase looks linear but to judge if this is the



(a) $\mathrm{d}N/\mathrm{d}y$ values



(b) $\langle p_\mathrm{t} \rangle$ values

Figure 5.4: Values of $\mathrm{d}N/\mathrm{d}y$ and $\langle p_\mathrm{t} \rangle$ calculated using different extrapolation scenarios. Lines indicate a 5% uncertainty which was assigned as the extrapolation uncertainty. A label with "+data" means that the calculation was done using data points and a fitted function in a range not covered by the measurement. A label without "data" means that the plotted value was calculated using a fitted function. Labels with PYTHIA or PHOJET mean that for the extrapolation a generated spectrum was used.

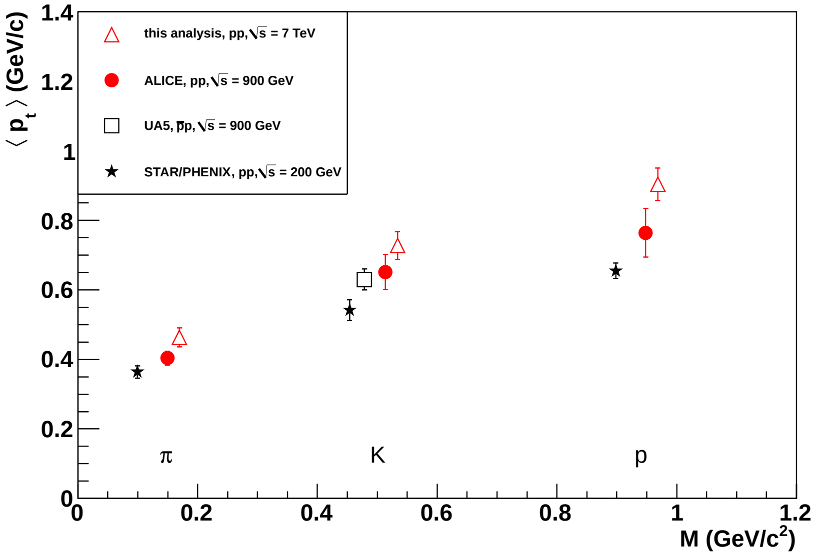


Figure 5.5: Comparison of $\langle p_t \rangle$ as function of mass to results of previous experiments ([51] and references therein).

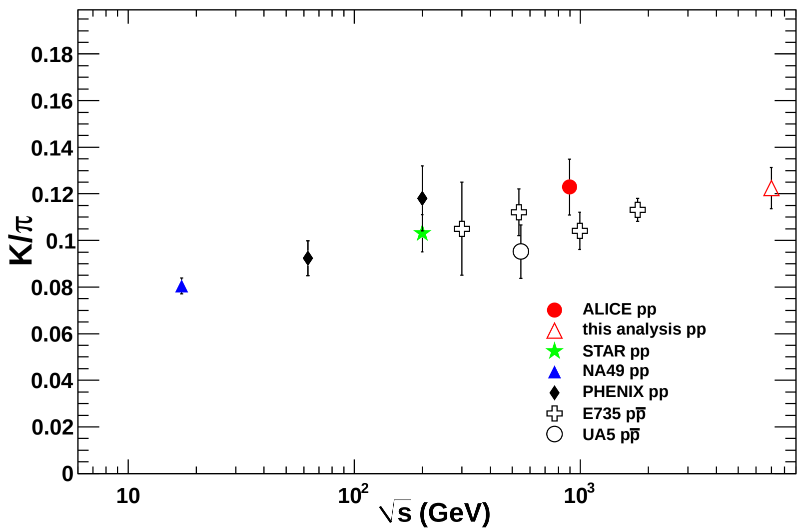


Figure 5.6: Ratio of the sum of the kaon integrated yields to the sum of the pion integrated yields at mid-rapidity as function of the energy of colliding protons ([51] and references therein).

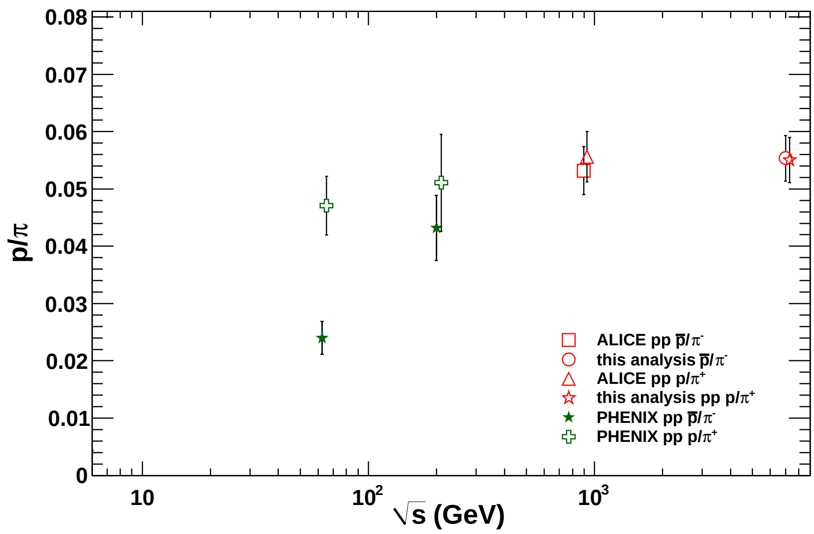


Figure 5.7: Ratios the proton and antiproton integrated yields to the pion integrated yields at mid-rapidity as function of the energy of colliding protons [52].

case, the results for other particles (e.g. Λ) are needed. The $\langle p_{\rm t}\rangle$ increases by a small amount with the collision energy for all the particles. This increase can be caused by the increases of $\mathrm{d}N/\mathrm{d}\eta$ between two collision energies as reported in [54].
As a second step the ratios of the integrated yields were compared with results at lower energies. For the ratios the normalization uncertainties cancel out. Figure 5.6 shows the energy dependence of the ratio $(K^+ + K^-)/(\pi^+ + \pi^-)$ at mid-rapidity ([51] and references there). Figure 5.7 shows the energy dependence of the ratios p/π^+ and $\bar{p}/\pi^-$ at mid-rapidity. Results of the PHENIX experiment are taken from [52]. Based on these results it can be concluded that: as the energy in proton-proton collisions increases from 900 GeV to 7 TeV, no significant change in the relative particle yields is observed. This observation is also valid for the lower energies with the exception of the antiproton yield.

The predictions from the thermal models [53] with $\gamma_s = 0.6$ and T=170 MeV for proton-proton collisions at $\sqrt{s}$ =10 TeV (predictions were available only for this energy) are:

- $(K^+ + K^-)/(\pi^+ + \pi^-)$=0.116,
- p/π^+=0.08, $\bar{p}/\pi^-$=0.074.

Assuming that particles ratios stayed constant from the 7 TeV to 10 TeV energy of colliding protons that would imply that the thermal model overestimates the proton and antiproton productions or underestimate the pion and kaon productions.

5.6 Comparison with the event generators

The combined spectra were compared to predictions of PHOJET and a few PYTHIA tunes. Including the PYTHIA tune 350 [55] (Perugia 2011), which was developed using the first LHC results concerning the particle production in the proton-proton collisions (e.g. [51]). Figure 5.8 shows the comparison between measured spectra and the event generator predictions for positive particles (the situation for negative particles is similar). For all figures the normalization uncertainty (8.3%) should be taken into an account. For the pion spectra, the PYTHIA tunes 320, 350 can describe the shape of the spectra for $p_{\mathrm{t}} > 0.5$ GeV/c. Tune 320 describes the yield after taking into account the normalization uncertainty, while tune 350 misses the yield by much more than the normalization uncertainty. The measured kaon spectra are in a good agreement with the PYTHIA 350 tune predictions. For proton spectra, the situation is similar to the pion case. The shape of the spectra for $p_{\mathrm{t}} > 1.0$ GeV/c is in agreement with PYTHIA tunes 320, 350, but in this case the yield is reproduced by tune 350. To study this more, the ratios of the sum of the kaon spectra to the sum of the pion spectra (K/π) and the sum of proton and antiproton spectra to the sum of the pion spectra (p/π) were calculated. These ratios are shown in Figures 5.9, 5.10, where it can be seen that models do not describe the K/π ratio at $p_{\mathrm{t}} > 0.5$ GeV/c. The p/π ratio is reproduced for higher p_{t} by the PYTHIA 109 tune , but this tune does not describe the individual spectra.

To see what helps to describe the kaon spectra, the PYTHIA hadronization parameters related with the strange quark production were compared between the tunes 320 and 350 (see Table 5.4).

Table 5.4: Hadronization parameters of the PYTHIA tunes 320 and 350 .

parameter	tune 320	tune 350	description
PARJ(1)	0.073	0.087	suppression of diquark production compared to quark-antiquark pair production;
PARJ(2)	0.2	0.19	suppression of s quark production compared to u, d quark production;
PARJ(3)	0.94	0.95	extra suppression of diquark production with at least one strange quark compared to normal suppression of the strange quark production;
PARJ(4)	0.032	0.043	suppression of diquark spin 1 states compared to the spin 0 states;
PARJ(6)	0.5	1.0	extra suppression of strangeness pair production for baryons in the popcorn model;
PARJ(7)	0.5	1.0	extra suppression of production of strange mesons in the popcorn model;
PARJ(12)	0.4	0.4	probability that a strange meson has spin 1;
PARJ(21)	0.313	0.33	spread of the Gaussian distribution which generates p_{t} of hadrons;

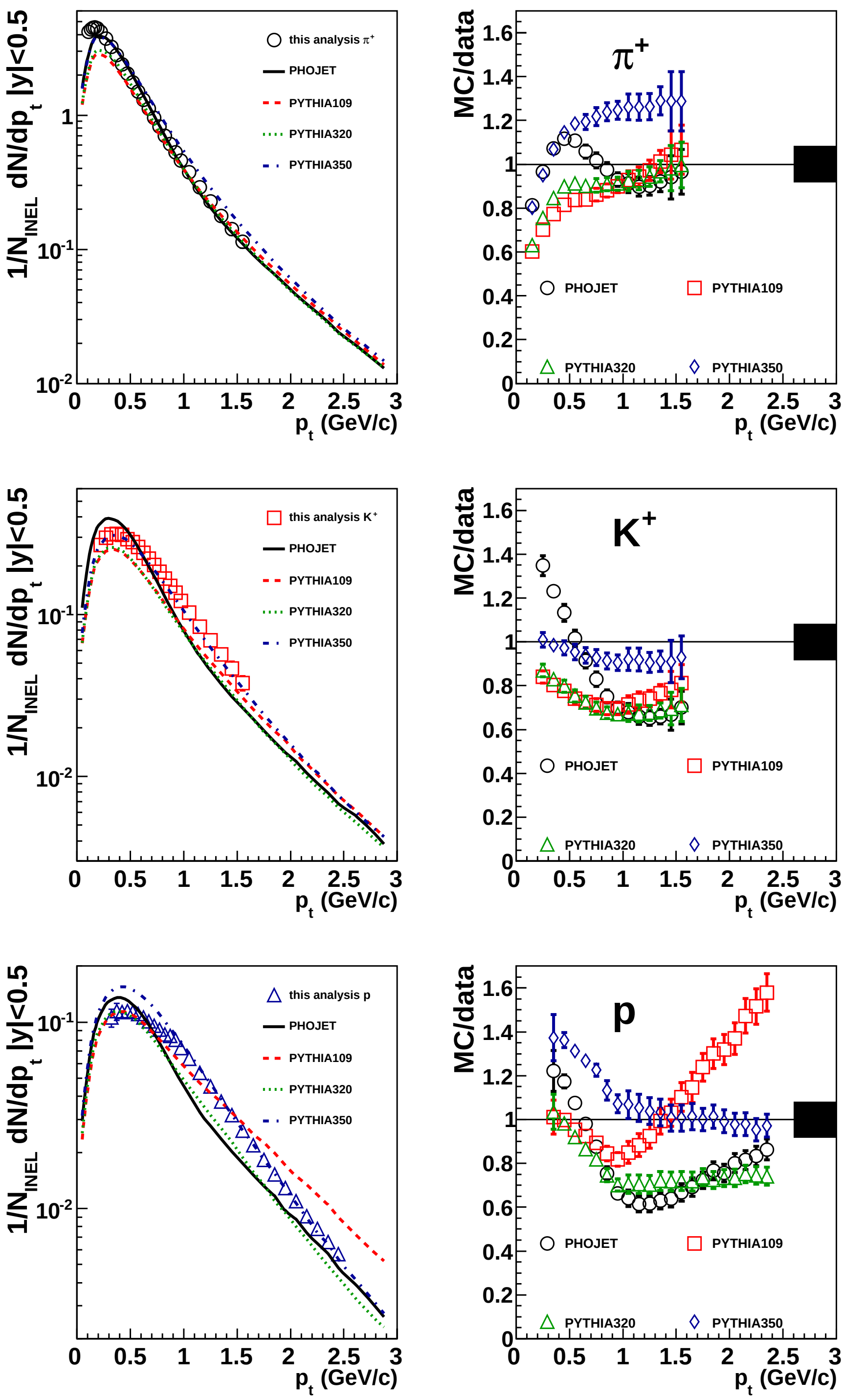


Figure 5.8: Comparison between spectra and predictions from different event generators for positive particles. The black boxes on the ratio figures indicate the size of the normalization uncertainty.

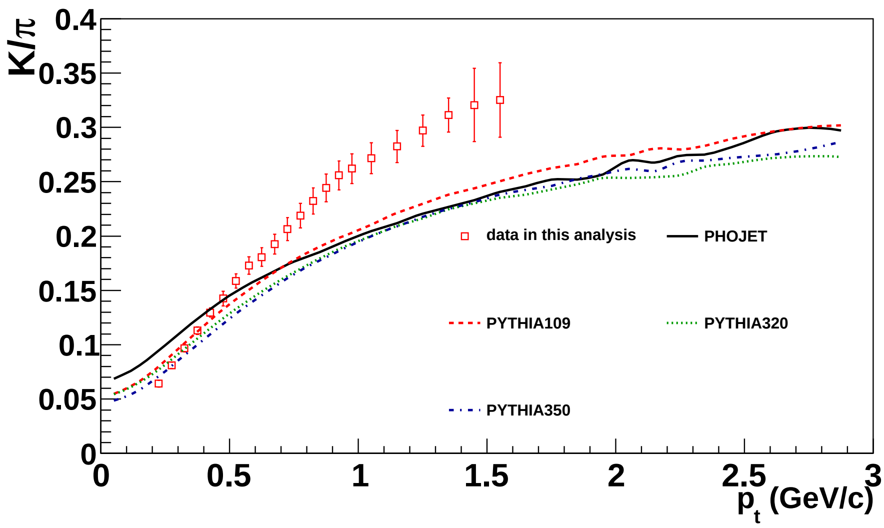


Figure 5.9: Ratio of the sum of kaon spectra to the sum of pion spectra as function of p_t compared with predictions from different event generators.

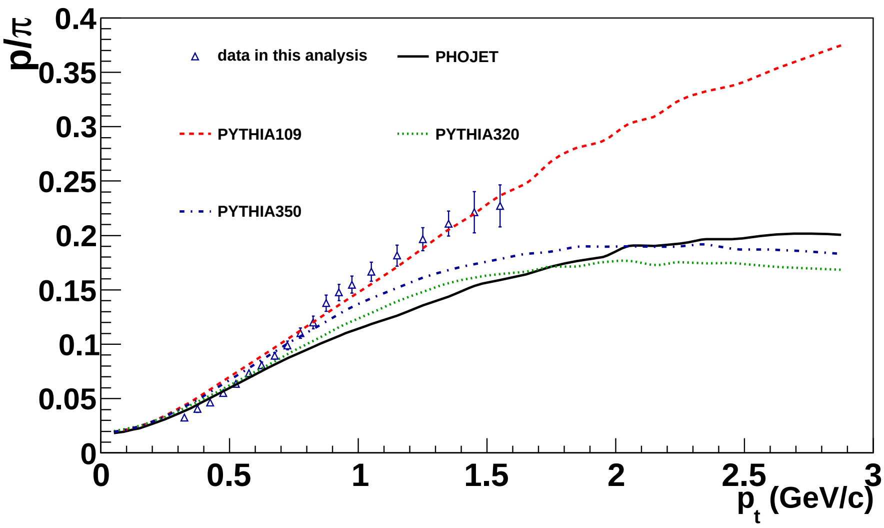


Figure 5.10: Ratio of the sum of proton and antiproton spectra to the sum of pion spectra as function of p_t compared with predictions from different event generators

Even though some parameters responsible for the strangeness production (e.g. PARJ(6) and PARJ(7)) are quite different between the tunes, the K/π ratio (see Figure 5.9) does not change much. This suggests that the better description of the kaon spectra in the tune 350 is achieved by increasing the total particle production and not by the change of the hadronization parameters.

5.7 Blast Wave fits

All six spectra were fitted simultaneously using the Blast Wave (BW) and Blast Wave Tsallis (BWT) functions given by Equations 1.13, 1.15. In case of the first form, the amplitudes, $\langle\beta\rangle$ and T were the fit parameters. For the BWT the amplitudes, T, and q were the fit parameters. The other parameters were constrained, $n = 1$ and $\langle\beta\rangle = 0$ ($\langle\beta\rangle = 0$ means that Equation 1.10 is fitted to all spectra, using common values for T and q). This is consistent with the approaches presented in [24, 56]. For both functions two scenarios of the fit were checked, using 6 amplitudes as fit parameters or only 3, a common amplitude for negative and positive particles of the same type. In the end it was decided to fit all 6 amplitudes. During the fit procedure the pion points with $p_{\mathrm{t}} < 0.5$ (GeV/c) were not used, which is a common procedure to avoid the effect of products of the resonance decays. Different starting points for the fits were used to probe the parameter space. Fits for many starting points converged to the same minimal values of the χ^2/ndf.

The results of the Blast Wave fit for positive particles are shown in Figure 5.11. Figure 5.12 shows ratios of data points to the fitted Blast Wave functions. Figure 5.13 and 5.14 show the same results for the Blast Wave Tsallis fit. The fitted parameters are summarized in Table 5.5 and compared to the fit results for proton-proton collisions at $\sqrt{s}$ =200 GeV (see Table 5.6) [56, 24] .

Table 5.5: Results of the Blast Wave fits.

BW	BWT
$T = 161 \pm 2$ MeV	$T = 104 \pm 1$ MeV
$\langle\beta\rangle = 0.303 \pm 0.003$	$(q - 1) = 0.103 \pm 0.001$
$n = 3.3 \pm 0.1$	
χ^2/ndf=0.57	χ^2/ndf=0.88

Table 5.6: Results of the Blast Wave fits to the measurement of the STAR collaboration of the spectra in proton-proton collisions at $\sqrt{s}$ =200 GeV.

BW	BWT baryons	BWT mesons
$T = 127 \pm 13$ MeV	$T = 97 \pm 10$ MeV	$T = 89 \pm 4$ MeV
$\langle\beta\rangle = 0.244 \pm 0.081$	$(q - 1) = 0.073 \pm 0.005$	$(q - 1) = 0.1 \pm 0.003$
$n = 4.3 \pm 1.7$		

The fit functions for pion and kaon spectra miss the high p_{t} part of the spectra. This does not have a significant influence of the value of χ^2/ndf, due to the overestimation of the uncertainties. To determine whether the spectra are

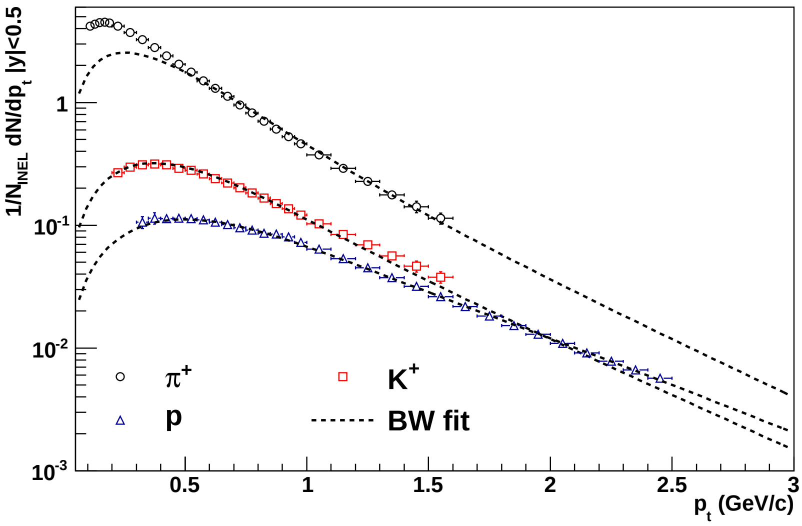


Figure 5.11: Result of the fit of the Blast Wave functions (Equation 1.13) to the positive particle spectra.

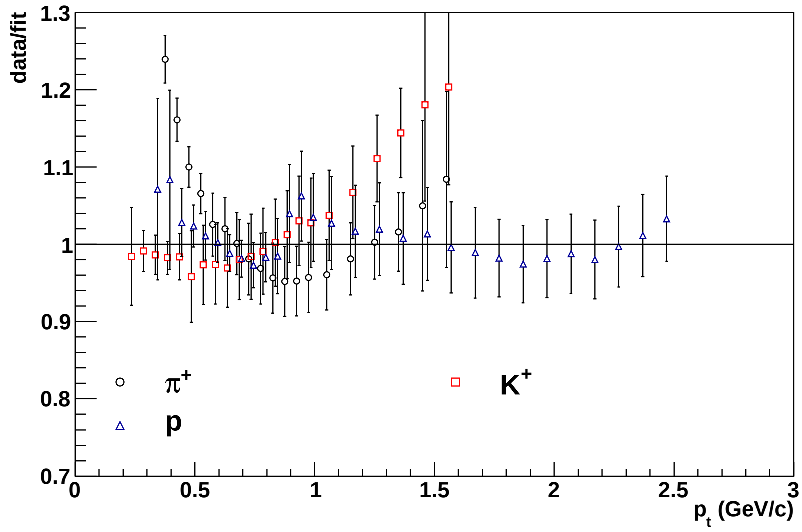


Figure 5.12: Ratios of data points to the fitted Blast Wave functions shown in Figure 5.11. Points are shifted from bin centres for a presentation purpose.

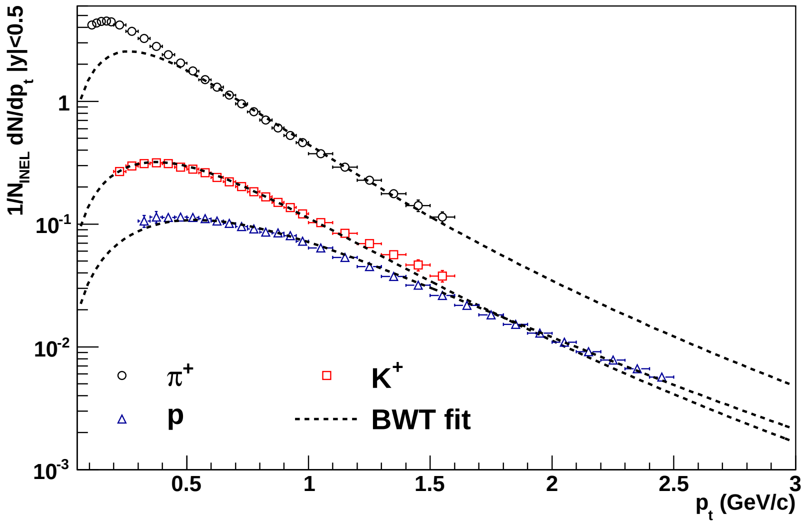


Figure 5.13: Result of the fit of the Blast Wave Tsallis functions (Equation 1.15) to the positive particle spectra.

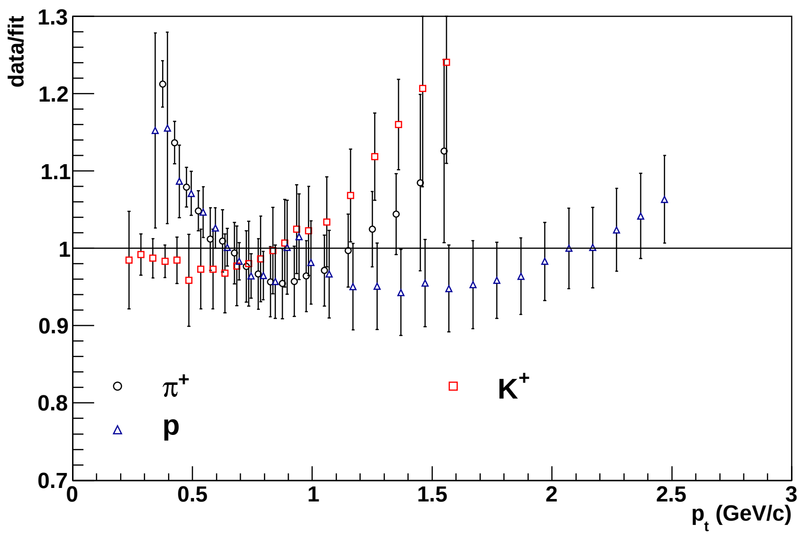


Figure 5.14: Ratios of data points to the fitted Blast Wave Tsallis functions shown in Figure 5.13. Points are shifted from bin centres for a presentation purpose.

described by the BW or BWT functions or not, this analysis should be repeated using spectra with a better estimation of the uncertainties. Comparison with the results for proton-proton collisions at $\sqrt{s}$=200 GeV currently shows an increase of T (for both fits) and $\langle\beta\rangle$ (for BW), while q is similar to the value for meson spectra in proton-proton collisions at $\sqrt{s}$=200 GeV.

5.7.1 Interpretation of the Blast Wave fits

The Blast Wave function was fitted to the spectra from the event generators, using the same p_{t} ranges as for the measured spectra. Figure 5.15 shows the result of the fit for positive particle spectra generated by the PYTHIA 109 tune. The obtained values of the parameters are: $T = 131.1 \pm 0.1$ MeV, $\langle\beta\rangle = 0.3503 \pm 0.1$, $n = 3.284 \pm 0.003$ (uncertainties are statistical). The value of $\langle\beta\rangle$ is comparable to the value from Table 5.5. The $\chi^2/\mathrm{ndf} = 28.4$ for this fit indicates that spectra are not described by the Blast Wave function. Figure 5.16 shows the ratios of data points to the fitted Blast Wave functions, which are similar as for data (see Figure 5.12). The deviation from unity of those ratios for data does not result in a large value of χ^2/ndf because the fit procedure includes systematic uncertainties. This studies show that with the current experimental precision the Blast Wave shape can describe spectra as generated by the PYTHIA tune 109. PYTHIA does not assume radial flow or thermalization, which is different from the Blast Wave assumptions. This means that the non-zero value of $\langle\beta\rangle$ (Table 5.5) cannot be directly interpreted as a signature of collective behaviour in the system created by proton-proton collisions at $\sqrt{s} = 7$ TeV.

5.8 Conclusions

The spectra measured by the 4 methods are in a good agreement but a better estimation of the systematic uncertainties is needed for all methods. The measured spectra are described by the Lévy-Tsallis function while the thermal motivated functions that have been used by other experiments at lower energies cannot describe the spectra. The relative particle yields are the same in proton-proton collisions at $\sqrt{s}$ =900 GeV and $\sqrt{s}$ =7 TeV. The $\langle p_{\mathrm{t}}\rangle$ for pions, kaons and protons slightly increases with collision energy. Commonly used event generators cannot describe all spectra at once. The Blast Wave model gives results consistent with those reported for the STAR data. The non-zero value of $\langle\beta\rangle$ which is the result of the Blast-Wave fit cannot be interpreted as a signature of collective behaviour.

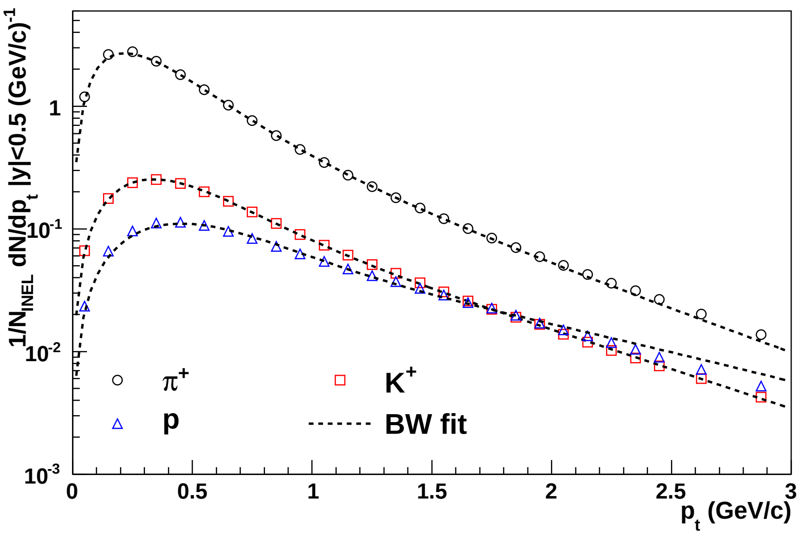


Figure 5.15: Result of the fit of the Blast Wave functions (Equation 1.13) to the positive particle spectra generated by the PYTHIA 109 tune.

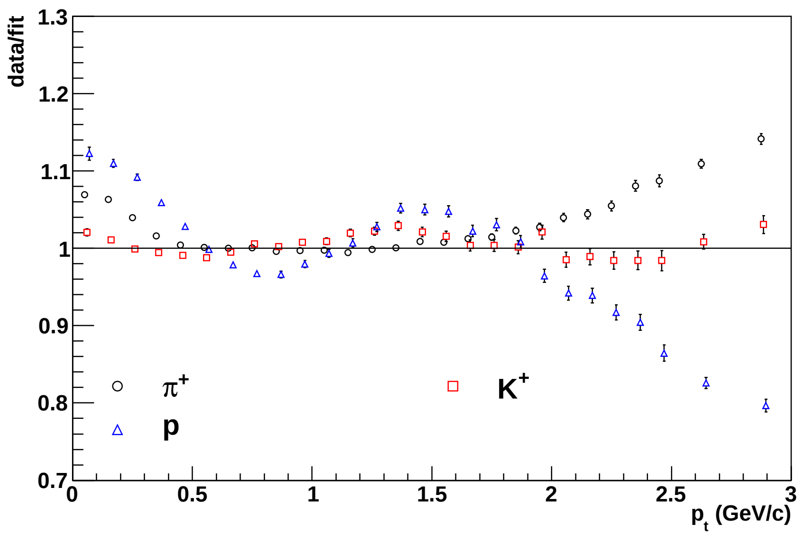


Figure 5.16: Ratios of the spectra generated by the PYTHIA 109 tune to the fitted Blast Wave functions shown in Figure 5.15. Points are shifted from bin centres for a presentation purpose.

Chapter 6

Event multiplicity dependence of spectra

6.1 Introduction

In the previous chapter the collective behaviour of the system created in the proton-proton collisions is not found in the minimum-bias data sample, but this behaviour could be a property of the high multiplicity collisions. In this chapter the method described in Chapter 4 is used to study spectra as a function of event multiplicity in proton-proton collisions. The correlation between the tracking efficiency and the multiplicity is investigated. A dependence of the spectra shapes on the multiplicity is observed. This analysis uses the same data and settings (track cuts and event selection) as the 7 TeV analysis in Chapter 4.

6.2 Two multiplicity estimators

Two independent methods to measure the event multiplicity (multiplicity estimators) are used in this analysis. The first estimator is the number of tracks fulfilling the standard ALICE track cuts (see Section 4.3.2) with $p_{\mathrm{t}} > 0.15$ GeV/c and $|\eta| < 0.8$ (ntracks estimator). The second estimator is the number of clusters in the second layer of the SPD corrected for the acceptance using the primary vertex z position (nSPD2 estimator). Figure 6.1 shows the correlation between those two estimators, used to study effects of the multiplicity estimation on the spectra measurement.

Five bins in multiplicity for both estimators were defined. First the bins for the ntracks estimator were defined in such a way that each of them has the same number of reconstructed tracks in data. The bins for the ntracks estimator were transformed to bins for the nSPD2 estimator assuming a linear correlation between the estimators. All the bins are shown in Figure 6.1 using lines. The exact values of bin limits are summarized in Table 6.1.
The resolutions of the estimators were compared using simulations. For each measured multiplicity bin a distribution of the generated multiplicity was extracted. The generated multiplicity is the number of primary charged particles

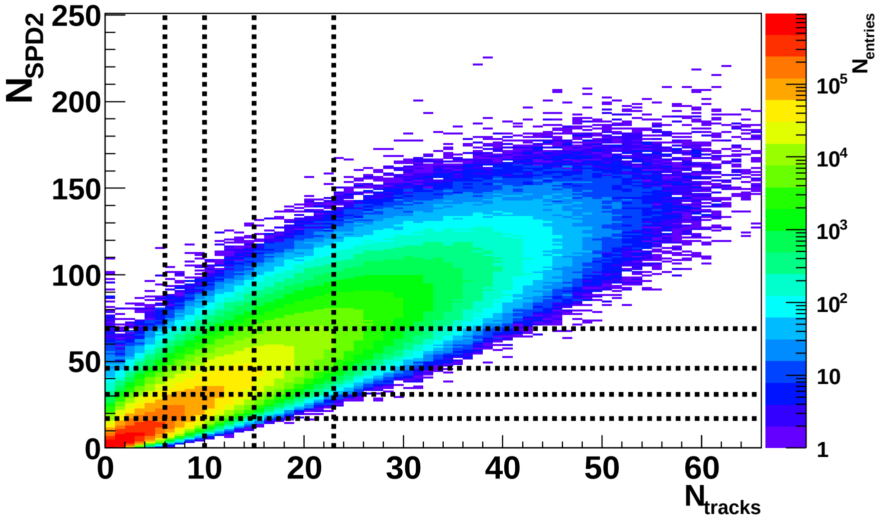


Figure 6.1: Correlations between the two multiplicity estimators (ntracks and SPD2 for the definition see Section 6.2) in data. Lines indicate multiplicity bins from Table 6.1.

Table 6.1: Multiplicity bins for both estimators.

	ntracks estimator	nSPD2 estimator
bin1	[0;5]	[0;16]
bin2	[6;9]	[17;30]
bin3	[10;14]	[31;45]
bin4	[15;22]	[46;68]
bin5	[23;∞)	[69;∞)

(according to the ALICE definition) with $|\eta| < 0.8$ and $p_{\mathrm{t}} > 0.15$ (GeV/c). The results obtained using the PHOJET simulation are shown in Figure 6.2, where distributions for both estimators are plotted. It can be seen that the ntracks estimator has better resolution than the nSPD2 estimator. In each bin the mean multiplicities for both estimators are similar, this observation confirms that the correlation between the two estimators is linear.

6.3 Selection bias

Due to statistical fluctuations of the tracking efficiency, events with a efficiency larger than the average efficiency tend to have a reconstructed multiplicity larger than the expected one. This effect introduces a correlation between the tracking efficiency and the multiplicity estimator. From now on the result of the correlation will be referred to as the selection bias. The selection bias is especially strong for the ntracks estimator because the multiplicity selection and the

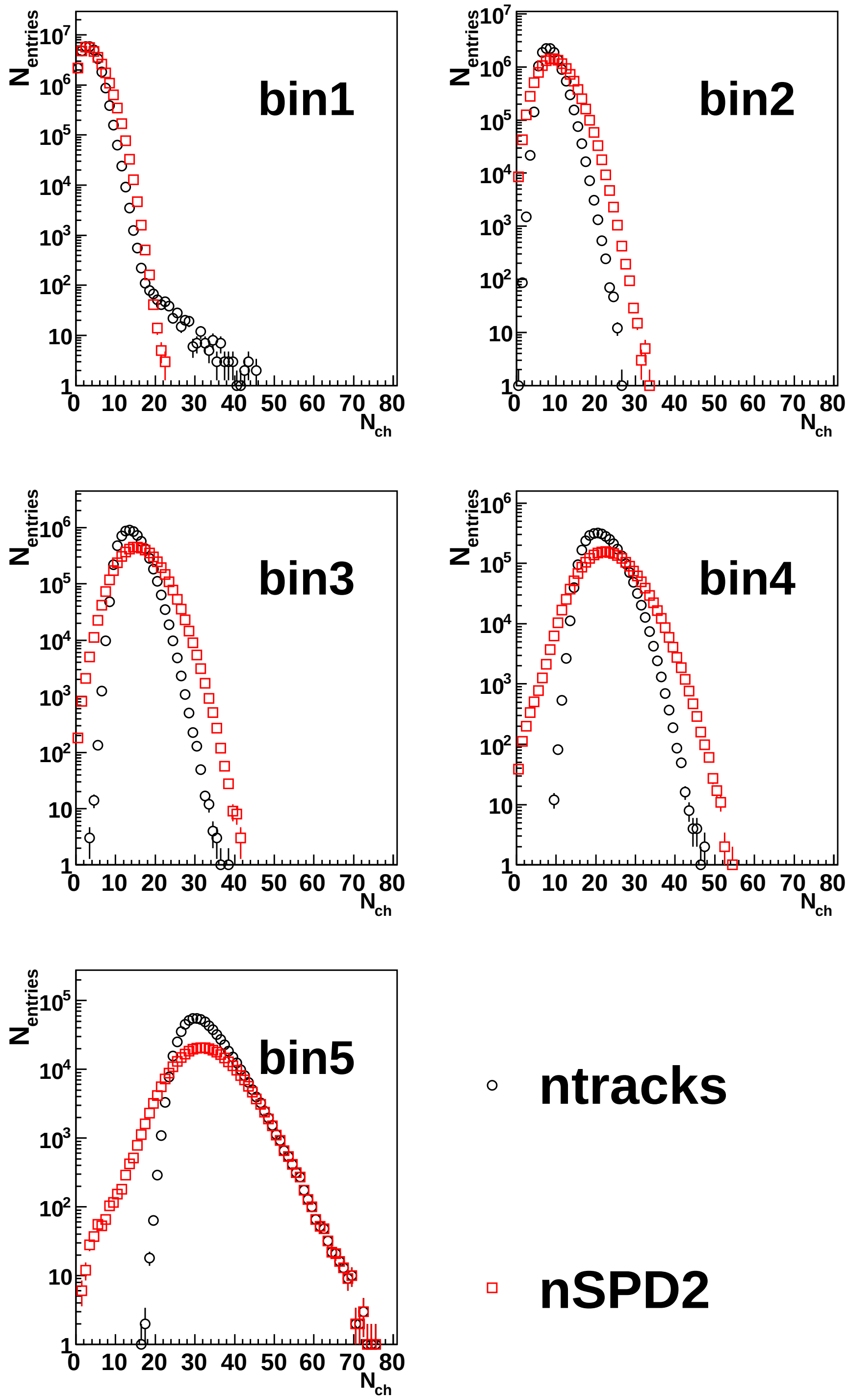


Figure 6.2: Distributions of the number of generated primary particles by PHOJET with $|\eta|$ <0.8 and $p_{\rm t}$ >0.15 (GeV/c) (N_{ch}) in multiplicity bins selected using ntracks and SPD2 multiplicity estimators.

spectra measurement are made using the same tracks.

The qualitative features of the selection bias can be understood using a simple model. If $N(j)$ is the probability to have an event with j generated primary particles and $p(i|j)$ gives the probability that i tracks are reconstructed for an event with j generated primary particles and that the influence of secondary particles can be neglected[1] the tracking efficiency for primary particles in a bin $[t_{min};t_{max}]$ of ntracks estimator is given by:

$$\varepsilon_{bin}(t_{min}, t_{max}) = \frac{\sum_{i=t_{min}}^{t_{max}} \sum_{j=i}^{\infty} p(i|j)N(j)i}{\sum_{i=t_{min}}^{t_{max}} \sum_{j=i}^{\infty} p(i|j)N(j)j}. \tag{6.1}$$

Under the assumption that the number of reconstructed tracks is smaller than the number of generated primary particles (no track splitting), the reconstruction of each track is an independent process and the probability to reconstruct a single track (ε) does not depend on the $p_{\rm t}$ and the particle type. The probability $p(i|j)$ can then be approximated by the binomial distribution:

$$p(i|j) = \binom{j}{i} \varepsilon^i (1-\varepsilon)^{j-i}. \tag{6.2}$$

In this approach ε is the efficiency in minimum-bias (MB) events:

$$\varepsilon = \varepsilon_{bin}(0, \infty). \tag{6.3}$$

The most important thing is the fact that $\varepsilon_{bin}(t_{min}, t_{max})$ depends on the multiplicity distribution ($N(j)$). This distribution can be modelled in different ways by different event generators and does not reproduce the true one.

A magnitude of the selection bias can be seen in Figure 6.3 where the efficiency in each ntracks multiplicity bin is calculated using the $N(j)$ distribution from PHOJET. Those calculations show a few important properties. The first multiplicity bin, which has an average number of reconstructed tracks below the average number of reconstructed tracks in the MB sample, has a $\varepsilon_{bin} < \varepsilon$ while the other bins have $\varepsilon_{bin} > \varepsilon$. For $\varepsilon \approx 0.7$ (the typical efficiency for pions) the selection bias is at the level of a few percent.

The efficiencies (Equation 4.8) for both estimators in all multiplicity bins were calculated using simulations based on both event generators. In all multiplicity bins the same fit ranges were used as for the 7 TeV analysis in Chapter 4. Figures 6.4 and 6.5 show the positive pion efficiencies for the PHOJET simulation. For the ntracks estimator the selection bias is present and is of the same magnitude as predicted (see Figure 6.3). For the SPD2 estimator the effect is smaller (< 0.01), as the result of a weaker correlation between the number of reconstructed tracks used to calculate the efficiency and the number of clusters in the second layer of the SPD. As a next check the positive pion efficiencies for the ntracks estimator calculated using both generators were divided by the MB efficiency from the PHOJET, and the corresponding ratios are shown in Figure 6.6. The efficiency for the same bin calculated based on the PYTHIA or the PHOJET simulations are different (e.g. bin1), conforming that the selection bias is model dependent.

[1] In the track sample selected by the ntracks estimator, which uses standard ALICE track cuts, contribution of secondary tracks is at the percent level.

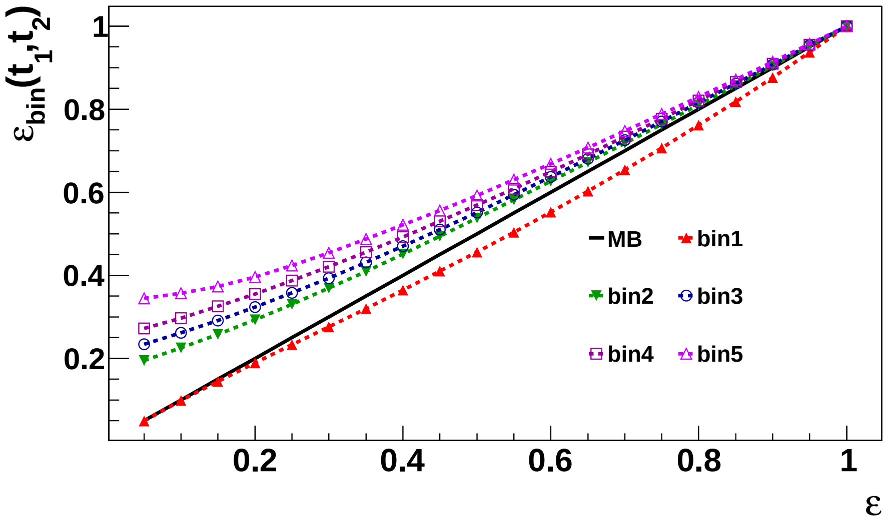


Figure 6.3: Efficiency in the ntracks multiplicity bins according to Equation 6.1 using PHOJET multiplicity distribution as $N(j)$.

6.4 Correction for the selection bias

The selection bias depends on the multiplicity distribution. If in a simulation this distribution does not match the true one, the spectra calculated using efficiencies from that simulation will not be correct. There are also second order effects which contribute to the selection bias and can influence the result, like the composition of an event. For example an event with two kaons and three pions has a different probability to have four tracks reconstructed than an event with five pions. This is a result of the difference between pion and kaon efficiencies.

The simplest way to correct for the difference between the multiplicity distributions of data and simulation is to assign to each simulated event a weight dependent on the generated multiplicity. Weights are defined as the ratio of the true multiplicity distribution to the generated multiplicity distribution. To calculate the weights, the multiplicity distribution in $|\eta| < 1$ for inelastic proton-proton collisions at $\sqrt{s}$=7 TeV as measured by ALICE (result from [57]) was used as the data distribution. This distribution was compared with the same distribution for both event generators (see Figure 6.7). For bins which are not covered by the measurement the value of the weight from the last common bin between data and generated distributions was used.

Using the weights the efficiencies in multiplicity bins were calculated for the both generators. Figure 6.8 shows the comparison of positive pion efficiencies for the weighted and the standard PHOJET simulation. The largest difference is in bin5 because the data and the PHOJET multiplicity distributions are most different for large multiplicities (see Figure 6.7). The large difference between the PHOJET and PYTHIA efficiencies in bin1 shown in Figure 6.6 is caused by

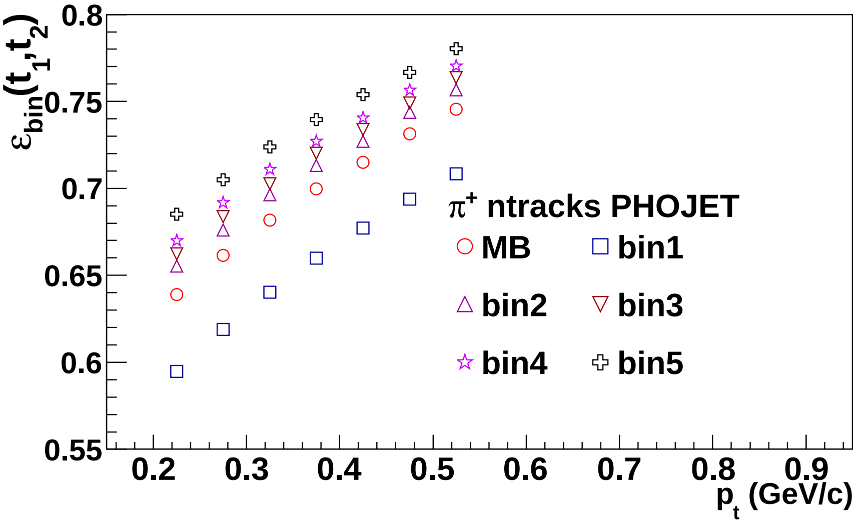


Figure 6.4: Positive pion efficiencies (Equation 4.8) in multiplicity bins for the ntracks estimator based on the PHOJET simulation.

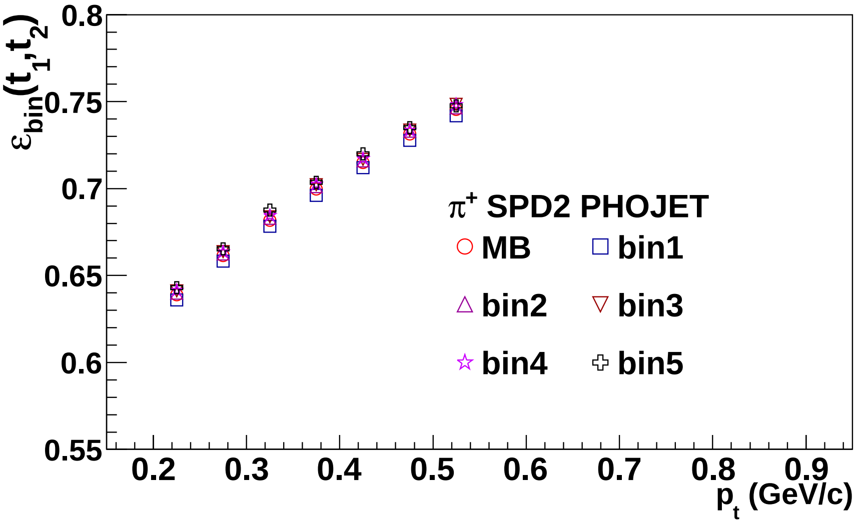


Figure 6.5: Positive pion efficiencies (Equation 4.8) in multiplicity bins for the nSPD2 estimator based on the PHOJET simulation.

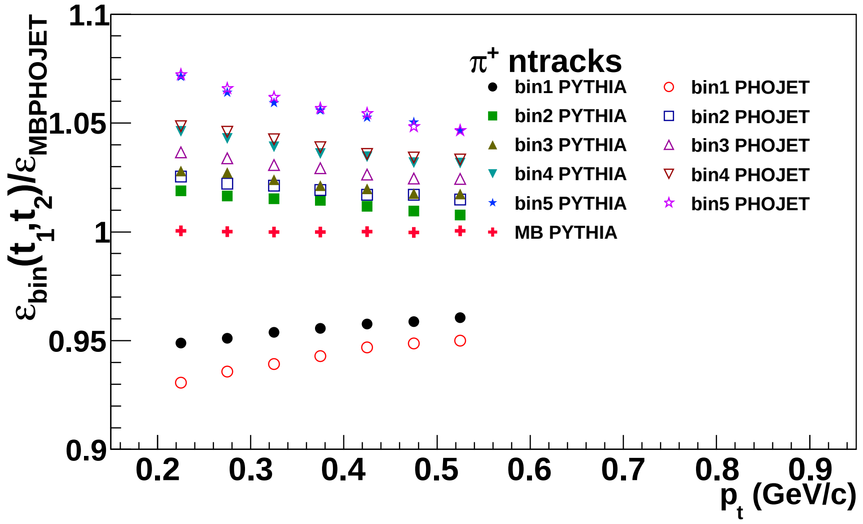


Figure 6.6: Ratios of the positive pion efficiencies (Equation 4.8) for the ntracks estimator calculated using the PYTHIA and PHOJET simulations to the MB efficiency in the PHOJET simulation.

the difference between PHOJET and PYTHIA multiplicity distributions in this multiplicity range. Figure 6.9 is analogous to Figure 6.6 but it shows ratio of positive pion efficiencies calculated after applying the weights. The difference between the efficiencies for both simulations is smaller than in the case of the standard simulations. The remaining difference can be caused by the second order effects. For proton (antiproton) and kaons efficiencies the same behaviour was observed.

6.5 Spectra in multiplicity bins

The corrected spectra were calculated in each multiplicity bin for both estimators using the procedure described in Chapter 4 for the 7 TeV data sample without the event efficiency correction. The event efficiency was not applied because for the SPD2 estimator events without a primary vertex cannot be assigned to any bin, the correction based on the vertex position cannot be calculated. So it is impossible to correct the spectra in each SPD2 bin for particles lost due to this. For the ntracks estimator the weighted efficiencies were used and the weighted dca$_{xy}$ distribution during calculation of the proton (antiproton) contamination. All spectra were normalized to the number of events in each multiplicity bin used for analysis. The same way of normalizing spectra for both estimators gives the possibility to compare the yields. Those spectra were then divided by the MB spectra, which were calculated without applying the event efficiency corrections and normalized in the same way as the spectra in the multiplicity bins. Results for positive particles are shown in Figures 6.10, 6.11 and 6.12.

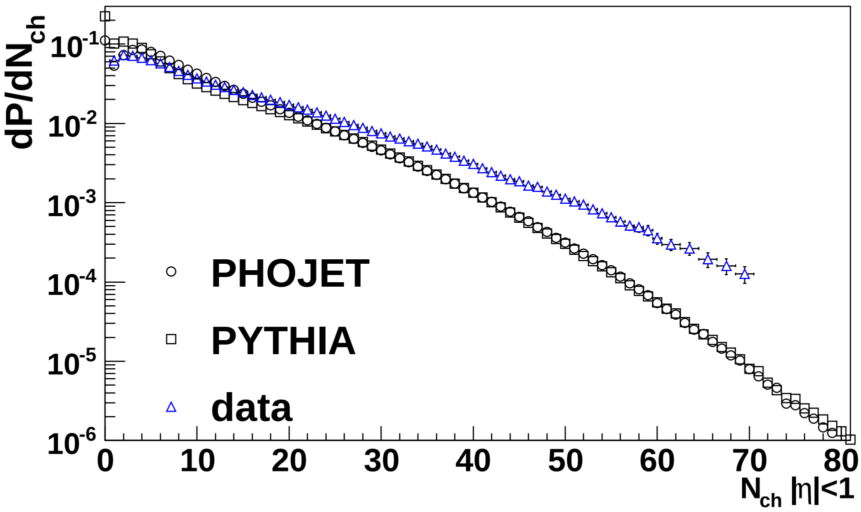


Figure 6.7: Comparison of the measured multiplicity distribution [57] to the results of PYTHIA (Perugia0) and PHOJETsimulations.

When ratios are calculated many of the systematic uncertainties cancel out, especially those connected with the detector material, which are an important contribution to the total uncertainty of the spectra measurement. The spectrum uncertainties which can be important for the ratios are those caused by the fit ranges used to calculate the raw yields. The influence of the fit range on the raw spectra can be multiplicity dependent due to a possible multiplicity dependence of the particle ratios. All the ratios were calculated using the changed ranges as defined in Section 4.4.1 for the 7 TeV analysis. As for the MB spectra, only ratios in the last used p_{t} bins were affected by the change of the ranges. This change increases with the multiplicity. Due to the changes of the fit ranges, the ratios of spectra have not changed more than 0.5% for pions, 3% for kaons and 3% for protons (antiprotons). The quoted value for each particle is the change for the last p_{t} in the largest multiplicity bin. Dependence of the ITS-TPC matching on the event multiplicity was not observed, suggesting that this part of the systematic uncertainties does not contribute to the ratio uncertainties. The corrections for secondary protons and antiprotons were checked as a function of multiplicity. It was found that they increase with multiplicity but the p_{t} dependence is not changed. As an additional check for all the particles, the MB spectrum was compared with the sum of all the spectra in the multiplicity bins, before normalizing the spectra. For both estimators it was found that the difference is below 1%, which can be interpreted as part of the systematic uncertainties of the ratio measurement. Pile-up can also contribute to the uncertainties, especially for the large multiplicities. The analysis was run with and without pile-up tagging and no effect on the p_{t} dependence of the ratios was found. All these studies give an estimate of the systematic uncertainties of the ratios. The total systematic uncertainties of ratios are smaller than 1.5% for pions, 4% for kaons and 4% for protons.

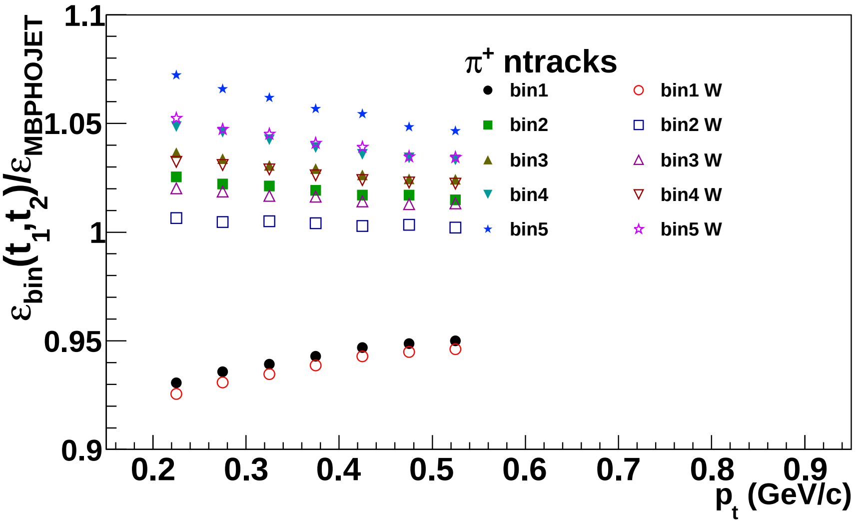


Figure 6.8: Comparison of positive pion efficiencies (Equation 4.8) in multiplicity bins for the ntracks estimator calculated using the standard PHOJET simulation and after re-weighting it (indicate by "W" in the legend).

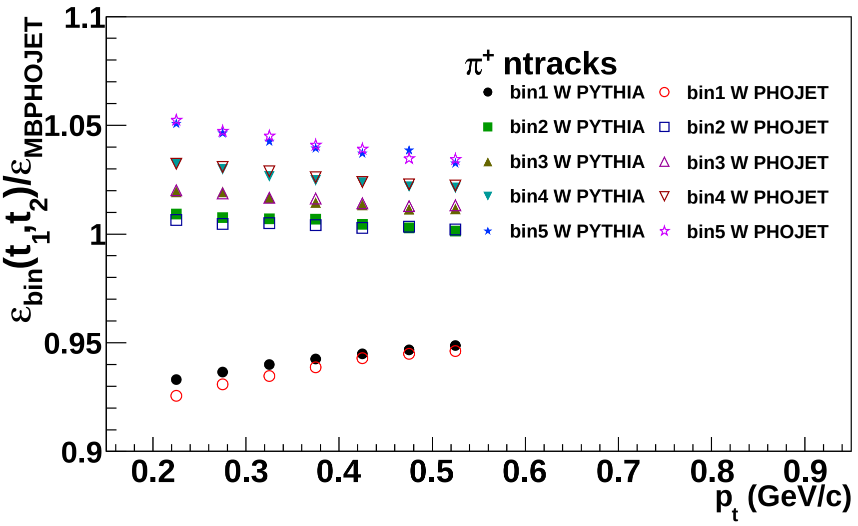


Figure 6.9: Ratios of the positive pion efficiencies (Equation 4.8) in multiplicity bins for the ntracks estimator calculated using the weights for the PYTHIA and PHOJET simulations to the MB efficiency (Equation 4.8) in the PHOJET simulation.

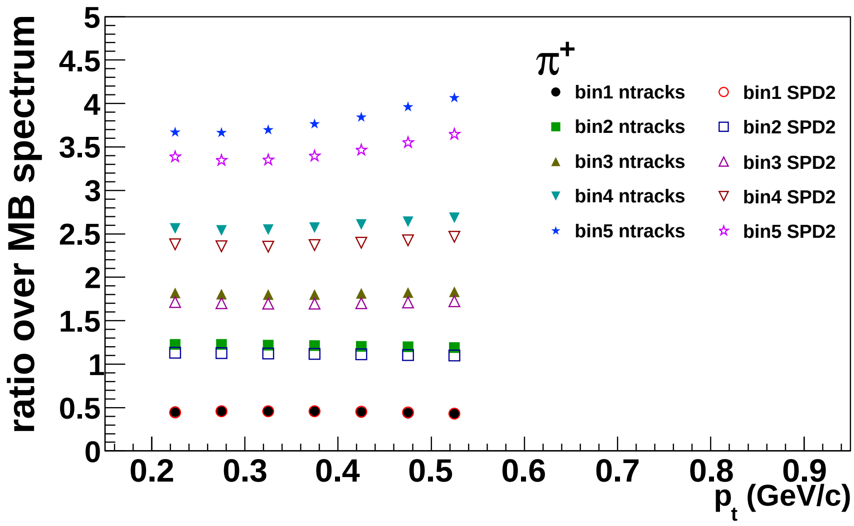


Figure 6.10: Ratios of the corrected spectra in multiplicity bins to the MB spectrum for positive pions.

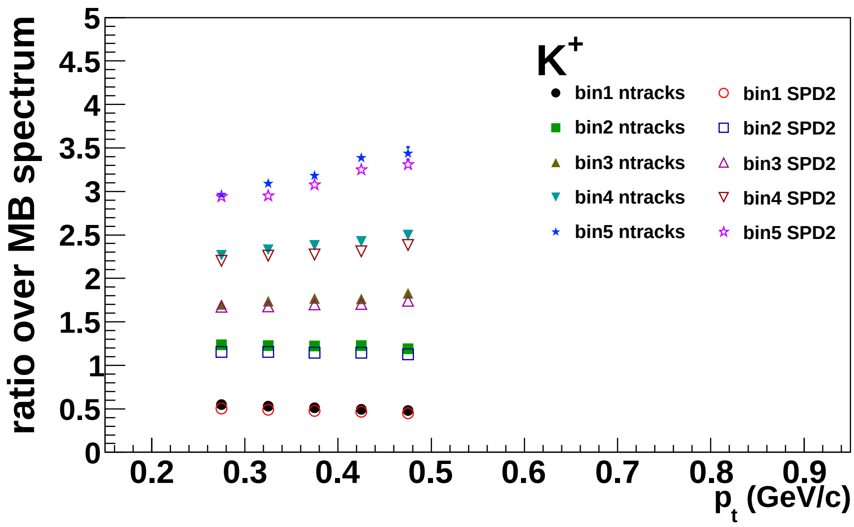


Figure 6.11: Ratios of the corrected spectra in multiplicity bins to the MB spectrum for positive kaons.

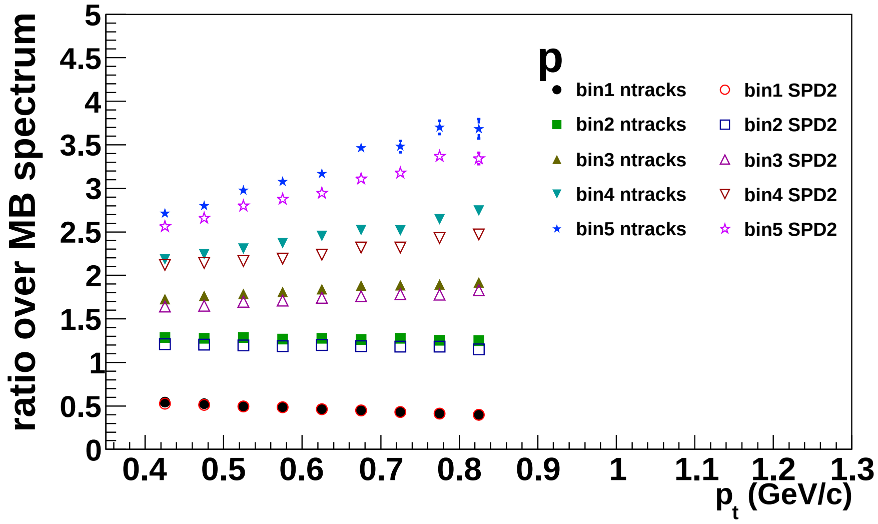


Figure 6.12: Ratios of the corrected spectra in multiplicity bins to the MB spectrum for protons.

The observed increase of the ratios with p_{t} for high multiplicities is larger than the evaluated systematic uncertainties. The increase is seen for both estimators. Therefore it can be concluded that the $\langle p_{\mathrm{t}} \rangle$ for all particles increase with increasing event multiplicity. This has been already reported in [54]. The integrated yields cannot be studied using only this method because the extrapolation uncertainty is too large.

6.6 Conclusions

Any multiplicity estimator using the number of the reconstructed tracks introduces a dependence of the tracking efficiency on the multiplicity. The effect depends on the multiplicity distribution. To correct for that the simulation has to have a multiplicity distribution which matches the true one. It is found that the $\langle p_{\mathrm{t}} \rangle$ of pions, kaons, protons and antiprotons increases with the event multiplicity in proton-proton collisions at $\sqrt{s} = 7$ TeV.

Chapter 7

Conclusions

Based on the analyses presented in Chapter 5 and 6, it is possible to comment on the issues raised in Section 1.5.

The shape of pion, kaon and proton p_{t} spectra in proton-proton collisions at $\sqrt{s}$ =900 GeV [51] and $\sqrt{s}$ =7 TeV cannot be described by thermal functions, especially the high p_{t} part. On the other hand, the Lévy-Tsallis function gives a good description of the shape in the measured range (see Section 5.3.2).

The same ratios of the mid-rapidity yields of different particles in proton-proton collisions at $\sqrt{s}$ =900 GeV and $\sqrt{s}$ =7 TeV suggest that the underlying mechanism of particle production is also the same for those energies. The predictions of the thermal production fail to describe the measured values (see Section 5.5). To rule out or confirm thermal production at 7 TeV more particle yields are needed.

The increase of particle $\langle p_{\mathrm{t}} \rangle$ with energy is consistent with the increase of $\langle p_{\mathrm{t}} \rangle$ with $\mathrm{d}N/\mathrm{d}\eta$ reported by previous experiments [54](see Section 6.5).

Proof of collective behaviour of the system produced in proton-proton collisions at $\sqrt{s}$ = 7 TeV is not found. The results of the Blast Wave fits on the 7 TeV spectra are consistent with previous results and show the increase $\langle \beta \rangle$ parameter. The non-zero value of $\langle \beta \rangle$ cannot be used as a proof of collective behaviour (i.e. radial flow) in proton-proton collisions because the observed effect can be produced by models (as PYTHIA) which do not have any collectivity (see Section 5.7). To judge whether the collective behaviour is observed, a scan of the properties of the spectra as a function of $\mathrm{d}N/\mathrm{d}\eta$ is needed, ideally for two energies of colliding protons. Those results should allow to answer two questions:

- How different is the soft part of proton-proton collisions in case of the mid-rapidity $\mathrm{d}N/\mathrm{d}\eta$=4 and $\mathrm{d}N/\mathrm{d}\eta$=30 ?
- Is this difference related to the energy of the colliding protons or the increased $\mathrm{d}N/\mathrm{d}\eta$?

The commonly used event generators cannot described the spectra. The newest PYTHIA tune 350 based on the LHC results does not provide a better description than the old tunes (e.g. 320).

As a first step, I suggest to further integrate the methods used to measure the spectra in the ALICE detector, reduce the systematic uncertainties and improve the agreement between the different methods.

Appendix A

Spectra at $\sqrt{s}$ =900 GeV using the 2010 data sample

In 2010 the ALICE experiment recorded a sample of proton-proton collisions at $\sqrt{s}$ =900 GeV. This data sample had around 6M events passing the ALICE physics selection. Using this data sample the new spectra were calculated and compared with the 900 GeV spectra presented in Chapter 4. During the calculations of the new 900 GeV spectra all settings of the 7 TeV analysis were used, because both data samples were collected during the same period. Ratios of the 900 GeV spectra (new and old) normalized to the inelastic collisions are shown in Figure A.1. Only statistical uncertainties are plotted because they dominate. Many of the systematic uncertainties cancel out in the calculation of the ratios.

Deviation of the pion ratios from unity can be explained by the uncertainties coming from the normalization procedure (2%) [50]. The kaon ratios suggest that in the old sample the spectra were underestimated (all points are above one). It was found that this underestimation is mainly caused by the calculation of the raw yields (determination of the fit ranges). Due to limited statistic it was hard to spot that during the analysis of the old data sample. The proton and antiproton ratios in their statistical uncertainties are consistent with one.

It is recommended to redo the 900 GeV analysis on the 2010 data sample, this will reduce the statistical uncertainties and improve the control on the systematic uncertainties.

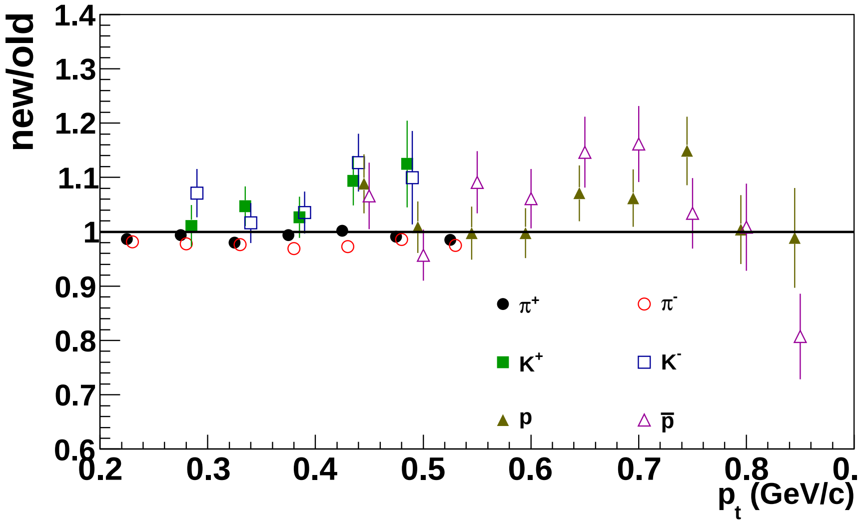


Figure A.1: Ratios of new spectra (2010 data sample) over old 900 GeV spectra (2009 data sample). The markers are shifted from the bin centres for a presentation purpose.

Appendix B

Collision kinematics

The rapidity is defined as follows:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{B.1}$$

where E is energy of the particle and p_z is a momentum along the beam direction. For momentum larger than the mass of the particle rapidity can be approximated by the pseudorapidity η:

$$\eta = \frac{1}{2} \ln \frac{p + p_z}{p - p_z} = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right], \tag{B.2}$$

where θ is the angle between particle momentum and the beam axis.

Samenvatting

In dit proefschrift is gebruik gemaakt van de data verzameld door de A Large Ion Collider Experiment (ALICE) gedurende het eerste operatiejaar van de Large Hadron Collider (LHC). Het eerste deel bevat een beschrijving van de calibratieprocedure van de Silicon Strip Detector (SSD), onderdeel van de Inner Tracking System (ITS). Deze calibratie is gedaan op het niveau van de individuele modules. Het heeft de fluctuaties op de gemeten energie verlies signaal gereduceerd en introduceert een uniformiteit van het energie verlies signaal in de ITS. De studie van de prestaties van de calibratie heeft een probleem laten zien voor het hoge-energieverlies signaal. Het probleem kan gerelateerd worden aan de instellingen van de SSD electronica en moet verder onderzocht worden.

Het hoofdonderwerp is de meting van pion-, kaon-, proton- en antiprotonproductie als een functie van hun transversale momenta (oftewel de p_{t} spectra) in proton-proton botsingen. De identificatie van de deeltjes was gebaseerd op hun energieverlies in de ITS. De ontwikkelde meetmethode is gebruikt om de spectra te meten bij twee energien van de protonenbotsingen ($\sqrt{s}$ =7 TeV en $\sqrt{s}$ =900 GeV). Het heeft een nauwkeurigheid van een paar procent, afhankelijk van het soort deeltje. Een van de belangrijkste onzekerheden op de meting wordt veroorzaakt door de beschrijving van de interactie van de deeltjes met het detectormateriaal, welke gereduceerd kan worden.

De 7 TeV resultaten zijn vervolgens gecombineerd met die van andere methoden voor deeltjesidentificatie. Deze vergelijking laat zien dat de resultaten van de verschillende methoden binnen een paar procent met elkaar overeenstemmen. Het gecombineerde resultaat is vergeleken met voormalige experimentele resultaten en theoretische modellen, waaronder diegenen die een collectief gedrag van het systeem gedurende een botsing beschrijven. De huidige nauwkeurigheid van de meting geeft niet de mogelijkheid deze modellen te bevestigen danwel uit te sluiten. Deze studie laat verder zien dat de relatieve deeltjesproducties voor 900 GeV en 7 TeV vergelijkbaar zijn.

De ontwikkelde methode is ook gebruikt om de spectra als functies van de botsingsmultipliciteit te meten voor proton-proton botsingen bij $\sqrt{s}$ =7 TeV. De focus is hierbij gelegd op het onderzoeken van de correlaties tussen 'tracking' efficientie en de multipliciteitsraming. Dit ondezoek heeft aangetoond dat deze correlatie afhankelijk is van het model gebruikt tijdens simulaties van de detectorprestaties. Deze afhankelijkheid kan tot een fout van een paar procent leiden in de spectrummeting. De methode om voor deze correclaties te corrigeren is ontwikkeld en getest. Het eindresultaat laat een toename van de gemiddelde p_t van de geproduceerde deeltjes zien bij een toename van de botsingsmultipliciteit.

Acknowledgements

Firstly, I would like to thank my parents for all support during all my studies and for blocking my football career in RKS Radomiak. They decided that I should work using my brain not my legs. Many thanks to my brother Adam and my sister-in-law Magda for support but mainly for a place to stay when I was in Warsaw. Magda good luck with your Ph.D.

For 4+ years of an unique work I would like to thank my co-promotes: Marco and Gert-Jan. It was a great honour for me to work with you. I have learnt a lot from both of you. The biggest thanks are for your patience to all my crazy ideas and my bad English. THANKS. Many thanks for my promoter Raimond, he accepted my as Ph.D. even-thought I have not worked on the *Iflow*. Big thanks for my first promoter Rene, who is not with us any more. I hope that my thesis is up to your standards. Many thanks go to Panos for all your help, officially you were not my co-promoter, but you were. Here also I would like to thank two people: my high school teacher Marek Golka for showing me physics and professor Jan Pluta my master thesis promoter, who had introduced me to ALICE.

The biggest thanks go to my friends: Marcin, Paweł and Sara; Sara for *Idziemy na fajka*, guys for sharing with me a hard life of a Ph.D. in Utrecht (e.g. beers in Cafe Belgie and football games in Mick), to all of you for all our night discussions. I hope that there is still a lot to come.

Many thanks also go to all colleges from ALICE group in UU/NIKHEF. Especially to those which have been with me in the group for a long time during my Ph.D: Alessandro (for all stories), Ante (for being so unique Sun of Croatia), Cristian (for a music taste), Deepa (for your laugh), Ermes (for being a STAR in the ALICE group), Marta (for showing me a real Spain), Martijn (for amplification of the dirty jokes), Mikołaj (for distance to work), Misha (for tolerating me as an office mate), Naomi (also for distance to work), Raoul (for the dirty jokes), and those who have been working with me for a shorter time: Alis, Andrea, Carlos, Emanuele, Grazia, Merijn, You, Niels, Piotr, Sandro, Sarah, Wilco. To all of you good luck with all your plans. Thanks also to the senior members of the group: Andre, Arie, Astrid, Kees, Michiel, Paul, Thomas, Ton.

Many thanks to all colleges from Utrecht: Asia, Erica, Elena, Henri, Ola, Remy, Steve, Teemu and many many more, for all parties.

Last but not least many thanks to all people from ALICE.

Curriculum Vitae

2007-2012	Ph.D. researcher Institute for Subatomic Physics, Utrecht University
2002-2007	M.Sc. Eng. Computer Physics Faculty of Physics Warsaw University of Technology
1998-2002	The Jan Kochanowski Upper Secondary Comprehensive School No 6 in Radom, Poland

Bibliography

[1] M.Chojnacki, J. Phys.G, Nucl. Part. Phys. **38** 12407 (2011).

[2] M.Chojnacki, J.Phys.Conf.Ser. **270** 012006 (2011).

[3] D.Griffiths, *Introduction to Elementary Particles.* WILEY-VCH (2004).

[4] C. Wong, *Introduction to High-Energy Heavy-Ion Collision.* World Scientific.

[5] S.Albino, Rev.Mod.Phys.**82** 2489-2556 (2010).

[6] T. Sjostrand, S. Mrenna, P. Z. Skands, JHEP **05**, 026 959 (2006).

[7] B. Andersson, *The Lund Model.* Cambridge University Press (2005).

[8] P. Eden and G. Gustafson, Z. Phys. **C75** 41 (1997).

[9] A. Donnachie and P.V. Landshoff, Phys. Lett. **B296** 227 (1992).

[10] G.A. Schuler and T. Sjostrand, Nucl. Phys. **B407** 539 (1993).

[11] G.A. Schuler and T. Sjostrand, Phys. Rev. **D49** 2257 (1994).

[12] R. Engel, J.Ranft, Phys. Rev. **D54** (1996) 4244.

[13] F.W. Bopp, R. Engel, J. Ranft, arxiv.org/abs/hepph/9803437.

[14] F. Becattini, U. Heinz, Z. Phys. **C76** 269-286 (1997).

[15] F. Becattini, P. Castorina, A. Milov, H. Satz Eur. Phys. J. **C66** 377-386 (2010).

[16] R. Hagedorn and J. Ranft, Nuovo Cim. Suppl. **6** 169 (1968).

[17] The STAR Collaboration, Phys. Rev. **C75** 064901 (2007).

[18] B. De, S. Bhattacharyya, G. Sau, S. K. Biswas , Int. J. Mod. Phys. **E16** 1687-1700 (2007).

[19] C. Tsallis, J. Stat. Phys. **52** 479 (1988).

[20] C. Beck, Phys. **A286** 164-180 (2000).

[21] G. Wilk, Z. Włodraczyk, Phys. Rev. **C76** 054903 (2009).

[22] R. Campanini, G. Ferri, Phys. Lett. **B703**(3) 237-245 (2011).

[23] F. Cooper, G.Frye, Phys. Rev. **D10** 186 (1974).

[24] Z.Tang *et al.* Phys. Rev. **C79**, 051901(R) (2009).

[25] The ALICE Collaboration J. Instrum. **3** S08002 (2008).

[26] The ALICE Collaboration CERN-LHCC-99-012.

[27] J. Alme *et al.*, Nucl. Inst. Meth. **622** 31636 (2010).

[28] The ALICE Collaboration CERN-LHCC-2000-012.

[29] The ALICE Collaboration CERN-LHCC-2002-016.

[30] The ALICE Collaboration CERN-LHCC-2004-025.

[31] W. Carena *et al.*, ALICE-INT-2005-015.

[32] http://root.cern.ch/.

[33] R. Brun, F. Carminati, CERN Program Library, Long Writeup W5013 (1993).

[34] R. Fruhwirth Nucl. Instrum. Methods Phys. Res., **A262** 444 (1987).

[35] Particle Data Group, Particle physics booklet (2004).

[36] M.G. Albrow *et al.*,(Tev4LHC QCD Working group), arXiv:hep-ph/0610012.

[37] P. Z. Skands, Contribution to the 1st International Workshop on Multiple Partonic Interactions at the LHC, Perugia, Italy, Oct. 2008, Fermilab-Conf-09-113-T, arXiv:0905.3418[hep-ph] and arXiv:1005.3457[hep-ph].

[38] K. Nakamura (Particle Data Group), J. Phys. G: Nucl. Part. Phys. **37** (2010).

[39] http://alimonitor.cern.ch/configuration/.

[40] The PHOBOS Collaboration, Phys. Rev. **C75** 024910, 2007.

[41] R. Barlow and C. Beeston, Comp. Phys. Comm. **77** 219-228 (1993).

[42] The ALICE Collaboration, Phys. Rev. Lett. **105**(7) (2010).

[43] P. Christakoglou, ALICE-INT-2010-006, (2010).

[44] A.S. Clough *et al.*, Nucl. Phys. **B76**(1) 4 15-28 (1974).

[45] D.V. Bugg *et al.*, Phys. Rev. **168** 1466.

[46] E. Friedmann *et al.*, PRC **55** 1304.

[47] The ALICE Collaboration, Eur. Phys. J. **C71**(3) 1594 (2011).

[48] The ALICE Collaboration, Eur. Phys. J. **C68** 954 (2010).

[49] The ALICE Collaboration, Phys. Lett. **B693** 956 (2010).

[50] M.G. Poghosyan, J. Phys. G Nucl. Part. Phys. **38** (2011).

[51] The ALICE Collaboration, Eur.Phys.J. **C71**(6) (2011) 1655.

[52] The PHENIX Collaboration, Phys. Rev. **C83** 064903 (2011).

[53] F. Becattini *et al.*, J. Phys. G Nucl. Part. Phys. **38** (2011).

[54] The E735 Collaboration, Phys. Rev. D **48** 3 984-997 (1993).

[55] P. Skands, Phys. Rev. D **82**, 074018 (2010), arXiv:1005.3457v4.

[56] The STAR Collaboration, Phys. Rev. **C79**, 034909 (2009).

[57] The ALICE Collaboration, Eur. Phys. J. **C68**, 345354 (2010).