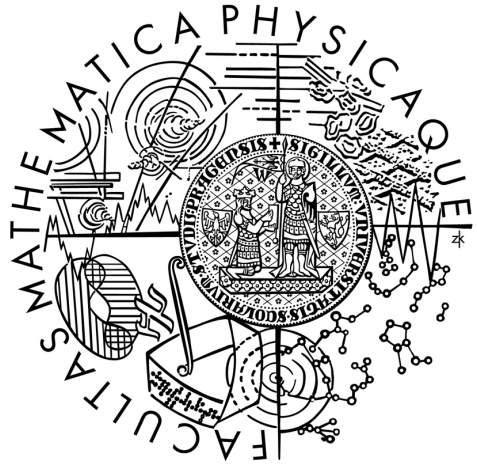
Charles University in Prague
Faculty of Mathematics and Physics

MASTER'S THESIS



Petr Balek

Study of pp and PbPb collisions at LHC

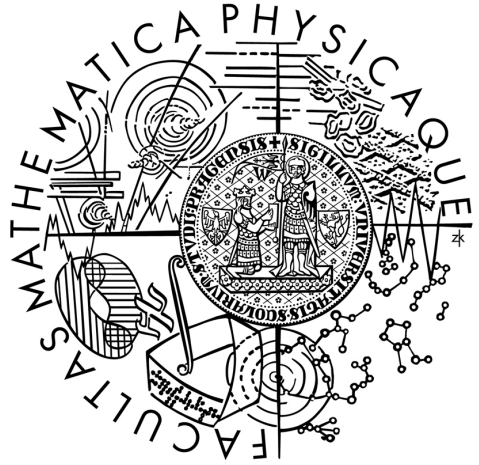
Institute of Particle and Nuclear Physics

Supervisor: RNDr. Jiří Dolejší, CSc.
Study programme: Physics

2010

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Petr Balek

Studium pp a PbPb srážek na LHC

Ústav částicové a jaderné fyziky

Vedoucí práce: RNDr. Jiří Dolejší, CSc.
Studijní program: Fyzika

2010

I would like to thank my supervizor Jiří Dolejší for a lot of comments, advices and continuous support.

I would like also thank to Martin Rybář who explained me issues related to run numbers, lumiblocks and BCIDs. Twice.

My thanks belong to Ája Matějková as well because she helped me with my English and her comments make this thesis a little bit more readable.

I would like also to thank Kačenka for employing my mother and sister, so I could write this thesis in calm.

Last but not least, I would like to thank Jitka Šívrová for help with drawing fig. 45 and 46.

I declare that I wrote my master's thesis independently a exclusively with the use of the cited sources. I agree with lending and publishing the thesis.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

In Prague, 5th August 2010 Petr Balek

Contents

1 Introduction **6**

2 LHC & ATLAS **7**
- 2.1 LHC . . . 7
- 2.2 ATLAS . . . 8
- 2.3 Inner detector . . . 9
- 2.4 Pixel detector . . . 10
  - 2.4.1 Deposited charge . . . 10
  - 2.4.2 Number of hits per track . . . 11
  - 2.4.3 Response of layers . . . 11

3 Deposited charge **13**
- 3.1 Deposited charge . . . 13
- 3.2 Deposited charge per unit length . . . 13
- 3.3 Distribution of Q_d . . . 17
- 3.4 Momentum dependence of the parameters of the function (4) 19

4 Particle identification **26**
- 4.1 Calculation of probability . . . 26
- 4.2 Shares of pions, kaons and protons . . . 28
- 4.3 Efficiency & contamination . . . 30

5 Monte Carlo vs. first data **35**
- 5.1 Events selection . . . 35
- 5.2 Tracks selection . . . 35
- 5.3 Comparing Monte Carlo with data . . . 38

6 Athena tool **46**

7 Conclusion **48**

References **49**

A Landau distribution **52**

B Efficiency & contamination **53**

C Track parameters **55**

Title: Study of pp and PbPb collisions at LHC
Author: Petr Balek
Department: Institute of Particle and Nuclear Physics
Supervisor: RNDr. Jiří Dolejší, CSc.
Supervisor's e-mail address: dolejsi@ipnp.troja.mff.cuni.cz

Abstract: In the present work we study the identification of the particle using deposited charge per unit length in the Pixel detector (part of the ATLAS detector) wich can offer an almost analogous read-out of the deposited charge. First, we parametrically describe the dependences of the deposited charge to reduce the volume of the data used in further analysis. Then we use these parameters as an input for particle identification. We elaborate the method of the identification that allows us to say in most of the cases whether the particle passing through the detector was pion, kaon, or proton. This method can be reproduced using Athena tool that has been written.
Keywords: Pixel detector, deposited charge, particle identification

Název práce: Studium pp a PbPb srážek na LHC
Autor: Petr Balek
Katedra (ústav): Ústav částicové a jaderné fyziky
Vedoucí diplomové práce: RNDr. Jiří Dolejší, CSc.
e-mail vedoucího: dolejsi@ipnp.troja.mff.cuni.cz

Abstrakt: V předložené práci studujeme identifikaci částic za pomoci náboje deponovaného v Pixelovém detektoru (součást detektoru ATLAS), který je schopen vyčítat deponovaný náboj téměř analogově. Nejprve popíšeme závislost deponovaného náboje na jednotkovou délku na hybnosti pomocí řady parametrů, abychom snížili objem zpracovávaných dat. Poté použijeme tyto parametry jakožto vstupní data pro identifikaci částic. Vypracujeme postup identifikace na jehož konci můžeme ve většině případů rozhodnout, zda částice prolétnuvší detektorem byla pion, kaon, nebo proton. Tento postup je snadno zopakovatelný pomocí Athena tool, který byl pro tento účel napsán.
Klíčová slova: Pixelový detektor, deponovaný náboj, identifikace částic

1 Introduction

Proton–proton or heavy ions collisions cover very wide range of studies and it is impossible for a single person to cover it. I focused on issues related to identification of particles. The identification is very useful. The particle identification will help us to suppress the background and to correctly analyze various physical processes.

As will be shown later, the identification is possible only for particles with momentum from 0.5 GeV to 1.5 GeV. The lower boundary comes from detector geometry whereas the upper boundary has physical origin.

In the next section a brief description of the LHC and the ATLAS detector will be provided. The Inner detector and mainly Pixel detector will be stressed out, because the data collected in Pixel detector will be used for particle identification.

In the section 3 we will deal with a deposited charge in the Pixel detector and we will try to parametrize it for further easy handling.

In the main part of this thesis, in section 4, we will depict the identification process with all consequences. In addition, we will discuss the efficiency of the identification.

In the next section, after short introduction concerning the selection of events and tracks, the first LHC data will be compared with Monte Carlo.

In the section 6 the description about Athena tool is given. The Athena tool enables identification in a very comfort way.

In the end, a short conclusion will be provided.

After conclusion, several appendices are attached. They deal with various topics that are considered to be too long to be in the main text, but too interesting to be left out.

2 LHC & ATLAS

2.1 LHC

LHC (Large Hadron Collider) is a huge accelerator at CERN that allows protons or heavy ions to be accelerated. It was installed in the same tunnel as LEP (Large Electron–Positron Collider) was. The circumference is almost 27 km and it is situated under Swiss-French border near Geneva, between Jura mountains and Lake Geneva. Its main purpose is to search for Higgs boson and other predicted particles (SUSY). Naturally, it will search for unpredicted phenomena as well.

The LHC was planned since early 1980's. The plans were affected by the proposal of Superconducting Super Collider (SSC) with center of mass energy 40 TeV that was planned to be built in Texas, USA. This energy was unreachable by a hadron collider installed in LEP tunnel. However, SSC project was canceled in 1993. LHC was approved by CERN Council at its 100^{th} session in December 1994 [1]. Several non-member states, including USA, were invited to participate in the LHC project. This helped to speed up the LHC construction. [2]

After overcoming problems with funding and with several delays of the date of commissioning, the first beam was injected on 10^{th} September 2008 [3]. The inauguration took place on 21^{th} October 2008 [4], despite the unhappy incident on 19^{th} September 2008 that resulted in serious destruction of tens of magnets and in a helium leak into the tunnel. This incident caused more than one year suspension of the project [5]. In November 2009, the LHC became operating again [6]. At the beginning the energy per beam was 450 GeV. Before Christmas 2009, the LHC reached 1.18 TeV per beam [7]. From the end of March 2010 the energy is 3.5 TeV [8].

Before the particles are injected into the LHC, they need to be preaccelerated. First, protons are accelerated in the LINAC 2 (linear accelerator). Then, they continue to the PS Booster and to the PS (Proton Synchrotron). Heavy ions will be firstly accelerated in the LINAC 3, stored in the LEIR (Low Energy Ion Ring) and then they will be injected to PS. From PS the particles (both protons and heavy ions) continue to the SPS (Super Proton Synchrotron). From the SPS they are finally injected in two beams with the opposite direction into the last part of the accelerator complex - the LHC. [9]

The designed maximum energy of protons in one beam is 7 TeV, so the total energy in center of mass frame would be 14 TeV. Currently the maximum of $\sqrt{s}$ is 7 TeV. [8]

Design luminosity is $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$. Current luminosity is the order of $10^{30}\,\text{cm}^{-2}\text{s}^{-1}$ [10].

The accelerator is also able to accelerate heavy ions, primarily lead ions Pb^{82+}. The designed energy per one nucleon is 2.76 TeV ($\sqrt{s_{NN}}$ = 5.5 TeV) with luminosity $10^{27}\,cm^{-2}s^{-1}$ [11]. First collision of the heavy ions are planned to the end of year 2010 with half of the energy and with luminosity $10^{25}\,cm^{-2}s^{-1}$ [12, 13].

2.2 ATLAS

ATLAS (A Toroidal LHC ApparatuS) is one of the four big experiments of LHC. It is a general purpose detector that allows various physical processes to be recorded and reconstructed. It was designed for proton-proton collision but it will be suitable for heavy ions collisions as well.

Letter of intent come in 1992, where an idea of general purpose detector was presented [14]. Several Technical design reports (e.g. [15, 16, 17]) were issued in following years. After years of construction, last piece of the detector found its place in February 2008 [18].

The design of the ATLAS is shown on fig. 1. The outermost part of ATLAS is made by muon chambers where momentum of muons is measured. Under muon chambers, there are hadronic and electromagnetic calorimeters where the particles deposit their energy. In the middle of the ATLAS, there is Inner detector. It is used for detecting tracks of all charged particles.

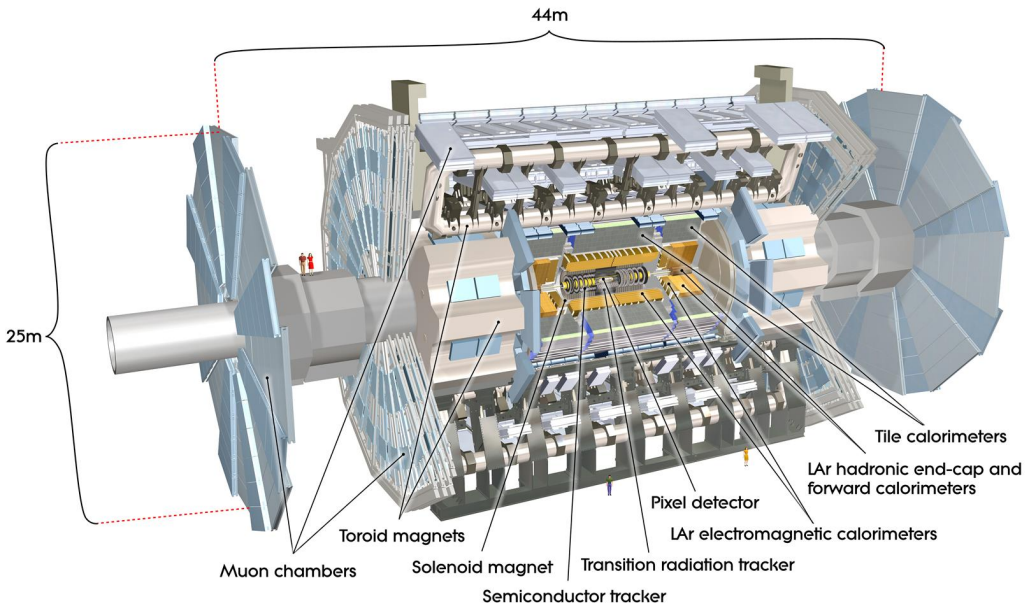


Figure 1: ATLAS detector

From the curvature of the track it is possible to determine its momentum unless the track is too straight. Tracks are also used for finding and fitting primary vertices of the collisions. Magnets that generate magnetic field are an important part of the detector - magnetic field makes the particles move on the helix according to the Lorentz force.

While combining data from all parts of the ATLAS, the precise image of the physical event is obtained.

2.3 Inner detector

In fig. 2 there is picture of the Inner detector. It occupies region $|\eta| < 2.5$. The main purpose of the Inner detector is to detect tracks of charged particles and to measure their momentum.

Inner detector consists of three parts:

- transition radiation tracker (TRT)
- semiconductor tracker (SCT)
- pixel detector

TRT is the outermost part of the Inner detector. It is made of almost 300 000 drift tubes (straws). In barrel, straws are 144 cm long and they are

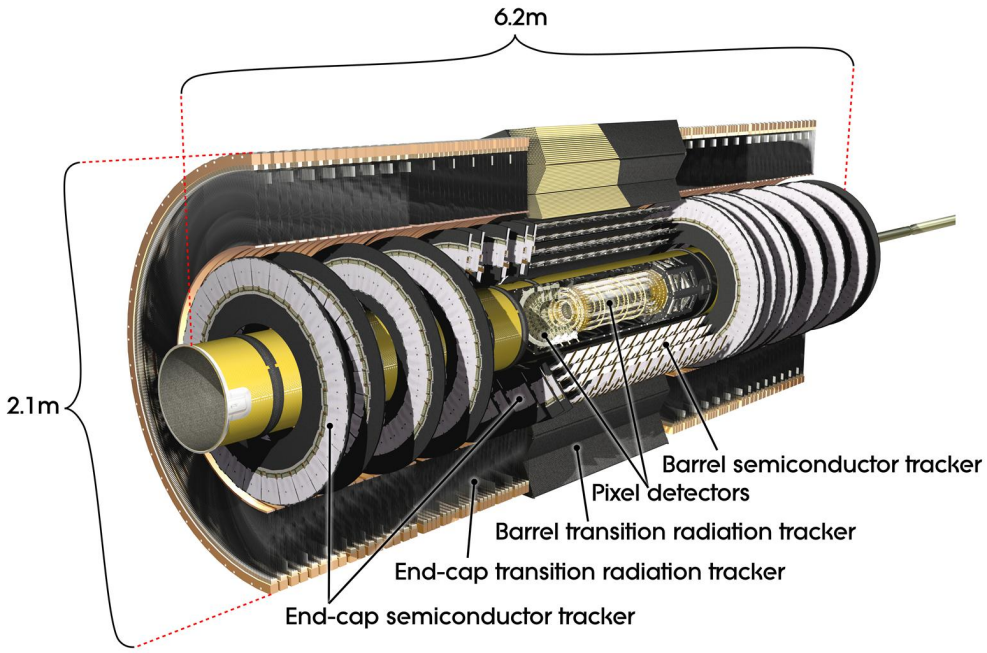


Figure 2: Inner detector

parallel with Z axis. In end-caps, straws are 37 cm long and they are arranged perpendicular with Z axis. Diameter of the straws is equal to 4 mm and the straws are filled with mixture of Xe, CO_2 and O_2.

SCT is located between TRT and Pixel detector. It has 4 barrel layers and 2×9 end-cap disk layers. Silicon strip detectors are on both sides of the layers. The stripes from the opposite sides of layers form an angle of 40 mrad (roughly 2.3°). [19]

2.4 Pixel detector

The requirement of the Pixel detector is to give at least 3 space points on each track.

Pixel detector has 3 barrel layers and 3 end-caps layers on each side. Barrel layers are numbered 0, 1 and 2. Layer 0 (the innermost) is also called B-layer because it allows b-tagging. The layers have radius from 50.5 mm to 122.5 mm. The layers are made from staves that are installed as shingles in a cylindrical surface, so the staves are parallel with Z axis. Number of staves vary from 22 (layer 0) to 52 (layer 2). Each stave consist of 13 modules.

End-caps have 3 identical disk layers on each side. On each disk there are 8 sectors (equivalent of the staves), one sector has 6 modules.

On every module there are around 46 000 pixels. The total number of pixels is $\sim 67 \cdot 10^6$ in barrel and $\sim 13 \cdot 10^6$ in end-caps. The size of the pixels is $50\,\mu\text{m} \times 400\,\mu\text{m}$, but small fraction of them has size $50\,\mu\text{m} \times 600\,\mu\text{m}$. Thickness of the pixels is $250\,\mu\text{m}$.

Each pixel has tunable threshold. Only pixels that exceed the threshold are read out. Typical threshold corresponds to deposited charge 4000 e.

2.4.1 Deposited charge

In comparison with SCT, the Pixel detector has an advantage of possibility of measuring a deposited charge in the detector. Pixel detector does not send only digital information if there was some hit (yes - no), but it sends *time over threshold* (ToT). ToT represents time when output signal is higher than a given threshold. This time is proportional to the deposited charge. The deposited charge can be measured in this way.

Deposited charge is used for suppressing noise and, which is more important, for improving resolution of the hit. When a particle passes through the module, it deposits some charge. This charge could be split among several pixels when the particle passes at an angle. The shares of the read-out charges are given by the trajectory of the particle. The knowledge of the shares can improve resolution of the hit, e.g. for particles with incidence angle 10° from 10.1 μm to 7.2 μm [20].

ToT is measured in the unit of bunch crossing clock (25 ns) while maximal value of ToT is 255 clock cycles. One cycle corresponds to approximately 700 charges of electron (e). Average ToT is almost 30 clock cycles. [21]

A typical deposited charge is around $20 \cdot 10^3$ e but it strongly depends on the kind of particle passing through the detector and on its momentum.

2.4.2 Number of hits per track

If particle traverses e.g. three layers of the detector, it will not mean at all that the number of obtained hits will be three. We can get less hits and also – quite surprisingly – more hits. Distribution of number of the hits per track is in fig. 3.

We could lose some hits because a track is incorrectly fitted or a deposited charge did not exceed the threshold.

Modules of the detector are overlapping to cover safely 2π in ϕ so one track is able to pass through two or more adjacent modules in one layer.

The requirement "at least 3 space points per track" is met in ∼90% of cases.

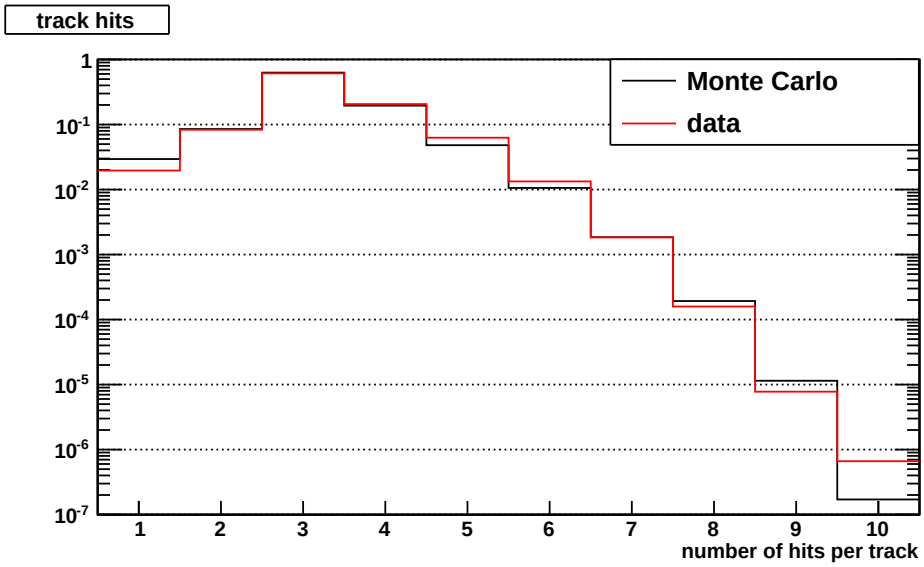


Figure 3: Number of hits per track

2.4.3 Response of layers

In further sections, an assumption that all hits are equal is adopted. Generally, one layer of the pixel detector could provide more deposited charge than other because of the different properties. In the fig. 4 and 5 the deposited

charge per unit length are displayed. Despite Monte Carlo still needs some improvement, data correspond to the simulations very well. The data are from December 2009 and it will be more discussed in sectiom 5. Q_d stands for deposited charge per unit length and it will be described in the next section. For now, we can say that particles deposit a very similar amount of charge in each layer of the Pixel detector, so our assumption works quit well.

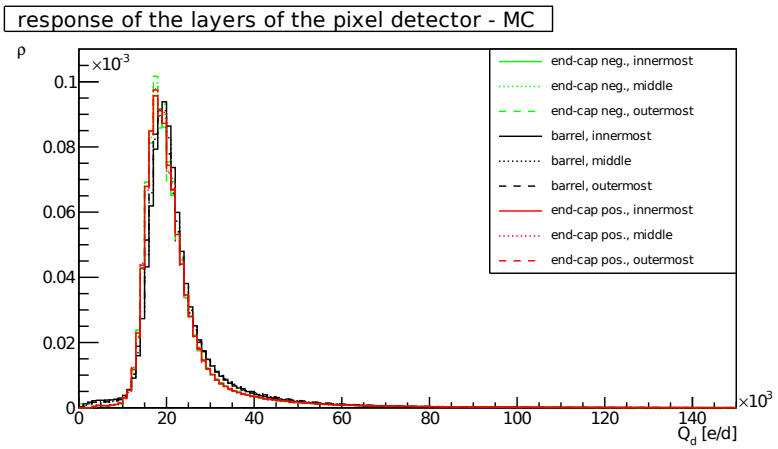


Figure 4: Response of the layers, Monte Carlo
Positive end-cap is the one with positive Z coordinate. Positive part of the Z axis is pointing to Geneva [1]. Negative part is pointing to Saint-Genis.

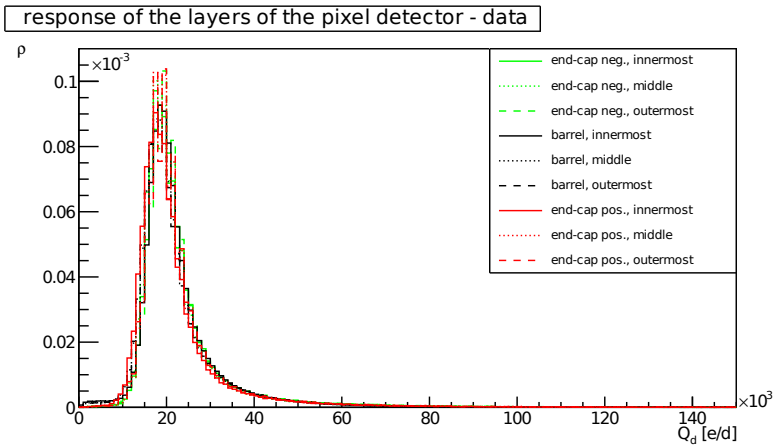


Figure 5: Response of the layers, data

[1] X axis is pointing to the center of the LHC ring. Y axis is pointing upward. Altogether they form right-handed coordinate system.

3 Deposited charge

3.1 Deposited charge

Deposited charge is strongly dependent on the type of a particle. Furthermore, some kinds of particles in any sample are frequent (e.g. pions), some of them are less frequent (kaons or protons) and some type of particles are rare (electrons, muons, ...).

In fig. 6 there is a number of hits for 10^6 Monte Carlo events as a function of deposited charge. As you can see, pions are the most common particles in the detector. Protons and kaons are less common but they have share very roughly 10% each. Electrons and muons are quite rare. Other particles are very rare. They are exotic baryons such a Ξ or Σ or particles that do not have appropriate Monte Carlo truth particle.

Electrons, muons and exotic baryons will not be considered any more. Their share is small and thus they will be neglected.

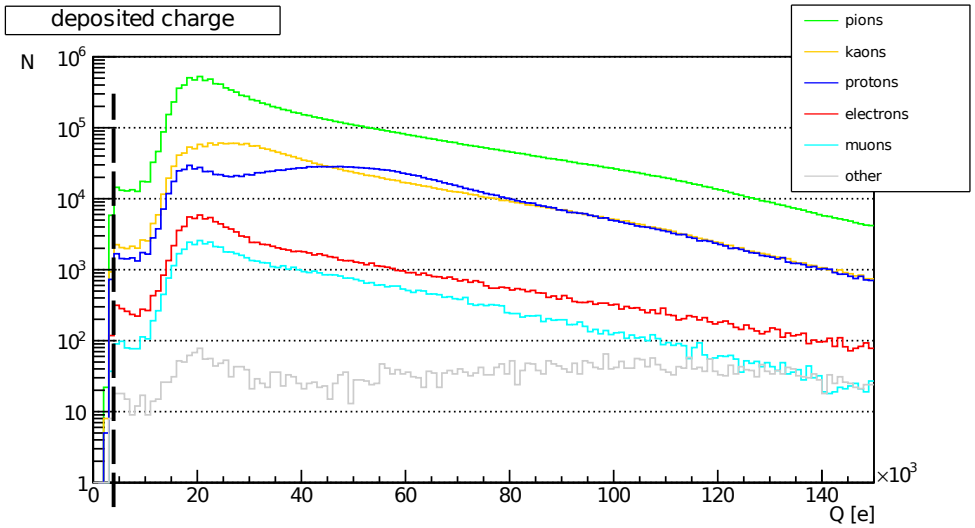


Figure 6: Distribution dN/dQ
Dashed line represents the typical threshold of the pixels ∼4000 e [22].

3.2 Deposited charge per unit length

Deposited charge is governed by the Bethe-Bloch formula (see e.g. equation (27.1) in [23]) and it depends on momentum and on type of the particle. Fluctuation can be roughly described by Landau distribution (see appendix A for more information about Landau distribution), but more accurate description is provided by a convolution of the Gauss distribution with the

Landau distribution [24].

Energy loss is mainly used for production electron–hole pair. The average energy needed to create electron-hole pair is $3.68 \pm 0.02\,\mathrm{eV}$ [22].

Deposited charge is proportional to the length of the track in the module. Situation is sketched in fig. 7. Straight line represents a track and its blue part means part of the track in the module. Length ℓ of this part depends on incidence angle α and thickness of the module d:

$$\ell = \frac{d}{\cos\alpha} \tag{1}$$

Curvature of the track in the module can be neglected because the modules are of very thin thickness.

Very useful variable is deposited charge per unit length. We choose the thickness of the module d as the "unit length". The meaning of it will be a charge that will be deposited if a particle passes throughout a module perpendicularly. In addition, it defines the scale for threshold.

Deposited charge per unit length (Q_d) is:

$$Q_d = \frac{Q_{tot}}{\ell} = \frac{Q_{tot} \cdot \cos\alpha}{d}, \tag{2}$$

where Q_{tot} is the total charge deposited in the module along the track. Dimension of Q_d is e/d, i.e. charge of positron over d.

Distribution of deposited charge per unit length is in fig. 8. It should be noticed that two "bumps" in proton distribution in fig. 6 disappear. The "bump" on the left come from the hits in end-caps, where particles have small θ (otherwise they could not reach end-caps) and thus the path of the particle is short. When Q_d is plotted, the difference is erased because the material used for modules in barrel and end-caps is the same.

These two "bumps" are not visible for distributions of other particles in fig. 6 because the differences between the two "bumps" were smeared due

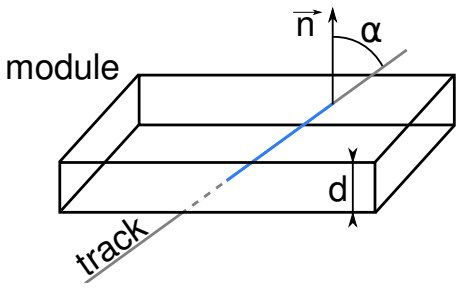


Figure 7: Incidence angle
$\vec{n}$ is perpendicular to the surface of the module. It is perpendicular to Z axis for barrel layers and parallel with Z axis for end-caps.

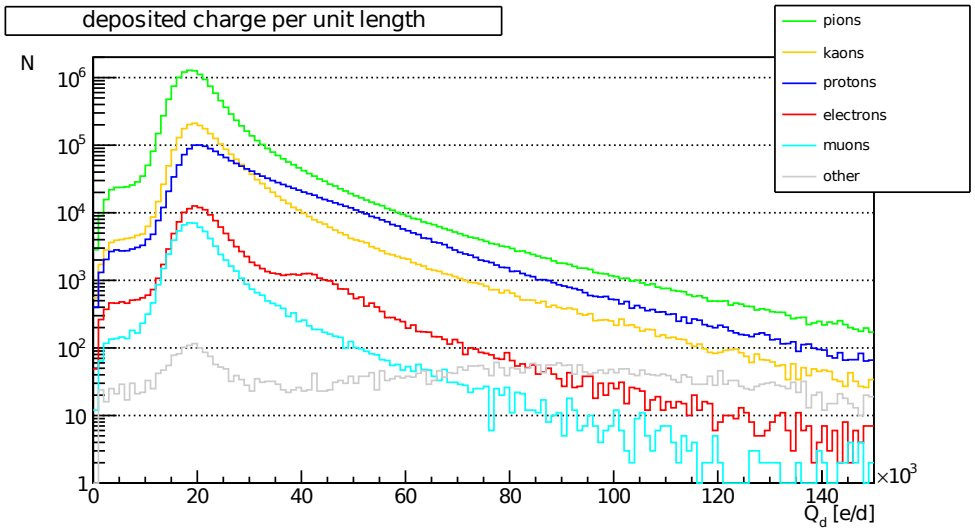


Figure 8: Distribution $\mathrm{d}N/\mathrm{d}Q_d$

to small deposited charge.

Another "bump" for electrons appears in fig. 8 at $\sim 45 \cdot 10^3\,\mathrm{e/d}$. Energy loss of electrons is roughly constant for different momenta, so this "bump" cannot be explained with high multiplicity of electrons with specific momentum. Unexpected number of electrons (and only electrons) has high Q_d. Origin of this phenomenon was not discovered. Fortunately, it does not bring any error in further calculation, since electrons have not been considered any more.

In fig. 9 there is energy loss of different types of particles. In fig. 10 there is Q_d for pions, kaons and protons. The distributions are correlated with fig. 9 and it can be seen that particles can be distinguished according to their deposited charge, at least particles of small momentum.

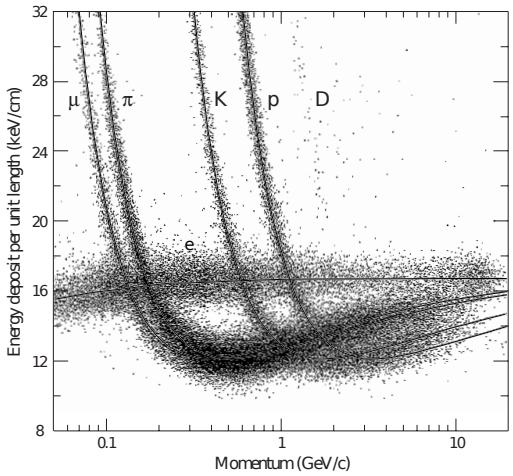


Figure 9: Energy deposit
Taken from Review of Particle Physics, fig. 28.15; [23]

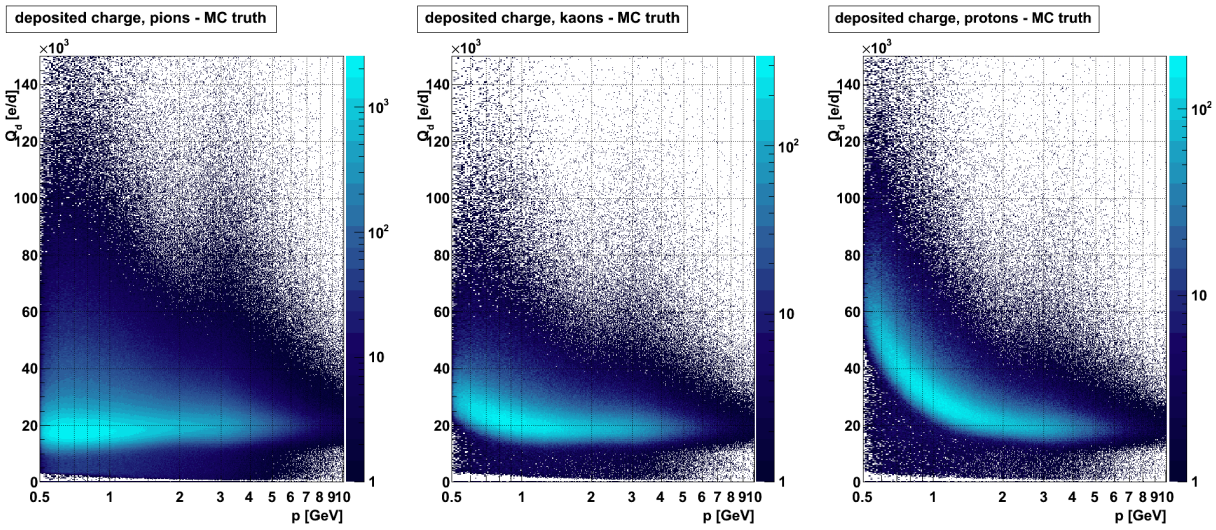


Figure 10: Deposited charge per unit length

3.3 Distribution of Q_d

For practical reason, it would be convenient to describe dN/dQ_d distribution with a function of few parameters and to determine the momentum dependence of these parameters.

While looking for the momentum dependence of the dN/dQ_d distribution we meet two conflicting requirements. We would like to have good statistics, so there have to be enough tracks in a momentum bin. However, we would also like to have distribution for well defined momentum, so the momentum bins should be as small as possible. It is not possible to meet both requirements at once, so the results will be the compromise between them.

The distribution of probability density ρ is showing in fig. 11. The probability that a hit will fall into a bin is:

$$\Delta\xi = \rho(Q_d) \cdot \Delta Q_d, \tag{3}$$

where $\Delta\xi$ is the probability, Q_d is deposited charge per unit length of the hit, $\rho(Q_d)$ stands for the probability density distribution and ΔQ_d is a width of the bin.

An integral of ρ over the whole range of Q_d is 1. ρ distribution can be easily obtained from $\mathrm{d}N/\mathrm{d}Q_d$ by normalization.

One can see the performance of the different fitting functions. The Gauss distribution is absolutely insufficient. The Landau distribution describes the distribution better. Convolution of the Gaussian and the Landau distribution fits the Monte Carlo very well except the small Q_d. Small Q_d is deposited when a particle passes through module at a high angle and deposited charge in one pixel does not exceed the threshold. Then only a part of deposited charge is recorded and it results in small Q_d.

The function that describes the distribution best is the convolution of

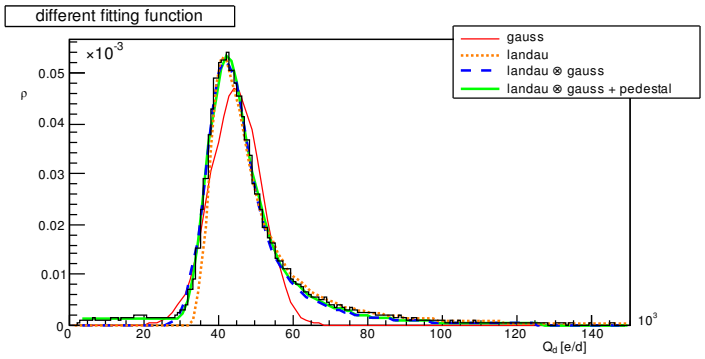


Figure 11: Distribution fitted by different functions
protons; $0.70\,\mathrm{GeV} < \mathrm{p} < 0.75\,\mathrm{GeV}$

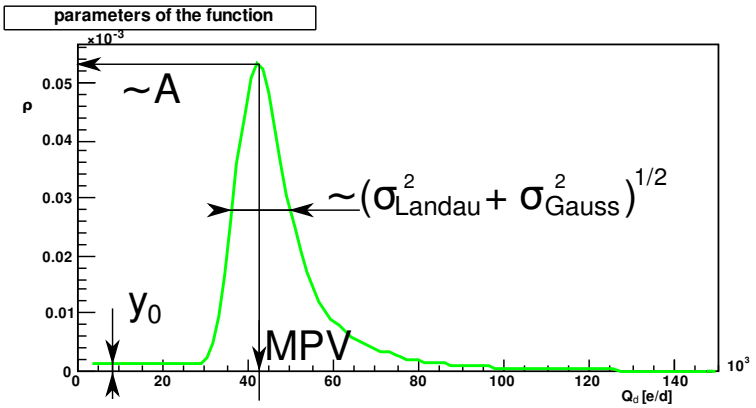


Figure 12: Parameters of the function (4)

the Gauss distribution and the Landau distribution for higher charges. If the value of this convolution is smaller that given value y_0 and x (Q_d) is small (i.e. it is not at the distribution tail), then y_0 is used as value of the function instead of value of the convolution:

$$\begin{aligned} f(x) &= (\text{landau} \otimes \text{gauss})(x) && \text{for} \quad x \geq x_0 \\ &= y_0 && \text{for} \quad x < x_0 \end{aligned} \tag{4}$$

where x_0 is the smaller solution of

$$(\text{landau} \otimes \text{gauss})(x_0) = y_0 \tag{5}$$

The function (4) has 5 parameters - see table 1.

Table 1: Parameters of the function (4)

#	name	meaning
1	MPV	approx. most probably value
2	σ_{landau}	width of Landau
3	A	amplitude
4	σ_{gauss}	width of Gauss
5	y_0	pedestal

The first three parameters come from the Landau distribution. The fourth one is from the Gauss distribution. The fifth parameter is the pedestal used for small values of x instead of convolution. See fig. 12 for illustration.

MPV and y_0 are exactly illustrated in the figure. Height and width of the function is only proportional to A and $\left(\sigma^2_{\text{landau}} + \sigma^2_{\text{gauss}}\right)^{1/2}$ respectively.

3.4 Momentum dependence of the parameters of the function (4)

Introducing function (4) make handling Q_d easier but the parameters of this function probably depend on momentum. We will try to describe their dependence on momentum by some series. This way we should be able to handle with Q_d in all region we are interested in, i.e. 500–1500 MeV.

The parameters are fitted with the series:

$$\sum_{i=-2}^{2} a_i \cdot p^i, \tag{6}$$

where p is momentum in GeV and a_i are constants to be determined. The fitting shows that no more powers in the series (6) are needed. Furthermore, we used Occam's razor so we excluded every term of (6) that is not necessary for description of the dependences.

The series (6) is continuous function of momentum, so the result are expected to be more accurate than in case of fitting isolated momentum range.

Fitted parameters of the function (4) are plotted in fig. 13–15. The parameters can be compared with each other in fig. 16.

Red point are obtained from distributions when width of the momentum bins was 50 MeV. This means that the first red points in each graph represent parameters of the fit when only tracks with momentum from 500 MeV to 550 MeV are used. Second red points are from fit with tracks that have momentum from 550 MeV to 600 MeV, etc.

Black points are analogic to the red ones, only the width of the momentum bins was 40 MeV.

Displayed errors of momenta (almost invisible) are RMS of the momentum distribution in the momentum bin. Vertical errors are errors of the fitted parameters.

Decreasing of the range of momentum bin ends in reducing statistics, so the parameters would be less accurate. Extending the range would smear the difference in deposited charge between higher and lower momentum tracks.

The goal is only to describe the dependence, there is no ambition to answer questions why the dependences have these shapes or what the physical variables are. The MPV distribution can be explained by the Bethe-Bloch formula. While pions are at the ionization minimum, MPV is more-less the same. Kaons and protons are heavier so the function is decreasing. Both Landau and Gauss distribution are normalized, but the normalization is corrupted by introducing pedestal y_0. Thus the amplitude A is smaller than it would be for convolution without pedestal. As pedestal grows, the amplitude has to decrease in order to preserve the normalization of the ρ distribution.

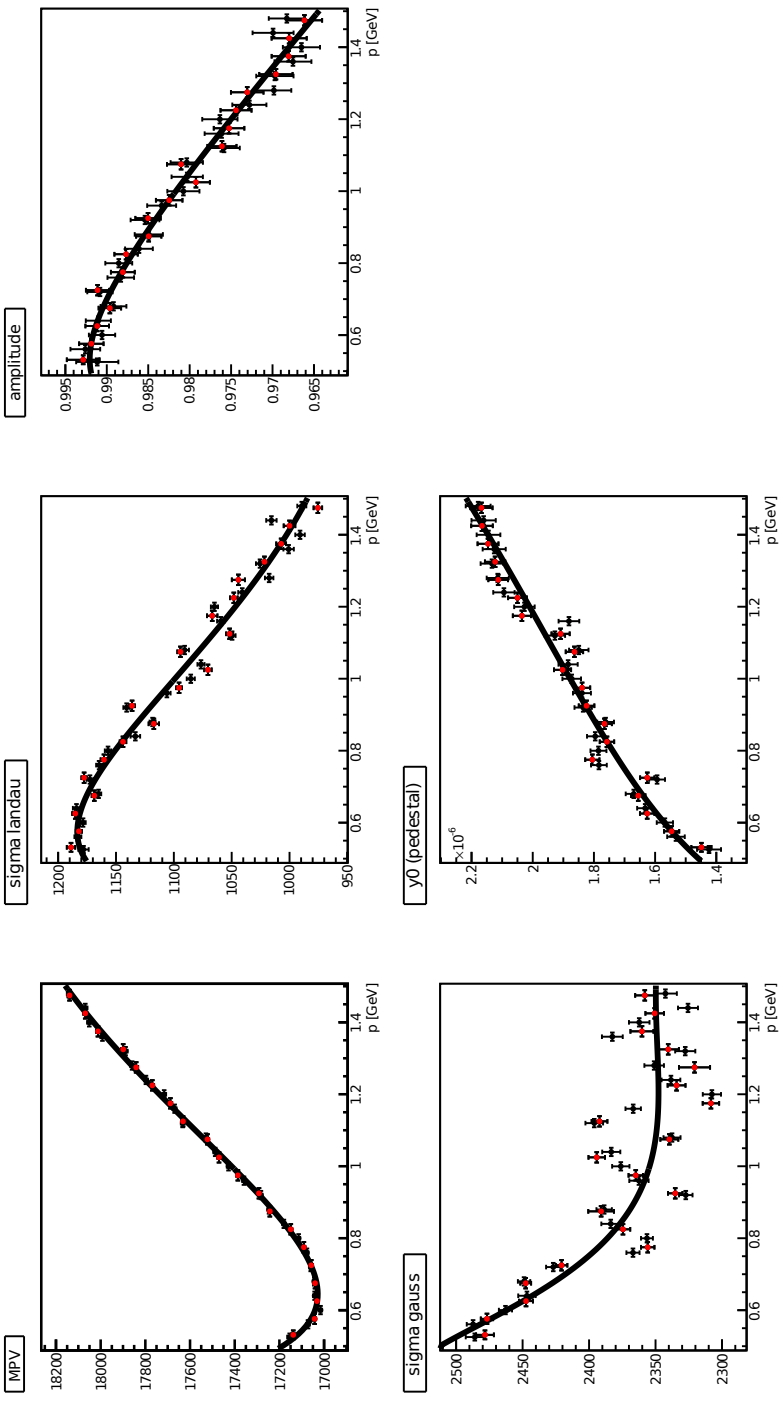


Figure 13: Parameters of pions

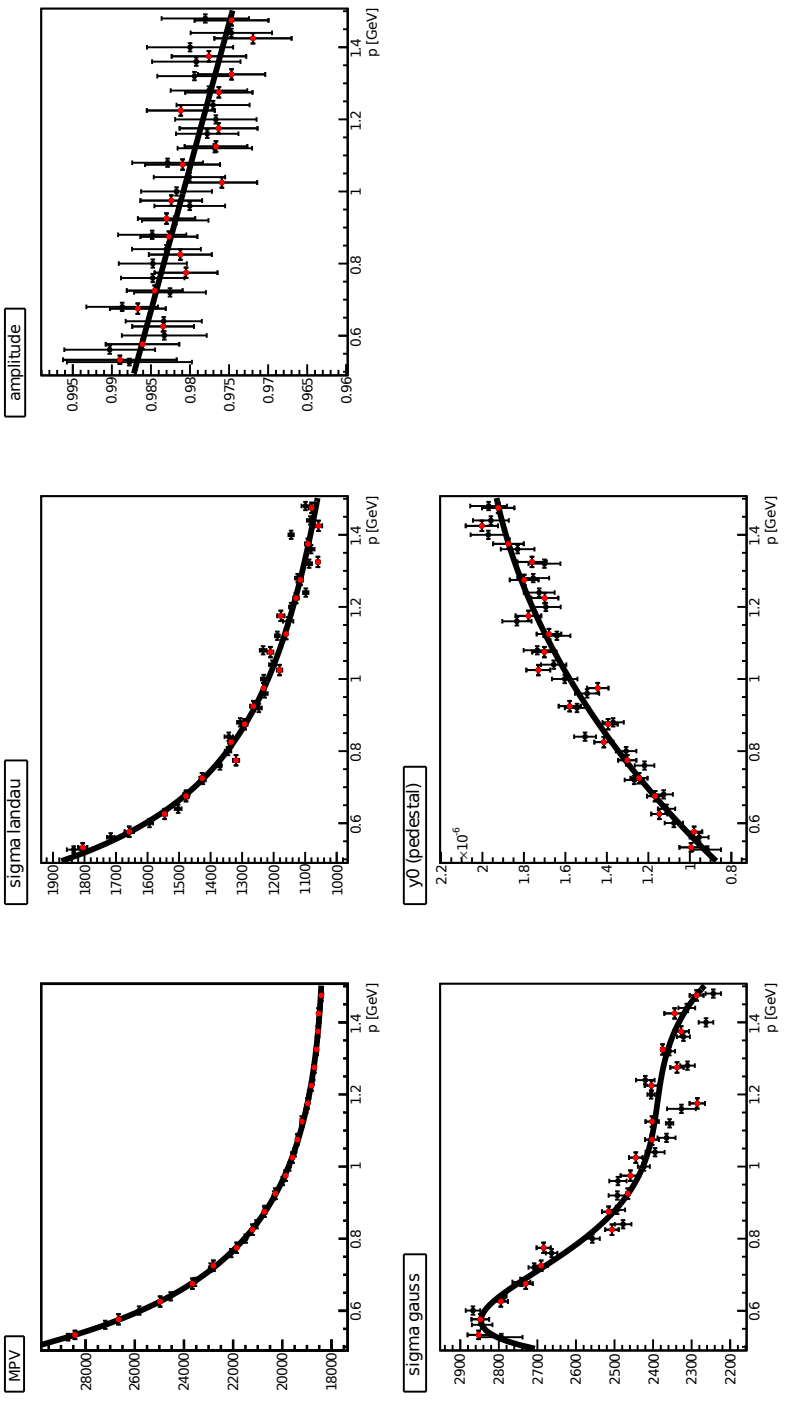


Figure 14: Parameters of kaons

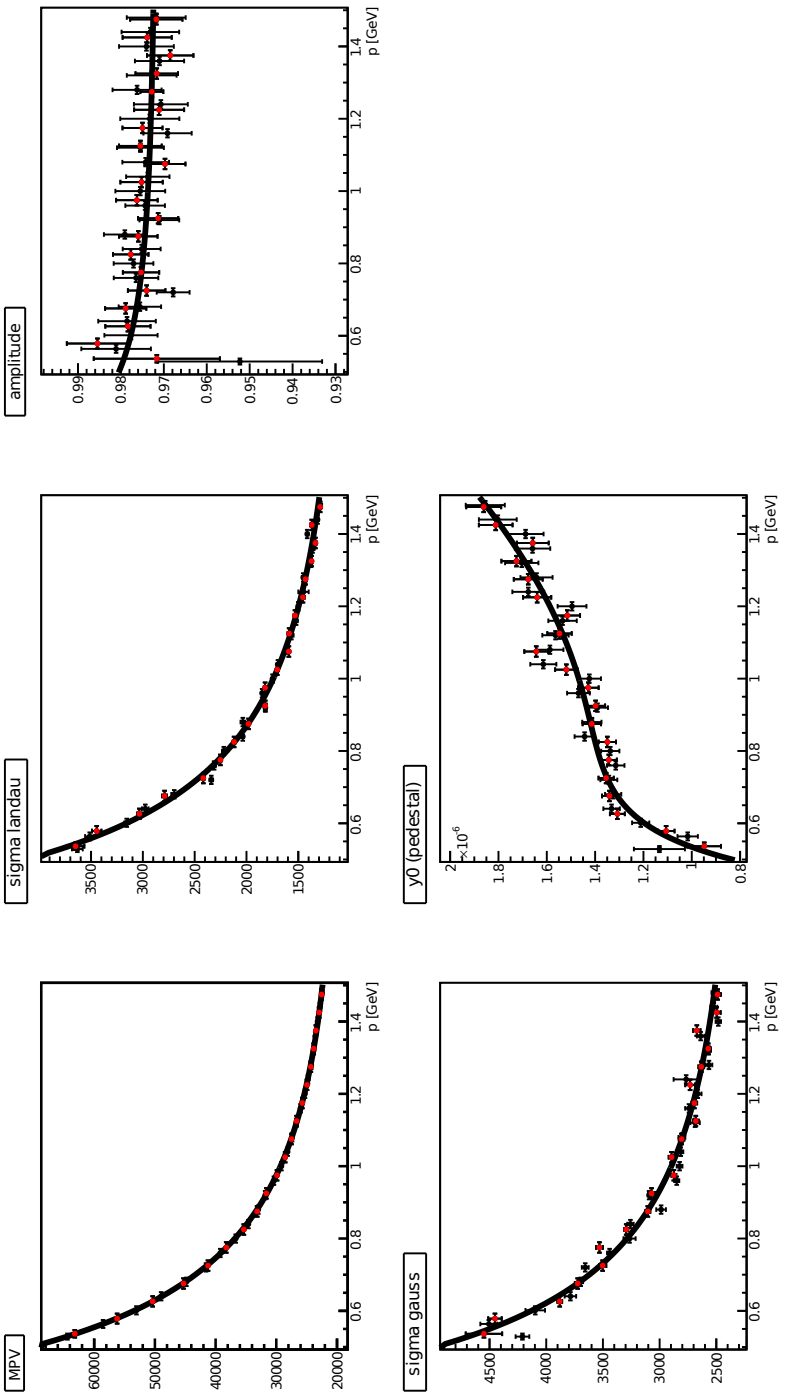


Figure 15: Parameters of protons

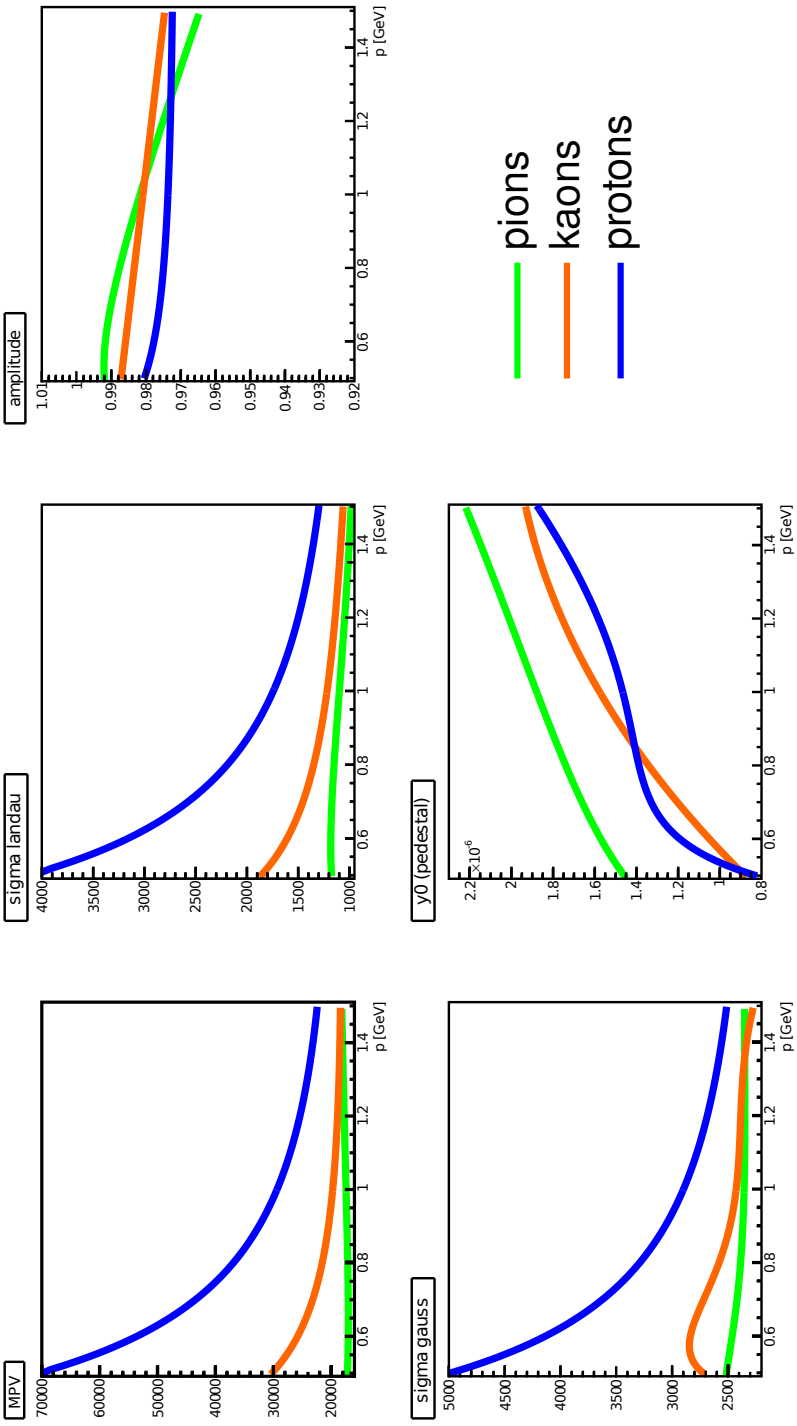


Figure 16: Parameters of pions, kaons and protons

Table 2: Parameters of pions

parameter	a_{-2}	a_{-1}	a_0	a_1	a_2
MPV	-	1880±50	10600±200	6400±200	-1480±60
σ_{landau}	-	-230±20	2070±70	-980±70	240±30
$A \cdot 10^3$	-2.8±0.8	-	1022±3	-38±3	-
σ_{gauss}	-250±20	1400±40	-	1550±50	-360±20
$y_0 \cdot 10^7$	-	-2.7±0.3	19.5±0.6	-	2.0±0.2

Table 3: Parameters of kaons

parameter	a_{-2}	a_{-1}	a_0	a_1	a_2
MPV	3720±10	-	14540±30	1480±20	-
σ_{landau}	198±4	-	1110±20	-93±11	-
$A \cdot 10^3$	-	-	993±3	-13± 3	-
σ_{gauss}	-3700±300	17700±1500	-26400±2400	19900±1800	-5000±470
$y_0 \cdot 10^7$	-	-0.4±0.3	-	23±1	-6.3±0.8

Table 4: Parameters of protons

parameter	a_{-2}	a_{-1}	a_0	a_1	a_2
MPV	13970±20	-	13640±60	1690±40	-
σ_{landau}	798±4	-	943±5	-	-
A $\cdot 10^3$	2.3±1.2	-	971±2	-	-
σ_{gauss}	695±7	-	2200±10	-	-
$y_0 \cdot 10^7$	-14.3±2.7	49±10	-45±11	25±4	-

So A and y_0 are associated in a reasonable way. However, the origin of y_0 as well as σ_{landau} and σ_{gauss} distributions is not so clear.

Note that Y scale of plots in fig. 13–15 are different for each kind of the particles. Thus discrepancy of points in σ_{gauss} in fig. 13 is comparable with discrepancy in fig. 14 or 15 – it is around 50 MeV – despite it is much less visible.

Parameters of the fits are in tables 2 – 4.

Finally, we have to store 75 float numbers. It is not a lot, since we consider that at the beginning we had tens of histograms, each with 150 bins. When the histograms were replaced by the function (4) we had the same number of histograms but each of them was represented only by 5 float numbers (MPV, σ_{landau}, A, σ_{gauss} and y_0). Finally, by finding dependence of these parameters on momentum, tens sets of 5 numbers is reduced to only 75 numbers (3 types of the particles, 5 parameters, 5 element of the series). It is possible to save these numbers in some practical way and use it in further projects.

In fig. 17 the difference between an exact fit and the function (4) when fitted parameters are used can be seen. The histogram is filled with pion tracks with momentum from 900 MeV to 950 MeV. The difference can be hardly found.

This range was not selected randomly. As you can see in fig. 13, parameters obtained from the fit have quite large discrepancy from fitted dependence, namely for σ_{landau} and σ_{gauss}. Compare appropriate red points with black line in fig 13.

Note that σ_{landau} is not comparable with σ_{gauss}. Full width at half maximum (FWHM) is roughly $4\sigma_{landau}$ for Landau distribution, while FWHM of the Gauss distribution is $2\sqrt{2\ln 2}\sigma_{gauss} \approx 2.35\sigma_{gauss}$.

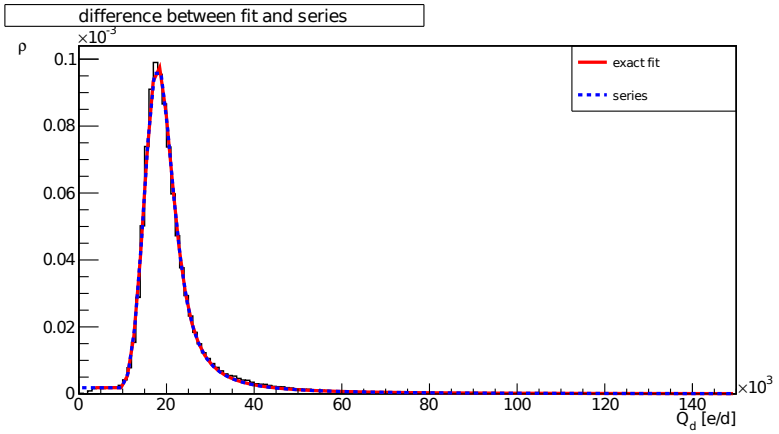


Figure 17: Comparison of the exact fit and the series (6)

4 Particle identification

This section will deal with identification of particles using deposited charge. It is hardly possible to say that a given track was a pion, but we aim at calculating the probability that it was a pion, a probability that it was a kaon and the probability of being a proton. No more particle species will be considered.

4.1 Calculation of probability

The method of the identification is based on the fact that deposited charge per unit length for given momentum p depends on the kind of the particle.

Here is an example of a simple reasoning: if the momentum is low and deposited charge per unit length is high, it will be probably proton.

However, there should be more precise method and we will try to find such a procedure.

We have already defined the probability density of depositing a charge Q_d in section 3.3. Consider a proton hit from the track with momentum e.g. 0.74 GeV and deposited charge per unit length e.g. $51.3 \cdot 10^3$ e. In fig. 18 there is the distribution of ρ in dependence of the deposited charge per unit length.

When we assume width of bin $\Delta Q_d = 1\,\mathrm{e/d}$, the probability that our hit falls into this bin is $\xi = 0.021 \cdot 10^{-3}$.

As we can see in fig. 3 (see section 2.4.2), only few tracks have only one hit, so we should consider general case of a track with N hits. Let's denote Q_{d1}, Q_{d2}, ..., Q_{dN} deposited charge per unit length for the first hit, the

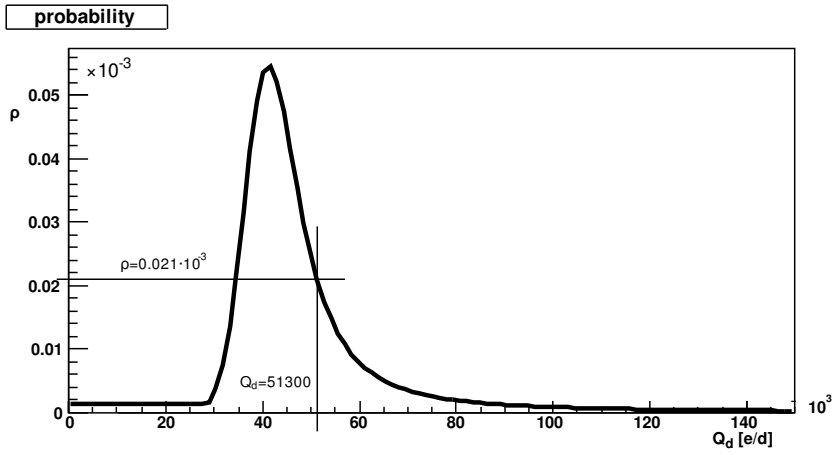


Figure 18: Probability ρ
proton, p=0.74 GeV

second hit, ..., the N-th hit along the track. Their probabilities according to (3) would be ξ_1, ξ_2, ..., ξ_N. The probability that we have a track with these deposited charges per unit length is:

$$\xi_{track} = \xi_1 \cdot \xi_2 \cdots \xi_N \tag{7}$$

ξ_{track} would be very small, but this is only the price for large spectrum of Q_d.

Now we will use the well-known theorem from the theory of probability:

$$P(B_\tau|A) \cdot P(A) = P(A|B_\tau) \cdot P(B_\tau), \tag{8}$$

where P(X) means probability that event X happened and P(X|Y) means probability that event X happened with requirement that event Y had happened.

In our case the event A is "we have measured deposited charges per unit length Q_{d1}, Q_{d2}, ..., Q_{dN}" and the event B_τ is "the particle passing through detector was particle τ ($\tau \in$ { pi, K, p, })".

When the momentum and the type of particle is known, we will have ρ distribution analogical to fig. 18. From the probability density distribution $\xi_{track,\tau}$ is obtained according to (3) and (7). Let's consider that $P(B_\tau)$ is 1/3 for all three particles, i.e. no type is preferred. We can use this assumption when the particles have the same multiplicity or when the information about their multiplicity is not known.

Right side of (8) is thus computable.

Because the hits must be made by one of the particle species, a condition has to be met:

$$P(B_{pi}|A) + P(B_K|A) + P(B_p|A) = 1 \tag{9}$$

Because P(A) is independent on kind of the particle and $P(B_\tau)$ is so far the same for all types of particles, we can compute the probability that the hits were made by particle τ:

$$p_\tau = P(B_\tau|A) = \frac{\xi_{track,\tau}}{\sum\limits_{\sigma=pi,k,p} \xi_{track,\sigma}} \tag{10}$$

Now let's explain the procedure to say something about the particle kind. When the probability p_{τ_0} is much higher than the other two, we can say that all particles with these probabilities are of the type τ_0. Naturally, this will introduce an error to the identification, because there is a small fraction of particles that are not of the type τ_0. We will deal with these errors in section 4.3.

The definition of "much higher" is crucial and it is represented by number $p_{\text{threshold}}$. When p_{τ_0} exceed it, the track is identified as τ_0. When no p_τ exceed it, the track is not identified as anything. It has no sense to set the threshold low. The value below 1/3 is absolutely meaningless. The reasonable values start at 1/2. In the section 4.3 this will be elaborated more on.

It is not always necessary to specify the kind of the particle. When we are interested in e.g. distribution of particles in a jet, the assignment of the track to the particle type is not necessary. The interesting thing is the distribution of the particles, which is obtained from a lot of collisions and whether one selected track was pion or kaon is not important and the probability from (10) is enough.

Because $\xi_{track,\tau}$ are compared with each other, their magnitude is not important. They can be multiplied by some constant without any consequences. This is practical in Athena tool (see section 6) when ξ_i are very small, the calculation would meet rounding error or even underflow and corrupt whole algorithm. To eliminate this danger, every ξ_i is multiplied by a high constant. This constant must be the same for all 3 kinds of particle and that is also the only condition.

4.2 Shares of pions, kaons and protons

However multiplicity of the kinds of the particles are not the same, as you can see on Monte Carlo sample in fig. 19. Fortunately, the extension of the described model is straightforward. The assumption of the same multiplicity will be omitted.

The shares of each particle $\mathrm{P}(\mathrm{B}_\tau) = s_\tau$ will be described by series:

$$\sum_{i=-3}^{2} b_i \cdot p^i, \tag{11}$$

where b_i are parameters to be determined and p is the momentum of the track. Fitted parameters are in table 5. No more powers of the parameters are needed as was shown by fitting.

The series describe the shares of the particles very well, however a problem occurs at share of protons. The series falls below 0 at approx. 515 MeV. This problem cannot be easily fixed because the first point is at 520 MeV. Any improvement was accomplished by reducing the range that the points represent and thus increasing number of points in fig. 19. The not fully satisfactory fix of this problem is to use value of 0 instead of unphysical negative values of (11).

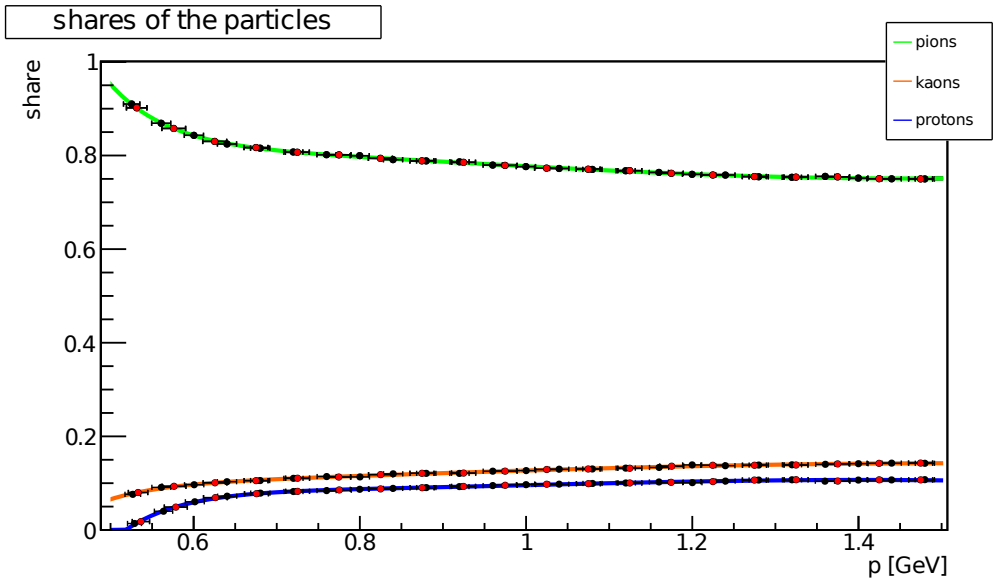


Figure 19: Shares of the particles

Table 5: Parameters of the shares of the particles

parameter	a_{-3}	a_{-2}	a_{-1}	a_0	a_1	a_2
s_{pi}	0.31±0.02	-1.30±0.09	1.9±0.1	-0.28±0.08	-	0.097±0.009
s_K	-0.093±0.007	0.41±0.03	-0.66±0.05	0.50±0.03	-	-0.034±0.004
s_p	-0.244±0.006	0.99±0.03	-1.42±0.04	0.85±0.02	-	-0.071±0.003

Equation (10) turns into

$$r_\tau = \frac{\xi_{track,\tau} \cdot s_\tau}{\sum_{\sigma=pi,k,p} \xi_{track,\sigma} \cdot s_\sigma}, \tag{12}$$

where r_τ is the probability that the hits were made by particle τ.

The identification process is done in the same way as in previous subsection. Since $r_{\text{threshold}}$ is given and r_{τ_0} exceed it, the track is identified as τ_0, however it is not necessary to specify the type of the particle if we are interested only in some particle distribution in jets or so.

As in the case of ξ_i, also s_τ can by multiplied by a constant that is the same for all three numbers.

It is very tricky to use information about multiplicity of the particles. On one hand, there is a fact that pions are more common than kaons or protons, so most of the particles should be identified also as pions. However, on the other hand, there is uncertainty of knowledge of each share. Shares presented in fig. 19 are based on all tracks in the sample. When we are

interested only in selected tracks (e.g. in those that are candidates to be a product of a decay of an unstable particle), the shares could be different.

It is important to carefully decide whether we will not use information about shares of particles and the identification will be based only on (10) or the information will be used and the identification will be based on (12). When the second option is chosen, the shares have to be meaningful, otherwise much more uncertainty would be introduced to the identification.

4.3 Efficiency & contamination

The performance of the identification process is described by two variables – efficiency ε and contamination c. In order to recognize their meaning follow your intuition. Efficiency is the share of particles that are correctly identified among all particles. Contamination is the share of particles that are "fare dodgers" among particles identified as specific kind. These variables could be appreciated in Monte Carlo simulation and considered as fair estimates for the real data.

More detailed description of efficiency and contamination can be found in appendix B.

As it is obvious from fig. 9 or 10, deposited charge strongly depends on kind of particle and also on its momentum. In fig. 20, 21 and 22 there are distributions of ε and c for pions, kaons and protons respectively while we assume the same share of all types of particles. Efficiency decreases with higher momentum, while contamination has opposite trend. This can be expected because energy loss are very similar for higher momentum and thus particles cannot be distinguished from one another.

In fig. 23, 24 and 25 there are similar distribution as described above, but during identification the information about relative abundance of different particle kinds is used. Distribution of kaons (fig. 24) ends at 1.15 GeV. The next bins have no entries.

The surprisingly low efficiency can be find in fig. 25 for momentum from 0.50 GeV to 0.55 GeV. This is consequence of the problem with share of the protons discussed above. When $s_p = 0$, there is no chance for a particle to be identified as proton.

The displayed errors are statistical. More about the errors can be found in appendix B.

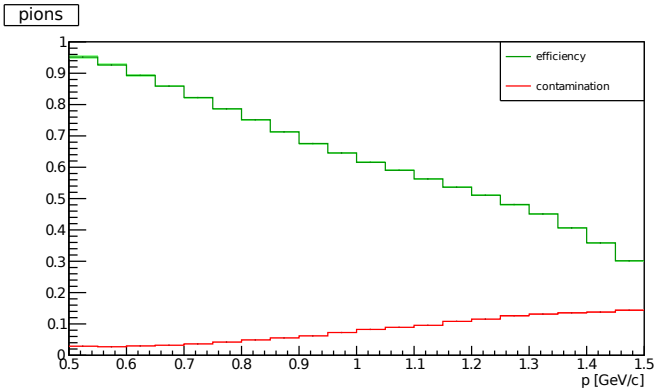


Figure 20: pions, $p_{\text{threshold}}$=0.5, formula (10) used

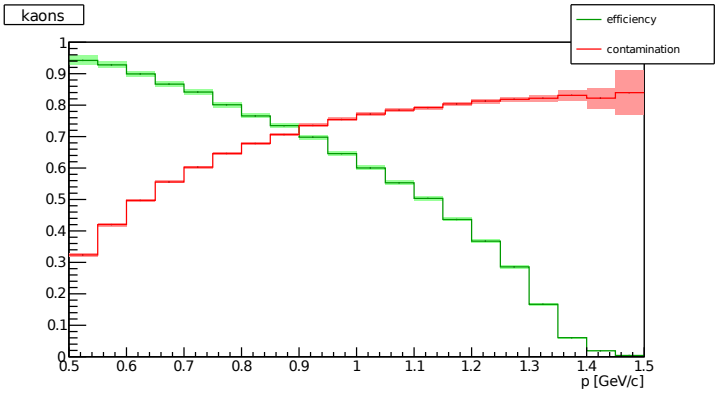


Figure 21: kaons, $p_{\text{threshold}}$=0.5, formula (10) used

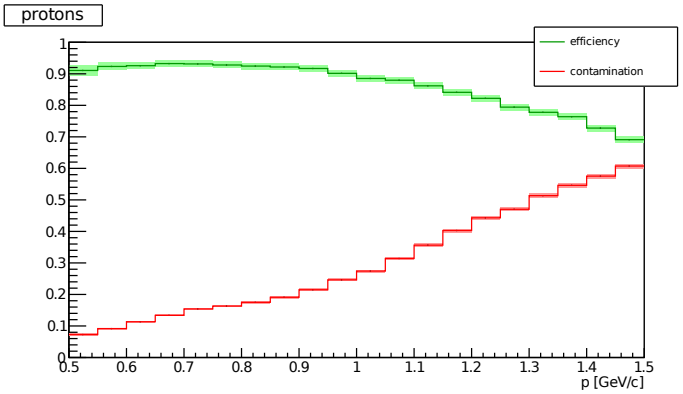


Figure 22: protons, $p_{\text{threshold}}$=0.5, formula (10) used

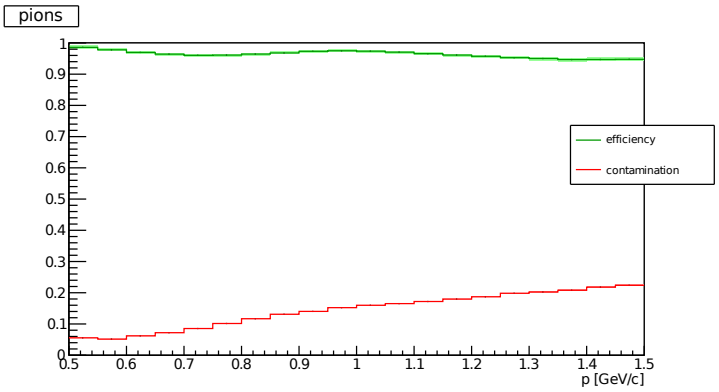


Figure 23: pions, $r_{\text{threshold}}$=0.5, formula (12) used

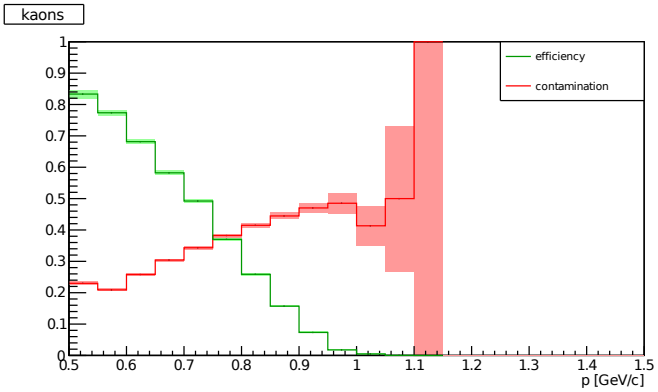


Figure 24: kaons, $r_{\text{threshold}}$=0.5, formula (12) used

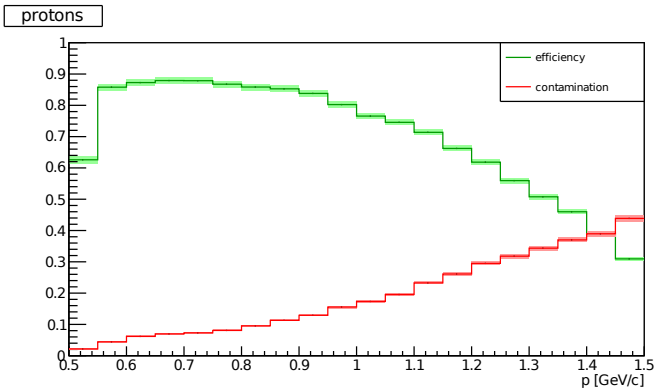


Figure 25: protons, $r_{\text{threshold}}$=0.5, formula (12) used

Efficiency and contamination also depend on number of hits per track. It is clear that the higher number of hits we have, the better result we expect. In fig. 26 there are efficiencies and contaminations for different numbers of hits per track. The result is not surprising and it is as expected.

The probability threshold is also one of the factors that modify distribution of efficiency and contamination. In fig. 27 there can be seen that efficiency goes lower with rising threshold. With rising threshold it is more difficult to reach this threshold and thus more particles are not identified as anything. Because they are not identified as *anything* and they are not identified wrongly, contamination is lower with higher threshold.

By setting the threshold high it is possible to get low contamination. The price you pay is less satisfying efficiency.

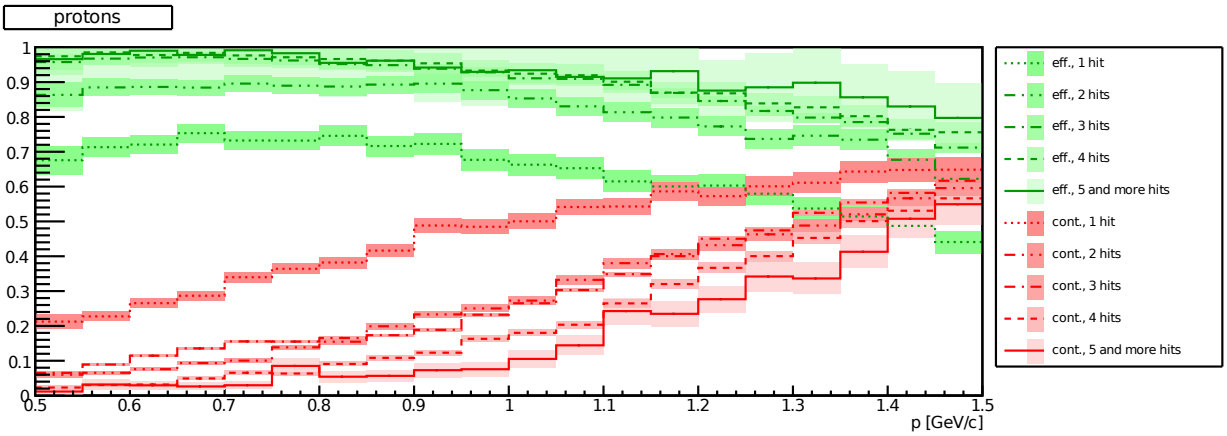


Figure 26: protons, $p_{\text{threshold}}$=0.5, formula (10) used

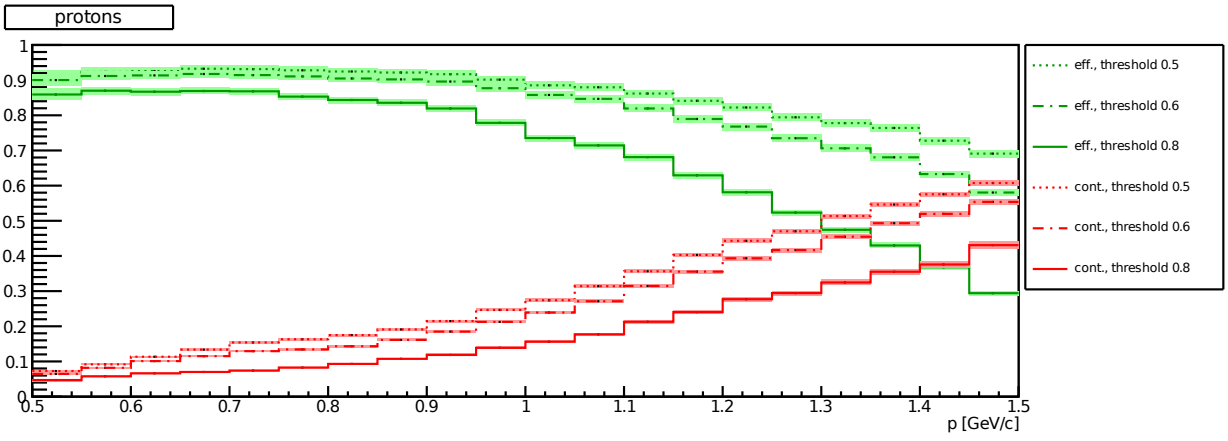


Figure 27: protons, formula (10) used

When the parameters of the momentum bins were replaced by the series (6), one would expect better performance. The series (6) create ρ distribution exactly for track's momentum. When parameters of momentum bins is used, ρ distribution would be a little bit inaccurate, because only values of the points in fig. 13–15 would be available, in contrast to continuous spectrum of values of the series (6).

In fig. 28 there is a distribution of efficiency. As you can see, the difference is hardly 0.5%. There is a difference between ρ distribution of the momentum bin and ρ distribution obtained in the (6), but this difference is usually not big enough to caused misidentification.

Situation of contamination distribution is very similar. When probability density ρ is computed according to (6), the performance is slightly better.

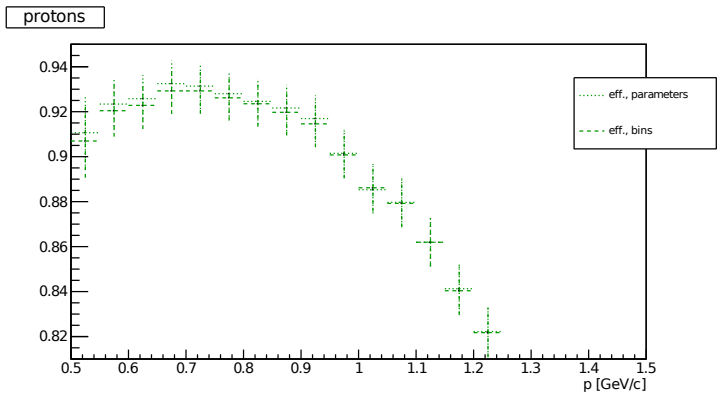


Figure 28: protons, $p_{\text{threshold}}$=0.5, formula (10) used

5 Monte Carlo vs. first data

The data used for the analysis are from the first runs in December 2009. The energy was 450 GeV per beam, i.e. $\sqrt{s} = 900\,\mathrm{GeV}$.

5.1 Events selection

It is quite practical when analysis is run only on selected events. E.g. those with some part of the detector switched off are not very usable. Moreover, analyzed events should originate from proton – proton collisions and not from proton – beam-gas collisions etc. The first problem is solved with runs and lumiblocks, the second one with bunch crossing identifier (BCID).

Each event is marked with run number, lumiblock and BCID.

Run number is associated with status of TDAQ system (trigger, data acquisition and detector control systems) and it specifies if the data are collecting or are not due to some testing or so. Only run numbers corresponding to the states useful for analysis were selected.

Runs are divided into lumiblocks. A lumiblock is a period of 120 s. Each lumiblock has assigned information about the status of the detector, i.e. which parts were on, which ones were off, about various thresholds etc. When some part of the detector is in stand-by mode, the data are being collected, but they are of low quality. Only lumiblocks with correctly operating detector were selected.

When every bunch crossed interaction point, BCID is assigned to the corresponding event. It is clear we are interested in "correct" BCID only, i.e. those assigned to bunch – bunch collision.

In table 6 there are selected events used for further analysis. The list was taken from [25].

5.2 Tracks selection

Not all tracks that are reconstructed are really interesting. ATLAS detector detects all particles passing through it and this turns to be problem. Cosmic muons did a good job for alignment but they are unwelcomed background for proton – proton (or heavy ions) collisions.

Due to malfunction, misalignment or noise, it is also possible that the reconstruction was wrong and the reconstructed track does not belong to any real particle. Those tracks are called fake.

It is necessary to distinguish tracks from collision from fake tracks, cosmic muons' tracks or tracks from beam gas events. For this purpose several cuts are used. Overview of these cuts is in table 7. The cuts are from [26].

Table 6: Good runs, lumiblocks and BCID

run	lumiblocks	BCID
141749	22–100	1, 2674
141811	126–165	1, 2674
142149	65–87	1, 2674
142154	22–35	1, 2674
142165	134–257	1, 101, 2774
142171	217–229	1, 101, 2774
142174	8–16, 18–47	1, 101, 2774
142189	140–147	1, 101, 2774
142191	7–36, 141–233	1, 101, 2774
142193	33–153	1, 101, 2774
142195	11–54	1, 101, 2774
142383	260–282	1, 101, 182, 992, 1073, 1883, 1964, 2774, 2855

Table 7: Track Cuts - 900 GeV

p_T	$>500\,\mathrm{MeV}$
$\lvert\eta\rvert$	< 2.5
♯ pixel hits	≥ 1
♯ SCT hits	≥ 6
$\lvert d_0\rvert$	$<1.5\,\mathrm{mm}$
$\lvert z_0 \cdot \sin\theta\rvert$	$<1.5\,\mathrm{mm}$

Parameters of the tracks are computed with the respect to the vertex. In appendix C there is more information about parameters of the tracks as well as figures illustrating the meaning of them.

These cuts have already been used in the previous sections to select Monte Carlo tracks, despite it was not emphasized.

The impact of cuts above can be seen in fig. 29 where the distribution of number of the tracks per one unit of pseudorapidity is presented.

The black line represents all tracks. One can see artificial peaks probably from cosmic muons. Blue, green and red line represents distribution after applying some cuts. Orange line represents distribution when all cuts in table 7 are applied.

The performance of the used cuts can be seen in fig. 30. It is similar distribution as in fig. 29, but the tracks from unpaired bunch crossing are used. Tracks are thus from background and obviously all of them should be removed by cuts. The orange line representing $\mathrm{d}N/\mathrm{d}\eta$ when all cuts are applied is a flat distribution with maximum below 0.05.

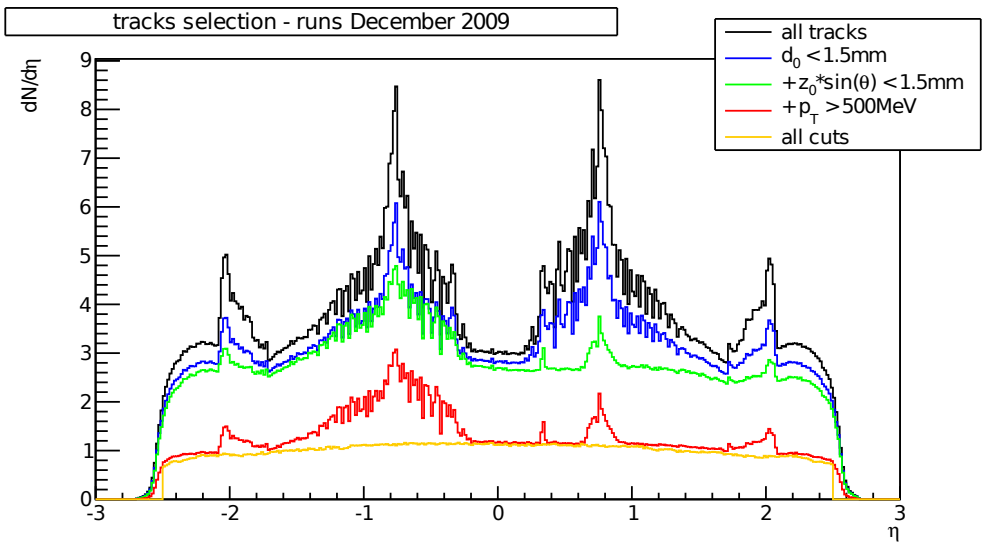


Figure 29: Distribution of number of the tracks per one unit of η
Note that cuts are applied cumulatively, e.g. green line is the distribution after applying d_0 and $z_0 \cdot \sin\theta$ cuts.

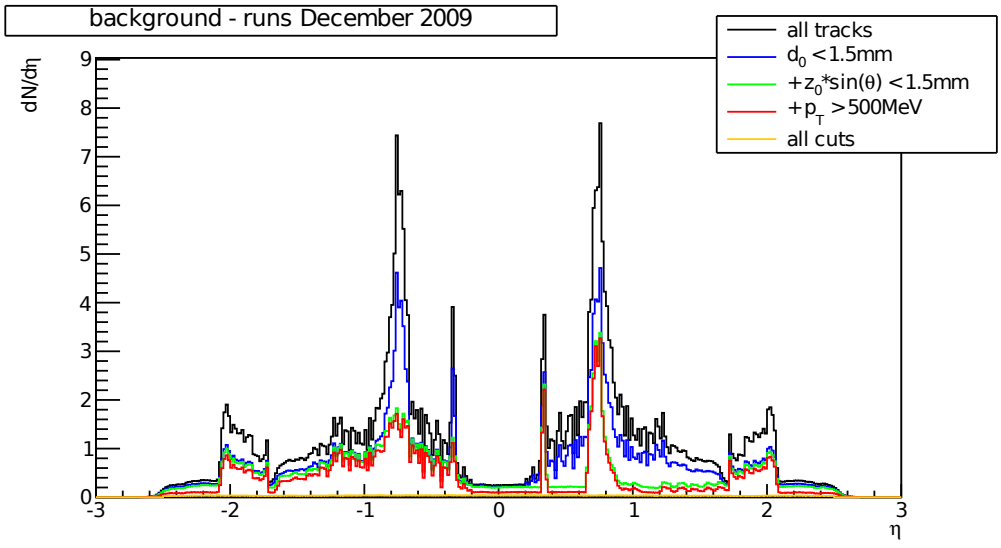


Figure 30: Distribution of number of the tracks per one unit of η
Distribution for tracks recorded in unpaired bunch crossing. Cuts are also applied cumulatively.

5.3 Comparing Monte Carlo with data

In this subsection, the equation (10) is used to compute the probabilities and the probability threshold is set to 0.5, unless otherwise stated.

In the fig. 31 and 32 there is deposited charge per unit length for all kinds of particles together. The first one is Monte Carlo and the second one is for data. While plotting distribution for data, the cut for p_T was dropped out to see the differences among particles more clearly. In the fig. 31 protons could be distinguished from other particles. In the fig. 32 the difference among particles is obvious, unfortunately only for lower momentum lower than 0.5 GeV that is not considered during the identification process. The difference for momentum above 0.5 GeV can be hardly seen.

In the fig. 33, 34 and 35, we can compare the distributions analogical to fig. 10.

If the identification was always correct (i.e. efficiency 100%, contamination 0%), the distribution of momentum vs. Q_d would look like fig. 33. Unfortunately this is impossible, so the Monte Carlo prediction of the distribution p vs. Q_d looks like in fig. 34. Finally, in the fig. 35, there are real data.

The most notable difference between perfect identification in fig. 33 and real identification (based on (10) or (12)) in fig. 34 and fig. 35 is missing hits just below the most probable Q_d. When a particle has such hit, it will be identified incorrectly. Kaons are thus identified as pion, because probability that this Q_d is deposited by pion is higher than the probability that it is deposited by kaon. The situation is the same for protons distribution. Protons are incorrectly identified as kaons.

Note that the single hits are plotted but for identification purpose all hits along the track are used.

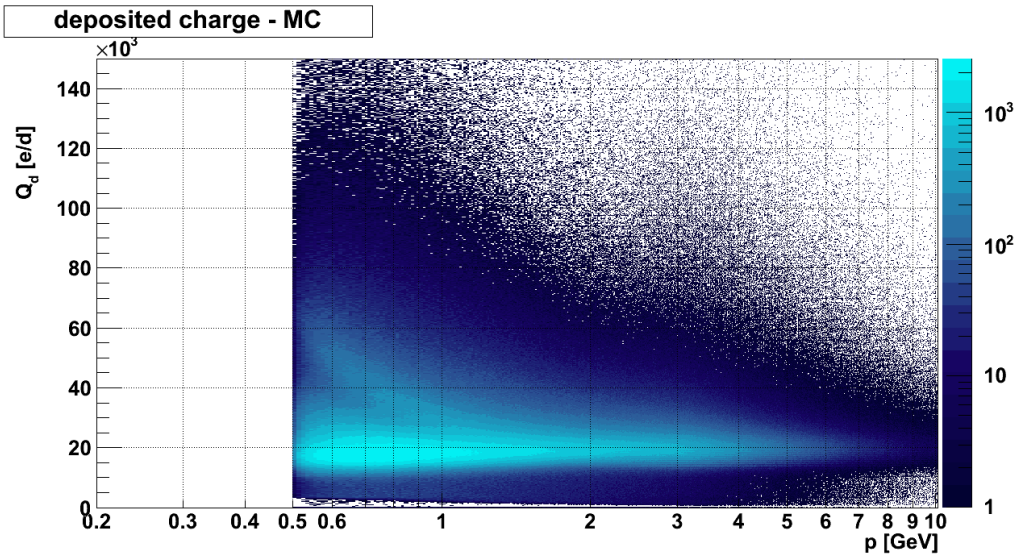


Figure 31: Q_d for all particles together, Monte Carlo

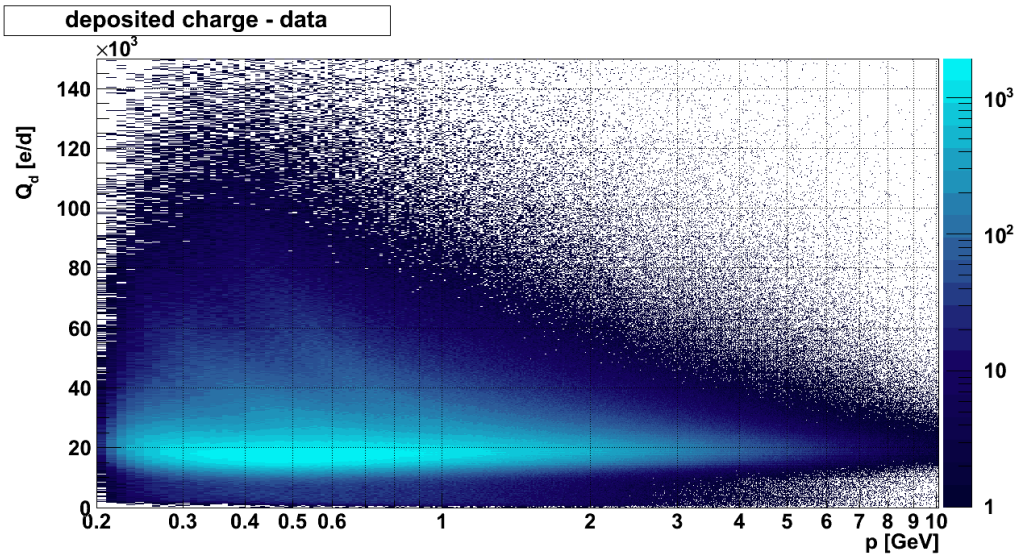


Figure 32: Q_d for all particles together, data

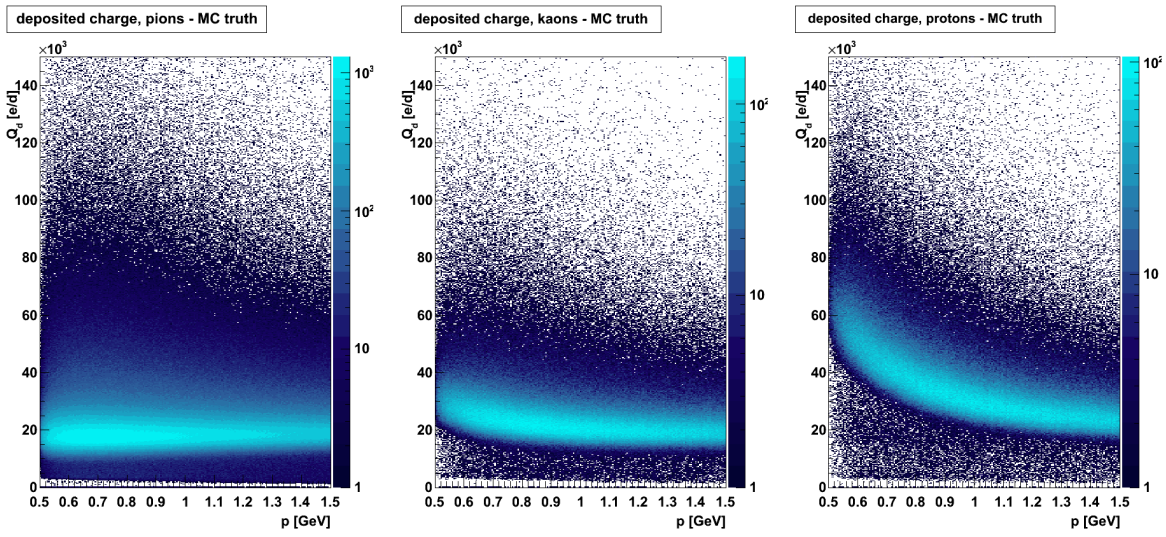


Figure 33: Perfect identification

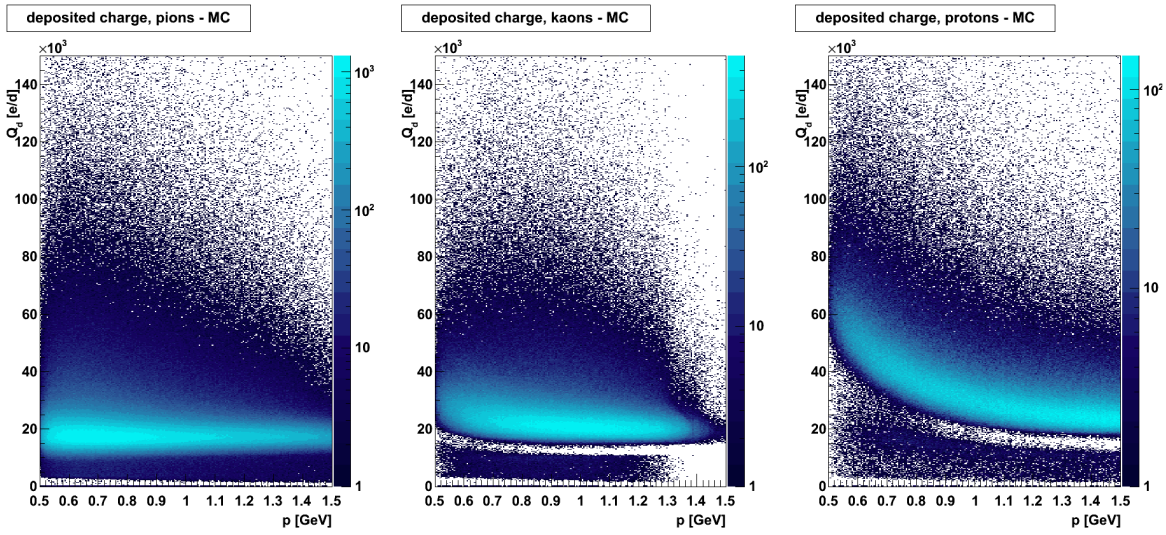


Figure 34: Identification of Monte Carlo particles

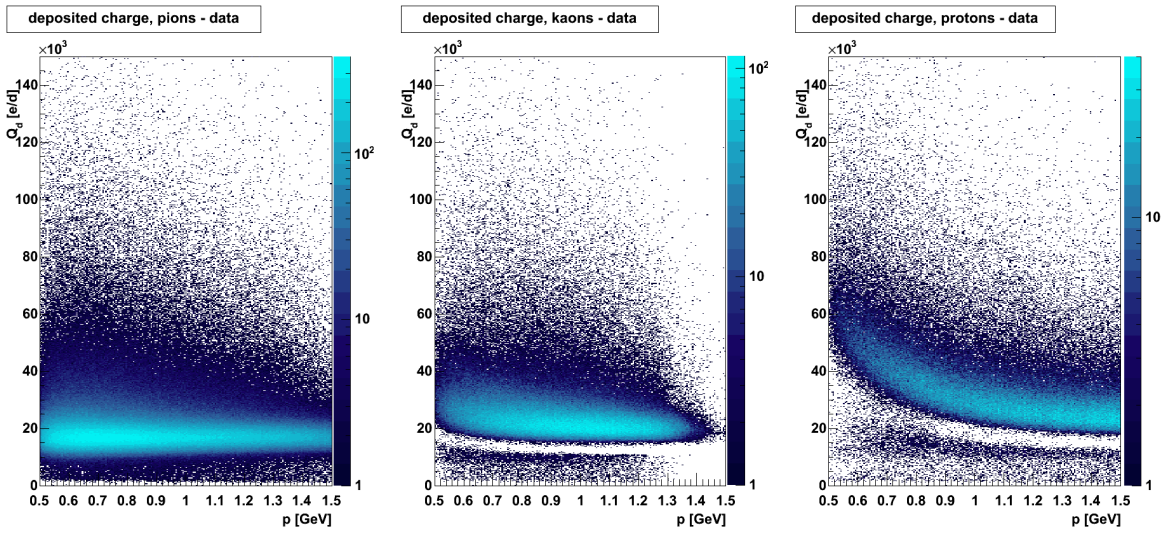


Figure 35: Identification of data

In the fig. 36 and 37 there are distributions p vs. probability obtained during identification process. The low momentum protons are identified quite easily which corresponds with the high probability. The identification of pions and kaons is not so unambiguous.

Especially in the fig. 36 in distributions for protons and kaons there are visible narrow lines. They come from tracks with only one hit. Those track cannot be identified very well. The data distributions (fig. 37) suffer from lower statistics but the lines are also visible, although with a bit of difficulty.

In the fig 38 and 39 there are similar distribution as described above, only formula (12) is used to compute the probabilities. The identification of kaons is rather unconvincing.

The line from proton tracks with one hit is visible in fig. 38. For low momentum these tracks suffer from low s_p. For higher momenta the difference in Q_d between protons and other particles smears and s_p relatively small compared with s_{pi} makes the situation worse. There is only indication of this line in fig. 39.

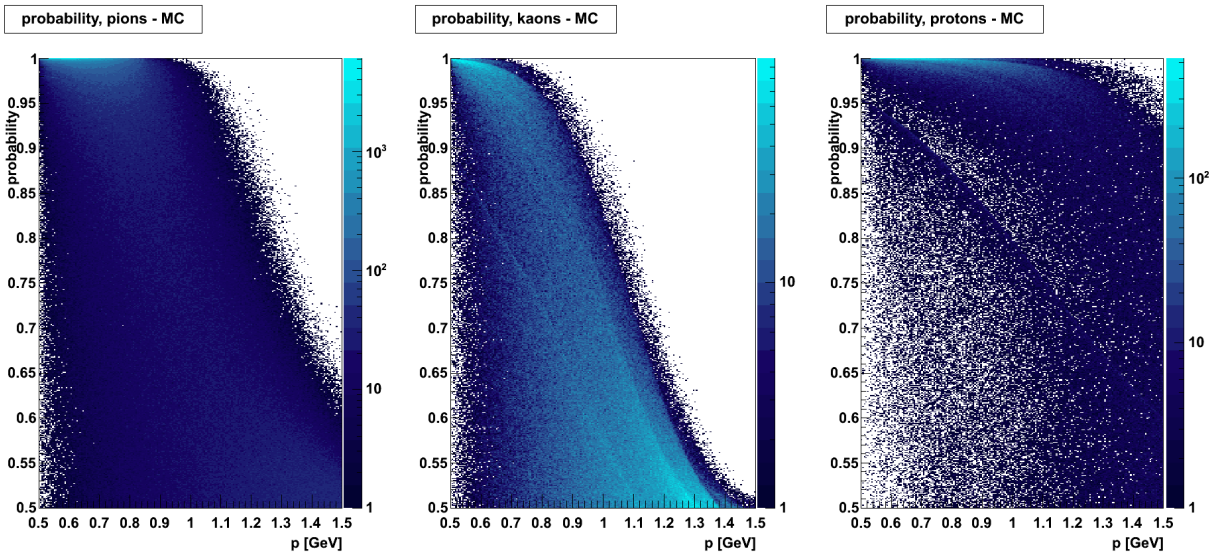


Figure 36: Probability, Monte Carlo, formula (10) used

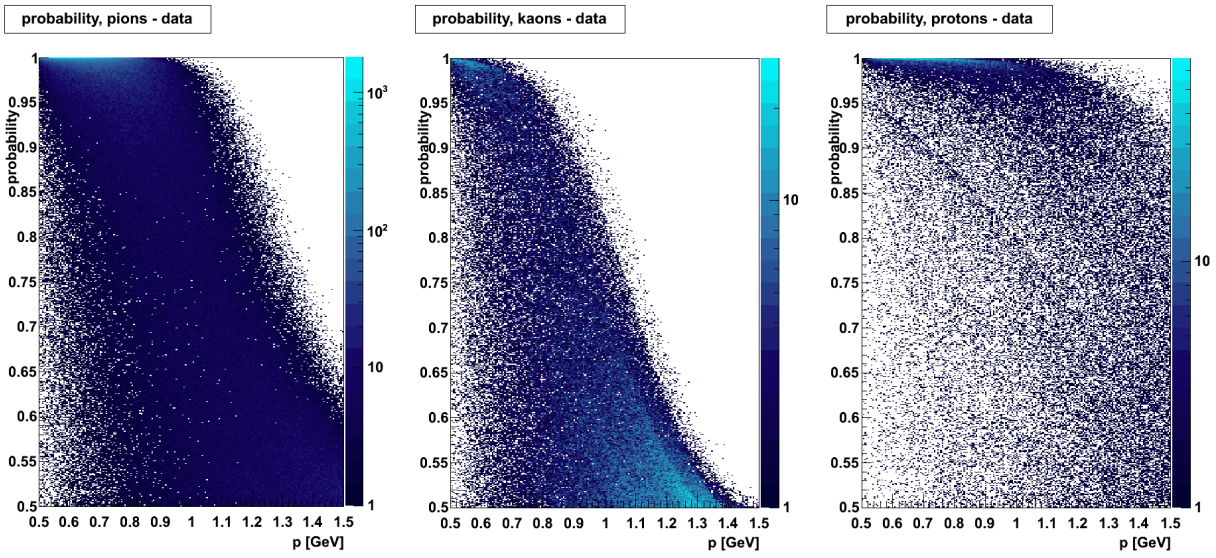


Figure 37: Probability, data, formula (10) used

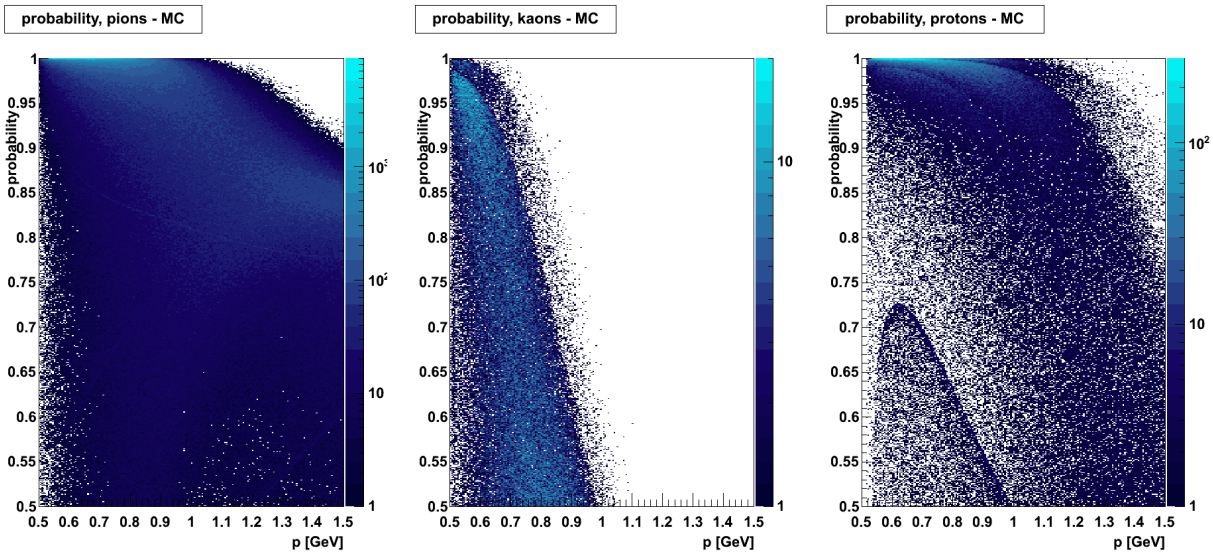


Figure 38: Probability, Monte Carlo, formula (12) used

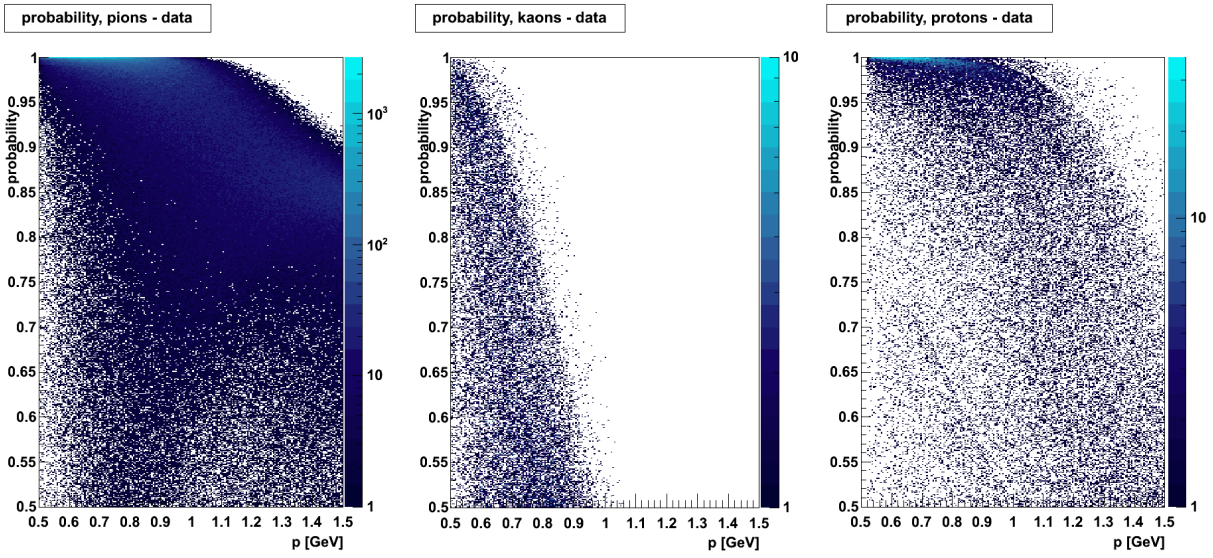


Figure 39: Probability, data, formula (12) used

It is also interesting to look whether the identification can restore the shares of the particles. In fig. 19 in section 4 the shares of the particles is shown when the identification was made by Monte Carlo truth. These shares are used in variables s_τ.

In the following paragraphs we will not identify each track, but we will use only probability p_τ or r_τ as discussed bellow equation (10).

In fig. 40 and 41 we can compare Monte Carlo and data when shares s_τ are used in calculation. In data there are more pions than in Monte Carlo, however the trends are the same.

In fig. 42 and 43 the shares based on (10) are shown. The information about Q_d of hits along a track is clearly not sufficient as the particles with higher momenta deposit very similar charges per unit length.

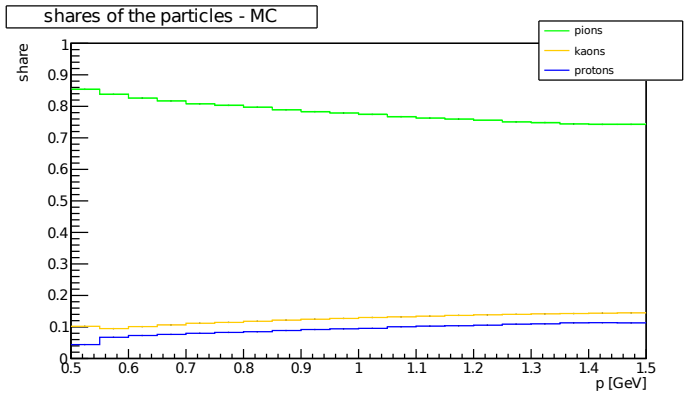


Figure 40: Shares of the particles, Monte Carlo, formula (12) used

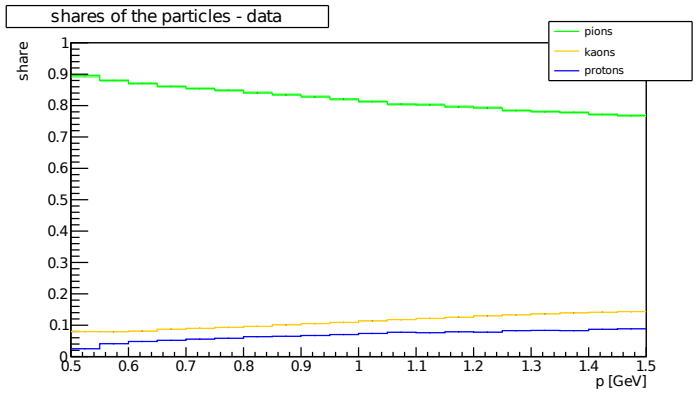


Figure 41: Shares of the particles, data, formula (12) used

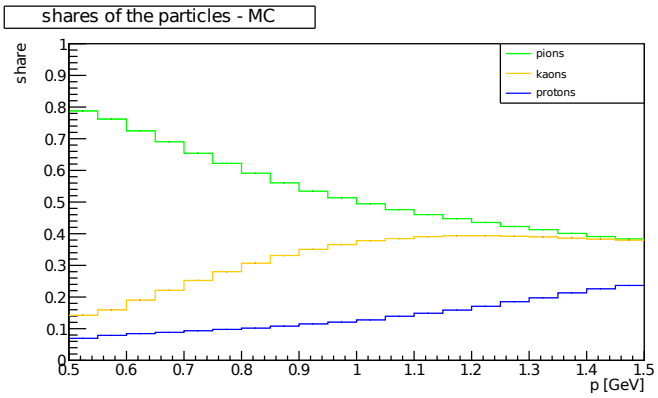


Figure 42: Shares of the particles, Monte Carlo, formula (10) used

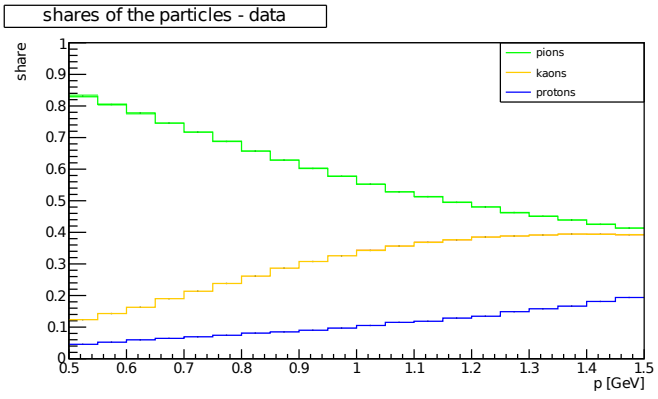


Figure 43: Shares of the particles, data, formula (10) used

6 Athena tool

Athena is a framework that is used to solve various physical problems at ATLAS experiment. It consists of many packages while each package deal with different task, e.g. in one package there are algorithms for finding jets, other package is responsible for validation of vertex fitting and finding, in another package there are defined some useful classes etc. It is also possible to write your own package where is your analysis. By logical concatenation of the tasks one could get desired results.

Some packages use so-called tools. Algorithms could delegate some tasks to tools. A tool has one certain purpose, e.g. find a vertex using FastFinder and a selected set of tracks. When you would like to use another algorithm to find a vertex, just change the name of the tool and Athena will do the rest.

For easy identification of the particles Athena tool has been written.

The main part is overloaded `identify` method. It is provided in two version:

- `Trk::ParticleHypothesis identify(const Trk::Track* track, float* prob = 0);`
- `Trk::ParticleHypothesis identify(const Trk::Track* track, std::vector<Trk::ParticleHypothesis>* partHypVec, std::vector<float>* probVec = 0);`

where `track` is a track we want to identify.

Both methods are returning most probable particle (`Trk::pion`, `Trk::kaon` or `Trk::proton`). When some error occurs (unsuccessful allocation on a heap etc.), the track has no pixel hits, the momentum of the track is out of range or none of the probabilities exceed the threshold, `Trk::undefined` is returned.

When `prob` is nonzero pointer, also the probability of the particle with the best match is saved in it. When any of the probabilities is not good enough, the best one is saved nevertheless, despite `Trk::undefined` is returned.

Since the information about the best matching particle is not enough, the second version of the method come on the scene.

`partHypVec` has to be valid pointer to `std::vector<Trk::ParticleHypothesis>` and thus this vector should be allocated in the heap before `identify` is called. Instead of using pointer, it is of course also possible to pass along the address of the object with the operator &.

In this vector all three possible particles are saved. The first position is reserved for the particle with the best match, the third place has the less probable one.

When `probVec` is a nonzero valid pointer, the appropriate probabilities are saved in it with the same order as in the previous vector.

If any of the probabilities is not high enough, `Trk::undefined` will be returned but also the vector(s) are filled correctly. When any other problem occurs, `Trk::undefined` is returned without filling any vector. You should always check the size of the returned vector.

There are two parameters that could be set by user:

- `ProbabilityThreshold`
- `UseShare`

`ProbabilityThreshold` is the threshold for probability of particles. It plays role of $p_{\text{threshold}}$ and $r_{\text{threshold}}$. Only when probability of some kind of particles exceed this threshold, the particle is identified. Otherwise `Trk::undefined` is returned. If `ProbabilityThreshold` is too small so more types of particles exceed it, the one with the highest probability will be returned.

When `UseShare` is set to nonzero, the algorithm will use fact, that the most of the particles are pions and the probabilities will be computed according to (12). This option decreases efficiency for kaons and protons with higher momentum. When `UseShare` is set to 0, the probability is based only on the deposited charges per unit length and the probabilities will be based on (10).

7 Conclusion

In this thesis we have studied the identification of the particles, especially the identification of pions, kaons and protons.

We have fitted the probability density ρ of deposited charge per unit length Q_d by convolution of the Gaussian with the Landau distribution, while this convolution was modified by pedestal to describe the distribution at low Q_d. Then we find dependence of the parameters of this function on momentum which allow as compute ρ distribution for any momentum. This resulted in slightly better performance in further identification, but the main motivation was to reduce amount of the processing data.

The identification method has been shown and it could handle tracks of momentum from 0.5 GeV to 1.5 GeV. It was introduced in two similar variations. The first of them takes into account only Q_d of the hits along the track, while the second one also uses multiplicities of the particles.

The performance of the identification has been shown in Monte Carlo. The performance is comparable with another particle identification [27]. However the method of the identification was not clearly shown in the cited presentation so the precise comparison is not possible.

The first data have been compared with the simulation with good agreement, despite the data has lower statistics than Monte Carlo.

The identification process has been written down as Athena tool, so further identification can be run directly in Athena.

Better performance could be accomplished by enlarging momentum range from 0.5–1.5 GeV to 0.1–1.5 GeV. Tracks above 1.5 GeV seem not to be able to be identified. Maybe small shift of this threshold could be done for protons, but the results will not be most likely very good. Tracks below 0.5 GeV are accessible in new 7 TeV data, however they are not accessible in 900 GeV data. The low momentum tracks are very promising as the differences in Q_d between the types of particles are more significant. Since the LHC should run at 7 TeV for a long time, the identification in this range would be useful. However the optimism is reduced by plot in [27] were the rapid decline of the efficiency for momenta below 300 MeV is shown.

References

[1] CERN Council: *Approval of the Large Hadron Collider (LHC) Project*, (1994), CERN/2075/Final

[2] Chris Llewellyn Smith: *How the LHC came to be*, Nature **448**, 281 (19 July 2007), doi:10.1038/nature06076

[3] *First beam in the LHC - accelerating science*, http://press.web.cern.ch/press/PressReleases/Releases2008/PR08.08E.html

[4] *CERN inaugurates the LHC*, http://press.web.cern.ch/press/PressReleases/Releases2008/PR16.08E.html

[5] *Incident in LHC sector 3-4*, http://press.web.cern.ch/press/PressReleases/Releases2008/PR09.08E.html

[6] *The LHC is back*, http://press.web.cern.ch/press/PressReleases/Releases2009/PR16.09E.html

[7] *LHC ends 2009 run on a high note*, http://press.web.cern.ch/press/PressReleases/Releases2009/PR20.09E.html

[8] *LHC research programme gets underway*, http://press.web.cern.ch/press/PressReleases/Releases2010/PR07.10E.html

[9] *The accelerator complex*, http://public.web.cern.ch/public/en/Research/AccelComplex-en.html

[10] M. Aleksa, B. Gorini: *Status Report Run Coordination - presentation at ATLAS Weekly*, (6 July 2010), http://indico.cern.ch/getFile.py/access?contribId=0\&resId=1\&materialId=slides\&confId=72820

[11] Lyndon Evans and Philip Bryant (editors): *LHC Machine*, Journal of Instrumentation **3**, (2008), doi: 10.1088/1748-0221/3/08/S08001

[12] M. Aleksa, B. Gorini: *Run Coordination Report - presentation at ATLAS Weekly*, (8 June 2010), http://indico.cern.ch/getFile.py/access?contribId=1\&resId=0&materialId=slides\&confId=72816

[13] Martin Rybář, *private discussion*

[14] ATLAS Collaboration: *ATLAS - Letter of Intent for a General Purpose pp Experiment at the Large Hadron Collider at CERN*, (1992), CERN/LHCC 92-4

[15] ATLAS Collaboration: *ATLAS Detector and Physics Performance - Technical Design Report*, (1999), CERN/LHCC 99-14/15

[16] ATLAS Inner Detector Community: *Inner Detector - Technical Design Report*, (1997), CERN/LHCC 97-16/17

[17] ATLAS Pixel Detector Community: *ATLAS Pixel Detector - Technical Design Report*, (1998), CERN/LHCC 98-13

[18] *ATLAS completes world's largest jigsaw puzzle*, `http://press.web.cern.ch/press/PressReleases/Releases2008/PR02.08E.html`

[19] G. Aad et al.: *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3**, (2008), doi: 10.1088/1748-0221/3/08/S08003

[20] G. Aad et al.: *ATLAS pixel detector electronics and sensors*, Journal of Instrumentation **3**, (2008), doi: 10.1088/1748-0221/3/07/P07007

[21] Nicoletta Garelli et al.: *The Tuning and Calibration of the Charge Measurement of the Pixel Detector*, (2009), ATL-COM-INDET-2010-017

[22] G. Aad et al.: *The ATLAS Inner Detector commissioning and calibration*, (7 Jun 2010), arXiv:1004.5293v2 [physics.ins-det]

[23] C. Amsle et al. (Particle Data Group): *Review of Particle Physics*, Physics Letters **B667**, 1 (2008)

[24] S. Hancock et al.: *Energy loss and energy straggling of protons and pions in the momentum range 0.7 to 115 GeV/c*, Physical Review A **28**, 615–620 (1983), doi: 10.1103/PhysRevA.28.615

[25] *Useful Runs with Good Data* `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/InDetPerformanceDQFirstRuns#Useful_Runs_with_Good_Data`

[26] *Minimum Bias Events Selection, Track Selection for 900 GeV analysis*, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MinimumBiasEventsSelection?rev=21#Track_Selection_for_900_GeV_anal`

[27] C. Gemme et al.: *Updates of Pixel dE/dx measurements - presentation at ID week*, (24 June 2010), `http://indico.cern.ch/getFile.py/access?contribId=50\&sessionId=2\&resId=0\&materialId=slides\&confId=74540`

[28] M. Marucho, C. A. Garcia Canal, Huner Fanchiotti: *The Landau Distribution for Charged Particles Traversing Thin Films*, (2006), arXiv:hep-ph/0305310v3

[29] *TMath namespace, function Landau()* `http://root.cern.ch/root/html/TMath.html#TMath:Landau`

[30] `https://svnweb.cern.ch/trac/atlasoff/browser/Tracking/TrkEvent/TrkParameters/trunk/doc/images/Perigee.png`

A Landau distribution

The Landau distribution describes roughly the fluctuation of the energy loss of particles in a thin layer of a material. It can be written as:

$$\lambda(x) = \frac{1}{\pi} \int_0^\infty \mathrm{e}^{-xy}\, y^{-y} \sin(\pi y)\ \mathrm{d}y \tag{13}$$

and it is drawn in fig. 44 [28]. The Landau distribution has no free parameters. They can be introduced by:

$$x \longrightarrow \frac{x - MPV}{\sigma_{landau}}, \tag{14}$$

where MPV stands for "most probably value".

An integral of (13) is 1. An integral of the distribution modified with (14) is σ_{landau}. Naturally, the integral can be changed by an amplitude A when (13) is multiplied by it.

A maximum of the Landau distribution is at $x = -0.22278$. A maximum of the modified distribution is at $x = MPV - 0.22278$. MPV is thus only *approximately* the most probable value, however it is introduced this way in ROOT [29] and it would be nice to be consistent.

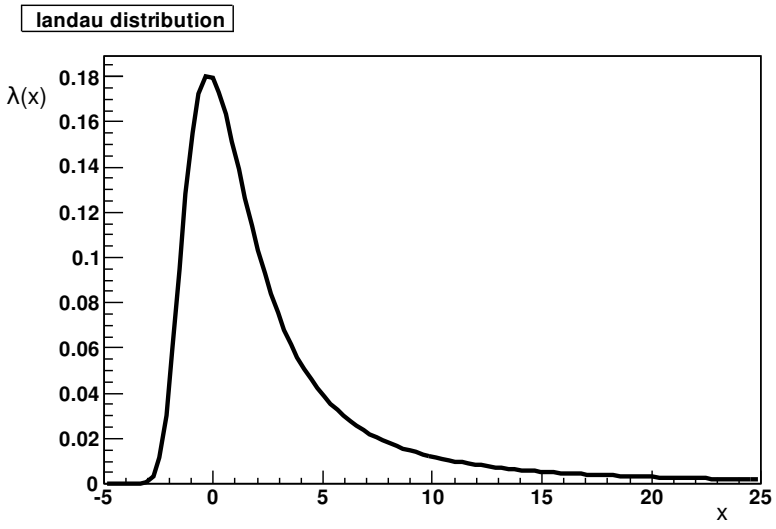


Figure 44: Landau distribution

B Efficiency & contamination

Performance of the identification can be described by two parameters – efficiency and contamination. Efficiency is fraction of particles that are identified correctly over all MC truth particles. Contamination is fraction of incorrectly identified particle over all identified particles.

Table 8: Efficiency and contamination

identified as → MC truth ↓	pion	kaon	proton	don't know
pion	A	B_1	B_2	B_3
kaon	C_1	D_{11}	D_{12}	D_{13}
proton	C_2	D_{21}	D_{22}	D_{23}
other	C_3	D_{31}	D_{32}	D_{33}

Consider that we are interested in identification of pions.

In the first row of table 8 there are all pions generated by Monte Carlo. Some of them are correctly identified as pions. This amount is A. Some of them are incorrectly identified as kaons or protons. These amounts are B_1 and B_2 respectively. And some of them are not identified as anything (B_3).

Efficiency of pion identification is:

$$\varepsilon_{pion} = \frac{A}{A + B_1 + B_2 + B_3} \tag{15}$$

In the first column of table 8 there are all particles identified as pions. Some of them are really pions (A), but some of them are kaons (C_1), protons (C_2) or another particles (C_3).

Contamination of pion identification is:

$$c_{pion} = \frac{C_1 + C_2 + C_3}{A + C_1 + C_2 + C_3} \tag{16}$$

Efficiency and contamination of other particles identification are defined in analogic way.

Statistical error of efficiency σ_ε or contamination σ_c can be computed as:

$$\sigma = \sqrt{\sum_i \left(\frac{\partial \varepsilon}{\partial x_i} \sigma_{x_i} \right)^2}, \tag{17}$$

where x_i stands for appropriate variables, e.g. for A, B_1, B_2 and B_3 in case of efficiency of pion identification. σ_{x_i} is estimated as $\sqrt{x_i}$. The statistical

error of ε_{pion} is:

$$\sigma_{\varepsilon_{pion}} = \left(\frac{A\left(2A + B_1 + B_2 + B_3\right)}{\left(A + B_1 + B_2 + B_3\right)^3} \right)^{1/2} \tag{18}$$

The statistical error of c_{pion} is:

$$\sigma_{c_{pion}} = \left(\frac{\left(C_1 + C_2 + C_3\right)\left(A + 2C_1 + 2C_2 + 2C_3\right)}{\left(A + C_1 + C_2 + C_3\right)^3} \right)^{1/2} \tag{19}$$

C Track parameters

Any track can be described by five parameters:

- d_0
- z_0
- ϕ_0
- θ
- q/p

It is very important to say, that these parameters are computed with the respect to the given point. When you change this point, some parameters will be changed as well.

The meaning of the parameters is illustrated in fig. 45. The given point is denoted P. Red line represents the track in the detector. It is a part of a helix with a axis parallel to Z axis. Black bold line is the XY projection of the track while Z position is given by Z coordinate of P. It is part of a circle.

Point A is so-called "point of closest approach". This is quite confusing because it is closest to P only in XY plane (compare point A with A').

The difference between P and A' (A in XY projection) is d_0. Difference between P and A in Z direction is z_0. Angle between X axis and tangent to the track projection at A is ϕ_0. Angle between Z axis and tangent to the track is θ. This angle is the same along the track so it does not matter at which point the angle is measured. Parameter q/p can be obtained from ρ (radius of the track projection) and magnetic field B while homogeneous field along Z axis is considered:

$$\frac{q}{p} = \frac{1}{\rho B}, \tag{20}$$

where q is charge of the particle and p is its momentum.

There are two usual ways how to select point P. When the vertex where the particle was created is not known, the nominal point of interaction is chosen. It should be geometrical center of the detector but the beam could be shifted.

When the vertex is known, it is chosen as P. Parameters d_0 and z_0 are then very small and we can use them in cuts.

Product $z_0 \cdot \sin(\theta)$ is often used for cuts as well. As seen on the fig. 46, it is the shortest distance between point A' and the track.

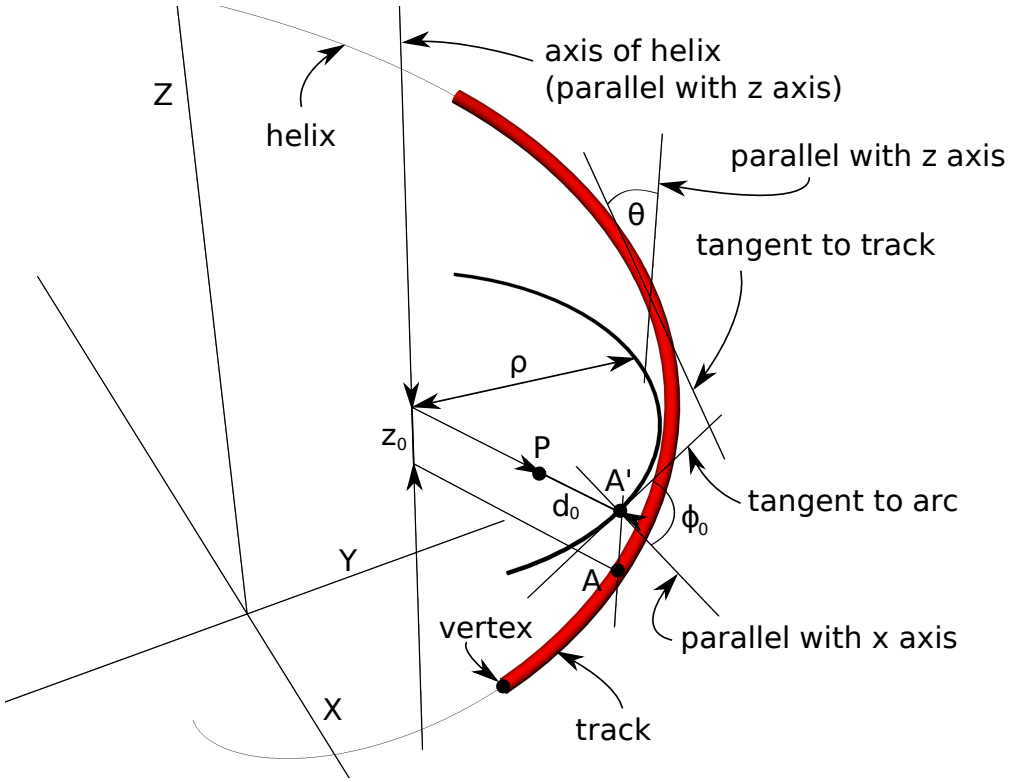


Figure 45: Track parameters
Figure is based on [30].

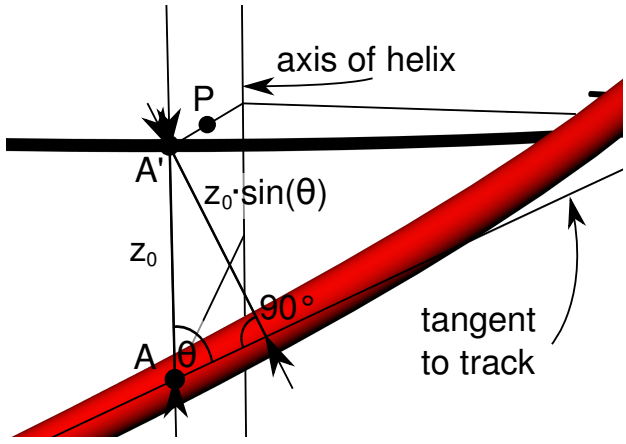


Figure 46: $z_0 \cdot \sin\theta$