CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Nuclear Sciences and Physical Engineering

Department of Physics

Diploma Thesis

Directional and spectrometric mapping of secondary radiation induced during hadron radiotherapy with miniaturized particle trackers

Bc. Lukáš Marek

Supervisor: doc. Ing. Carlos Granja, Ph.D.

Prague, 2020

Declaration:

I declare that I wrote my diploma thesis independently and with help of the cited bibliography.

I agree with usage of this project in the purport of the Act 121/2000 (Copyright Act).

Prohlášení:

Prohlašuji, že jsem svoji diplomovou práci vypracoval samostatně a použil jsem pouze podklady (literaturu, software, atd.) uvedené v přiloženém seznamu.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu 60 Zákona .121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

Název práce:

Směrové a spektrometrické mapování sekundární radiace indukované během hadronové terapie za pomocí miniaturizovaného dráhového detektoru

Autor: Bc. Lukáš Marek
Obor: Experimentální jaderná a částicová fyzika

Druh práce: Diplomová práce

Vedoucí práce: doc. Ing. Carlos Granja, Ph.D., ADVACAM, U Pergamenky 12, 170 00 Praha 7, Česká Republika

Konzultant: ———————————————————

Abstrakt: Tato diplomová práce pojednává o využití polovodičových detektorů v hadronové terapii pracující s uhlíkovými ionty. Hlavní aplikace spočívá v detekci sekundární radiace vytvořené interakcí primárních uhlíkových iontů. Tento fenomén naznačuje možné využití pro on-line monitorování pozice svazku přesněji jeho Braggova peaku. V rámci této práce jsou uvedeny v současnosti používané metody a jejich možné modifikace, především koincidenční procedura pro případ teleskopické architektury. Práce se soustředí na aplikaci metodologie při rozpoznání částic, která může sloužit ke zlepšení monitorování svazku. Byl navržen a testován zjednodušený algoritmus, podobný kNN algoritmu strojového učení, vykazující dostatečnou přesnost s možným on-line použitím. Závěrečná část práce se věnuje zkoumání a kvantifikaci nejistoty směrového mapování částic pocházející z jejich propagace médiem. K tomuto účelu je využit simulační nástroj FLUKA.

Klíčová slova: Hadronová terapie, polovodičové Timepix detektory, rozpoznání částic, FLUKA, nejistota směrového mapování

Title:

Directional and spectrometric mapping of secondary radiation induced during hadron radiotherapy with miniaturized particle trackers

Author: Bc. Lukáš Marek

Abstract: This diploma thesis deals with the usage of the semiconductor pixel detectors Timepix in the hadron therapy, namely carbon ion therapy. The main application lies in a detection of the secondary particles created during interactions of the primary ion carbons. This phenomenon suggests a possibility for an on-line monitoring of the beam's Bragg peak. In the frame of this work, an introduction of the currently used methodology is performed together with propositions of modifications especially in the case of the coincidence procedure for the telescope architecture. The thesis is concentrated on an application in a particle recognition which could be beneficial for the beam monitoring. A simplified methodology similar to the machine learning kNN algorithm is developed and tested, performing sufficient accuracy and offering a possibility for an on-line utilisation. The last part of the thesis is dedicated to an investigation and quantification of the particle propagation through a medium and the resulting uncertainty of the monitoring methods with the simulation tool FLUKA.

Keywords: Hadron therapy, semiconductor Timepix detectors, particle recognition, FLUKA, uncertainty of tracking

Acknowledgement

I would like to express my gratitude to my supervisor Carlos Granja for his patience and helpful advices. Furthermore, my special gratefulness goes to the whole Heidelberg team and Jan Jakubek. Last but not least, sincere thanks belong to my lovely wife Mirjam Marková.

Contents

Introduction

The hadron therapy, especially carbon ion therapy, offers a possibility how to efficiently treat cancer. The main beneficial feature of the carbon ions is hidden in the comprehended deposition of dose to a targeted medium and sparing a possible vital structures in its vicinity. The important part of the procedure is the adjustment of the beam characteristics, energy and aiming position, to perform accurate irradiation. Despite the fact that a careful planning precedes the therapy an unexpected altering of the geometry might appear mainly due to a patient-related sources (organ movements, creation of air cavities etc.). Therefore, an on-line monitoring is desired to obtain information about the current stopping beam position.

The most recent work proposed to utilize emerging particles, so-called secondary particles, from the medium during the therapy. Their origin can be tracked to the interaction of the ion carbons with a nuclei of the target via process of the fragmentation. This causes a disappearance of the primary particle, a carbon ion, and a creation of a new lighter ion as its fragment possessing very similar velocity. This features open a possibility to escape from the medium and consequently a detection is possible.

In this stage semiconductor pixel detectors of the Timepix family can be utilized. They possess characteristics suitable for the needed task: sufficiently accurate tracking of particles together with delivering a spectrometric information about the deposited energy. The current approach uses two detectors stacked on each other, the so-called telescope architecture.

The goal of this thesis is to introduce the current methods used for the data processing acquired from the Timepix detectors for the purposes of the beam monitoring and also if possible to propose modifications or a new methodology. It focuses on the secondary particles propagation through the medium and its effects on the monitoring. It is also suggested to develop a particle identification algorithm which could be potentially utilized for filtering of particles in the hadron therapy with a possibility of on-line processing (directly implemented on the measuring device).

The first chapter includes descriptions of the semiconductors application for detection of ionizing particles. The semiconductor features together with characterisation of the Timepix family detectors are summarized there. The seconds chapter comprehends the methodology used for the data analysis acquired during measurement with the Timepix detectors, namely a clusterisation and a coincidence or matching process. The chapter is concluded with a development of a particle identification algorithm which could be beneficial for an enhancement of the monitoring. The last chapter includes the investigation and quantification of the tracking uncertainty of secondary particles during their propagation through medium. The application of the suggested methodology for the data evaluation is included, especially the particle recognition.

Chapter 1

Position-sensitive Semiconductor Pixel Detectors

In the recent times semiconductor detectors are one of the most common measurement devices for the purposes connected with particle and nuclear experiments. Their features allow to perform combined measurements which would be hard to achieve with any other detectors. As examples, see the following brief list: a parallel precise position and energy measurement, high read-out frequency, low power consumption, direct electronic signal and many others [1].

The following chapter introduces the common principles and characteristics of the semiconductors and their application for the detection of charged particles. The second part includes description of the hybrid active pixel detectors of the Timepix family and their application for the measurement of the mixed radiation fields.

1.1 Semiconductor Features and PN junction

A charged particle traversing a medium is gradually losing energy mostly via electromagnetic interaction with the atomic electrons of the medium. This energy dissipation can manifest in electron liberation, creation of free charge. The transfer from a passive medium to a detector can be performed with a noticeable collection of this charge. A similar approach can be selected in the case of semiconductors as the medium. To be aware of the free charge gen-

eration a certain read out electronics is utilized. This is a brief explanation of the concept how the semiconductors are exploited as detectors of charged particles.
Semiconductors advantages mainly originate from their atomic structure. Their atoms are placed in a diamond lattice which is periodically repeated creating a crystal configuration, see the figure 1.1. The most common used semiconductors are silicon Si and germanium Ge. Each of the atoms have placed in its vicinity four neighbouring atoms with which it share four covalent electrons creating a covalent bond (the final shape resembles to tetrahedron).

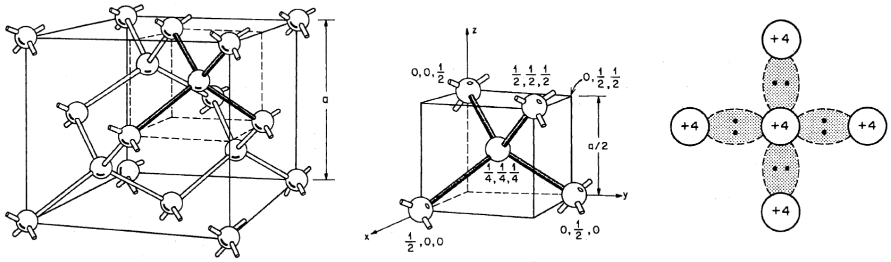


Figure 1.1: Diamond lattice, tetrahedron bond and its schematic representation [1].

The atomic energy levels in the lattice structure are modified by surrounding atoms. Instead of separate energy electron levels, they possess high density with small energy differences. More precise description can be made by a means of continues energy bands: a valance and conduction band. Between these bands, there is so-called band gap which for an intrinsic semiconductor does not include any energy level (some energy levels occur in presence of impurities). If an electron has sufficient energy, it can overcome the band gap and transfer to the valence one. This is manifesting as free charge carrier. A vacancy is produced in the valence band and this phenomenon is described as hole or effective free charge[1]. There is a difference between energy of the band gap and energy needed to create free electron-hole pair distinguishing semiconductors into an indirect and a direct class.The indirect semiconduc-

---

[1] It is not treated as particle with mass of an electron m_e but from quantum point view some mass can be incorporated with the hole and its electron but different from m_e [1]

tors are associated with not only demand on energy transfer to an electron in the valance band but also on some momentum transfer which is usually converted to the lattice as an excitation of lattice vibrations. An approximate relation can be derived for the energy required for electron-hole pair creation:

$$E_{\text{e-h creation}} \approx 2.8E_{gap} + 0.6 \text{ eV}, \tag{1.1}$$

where silicon possess $E_{\text{e-h creation}} = 3.6$ eV and the band gap is 1.12 eV. An important role in the real semiconductor applications plays a so-called doping. This processed is used to enhance a concentration of impurities which can effect the amount of created free charge (usually around from 10^{12} to 10^{18} cm^{-3} impurities). Two types of doping can be distinguished:

- n-type:
  Also called donor is a type of semiconductor which includes an impurities such a phosphorus or arsenic with additional valence electrons. Almost automatically one electron belonging originally to the impurity is freed and combination of negative ion and free electron is generated. The manifestation can be observed in the energy band structure when a new energy level is created close to the conduction band, see the figure 1.2.

- p-type:
  A term acceptor si also used for this kind of semiconductors. Creation of the acceptor is opposite to the donor case. An impurity with one smaller count of the valence electrons is inserted into the silicon crystal structure, for example boron or aluminium. The result is new energy level close to the valence band a one free hole per one impurity.

1.1.1 PN Junction and Carrier Transport

The connection of two oppositely doped semiconductors, called a p-n diode or PN junction, will generate a depletion region. The cause can be found in the composition, whereas the n-dopant includes positive ions the p-dopant negative ones. In the first stage after their connection, a free charges are moving from the place with higher concentration where an opposite free charge is present. This is resulting in the recombination process and only firmly fixed ions of doped semiconductors remain at their original positions. A space at

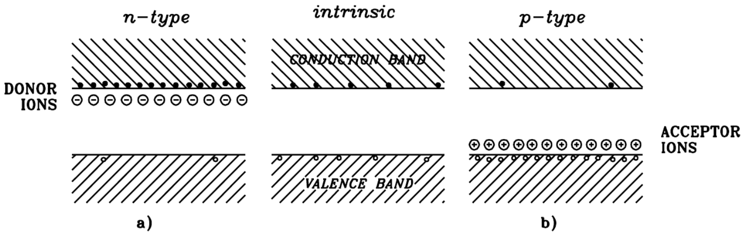


Figure 1.2: Energetic bands of intrinsic semiconductors together with example of the structure in the case of a n-type a p-type semiconductor [1].

the connection without free charge is called the depletion region. Due to an opposite charges of present ions, an electric field ε is crated in a direction acting against the previous free charge movement. This can be observed in the figure 1.3. The electric field can be characterized by electric potential difference between the two parts, voltage V_{bi}.

The usage of the p-n diode can be performed by an application of the external voltage to enlarge the depletion region. If a charge particle would traverse through this part of the diode, free charge is generated and collected by the external voltage on electrodes placed at the end of the diode.
The final signal consist of two parts: charge collection and current inductions (described by the Schockley-Ramo theorem [2]). The first of the mentioned phenomenon possesses two most important components: drift and diffusion. The drift is caused by the application of the external voltage described with a drift velocities $\boldsymbol{v_{n/p}}$:

$$\boldsymbol{v_n} = -\mu_n \boldsymbol{\varepsilon} \quad , \quad \boldsymbol{v_n} = -\mu_n \boldsymbol{\varepsilon} \tag{1.2}$$

where ε is the created electric field and μ is called mobility (its values can be found for example in [1]). The resulting velocities are linearly proportional to the applied field but for stronger fields saturate. The diffusion contribution is caused by a random thermal motion of free charge due to different spatial concentrations. This can be described with a carrier flux $\boldsymbol{F}$ and gradient of the carrier densities n and p (concentrations):

$$\boldsymbol{F_n} = -D_n \nabla n \quad , \quad \boldsymbol{F_p} = -D_p \nabla p \tag{1.3}$$

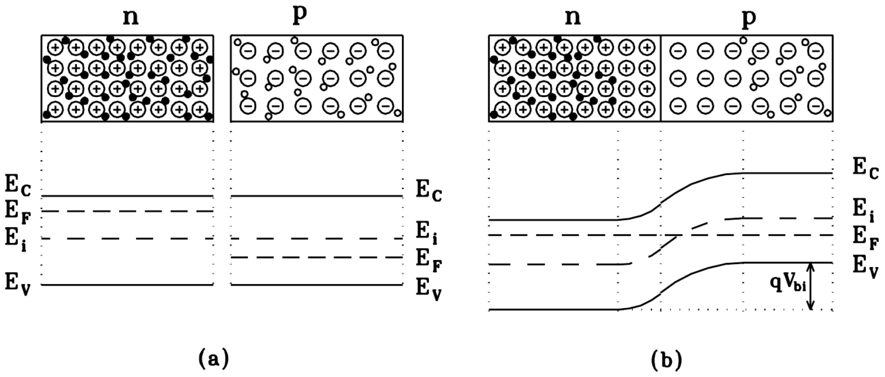


Figure 1.3: A p-n diode junction with shown energy band structure. The E_F is the Fermi energy and E_i is semiconductor intrinsic energy. If two differently doped semiconductors are connecte, they Fermi energies should coincide which is resulting in the occurrence of the electric field. Detailed description can be found in Lutz et al. [1].

where n, p are the hole and electron densities and D stands for diffusion parameters. It can be shown that the diffusion parameter and mobility are related by a so-called Einstein formula:

$$D = \frac{kT}{q}\mu. \tag{1.4}$$

A inseparable part of the problematic lies also in losing mechanisms, mainly recombinations. It can be shown that a full consideration of all contributors (generation, recombination, charge movement) can not be solved analytically without an additional conditions and solution has to be find via a numerical methods [3].

1.2 Charged Particles Interactions

The goal of this section is to briefly discuss the main interaction through which a majority of the charged particle energy dissipates in the semiconductors. A two main categories are included: heavy and light charged particles. A description of neutral particles interactions as photon and neutron can be found in in [4] and it will not be part of this text due to their low importance for the hadron therapy in this thesis. A partial continuation of this section can be found in the chapter 3 (the Multiple Coulomb scattering).

Heavy charged particles The heavy charged particles possess significantly larger mass than electrons (for example protons and heavy ions). The main cause of the energy dissipation lies in the interaction with the atomic electrons of the medium[2]. The mean value of the energy loss can be expressed with the well-known Bethe-Bloch formula [4]:

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \tag{1.5}$$

where $K = 4\pi N_A r_e^2 m_e c^2$ (N_A is Avogador's number and r_e is classical electron radius), z is the charge of the incident particle, Z, A is atomic number and atomic mass of the absorber, $W_{\max}$ is the maximal energy possible during the interaction, I represents a mean excitation energy and $\delta(\beta\gamma)$ is a density effect correction. The last mentioned correction originates from the case when heavy charged particles possesses very high velocities. The particle is polarising the medium and effecting the electric field with which it is interacting. An illustration of the total stopping power can be found in the figure 1.4

The heavy charged particles can also cause a creation of so-called delta ray, δ-ray. In this case, an electron acquires sufficiently high energy to be liberated from an atom and also to ionize the medium itself. This effect can be successfully utilized for a particle recognition as a characteristic feature for the heavy ions.
A very important phenomenon is realizing during a total absorption of a heavy charged particle. It was discovered that they tend to deposit large fraction of energy at the end their path, so-called Bragg peak. This is a leading principle in the hadron therapies when concentrated energy deposition is utilized for a precise irradiation of cancer tissue.

Light charged particles Electron and positron are usually considered as the light charged particles. A new possibility is opening to them via which

---

[2] A radiation losses are not considered due to the most probable momentum range with an upper limit of the order of 100 MeV/C.

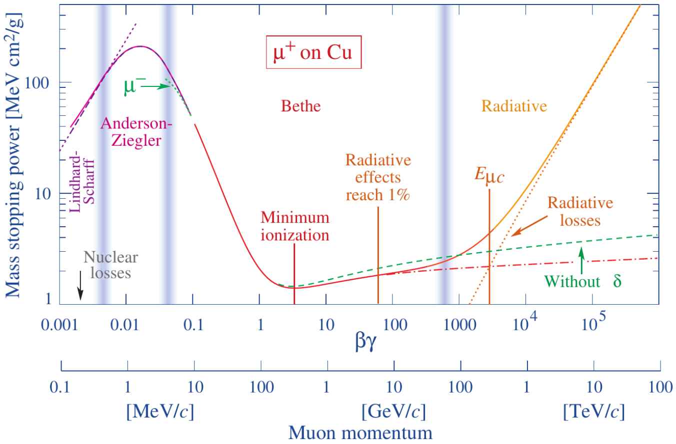


Figure 1.4: Stopping power for a positive muon in copper as a function of its energy or momentum [4].

they can deposit their energy, radiation. The raise of the importance is based on the significantly lower mass than in the case of the heavy charged particles. Several concepts have to be additionally considered in a comparison with the heavy charged particles to derive their stopping energy (spin, indistinguishability of incident and target particle etc.). The final formula for the stopping power of electrons follows equation of the Moller's [4]:

$$\left\langle -\frac{dE}{dx}\right\rangle = \frac{1}{2}K\frac{Z}{A}\frac{1}{\beta^2}[\ln\frac{m_ec^2\beta^2\gamma^2(m_ec^2(\gamma-1)/2)}{I^2} + (1-\beta^2) - \frac{2\gamma-1}{\gamma^2}\ln 2 + \frac{1}{8}\left(\frac{\gamma-1}{\gamma}\right)^2 - \delta], \tag{1.6}$$

where δ represents density effect correction. The final energy loss is a completed after adding also the contribution from radiation loss. This can be seen in he figure 1.5

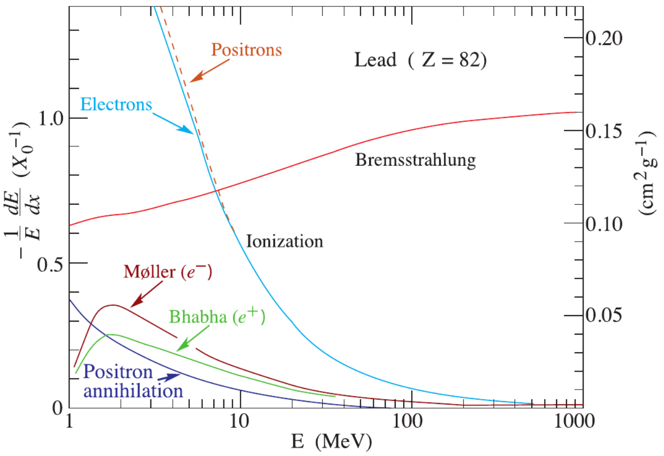


Figure 1.5: Fractional energy loss for electrons and positrons in lead dependent on the particle energy. [4].

1.3 Hybrid Active Pixel Detectors of Timepix Family

This sections includes description of the Timepix family detectors: the Timepix[5] and Timepix3[6]. Both of them are hybrid active pixel detectors and possess similar structure as it is illustrated in the figure **??**.

The first part of the detector is a sensitive medium composed of the doped semiconductors. According to the usage, several materials (Si, GaAs, CdTe etc.) can be exploited with different thickness (commonly from 100 to 2000 μm). The sensitive layer is surrounded by two types electrodes: common (back-side) and pixeled. The task of the common electrode is to mediate an application of the external voltage (also bias voltage). Each pixeled electrode or pixel is utilized for the charge collection after a particle crossing of the sensor. A matrix is created from 256 × 256 pixels with 55 μm pitch. These pixels are connected with the actual readout electronics (ASIC[3]) by the mean of so-called bump bonds. The collected and induced electric charge is processed by an individual CMOS readout for each pixel.

The Timepix chips are successors to the oldler Medipix2 model and both of them were developed by the Medipix collaboration at CERN. The main

---

[3] English abbreviation: Application Specific Integrated Circuit.

differences between the Timepix and Timepix3 lies in the individual readout electronics.
Both presented detectors can be labelled as miniaturized devices due to their small dimensions (together with the readout electronics in optimal set-up, the maximal dimension is approximately 20 cm).

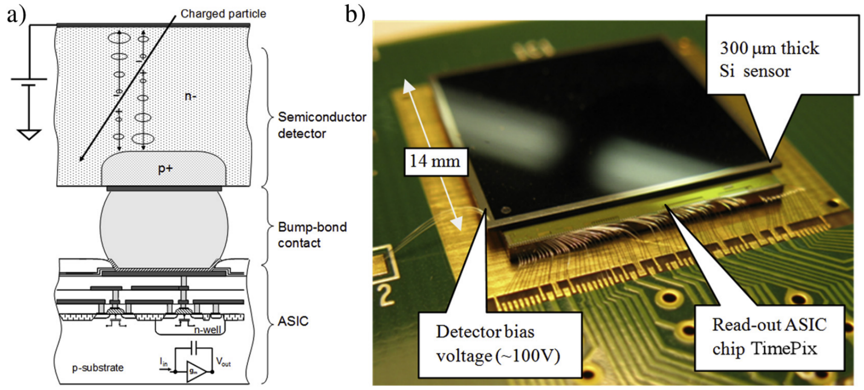


Figure 1.6: The right figure: a cross-sectional illustration of the hybrid pixel detector with a particle interaction within its sensor. The left figure: actual image of the Timepix detector and the peripheral electronics. Reprinted from [7].

1.3.1 Timepix

After the ASCI receives a collected and induced charge, a conversion into the to an analogue voltage pulse is performed which is subsequently shaped and amplified by the charge sensitive amplifier. Resulting signal is compared with an global threshold level and if it is exceeded, a 14-bit shift register is increased. The total counter depth reaches 11 810 counts. An external referential clock (maximal frequency is 80 MHz) is utilized generally for all pixels to increment the counter which depends on the selected operation mode:

- Event counting mode:
  The counter is increased only by one if the event is above the threshold.

- ToA mode (Time of Arrival):
  The counter is increased at every clock cycle after the first crossing of the threshold level to the end of the acquisition.
- ToT mode (Time over Threshold):
  The counter is increased at every clock cycle if the signal is over the threshold.

To switch between individual modes, each pixel possesses a 8-bit pixel configuration register which can also perform a masking of the pixel. All pixels are controlled by a shutter signal to operate in two states: in the first one the individual counters are incremented according to the given mode and in the second one the pixels shift their information to the readout logics. This is usually labelled as a frame mode architecture and an adjustable acquisition time is introduce to control the period after which the readout should be done.

1.3.2 Timepix3

The motivation for a development of the newer version to the Timepix can be found in its limitations:

- Frame mode: If for example in the ToT mode with open shutter a particle strikes an already hit pixel the signal is accumulated.
- Exclusive operations modes: The Timepix can not provide parallel measurement in the ToT and in the ToA mode.
- Readout insensitivity: During the readout, the matrix is insensitive to any other hit defining the dead time of the detector.

The Timepix3 meets most of the above criteria for the improvement. A new data driven architecture was implemented which can provide an immediate readout of a hit pixel. The operation ability of the detector were extended to allow a parallel measurement in the ToT an in the ToA modes. This improvement was paid by a lower capacity of the added register which is only 10-bit and significantly effects the measurement of the deposited energy. The approximate value of the detector dead time is 475 ns which is a time needed for the clearing the digital information from the pixel electronics. An enhancement of the time resolution was also achieved with an additional counter and a voltage control oscillator (with frequency 640 MHz) for the fine time stamp fToA. This defines the maximal time resolution to 1.526 nanoseconds. The conversion of the ToA to time is following: $t = 25 \times \mathrm{ToA} - 1.526 \times \mathrm{fToA}$ nanoseconds which reflects the mentioned frequencies. The final read out can be performed with the USB 3.0 with the total speed of 5 Gbps [8].

Both of the presented chip architectures found a large field for an application from scientific to commercial purposes. Only limited list of the their exploitation is presented and further information can be found in the article by Ballabriga et al. [9]:

- Space dosimetry and space weather
- Materials studies
- Spectroscopic X ray, gamma ray and particle imaging
- Electron microscopy
- Anti-matter search

More examples with detailed discussion can be seen in the next chapters.

Chapter 2

Detector Response and Data Processing

The acquired data from the Timepix detectors has to be processed with chain of several procedures to extract the information of interest. This chain can be observed in the schematic figure 2.1. The first two parts are dedicated to a detector preparation and the data measurement and they will be only partially discussed in this chapter. A preprocessing part continues with the main task to convert the stream of pixeled information to clusters via a clusterisation process and in the case of the telescope it should be able to match together only information obtained from one incident particle across the used detectors. Important segment is also energy calibration and description of the additional detector signal effects. The last part of the chain is an usage of the clustered information for a certain purposes and with respect to the hadron therapy, the section will be mainly focused on the particle identification and partially on its tracking. All presented algorithms were implemented with an usage of the C++ programming language with ROOT tool [10]. This option offers possibility for a fast data evaluation.

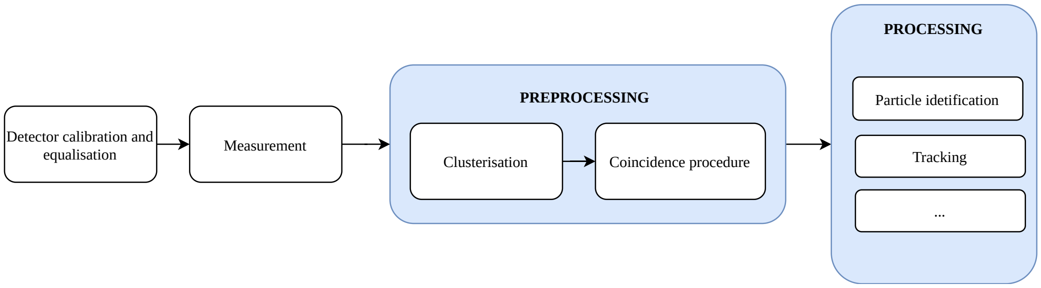


Figure 2.1: The scheme of the data processing chain. The coincidence process is optional and depends on the chosen detector set-up.

In the first part of this chapter the preprocessing steps are described: clusterisation and a calculation of the cluster variables and a coincidence process in the case of the stacked detectors. The preprocessing section is followed by the application in the particle identification task.

2.1 Preprocessing

The first step in the data evaluation chain is the preprocessing phase. Its form depends on the used detector set-up and can be distinguished into two basic classes:

- Single layer set-up
- Telescope: several detectors are stacked on each other and synchronized to provide several sensitive layers for a measurement

In both cases the first stage of the preprocessing has to be the **clusterization** or the cluster analysis. This process convert a stream of pixels to a clusters which should comprehend most of the information left by the traversing particle. An additional stage for the telescope architecture could to be ordered after the clusterisation, so called matching or **coincidence procedure**. The task of this process is to match those events from all telescope detectors which belongs to one particle.

The following section discuss both of these preprocessing steps together with the detector calibration and the additional detector signal effects. It is mainly focused on the Timepix3 detectors due to its usage in the current methods for the secondary particles tracking during the hadron therapy.

2.1.1 Per-Pixel Energy Calibration and Detector Signal Effects

After receiving the information from the detector, the energy and time for the Timepix3 detector is in corresponding ToT and ToA counts. These variables has to be calibrated and converted to to obtain results in energy and time. Time conversion is already described in the previous chapter 1 and in this subsection the ToT to energy calibration is discussed.
Same approach is used for the Timepix and Timepix3 ASICs. This work was published by Jakubek in [11] and it is based on the following calibration function:

$$\mathrm{ToT}(E) = aE + b - \frac{c}{E - t} \tag{2.1}$$

where a, b, c and t are parameters obtained from the calibration procedure. The calibration function can be observed in the figure 2.4 and its two main parts can be distinguished: non-linear which is close to the energy threshold and linear after approximately several tens of keV. During this process the sensor of a detector is irradiated with a fluorescence X-ray and gamma source. Each pixel of the sensor is aimed and a characteristic peaks can be observed in resulting ToT distributions corresponding to an expected and known energy peaks[1]. Subsequently, a fits are performed for each pixel to model an individual response according to the calibration function 2.1.

A detector response is accompanied by several signal distorting effects mainly caused by the electronics. The followings list includes the most important ones:

---

[1] The source is compound from a ^{241}Am which emits 59.5 keV gamma rays. They cross a iron and indium plate and generates additional fluorescence X-rays of 24.1 keV and 6.4 keV.

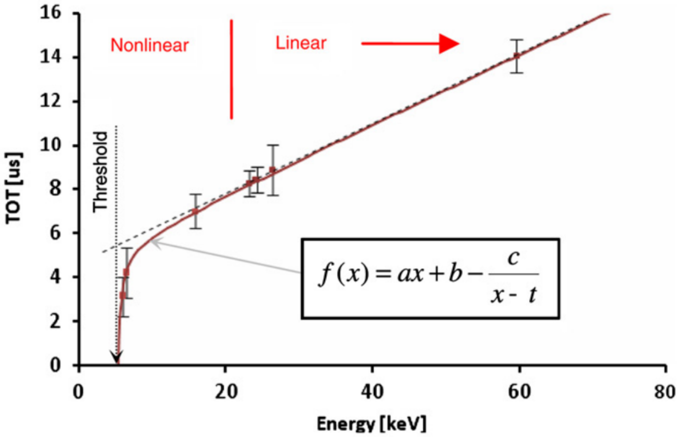


Figure 2.2: The calibration function shown with the highlighted linear and non-linear part [11].

- **Volcano effect**:
  Upon the crossing of heavy ionizing particle trough the detectors sensor, a large amount of energy (an order of MeV per pixel) can be deposited into several pixels. This is resulting in non linear response of the detector due to the electronic distortion [12].

  This distortion can be so dominant that it can result in a complete loss of signals for some pixels. The resulting pixel energy distributions for a sectional view can be seen in the figure 2.4. It was investigated that the amount of the volcano effect depends on two main parameters: applied bias via the charge sharing effect and features of the electronic pixel signal which can be tuned by the threshold and baseline value. The previous work was done for the Timepix detectors and in the case of the Timepix3 detectors, the volcano effects is much more profound, see the figure 2.3 and compare with the 2.4. One cause of the higher volcano effect is also caused by the lower capacity of ToT stamp register which has to be supplemented by higher values of I_{krum}[2] to achieve a reasonable energy resolution. The maximal detectable energy for the Timepix is approximately 2 MeV and in the case of the Timepix3 it is

---

[2] Parameter responsible for the signal discharging by raising its value.

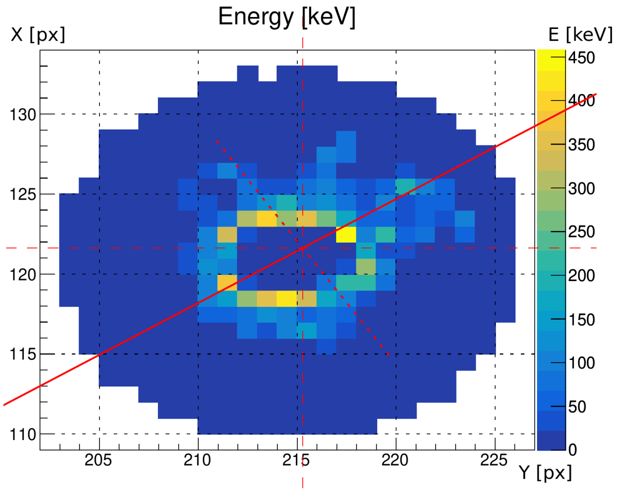


Figure 2.3: The figure of a heavy ion cluster in a energy plot with a direction almost perpendicular to the sensor plane. Two additional sensor effects can be observed: the volcano and halo effect.

maximally 1 MeV but always for the price of the energy resolution's distortion.
This effect has large consequences on the measurement of deposited energy and has to be included into consideration if the heavy ions are of interests. A possible solution could be proposed with sufficient description of the charge sharing effect under given conditions.

- **Time-walk effect**:
  This effect occurs during a time measurement and it was describe in the work of Turecek et al. [**?**, Turecek2016]nd Bergamn et al. [13]. The illustration can be seen in the figure 2.5. The issue occurs if the raising edge of the accumulated signal is insufficiently steep which is mainly valid for values less than 100 keV per pixel. It can be demonstrated on two pixel hits incoming in the same moment, see the read and grey lines in the figure 2.4. The red low signal originating from a deposited energy of 5 keV and the grey higher one comes from 30 keV. Therefore, the signal from higher energy is much steeper and it will cross the threshold level sooner than the lower signal despite the fact that they were detected at the same time. This behaviour is called a time-walk

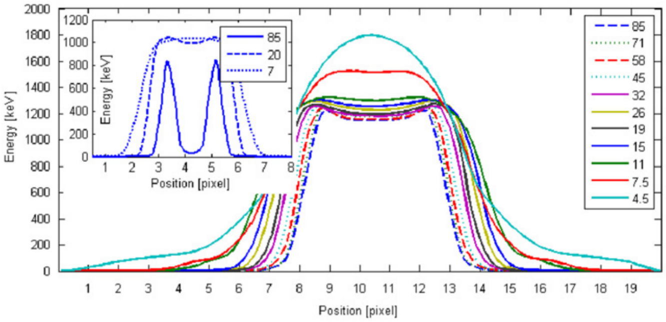


Figure 2.4: Sectional view profile of a cluster detected by Timepix. The main figure is with fine-tuned baseline and the inner one is for conventional baseline. Reprinted from [12].

effect and it can be concluded that its main dependence is on the deposited energy per pixel. An additional effect has also the threshold and its uncertainty. A correction can be made with a source of low energy events and utilizing the charge sharing effects which will spread originally one pixel event into several pixels with the same time. The difference in the time stamp of these pixels corresponds to a time-walk delay. The correction can be seen in the right figure 2.5 and it can be described with the function:

$$t_{\text{time-walk}}(E) = \frac{a}{(E-b)^c} + d \tag{2.2}$$

where a, b, c and d are free parameter fixed by the correction measurement. The final time with included tome-walk correction can be computed according following formula: $t_{\text{corr}} = t - t_{\text{time-walk}}(E)$[3]. For further work, the correction was not utilized because a precise measurement (less than 12.5 nanoseconds) of time is not of great importance for most of the cases.

- **Halo effect**:
  The figure 2.3 can be exploited as an illustration of the halo effect. The mostly low energetic part is labelled as a cluster halo (shown as

---

[3] If the correction is higher than the precision of the time measurement, 12.5 ns belonging to the coarse time then this value should be taken.

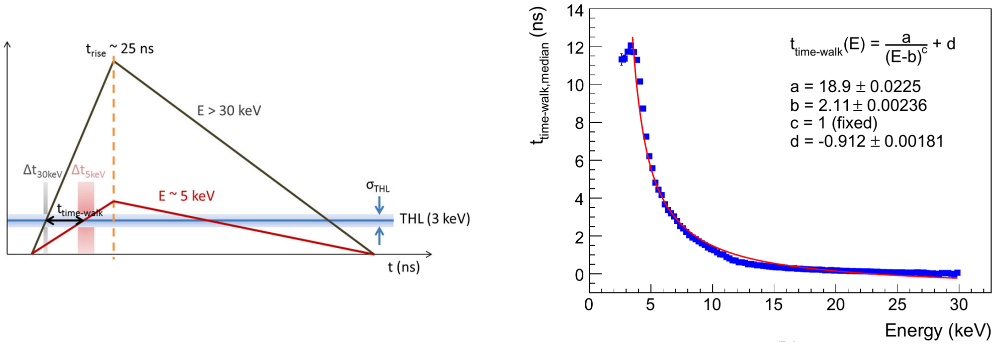


Figure 2.5: The left figure of the charge collection by the pixel electronics with highlighted effect of the time-walk. A lower signal (red) obtains a lower time stamp despite the fact that it occurred at the same moment as a higher signal(grey). The right figure shows a correction measurement and function. Reprinted form [13].

a blue part of the cluster in the figure). The origin of this effect is not so far fully described and it is still subject of research. Partial description can be offered by the means of the electric induction when high concentration of charge is moving towards the pixel electrodes causing a generation electric field. There is direct correlation between the amount of the deposited energy and size of the halo. It can be observed that rising the deposited energy is resulting in increasing of the number of the halo pixels and also its total energy. Disturbing the energy measurement, it can be utilized in the particle identification as a certain mark of high ionizing particles.

2.1.2 Clusterisation

The clusterisation converts a frame or a stream of received pixels from the Timepix and Timepix3 detectors to sets of pixels which belonging to individual particles. A distinction is needed between the two chip architectures. In the case of Timepix detector, the obtained data format is in a form of frames. The information is encoded into a sensor matrix of only hit pixels where each of them contain energy deposited during the acquisition time (if a measurement in ToT is performed). The same result can be also achieved with the Timepix3 detector in the same frame mode. If the data-driven mode is required, the final output data resembles a stream of pixels where time is stored together with energy. This significantly improves the ability for the restoration of the individual particle information.

The algorithm used in the case of the frame mode rely only on the coordinate information. The following process is used for the clusterisation of the frame when gradually each pixels is examined:

1. An iteration over whole list of created clusters is done (this set is empty for the first pixel). Each cluster is probed whether the current pixels is its neighbour expressed with following condition:

$$|x_i - x_{\text{curr-pix}}| \leq 1 \text{ and } |y_i - y_{\text{curr-pix}}| \leq 1 \quad ; \quad i \in \{\text{pixels of cluster}\} \tag{2.3}$$

   I [Yes] The cluster is added to a list of neighbour clusters.
   II [No] Next cluster is investigated.

2. An iteration over the list of neighbour clusters is performed and its size is conditioned:

   I [Only 1 cluster] Pixel is added to the only cluster
   II [2 and more clusters] All clusters are merged together also with the current pixel and the result is moved to the list of the created cluster. Consequently, individual merged clusters are deleted.
   III [No cluster or empty list of created clusters] Pixel is added to the list of the created clusters as a new cluster

When the list of pixels is exhausted then the clusterisation procedure is ended with the list of created clusters as its result.

In the case of Timepix 3 detector, an additional time information is exploited and the neighbour pixels has to also satisfy additional time-related condition. This constrain is expressed by free parameter called a time window t_{win} and it can be reformulated to the form:

$$|t_i - t_{\text{curr-pix}}| \leq t_{\text{win}} \quad ; \quad i \in \{\text{pixels of cluster}\} \tag{2.4}$$

The values of the free parameter t_{win} varies according to usage. The main effecting factor is an overall flux of particles but a generally its value is of the order of 100 nanoseconds. This choice also reflects the collection time of the created charge in the sensor by a one particle as a minimal value. The time for the collection is approximately of the order of 10 nanoseconds but it depends on the applied bias and subsequently on the sensor thickness.

An inseparable part of the clusterisation process is calculation of the cluster variables applied on the created and finished clusters. These features are selected according to the final purpose of the analysis. In this case the main inspiration is the particle identification and the hadron therapy.
Before the introduction of the variables definitions, a basic nomenclature has to be stated. To each pixel of a cluster is assigned a vector of coordinates: $\mathbf{r_i} = (x_i, y_i, z_i)$ and deposited energy E_i. The total number of pixels in the cluster is labelled as N_{pix}.

Energy sum This variables comprehends the total deposited energy (detected):

$$E_{\text{sum}} = \sum_{i=1}^{N_{\text{pix}}} E_i \tag{2.5}$$

Maximum energy/height Energy of the pixel with maximum deposited energy:

$$E_{\text{max}} = \max\{E_i | i \in N_{\text{pix}}\} \tag{2.6}$$

Size Variable comprehending the number of pixels in the cluster:

$$S = N_{\text{pix}} \tag{2.7}$$

Weighted center The first variable which combines the energy and coordinate information. The weighted center $\mathbf{R}_{\text{center}}$ of a cluster is computed according a idea of the center of mass of certain object where pixel deposited energy is used instead of mass:

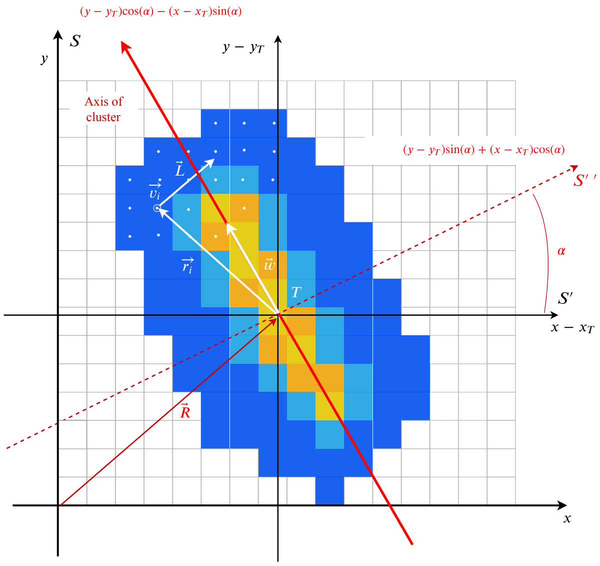


Figure 2.6: Depiction of a cluster in a sensor plane. The main variables used for the cluster axis are shown.

$$
\begin{aligned}
\mathbf{R}_{\text{center}} &= \frac{\sum_{i=1}^{N_{\text{pix}}} E_i \mathbf{r_i}}{E_{\text{sum}}} \\
R_{\text{center},j} &= \frac{\sum_{i=1}^{N_{\text{pix}}} E_i r_{i,j}}{E_{\text{sum}}}
\end{aligned}
\tag{2.8}
$$

Axis and polar angle The cluster axis plays a crucial roles for definition of further variables. There are several possibilities how to view on a cluster and one of them is the rigid body approximation. Each center of a pixel is taken as one point on a flat rigid body with mass/weight equal to energy or its power: $w_i = E_i^n$ (it can be exploited that the optimal value is around n = 1/2). The coordinate system with an origin in the weighted center rotated to the axis of a cluster is chosen for this case. The weighted center has coordinates $\mathbf{R}_{\text{center}} = (x_C, y_C, z_C)$. The z-coordinate is not considered for further calculations and is always neglected. Most of the steps will be skipped and only definitions, important relations and final formulas are introduced.

Three weight-coordinate sums has to be defined:

$$
\begin{aligned}
E_{x^2} &= \sum_{i=1}^{N_{\mathrm{pix}}} w_i (x_i - x_C)^2 \\
E_{y^2} &= \sum_{i=1}^{N_{\mathrm{pix}}} w_i (y_i - y_C)^2 \\
E_{xy} &= \sum_{i=1}^{N_{\mathrm{pix}}} w_i (x_i - x_C)(y_i - y_C)
\end{aligned}
\tag{2.9}
$$

It can be derived that the final variable/function has following prescription (it represents the momentum of inertia in the x-axis):

$$
f(\alpha) = L_(x) = E_{x^2} \sin^2(\alpha) E_{y^2} \cos^2(\alpha) - 2E_{xy} \sin(\alpha)\cos(\alpha) \tag{2.10}
$$

and this function has to be minimised to find the optimal angle. Final formula for the angle α of a cluster as a stationary point:

$$
\begin{aligned}
&\text{a) if } E_{x^2} \neq E_{y^2} \rightarrow \alpha = \arctan\left(\frac{2E_{xy}}{E_{x^2} - E_{y^2}}\right) + k\frac{\pi}{2}, \quad k \in \{\pm 1\} \\
&\text{b) if } E_{x^2} = E_{y^2} \rightarrow \alpha = \frac{\pi}{4} + k\frac{\pi}{2}, \quad k \in \{0, -1\}
\end{aligned}
\tag{2.11}
$$

It can be derived that the angle lies within this interval: $[-\frac{\pi}{2}, \frac{\pi}{2})$ (it is π-periodic function and therefore $\frac{\pi}{2}$ is neglected). To complete the calculation, the stationary point has to be investigated with second derivation of $f(\alpha)$ in the this point. The type of a local extreme (maximum or minimum) can be stated based on the following equation:

$$
\frac{\partial^2 f(\alpha)}{\partial \alpha^2} = (E_{x^2} - E_{y^2})\cos(2\alpha) + 2E_{xy}\sin(2\alpha) \lessgtr 0 \tag{2.12}
$$

and according to its value the angle in the case of a maximum is changed with a multiple of $\frac{\pi}{2}$ (plus or minus with respect to the original tested value to remain in the defined interval $[-\frac{\pi}{2}, \frac{\pi}{2})$). As an example of a successful assignment of the cluster axis, see the figure **??**.

Weighted standard deviations in x and y coordinate Both standard deviations are computed with respect to the coordinate system with the origin in the weighted center rotated to the cluster axis with following rotation:

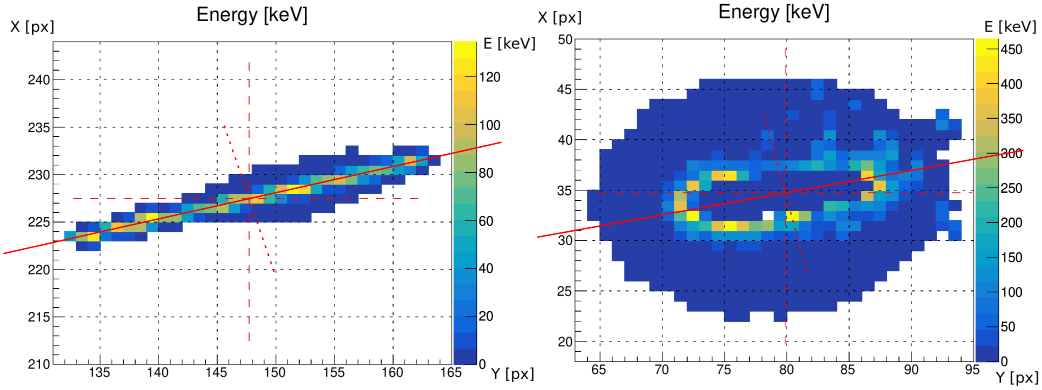


Figure 2.7: The calculation of the axis for a high energy proton on the right and heavy ion on the left. Rhe axis is shown with red full line. The dashed lines serves as the axis of the coordinate system with the origin in the cluster weighted center.

$$\begin{aligned} x_i'' &= (x_i - x_T)\cos(\alpha) + (y_i - y_T)\sin(\alpha) \\ y_i'' &= -(x_i - x_T)\sin(\alpha) + (y_i - y_T)\cos(\alpha) \end{aligned} \tag{2.13}$$

The weighted standard deviations is computed according equations:

$$\sigma_{\text{perp}} = \sqrt{\frac{\sum_{i=1}^{N_{\text{pix}}} E_i x''^2}{V_1 - \frac{V_2}{V_1}}} \tag{2.14}$$

$$\sigma_{\text{along}} = \sqrt{\frac{\sum_{i=1}^{N_{\text{pix}}} E_i y''^2}{V_1 - \frac{V_2}{V_1}}} \tag{2.15}$$

$$V_1 = E_{\text{sum}}\ , \quad V_2 = \sum_{i=1}^{N_{\text{pix}}} E_i^2 \tag{2.16}$$

This result is naturally biased but it should diminished for large N_{pix}. Main benefit arises form the information about the energy layout around the cluster axis. This is used for definitions of other variables which comprehends morphological features of the cluster.

Linearity The linearity should express the cluster concentration along the cluster axis. For this purposes, the standard deviations in the perpendicular σ_{perp} and along σ_{along} directions are used:

$$\text{linearity} = 1 - \frac{\sigma_{\text{along}}}{\sigma_{\text{perp}}} \tag{2.17}$$

An interval of possible values is: $(0, 1)$ with limit cases 0 when a cluster is most likely round and 1 when a cluster has a linear shape.

Border and inner pixels The count of border pixels N_{border} and of inner pixels N_{inner} can be also valuable information. Their ratio $\frac{N_{\text{inner}}}{N_{\text{border}}}$ is exploited for definition of other variables connected with the morphological cluster features.

Another usage lies in an identification of so-called cluster completeness. If the Timepix3 detector is overwhelmed with a high flux of particles it might happen that only certain parts of the clusters has similar time stamps and other are rapidly time shifted. If the cluster is large but still the ratio is very small $\frac{N_{\text{inner}}}{N_{\text{border}}}$ or it could indicate that the detector overwhelming is present.

Thickness and thinness These opposite variables should express the cluster shape in the sense of its width. They are defined with the cluster size S and the number of border N_{border} or inner pixels N_{inner}:

$$\text{thinness} = \frac{N_{\text{border}}}{S} \tag{2.18}$$

$$\text{thickness} = 1 - \text{thinness} = 1 - \frac{N_{\text{border}}}{S} = \frac{N_{\text{inner}}}{S} \tag{2.19}$$

The range of the values is from 0 to 1 where the variable equal to 1 reflects the highest quality in the feature.

Curliness This feature should reveal the curly nature of a particle track. An assumption is made for its definition: it is unlikely that high ionizing particles would undergo a high deflection in a detector sensor. Therefore, majority o fthe tracks should be also thin. The curliness can be defined as an product of the thinness and an opposition of the linearity:

$$\text{curliness} = (1 - \text{linearity}) \times \text{thinness} = \frac{\sigma_{\text{along}}}{\sigma_{\text{perp}}} \frac{N_{\text{border}}}{S} \tag{2.20}$$

Roundness The roundness should comprehend the cluster round shape likeness, see the figure **??**. The leading idea is a comparison of a circle content with the cluster size resulting in an ideal radius R_{circle} of the cluster if it would be a perfect circle with a center in the cluster weighted center:

$$R_{\text{circle}} = \sqrt{\frac{S}{\pi}} \tag{2.21}$$

A deviation from this shape can be expressed with the maximal distance of a periphery pixel R_{max} and its ratio with the above ideal radius:

$$\text{Roundness} = \frac{R_{\text{circle}}}{R_{max}} \tag{2.22}$$

This definition ensures that the range of the roundness do not exceed the 1 and is higher than 0.

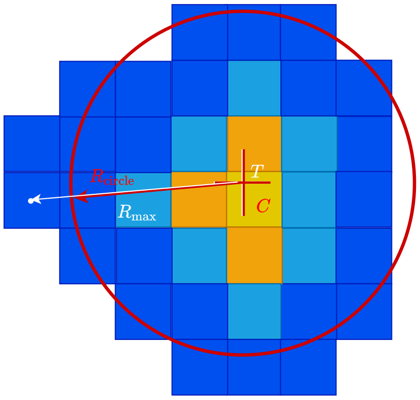


Figure 2.8: The concept of the cluster variable roundness. The point T is the weighted center of the cluster and C is a position of the ideal cluster circle with a borders illustrated with the red line.

Length The length is defined in the sensor plane (alias length 2D or L_{2D}). To determine its value, it is required to estimate the most probable trajectory of the particle which is a complex task. As an introducing example, a two tracks of fully penetrating particles are shown in the figure 2.9. A heavy ion is present in the first case and it is known that this particle crossed full width of the sensor with a minimum deflections but due to the large deposited energy and charge effects the final true trajectory is masked. In the second case of electrons with energy of several MeV, the final track possesses a curly shape in the sensors plane but also from the z coordinate (depth) point of view. Therefore, a need for a certain particle pre-classification is present to state when the length determination is feasible with sufficient precision.

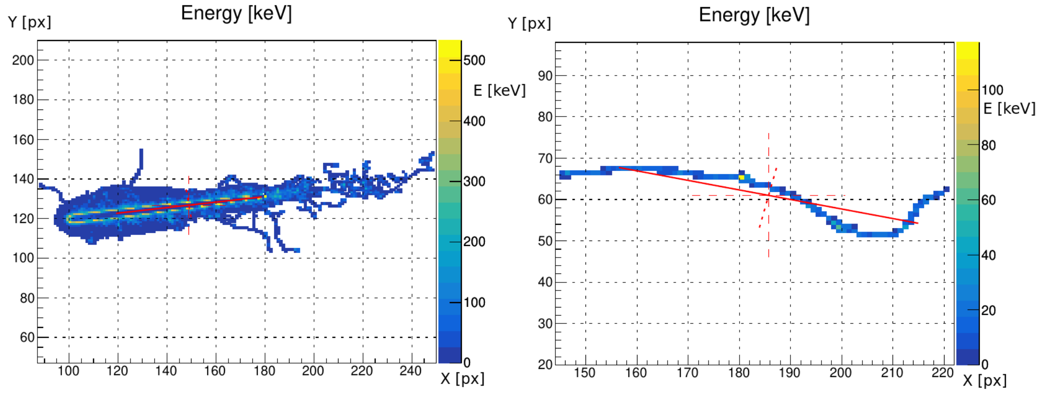


Figure 2.9: The right figure contains the track (deposited energy) of a heavy ion with approximately 70 degrees elevation angle. In the left figure is the track of a high energy electron also with high elevation angle.

The two main classification can be made on the basis of the particle capability to **cross** the sensor and also on its **morphology**. The sensor crossing capability:

- **Full-crossing:**
  A particle which is capable to fully traverse the sensor thickness is the most optimal case. A special case when particle ejects in site of the sensor has to be discussed but from the geometrical point of view, this trajectory is not so probable.

- **Partial-crossing, absorption:**
  If a particle is stopped in the sensor than the tracking possibilities are very limited and in some cases the track can not be restored from the detector information.

The morphological point of view was introduce in the article of Holy et al. [14] and this approach with several merged classes was assumed for further work:

- **Small blobs and dots:** photons and low energy electrons
- **Straight thin tracks:** minimum ionizing particles
- **Curly track:** high energy electrons
- **Heavy track and heavy blobs:** heavy ionizing particle

Tracks originating from the high energy electrons are also not describe within this work which is mainly caused by there high probability for large deflection within the sensor. This can be seen as a possibility to go "up and down" within the sensor depth which can not be sufficiently described even if the time stamp is considered. A testing of this type of length determination is also obstacle because the time concept is hardly implemented into simulations.

From the above mentioned morphological features and possibility to cross the sensor, only the particles capable to fully cross detector without significant portion of deflections can be with sufficient precision reconstructed by the proposed method. This should mainly corresponds to the heavy tracks, straight thin tracks and partially to heavy blobs.

- **Weighted standard deviations:**
  The cluster axis will define the most border opposite pixels with coordinates $\vec{r_1}$, $\vec{r_2}$ and the length L_{2D} can be derived as:

  $$L_{2D} = \sqrt{\sum_{i=1}^{3}(r_{1,i} - r_{2,i})^2} \tag{2.23}$$

  It is also possible to get more precise estimation of true particle track length $L_{2D,\text{prec}}$ with a subtraction of n the along standard deviations:

  $$L_{2D,\text{prec}} = L_{2D} - n \times \sigma_{\text{along}} \tag{2.24}$$

  where n can be from 1 to 3.

- **Membership function:**
  This calculation of length utilizes an approach presented in the article by Hoang et al [15] with certain modifications. This method exploits a projection (accumulation) of the pixel energy on two pixel sized bin axis: the cluster axis and perpendicular axis to the cluster one. These projections are tested by a member ship function which gives each bin a value from 0 to 1 according to its value of accumulated energy E_i. If the bin has value higher that the average of the projection axis E_{av} then it is assigned with 1 otherwise it will receive: $\frac{E_i}{E_{\text{av}}}$. The sum of these member ship values defines two lengths: major L_{major} in the case of the cluster axis and minor L_{minor} for the other one. The total length is extracted according to the final formula:

  $$L_{2D,\text{mem}} = L_{\text{major}} - A(L_{\text{major}}, L_{\text{minor}})L_{\text{minor}} \tag{2.25}$$

  where the constant A was empirically defined as:

  $$A(L_{\text{major}}, L_{\text{minor}}) = \frac{1}{L_{\text{major}} - L_{\text{minor}}} + \frac{C}{L_{\text{major}}}, \tag{2.26}$$

  where C is constant set to 5 and with an additional condition that the $L_{2D,mem}$ is always higher or equal to 0 [4].

The final method for the length computation is a combination of both approaches. It was investigated that the first simplified calculation when only L_{2D} is derived can be very accurate for the minimum ionizing particles. On the other hand, a large error emerges for the tracks of heavy ions (large contribution from charge sharing effect). The main distinction between those types of tracks can be found in their value of the thickness. Therefore an empirical condition on the thickness is implemented to the calculation: if thickness is less than 0.6 the L_{2D} is used otherwise the length $L_{2D,\text{mem}}$ originating from the member ship function is exploited.

The testing of the length calculation was performed with measured data of heavy ions (iron) and medium/high energy protons and it can be observed in the figure 2.10. A comparison between the true length computed from the known elevation angle and calculated one is shown as a function of the linearity and deposited energy. It can be seen that the weighted standard deviation approach with fixed constant to 2 is inaccurate for most of cases

---

[4] The $L_{2D,\text{mem}}$ can be negative which is replaced by 0.

but a possibility for a correction arises from the dependence on the linearity. If a correction factor dependent on the linearity would be included than a better result should be achieved. The results for the combine calculation of the length show a large bias in some cases, especial for the medium energy protons. This is mainly ascribed to an inaccuracy during this certain measurement. The large differences for small linearities which mostly reflects a round clusters are tolerable for the next calculation of the length in the volume of the sensor.

Elevation angle and total length The cluster elevation angle β is between a vector perpendicular to the sensor plane and a cluster axis:

$$\beta = \arctan\left(\frac{L_{2D}}{d}\right), \tag{2.27}$$

where d is the thickness of the detector and L_{2D} is converted to micrometers. The total length in the sensor volume follows the Pythagorean theorem under consideration that the given particle fully crossed the sensor:

$$L_{3D} = \sqrt{L_{2D}^2 + d^2} \tag{2.28}$$

Linear energy transfer - LET The linear energy transfer serves for an indication of the particle ionization within a sensor of a detector:

$$\mathrm{LET} = \frac{E_{\mathrm{sum}}}{L_{3D}}; \quad \left[\frac{\mathrm{keV}}{\mu\mathrm{m}}\right] \tag{2.29}$$

The variable has broad filed of usage, e.g. during a precise dose calculation or for the particle identification [15].

Energy pixel variables The following variables are connected with individual energy values of the cluster pixels:

- **Average energy**:

$$E_{\mathrm{pix,aver}} = \frac{E_{\mathrm{sum}}}{N_{\mathrm{pix}}} \tag{2.30}$$

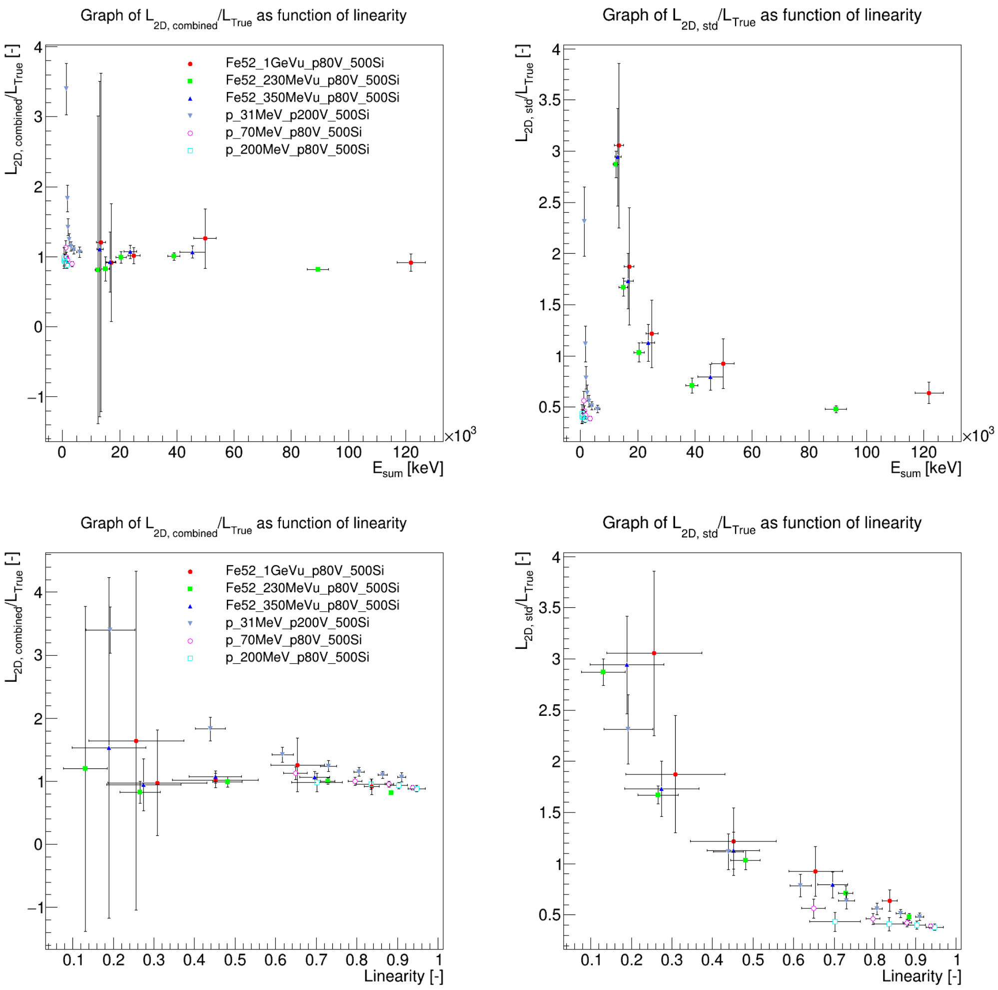


Figure 2.10: The comparison between true projected length L_{true} and calculated one as a function of linearity and deposited energy E_{sum}. The computation was performed with two approaches: the right plot is based on the combine way and the left plot is for weighted standard deviations lengths with fixed parameter n to 2. Several angles for each measurement in the legend are shown.

- **Variables within energy bins**:
  Very valuable information is encoded in the energy pixel distributions $f(E_{\text{pix}})$, see the figure 2.11. Heavy ions will serve as an example. It is known that upon their crossing of the sensor an occurrence of delta rays is possible together with the halo effect. A parallel can be made between the ion charge and the count of the delta rays and the size of the halo effect. This can be comprehended by suitably defined energy pixels variables. A similar approach can be fount in the article of Vilalta et al. [16].
  An simplification is exploited to make efficient extraction with respect to the calculation time. The energy pixel distribution is separated into n energy bins according to $n+1$ energy boundaries $\epsilon_1, ..., \epsilon_{n+1}$. This was used to define several so-called energy pixel variables:

  - Sums of energy - $\{E_j\}$:

  $$E_j = \int_{\epsilon_j}^{\epsilon_{j+1}} E_{\text{pix}} f(E_{\text{pix}}) dE_{\text{pix}} \tag{2.31}$$

  - Counts of pixels - $\{N_j\}$:

  $$N_j = \int_{\epsilon_j}^{\epsilon_{j+1}} f(E_{\text{pix}}) dE_{\text{pix}} \tag{2.32}$$

  A several possibilities arising also from a mutual combination of above variables or their ratios with the total sums (E_{sum}, N_{pix}). The final count of the energy bins was set to 3 with following energy boundaries:E_{thr}, 30 keV, 330 keV, E_{max}, where E_{thr} is the threshold energy (all for the Timepix3 usage). This options originates form the particle identification in the sections 2.2 and resolution power to distinguish between particle types.

2.1.3 Coincidence Process

The last part of the preprocessing discuss the telescope architecture. Several detectors are place in a stack geometry where all detector sensors are overlapping each other. This can be beneficial for the particle tracking due to the additional collected information. For further discussion mainly the Timepix3

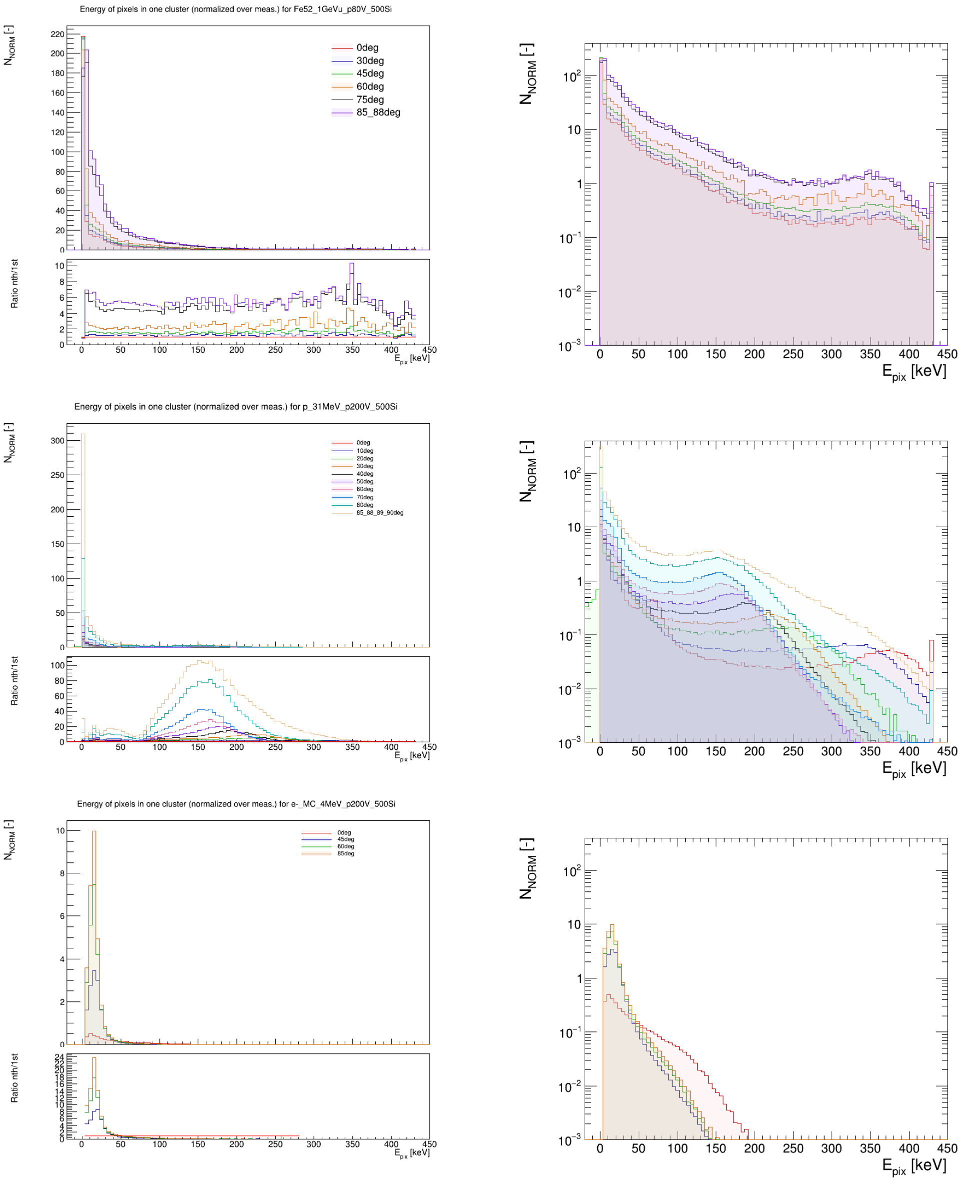


Figure 2.11: The energy pixel distribution for different particles and angles (averaged over measurement). The right part of the figures is in the linear scale with a ration plot at bottom (ration between the first and other distributions) and left part is in the log scale (same range for all three examples). The top figure includes the distribution of 1GeV/u ^{52}Fe ion with the elevation angles: 0, 30, 45, 60, 75, 85-88 degrees. The middle figure comprehends the distribution of 30 MeV protons with the elevation angles: 0-80 with 10 degrees step and 85-90 degrees. The bottom figure includes data from 4 MeV electrons with the elevation angle: 0, 45, 60, 85 degrees.

detectors are considered and possible generalisation for the Timepix detector is mentioned.
The key component of the telescope architecture for the Timepix3 detectors is their synchronization to obtain almost the same time stamps for particles interacting in both detectors. Exploiting this set-up plays crucial role for the beam monitoring during the hadron therapy.

The goal of this section is to propose methods for events matching across the telescope detectors. Two algorithms were developed:

- **Time based**
- **Unsimilarity based**

Two concepts have to be introduced for the problematic description. The telescope consists of n detectors. For the purposes of the coincidence/matching procedure, the first detector is labelled as the main one and the others as subordinate. This assigning of roles is important for the work with $n+m$ events which are closely time related and a need for their matching is present[5]. These events can not be distinguish with the basic time conditions due to the synchronization uncertainty propagated to the time stamps (approximately of the order of 10 nanoseconds). They are called multi-coincidence events and an additional sorting algorithm is required. The opposite cases when n events can be matched by the time algorithm are labelled as single-coincidence events.

Time based algorithm This approach exploits only the time stamps of the detected particles. Shall be assumed that the set-up is composed of only two detectors 1 and 2. An event from the detector 1 with time t_1 is matched together with an event originating from the detector 2 with time t_2 if following condition is satisfied:

$$|t_1 - t_2| \leq t_{\text{coinc-win}} \tag{2.33}$$

where $t_{\text{coinc-win}}$ is a free parameter which comprehends a tolerable time difference between two coincidence events. Its value has to be tuned according to specific case of usage but an average value is of the order of 100 nanoseconds.

---

[5] The number $m > 0$ therefore there are more than one vent coming from one detector.

One of the detectors is always labelled as the main one and serves as a referential time for others. Therefore in the case of more than 2 detectors in the telescope, each subordinate detector is tested with the main one via the above condition. If the time difference is higher than $t_{\text{coinc-win}}$ then this combination is prohibited and another event from the current subordinate detector is tested. This can result also into no found event at all and the current main event is skipped. Therefore only a main event which can be matched with all others subordinate detector events is considered as matched one, the output of the coincidence process.
As was already mentioned, if for example this condition would be fulfilled for two or more events from the detector 2 than a sorting algorithm is needed. A modification can be proposed with an usage of certain cluster variables which solve a problem with large portion of multi-coincidences with sufficient precision.

Algorithm based on unsimilarity This algorithm is based on principal that particles which should be matched together are the most similar ones. Nevertheless, the similarity is not a directly investigated feature of two events but instead an opposite one, unsimilarity, is computed and compared. The algorithm input is sets of cluster variables provided by the clusterisation. In the case of comparison of two clusters, each of them is characterized by this set $\mathbf{X}$, $\mathbf{Y}$ and following equation is used for calculation of their unsimilarity:

$$\text{UNS}(\mathbf{X},\mathbf{Y}) = \sqrt{\sum_{i=1}^{n} w_i \frac{[(X_i - Y_i) - \mu_i]^2}{\sigma_i^2}}, \tag{2.34}$$

w_i serves as weights (commented later in the text) and σ_i^2 and μ_i is estimated variation and estimated mean value for each utilized variable provided from single-coincidence events, computed according these equations:

$$\begin{aligned} \mu_i &= \frac{\sum_{i=1}^{N_{\text{single}}} x_i}{N_{\text{single}}} \\ \sigma_i^2 &= \frac{1}{N_{\text{single}}-1} \sum_{i=0}^{N_{\text{single}}} (x_i - \mu_i)^2 = \\ &\qquad \frac{1}{N_{\text{single}}-1} \left(\sum_{i=0}^{N_{\text{single}}} x_i^2 - N_{\text{single}} \mu_i^2 \right), \end{aligned} \tag{2.35}$$

where x_i is given variable. As was mentioned before, it is the single-coincidence events which mainly determine what should be the optimal difference of two possibly matched clusters. This can be marked as a huge benefit that the values according to which the unsimilarity should be decided are not input ones but derived within the algorithm itself.

This can be also a disadvantage of the method. Shall be assumed a measurement of a mixed radiation field with an equal composition of different types of particles with various energies. It is highly probable that for many cluster variables: energy, size etc. their variances will possess large values (these particles will deposit different amount of energy).It can be seen from the equation 2.34 that contribution from these variables will be suppressed by their large variances without any valuable contribution to the unsimilarity determination.

A partial solution can be found in variables which are independent of a particle features and they can be highlighted by the weighting system. The most profound example is time connected with a particle arrival to a detector. It can be observed in the figure 2.12 final distributions of single-coincidence events (it is standardize difference) for a measurement of phantom radiated by carbon beam. In this case the time difference distribution would have same effect as energy or size but instead in the process its value of weight was raised to higher value than those of energy and size. It must be stressed that the weights for all variables should be always decided according those single-coincidence distributions for each separate usage and only some hints how to do it can be provided (time is universal etc.). A possible generalisation can be performed to apply this algorithm for the Timepix telescope with omitting the time of cluster detection.

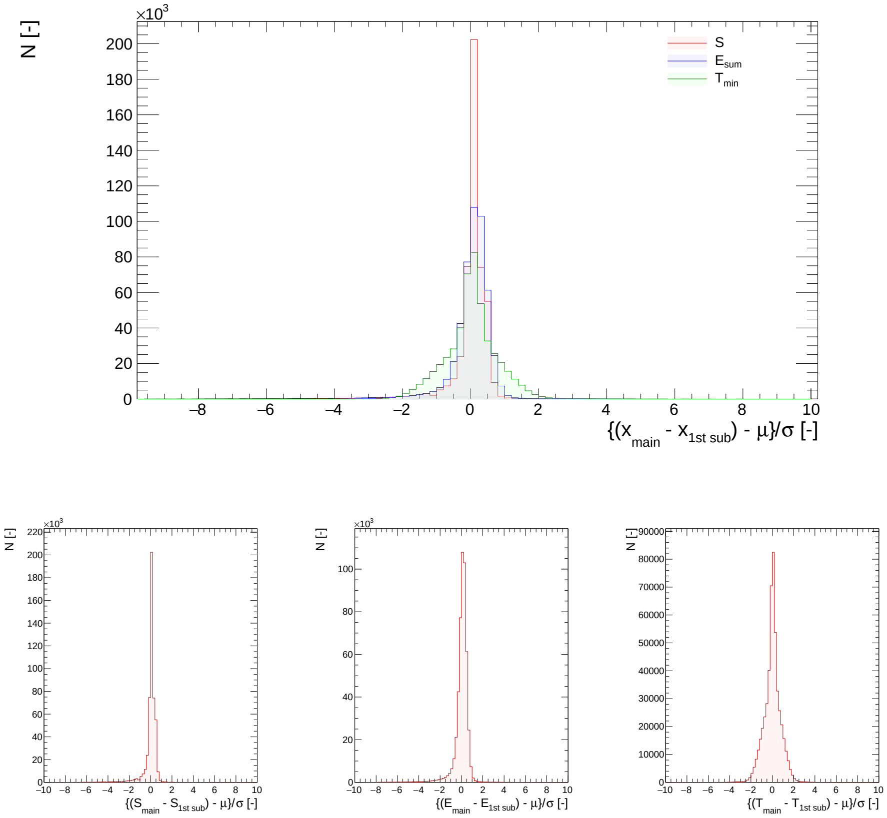


Figure 2.12: Standardise distributions of energy, size and time differences between main detector event and subordinate one in the case of single-coincidence events. In the bottom figure, distributions are in separated canvases.

2.2 Particle-type Recognition

The spectrometric and trajectory response of the detector can be utilized for the task of the particle recognition. This could be beneficial for overall usage and also for the hadron therapy especially in the case of the lateral beam monitoring.

In previous works with Timepix detectors were demonstrated capability of the particle recognition with a single layer set-up. One of the first efforts in this field was done by Holy et al. [14] with Medipix2 detector. The proposed classification on the morphological features mentioned in the previous sections. The resulting capability of the recognition method was demonstrated with success on the decreasing flux of particles with larger distance from the radiation source (namely the α, β and γ sources). A quantitative comparison with an actual source activities was performed in the of Bouchami et al. [17] with a final uncertainty of 8 percent (comparison of a measured and calculated activities). The disadvantage of the proposed methodology was mainly in the fixed values of used conditions which was very effected by the settings of the detector, Therefore change of for example detector threshold could significantly change the method. The first use of the Timepix detectors fro the particle recognition can be found in the work of Vilalta et al. [16] which was motivated by the previous works. In the presented methodology, a machine learning algorithms were utilized with a set of morphological and spectrometric features (including the certain type of energy pixel variables). The work is concentrated on the recognition of the heavy ions charge which is achieved with sufficient efficiency approximately 80 percent.

An important works was presented by Hoang et al. [18, 15]. The main contributions were separation of the heavy ions cluster into three parts: the central energetic region, the δ-rays and the skirt which are described by new cluster features (count of δ-rays, their energy etc.) together with the LET and other more commonly used cluster variables. The machine learning methods were included into the analysis with sufficient efficiency.

A possibly more general classification were offered by Granja et al. [19] who separates the wide field of particle into 8 classes based on particle specie, direction and energy. A morphological and spectrometrical cluster features are used for the particle recognition with fixed conditions which therefore possess the same non-invariance as the work of Holy et al in the case of different settings of detector. The main benefit of the work is more detailed view on the particle recognition in mixed fields.

One the most recent works was created by Mánek et al [20]. The conventional approach via a cluster features is replaced by more comprehensive analysis of clusters with randomizes computer visions. The work is concentrating on the recognition of heavy ions charge and protons. The authors propose to segment each cluster into several parts with energy accumulation of included pixels and train a machine learning algorithms.

The work is more concentrated on the utilizing of the Timepxi3 detectors. This decision was made based on the experience and current usage of the Heidelberg group. It will be also shown that the presented methodology is universal in the ASIC type with several adjustments of the algorithms.
The particle identification was approached by means of so called similarity algorithm combined with a decision tree. A large parallel can be found between the machine learning kNN algorithm[21] and similarity one. The main function of the decision tree is to provide a certain pre-classification of the detected events with fixed conditions. This approach arises from the fact that for some given particle specie the probability to achieve certain range of variable values is almost non-existing therefore a fixed conditions can be implemented. They can automatically omit this class within the classification process[6].
To describe the proposed methodology it is needed to firstly introduce some elemental terms used within this problematic inspired by the machine learning [22]. The overall task of the particle identification can be reformulated as an assigning procedure of predefined classes to detected events. In other words, certain set of classes, e.g. electrons and protons, are of interest and the particle identification algorithm should analyse the measured particles and separate them between these classes. To perform certain recognition it is needed to describe each event by set of attributes which are in this case a cluster variables. These attributes should be maximally significant for the particle recognition into the predefined classes.
In this point of view, this work is mainly focused on the following basic classes: **electrons**, **protons** and certain **heavy ions**. An investigation were performed to obtain the most significant attributes with respect to this set of classes. All previously mentioned cluster variables were probed and following

---

[6] A heavy ions can serve as an example. It is very improbable that their tracks are curly for most of the conditions therefore it is possible to skip this class for event which possesses high value of curliness.

list of features showed the highest significance:

1. Size S
2. Energy sum E_{sum}
3. Maximum energy E_{max}
4. LET
5. Standard deviation in the perpendicular direction σ_{perp}
6. Energy pixel variables - E_I and E_{III}
7. Thickness
8. Ration of the border and inner pixel counts $\frac{N_{\text{inner}}}{N_{\text{border}}}$

The significance were examined by means of several tools reflecting the separation of particles into species which will be the main task of the algorithm. This ability can be seen in the figure 2.13 for the LET and the thickness.

An effort was also dedicated to a development of variables which would be universal in sense of the detector settings. This is uneasy task because most of the cluster features are strongly bound with the detector settings as was mentioned in the previous section 2.1.2. There is now variable which would meet both criteria for the universality and also significance. Therefore it was suggested to omit the condition for universality and develop in the future a conversion model which would be capable to provide this service.
One of the minor constraints for the algorithm was a possibility for an implementation directly on the Timepix3 device into microprocessor and provide an on-line particle recognition, so called on board processing.

2.2.1 Similarity Algorithm

The second step in the particle recognition process is the similarity algorithm. Its main purpose is to provide particle classification based on the similarities between an unknown measured state and stored label states in a database. This task is approached by means of the distance in the significant variables space. To enhance the algorithm computation performance, the simplification has been utilized and stored label states were grouped together to a

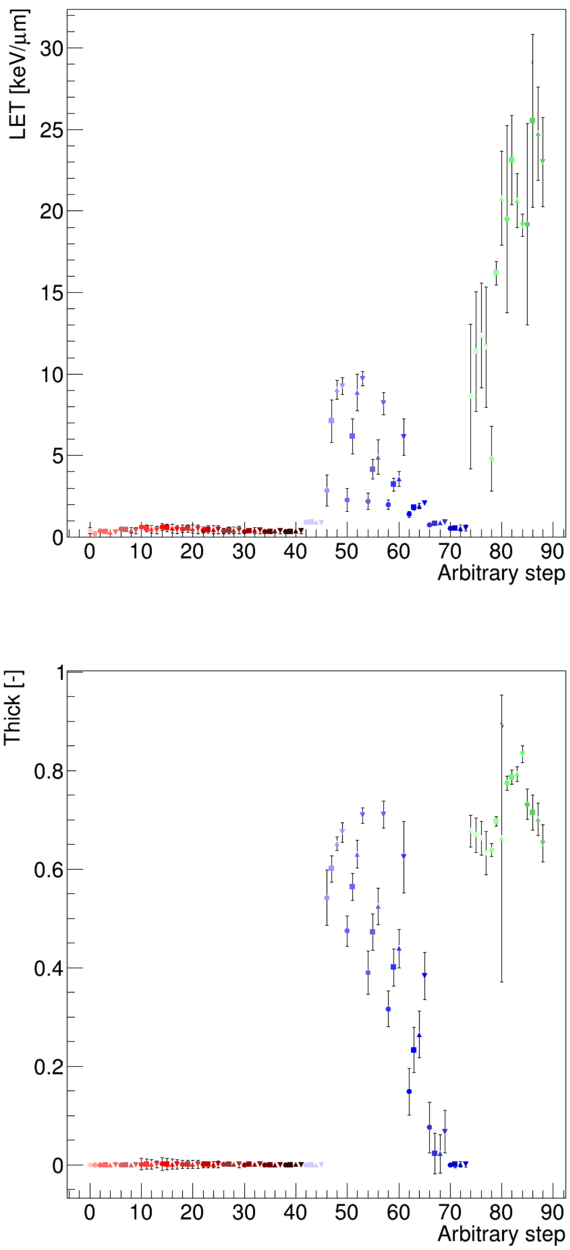


Figure 2.13: The right figure includes the LET for the data samples indicated in the subsection 2.2.2. The left figure is for the thickness and the same data samples. The red points illustrates electrons, blues points protons and green points heavy ions. Different shades of colors serve as a distinction between different energies. The elevation angle is repented with different shapes of the markers for one shade of color. The axis is arbitrary and can used to separate particle species: from O to app. 40 are electrons, from app 40 to 75 are protons and the rest of the range is dedicated to heavy ions.

several mean values with dispersions. Each part of the above algorithm definition is discussed in more details in the following text.
The main task of the similarity algorithm is to identify some unknown measured state X. This can be done by means of the distance in some cluster variable space of $\{R_I, ..., R_{N_{\text{var}}}\}$ where known labelled states $\{k\}$ are placed. The dimension of this space is equal to the number of the used variables, N_{var}. The known labelled states represent a very well defined measurements when all necessary free parameters are known, e.g. measurement of protons with energy of 30 MeV at the elevation angle of 30 degrees. In the first step, algorithm computes the distances between the set $\{k\}$ and X (an example of state m is used in the following equation):

$$\Delta R_{X,m} = \sqrt{\sum_{j}^{N_{\text{var}}} (R_{X,j} - R_{m,j})^2} \tag{2.36}$$

A schematic illustration of the algorithm can be seen in the figure 2.14 where a simplification with only 3 significant variables R_I, R_{II}, R_{III} is used are this will be hold for the rest of the description. As can be seen in the figure, each state k is surrounded by certain so called state dispersions $\sigma_{R_I,m}, \sigma_{R_{II},m}, \sigma_{R_{III},m}$ which defines some validity space for each state (the dashed lines). This means that if the unknown particle X lies within this validity space of the known state k it is possible for X to be the state k and on the other hand if the state X is outside of this space than it is not possible that X is state k. The validity spaces or state dispersions introduces a new class in the classification process. It be derived that when X particle does not lie within any validity space of the set $\{k\}$ than the algorithm is unable to decide which class should be assign to this unknown state X and therefore is is labelled as a class of **others**.

This is propagated to the first step of the algorithm as a limitation whether the distance between the X and k should be computed. The final output is a set of distances $\{\Delta R_{X,k}\}$ only these states which possible for the unknown state X, in other words the X lies within their validity spaces. An additional condition had to be implemented to the algorithm to avoid an overestimation of some final class by its representation in the set $\Delta R_{X,k}$ by more than one element. It will be derived in the further paragraph that the classification is done by the set of the distances (and their values) therefore if some class is represented by more than one element then it would obtain

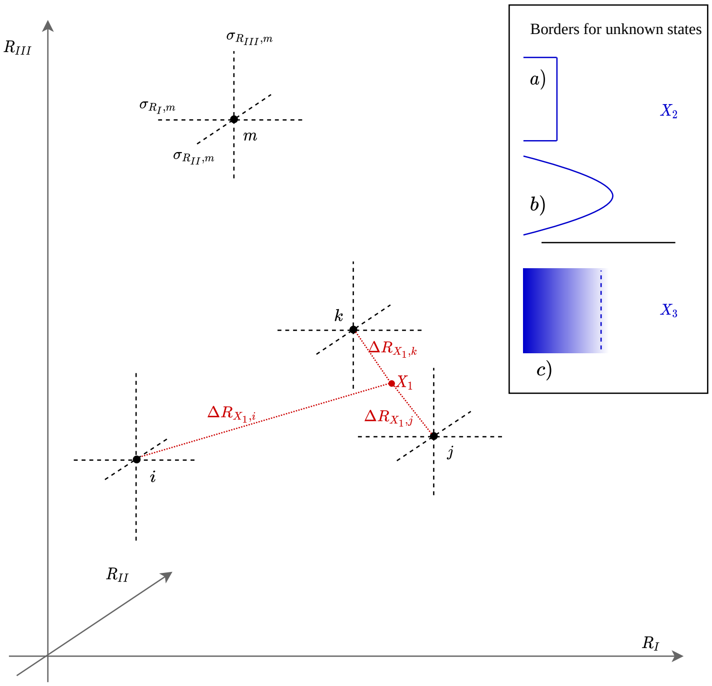


Figure 2.14: The scheme of the similarity algorithm for general 3D dimensional space of R_I, R_{II}, R_{III} variables. Algorithm probes the distances between referential states k, j, i and the unknown state X_1. In the right top part of the figure, the different approaches to the state borders are depicted.

automatically higher importance than other classes. This is eliminated by a condition that the class can be included only once and it should be the state with the minimal distance $\Delta R_{X,k}$.
In the second step, it has to be decided therefore a conversion between the distance $\Delta R_{X,k}$ and probability $p_{X,k}$ is needed. A 3 main cases are distinguished for this purposes based on the number of elements in the set $\{\Delta R_{X,k}\}$:

- **0 elements:**
  In this case, the algorithm was not capable to find any state which would comprehend in its validity space the unknown state X. In other words, the unknown state is ascribed to the class of **others** and neither state get any contribution.

- **1 element:**
  The number of found and possible state is only one, lets assume that it is state k. This means that the unknown state X belongs to its class:

  $$p_{X,k} = 1 \tag{2.37}$$

- **2 and more elements:**
  There are several states possible to the unknown state X. A redistribution rule is needed in this case to assign each state a appropriate probability. A several features has to be preserved when defining the rule and the main one is disproportion to the distance. In other words, if the unknown state is closer to the referential state then the probability should increase. The prescription of the rule is in the following way:

  $$p_{X,m} = \frac{1}{\Delta R_{X,m}} \frac{1}{\sum_i \frac{1}{\Delta R_{X,i}}} \leq 1 \tag{2.38}$$

  where two main parts can be observed. The first fraction reflects the distance disproportion and the second stands for the normalization therefore the sum over all assign probabilities $p_{X,m}$ is 1. In the figure 2.15, the redistribution rule and mainly its effects can be seen. An additional condition had to be constrained on the the rule to avoid an undesirable behaviour in the cases when the unknown state possesses almost the same value of R as the referential state. This leads to a redistribution of the probability in a way that this close state achieves almost 1 and other states close to 0 despite fact that all the states can

> be very similar. They likeness is defined with the distances between their R values (the points themselves) and also with their spreads. If these numbers are very similar[7] then each state should obtain the same amount of the probability.

One of the issues occurring in the algorithm is an exploitation of variables with different units and ranges. Therefore each of them have to be scaled to avoid an undesirable over or underestimation. The most simple way was preformed for this task and every variable was scaled to correspond the cluster size with a linear function: $R(S) = \frac{1}{w_i}S + b_i$ (R is one of the variables). This approach acquires a set of scaling weights $\{w_i\}$ which are applied during calculation of the distance by an additional multiplication in the sum with factor w_i^2 (the shift b_i is suppressed by the difference):

$$\Delta R_{X,m} = \sqrt{\sum_{j}^{N_{\mathrm{var}}} w_j^2 (R_{X,j} - R_{m,j})^2} \tag{2.39}$$

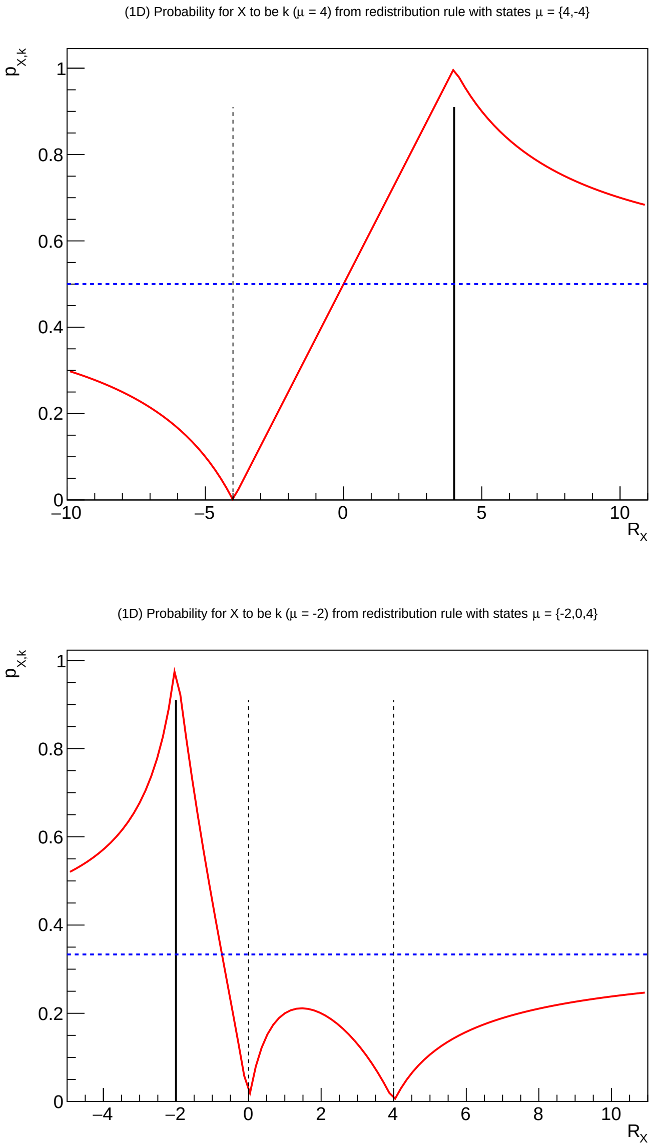
2.2.2 Database and Implementation

The referential states represent a well defined measurements as was already mentioned. In the case of for example protons with 30 MeV and the elevation angle of 30 degrees, a measurement with this settings is performed and statistically sufficient sample is taken. The original idea of the kNN algorithm is to store this data for each detected particle[8] which creates a set of approximately thousands of points in the variable space. A simplification is applied in this phase and the whole sample is replaced by a single mean value and appropriate dispersions which comprehends certain portion of the original sample. In the current implementation a 95 percent of the data is preserved in the state dispersion. This concept also offers a possibility to change this fixed approach by a probability densities if the given variable is reproducible by a certain function. For example the deposited energy follows the approximately Landau distribution for a thin medium. This idea is not

---

[7] This is parametrised and expressed with their distance and also with ratio of their spreads.

[8] Some upper limit for the number of particles is stated and only the significant points are considered.


(1D) Probability for X to be k (μ = 4) from redistribution rule with states μ = {4,-4}

(1D) Probability for X to be k (μ = -2) from redistribution rule with states μ = {-2,0,4}

Figure 2.15: The figure of the probability redistribution rule only for one variable R. The figures shows the probability $p_{X.k}$ that the unknown state X is a label state k with given mean value μ_k (the bold black line) including other possible states (the dashed black lines reflects their mean values) as a function of the R variables for the state X, R_X. In top plot the investigated label state is with a mean value of variable R equal to 4 and the other state is at -4. In the bottom plot, the investigated state is at -2 and the other possible states are at 0 and 4. The blue dashed line illustrates a limited case when variable R for the unknown state X reaches large values with respect to the differences between label states values of R therefore the probability is equally divided between all possible states.

exploited in the implementation due to the more demanding nature of this improvement and it will be discussed at the end of this section.
One of the most important part of the particle identification algorithm is the database where all the referential states are stored. Its composition has one the great effects on the method performance.
The current implementation uses the following set of data samples (the origins are in the squared brackets):

- **Protons:**
  1. High energy protons - 70, 100, 200, 300, 400 MeV [Cyklotron-Krakov, MC]
  2. Medium energy protons - 8, 13, 20, 30 MeV [Cyklotron-Řež]
  3. Low energy protons - 500 keV [MC]
- **Electrons:**
  1. High energy electrons - 16, 23 MeV [Microtron]
  2. Low energy electrons - 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.9, 1.2, 2, 4, 7 MeV [MC]
- **Heavy ions:**
  1. He - 20, 35 MeV [Cyklotron-Řež]
  2. Si - 400 MeV/u [HIMAC]
  3. Fe - 1 GeV/u, 350 MeV/u, 230 MeV/u [HIMAC]

All of the mentioned samples were measured or simulated with following set of elevation angels: 0, 30, 45, 60, 85-90 degrees[9]. The measurement were performed at the following list of accelerator facilities:

- [HIMAC] - Heavy Ion Medial Accelerator of the National Institute of Radiological Sciences NIRS, Chiba in Japan
- [Cyklotron-Řež] - U-120M isochronous cyclotron at the Nuclear Physics Institute, Czech Ac. of Sci. Řež

---

[9] In some cases the last value of the angle fluctuates between 75 to 90 degrees.

- [Microtron] - microtron accelerator at the Nuclear Physics Institute, Czech Ac. of Sci. Řež
- [Cyklotron-Krakov] - cyklotron at the Inst. Fyz. Jadrova, Polish Ac. of Sci. Krakow

The experimental set-up used in the measurements consist of a Timepix3 detector with sensor thickness of 500 μm and applied voltage of 80 or 200 V. The detector was placed on the rational stage with a possible step of 1 degree which serve for the altering of the elevation angle of detected particles. The position of the set-up with respect to the beam was checked by the means of a laser alignment system. During measurements the detector was rotated to above mentioned angles. This can be performed via distant control of the whole set-up with two computers, one in the vicinity of the detector (connection with USB3 and special lead shielding) and the other one in a control room (connection via ethernet cable). From the software point of view, each step of the measurement is configured and controlled by the PIXET software [23].
The last source of the data samples is Monte Carlo simulations with the Geant4 simulation tool [24], the abbreviation MC in the list. This data was provided within a cooperation with the Institute of the Experimental and Applied Physics by Stefan Göhl[10]. Majority of the charge sharing effects are included with an exception in the volcano effect and the halo effect due to their so far undescribed nature.

Most of the acquired data includes unwanted background events which have to reduced before the saving into the database. Certain part of the filtering was done with the deposited energy combined with the known direction for most of the samples which can be reflected in the values of morphological features of clusters. In the majority of the cases, the portion of the background in samples was minimal which was easing the filtering procedure. Data suffering with large background were excluded from the database because it could introduce into the methodology undesirably large uncertainties[11].

---

[10] Contact: Stefan Gohl, MSc., Stefan.Gohl@utef.cvut.cz.

[11] For example a measurement of cosmic rays at Earth ground as a source of muons. It is very hard to separate sample which would be as clear as possible.

2.2.3 Performance

The algorithm performance was tasted with above mentioned data set which was divided in to two groups according to its purpose: training and testing. The training part of samples is stored in the database and the testing part serves fro the probing of the algorithm efficiency. This separation is unequal with ratio 1:5 where 5 portions are for the training set. Also an unused data for the training, certain values of angles or energies, are utilized for the testing purpose.

The selected classes are following: **electrons**, **protons** and **others**. The class of others includes the undecided events and also the class of heavy ions. Motivation of this option can be found in its usage for the hadron therapy, especially the lateral beam monitoring. based on the previous works, it is assumed with confidence that the electron class will include photons which gives more importance to the idea of the others class as mainly created from heavy ions with high ionization.

The first presented results are fluxes of particles form the testing set and unused one, see the figure 2.16. The right figure shows a testing data set of protons with 30 MeV and 45 degrees[12]. It can be observed that the accordance with the total flux is almost 100 percent. The left figure includes fluxes for the unused set of protons with energy 30 MeV and 20 degrees. It can be deduced from the bottom ratio figure that the protons are on average represented by 95 percent of detected particles. These fluxes only serves as an examples of the algorithm behaviour and the crucial feature is its final efficiency which is discussed in the next paragraph.

The efficiencies of the classes can be expressed with so called confusion matrix which is a common tool to express the performance of the classification in the machine learning. This matrix can be seen in the figure 2.17 for the currently used set of classes and the database mentioned above. The column label expresses the input data and the row label reflects the output classification. The diagonal elements are portions of the detected particles if the same particle is as the input therefore they are proportional to the final efficiency. The off diagonal elements serves as indicator of the unsuccessful classification and they include the information about the method uncertainty. In the case of electrons, the mixed MC sample can be exploited to reveal the matrix elements if they are as the input data. This was not possible for

---

[12] All stated angles are elevation ones.

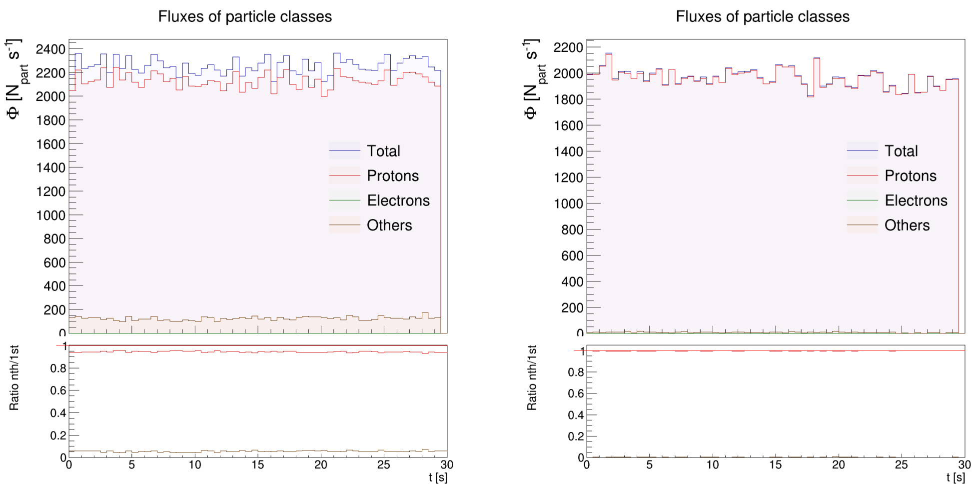
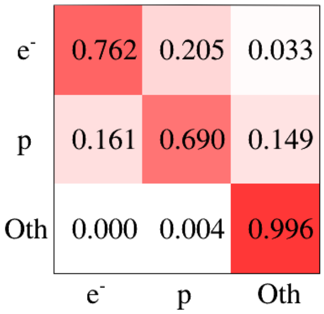


Figure 2.16: The fluxes of particle classes after processing via the particle classification algorithm. the right figure includes a testing data of protons with 30 MeV and 50 degrees. The left figure is for a protons with 30 MeV and 20 degrees. Both of the figures includes in the bottom part the ratios of the total flux with other fluxes.

protons because the present state of the simulation is not capable to simulate all needed detector effects therefore mixed samples can not be easily generate. A certain interval procedure had to be implemented for the used data. Protons stored in the database possess energy from 500 keV to 400 MeV. This energy range was separated into two equal intervals from 0 to 200 MeV and from 200 MeV to 400 MeV and for each of this interval an average efficiency was calculated. The overall efficiency was obtained as their average value. Similar procedure was utilized for the elevation angles which are also approximately desecrate.

It can be observed in the figure that the there is certain overlap between class of electrons and protons. This can be partially solved by an introduction of new classes which are subordinated or a merge of the original ones. It is namely the class of high and low energy protons with energies with energy limit 200 MeV, above this condition the proton is labelled as high energy one. It was investigated that mainly the tracks of these high energy protons resemble the electron ones and causes large uncertainties during their identification.

The resulting confusion matrices for this adjustment can be seen in the figure 2.18. An significant improvement can be observed int the left figure with low energy protons as a class. The right figure shows the current limitation of the algorithm where high energy protons generates ambiguities and discrimination with electrons is very coarse.

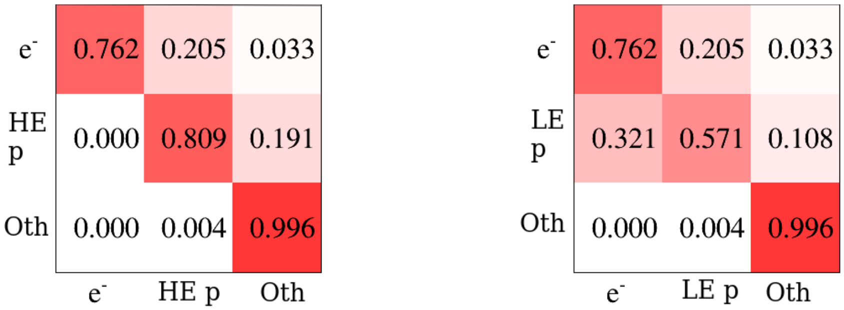


	e^-	p	Oth
e^-	0.762	0.205	0.033
p	0.161	0.690	0.149
Oth	0.000	0.004	0.996

Figure 2.17: The confusion matrix for the above mentioned database. The column labels reflects the input data and the column labels the output classification.

To conclude this section, it possible with presented algorithm perform particle classification with certain precision which depends on the selected set of classes expressed via the confusion matrices. The main source of uncertainties originates from similar features of the described particle classes, namely high energy protons and electrons. An additional error can be caused by the filtering procedure but it was assumed that it is a minor one. To achieve a higher resolution power, the set of classes can be modified and some them separated and merged which should be done according to a usage of the algorithm.

One of the biggest advantages is robustness of the algorithm. Changes or completion of the database or used variables allow a universal application. Important feature of the algorithm is the possibility to be implemented on the device due to its not demanding nature in the sense of the calculation time. This provides an option for the on-line particle identification.



	e^-	HE p	Oth
e^-	0.762	0.205	0.033
HE p	0.000	0.809	0.191
Oth	0.000	0.004	0.996

	e^-	LE p	Oth
e^-	0.762	0.205	0.033
LE p	0.321	0.571	0.108
Oth	0.000	0.004	0.996

Figure 2.18: The confusion matrices for two separated proton class. The right figure includes the high energy proton class and the left one the low energy proton class.

Chapter 3

Application in Hadron Therapy

3.1 Introduction

Radiotherapy is one of the most common approaches for a treatment of cancer. The main idea is to deliver the dose to the target volume and maximally spare the structures in its vicinity. This can be possible utilizing charged particles which due to their localized energy deposition meet this criterion.

The firs proposition to use high energy charged particles was done by R. R. Wilson in 1946 [25] during his investigation of proton beams and their dose characteristics. In comparison to other methods, e.g. photon radiotherapy, charged particles offers a delivery of adjusted and concentrated high dose to the target volume. It can be demonstrated in the figure 3.1 where relative biological effectiveness[1] is illustrated for different particle utilized for radiotherapy. This capability enables a sparring of a surrounding tissue which is highly appreciated especially for cases when a tumor can be located in a critical organ (brain, optical nerves etc.).

A crucial part during radiotherapy is a monitoring of a particle beam and its effects on the irradiated volume and consequent adjustment to achieve the most optimal dose delivery. This is mainly done by altering the energy of a beam which changes depth of the Bragg peak and together with the aimed position, a spatial application is possible. Despite a careful planning before the procedure, an additional patient-related sources of uncertainties emerge

---

[1] Relative biological effectiveness, RBE, quantifies actual effect of a dose on a biological tissue caused by a certain type of radiation. This variable is introduced due to a difference in the biological effect after absorption of the same amount of dose from different types of radiation [26].

during the application of the treatment and cause an undesirable misplaced irradiation. Those sources are mainly of geometrical nature e.g. creation of air cavities, movement of body organs. Consequently, this has to be include it into the treatment and corresponding adjustments should be performed to conform dose.

For this purposes one characteristic feature of heavy ion beams can be exploited. Upon a traversing of a primary particle, it undergoes several interactions with the volume particles, mainly atomic electrons and nuclei. One of the possible interaction is also inelastic scattering with nucleus resulting in so-called fragmentation. This causes a release of new particles which are mostly in the direction of the extinct primary one and with similar velocity [27]. The key feature is their ability to escape from the targeted volume and their detectability (this will be more discussed in subsection 3.1.1). It has been firstly proposed in the work of Dauvagne et al. [28] and Amaldi et al [29]. An unquestionable advantage of the approaches is its non-invasive way.

One possibility how to advance in this problematic is exploitation of the Timepix (Tpx and Tpx3) detectors which offer sufficient detection sensitivity within the scope and spectrometric ranges of escaped secondary particles. First clues towards the possibility of the Timepix detectors usage was done by Martišíková et al. [30], Jakubek et al. [31] and Henriquet et al. [32] . The main conclusion originating from this previous work should be that the monitoring of a beam dose delivery with semiconductor pixeled detectors in telescope architecture is feasible and promising. The first application for the monitoring of the beam range was done by Heidelberg group described in the article [33]. Contribution of this work lies in the introduction of a precise methodology which is based on the back-projection of a direction reconstructed from detected particles in a telescope. This will be discussed in more details in the subsection 3.2. The most recant work done in this summarized in the article [34] where a precise methodology for a lateral scanning of a beam is described with achieved precision from 0.84 to 2.59 mm.

The main contribution of this chapter should be quantification of effects during particle propagation through irradiated volume. It is focused on the horizontal determination of secondary particle origin which is of interest for final evaluation of the dose distribution. This should be helpful for more accurate predictions based on the particle tracking. An additional effort is also dedicated to a possible application of the particle identification to improve the lateral beam monitoring with the methodology mentioned in the section 2.2.

Content of this part is following: in the first section then main interactions and their effects are discussed then a description of experimental and simulation set-up is mentioned. This is continued by quantification of the final uncertainty with an additional interest in the individual contributions: Multiple Coulomb scattering, inelastic interaction especially additional fragmentations and contribution from a beam width. At the end of the chapter, a possible usage of particle identification is discussed to filter more affected particle species by propagation trough medium.

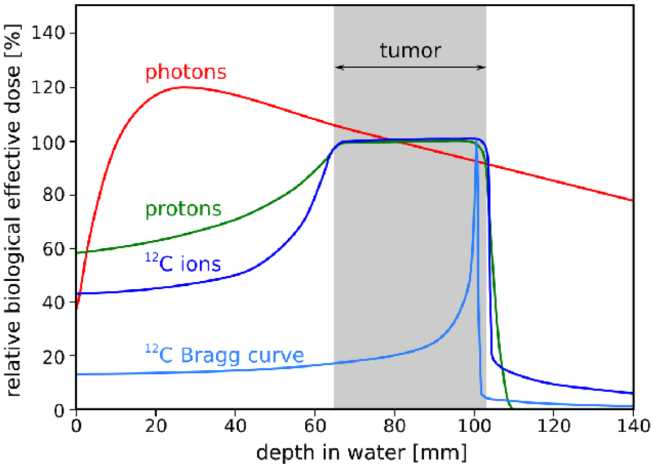


Figure 3.1: Biological effectiveness as fuction of a depth for protons, carbon ions and photons. The spread-out-Bragg-peak for protons and ion carbons (different energies) is shown [35].

3.1.1 Main Interactions of Ions in Medium

During the propagation of primary carbon ion trough medium several interactions are feasible with respect to its features. Considering the energy range used for the treatment purposes, carbon ions with energy approximately 200 MeV/n are exploited. The interaction with atomic electrons dissipates most of the energy of a primary particle but with respect to an altering its direction it is the least significant one. The biggest contributions are mainly from multiple Coulomb scattering and fragmentation alias interactions with the target nuclei.

- **Multiple Coulomb scattering (MCS)**:

  Charge particle crossing a medium is gradually deflected by Coulomb scattering with a medium nuclei. The resulting deflection angles are relatively small and this was successfully described by the Rutherford cross section (all additional contributions are neglected) [4]:

  $$\frac{d\sigma_R}{d\Omega} = \left(\frac{1}{4\pi\epsilon_0}\frac{zZe^2}{4E_{kin}}\right)^2 \frac{1}{\sin^4(\Theta/2)}, \tag{3.1}$$

  where E_{kin} is the kinetic energy of the projectile, Z is charge of the nuclei in the target, z is the charge of the projectile and Θ is the deflection angle. To express a net effect of multiple Coulomb scatterings which causes an overall deflection of a particle, the theory by Molière's was developed [36] (see figure 3.2). Through central limit theorem it can be derived that overall deflection distribution for thick target or large number of small scatterings should be Gaussian like and a contribution from possible "hard" scatters[2] cause a non-Gaussian tails [4]. An empirical formula for standard deviation of the deflection can be deduced for most of the cases when small angle deflection is caused, so-called Highland's formula [37]:

  $$\theta_{\text{plane}} = \frac{13.6\ \text{MeV}}{\beta cp} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \ln\left(\frac{xz^2}{X_0\beta^2}\right)\right] \tag{3.2}$$

  where result is in radians (solid angle), z, p and βc are the incident particle charge, momentum and velocity, x/X_0 is thickness of the medium

---

[2] Scattering with large angles.

expressed it the radiation length therefore it determines the material features of the medium. It can be deduced from the formula that heavier charged particles experiencing smaller effect of MCS [3]. Decreasing the momentum of incident particle will cause enlargement of the deflection. Regarding the medium effects, those with heavier nuclei will cause larger spread. [27].

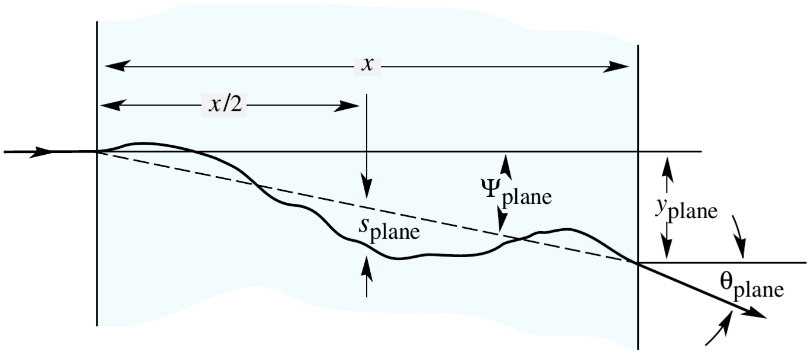


Figure 3.2: Schematic figure to visualize the effect of the multiple Coulomb scattering. The main deflection quantity is also shown: θ_{plane}.

- **Fragmentation:**
  The projectile can also interact via nuclear force with the target nuclei. This process has significantly lower probability than the Coulomb interaction but it possesses a large importance for the overall problematic of the beam monitoring as was already mentioned in the introduction to this chapter [27].

  Upon this interaction, the projectile nucleus inelastically interacts with the target nucleus which can be with respect to the projectile energy (several hundreds of MeV per nucleon) treated as staying at the rest. The characteristic of interaction and the results can be described with the help of an abrasion-ablation model [38], also discussed in Gaimard et al. [39], Gunzert-Marx et al. [40]. Form the geometrical perspective, it can be deduced that most of the interactions are happening peripherally (for an illustration, see the figure 3.3). In the first stage

---

[3] In a comparison between proton and carbon radiotherapy with the same beam range, the deflection can be 3 times larger for protons than for carbon nuclei [27].

of this process an overlapping nucleons of projectile and target nucleus are abrated into the so-called ”fireball”. The rest of the projectile and target are less effected. In the case of the projectile its fragment continues with almost unchanged velocity oppositely to the fragment of the target nucleus which is almost motionless. The consequent step is deexcitation of the fragments and the fireball by means of the evaporation of nucleons or light nuclei (this can be also accompanied by a photon radiation)[40].

Regarding the directional distribution of newly escaped particles, so-called secondary[4] particles, it is in the case of fragments steeply falling with a rising angle from the direction of the primary particles. A different behaviour can be observed for the evaporated particles which are isotropic and with much lower velocity therefore their detection is very unlikely. Very important consequence should be highlighted from the fact that the projectile fragment possesses similar velocity, if the energy deposition is taken into consideration it can be shown that such a lighter particle has capability to emerge from the medium [40]. One of the net effects can be seen in the deposited energy in the medium where tail-part can be seen after the Bragg peak of the primary carbon ion caused by the secondary particle (see the figure 3.1).

3.1.2 Experimental and Simplified Simulation Set-up

A following experimental setup was proposed and build to measure secondary particles induced during carbon radiotherapy and especially to reconstruct information about their direction, see the figure 3.4. The main part are two stacted dectectors Tpx3 (the telescope architecture) placed in the vicinity of an irradiated volume. The sensor thickness for both detectors is 300 μm and the applied bias is around 30V to maximally avoid the volcano effect. In this case (the figure 3.4), an anthropomorphic Alderson head phantom inhomogeneities has been utilized as the target medium [34]. The distance of detectors from the phantom isocenter is 12 cm (the closest possible distance) and the angle is set with respect to the beam axis (30 degrees in the depicted

---

[4] This term will be used through mots of the text for all particles created during the fragmentation process inside the target volume despite the fact that also the secondary particles undergo the fragmentation which is resulting in creation of tertiary particles, higher generation particles.

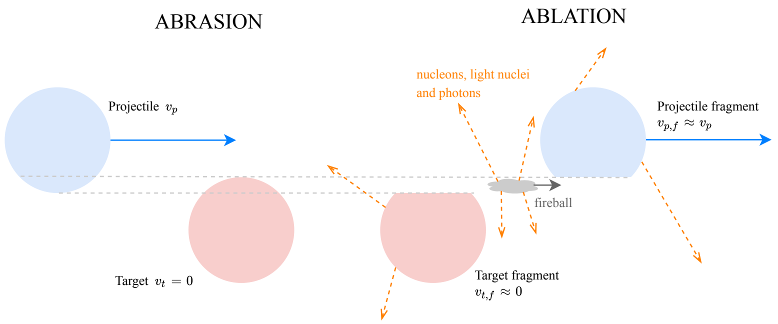


Figure 3.3: A schematic figure of the peripheral interaction between a projectile and a target nucleus with depicted fragmentation and evaporation/deexcitation processes.

case).

This experimental setup is replaced in a simulation with a simplification which concerning the irradiated volume, see the figure 3.5. A homogeneous PMMA cylinder substitutes original phantom (body tissue in the real scenario) to obtain sufficiently similar results but easing the simulation. The distances used in the experiment are preserved and only angle between telescope and the beam axis is altered.

The primary beam energy is 213.41 MeV/u which was calculated to obtain a range of the beam in the center of the PMMA cylinder, isocenter. The width of the beam was simulated with a 2D Gaussian fuction in the x and y axis with FWHM of 5.02 mm. To simulate similar broadening of beam energy (the figure 3.1), a so-called Ripple filter was placed in front of the source of carbon ions.

The choice of energy also motivated an option for the coordinate system origin which was set to the center of the cylinder.

3.2 Uncertainty of Particle Tracking

The main goal of this section is to investigate a **deflection** of a secondary particle(s) after its creation in a primary vertex[5]. This should be very beneficial to the final monitoring of the beam range and subsequently to an

---

[5] The fragmentation interaction of a primary particle.

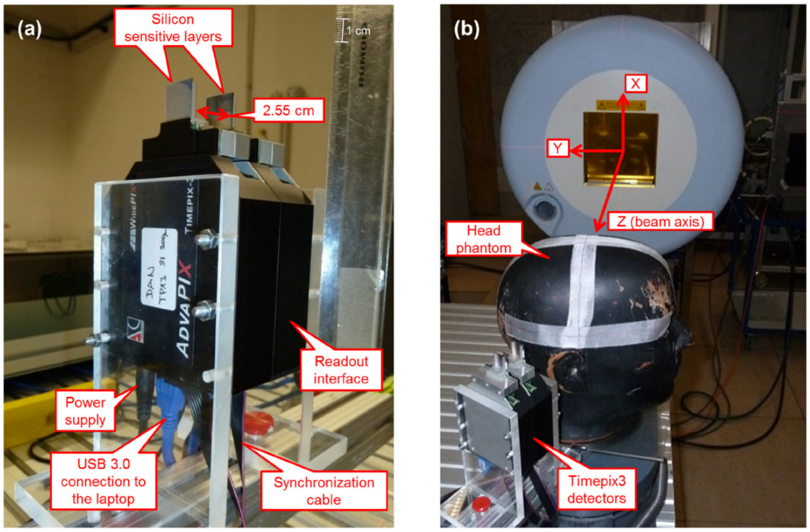


Figure 3.4: Photography of used experimental set-up: a) detailed figure of stacted Tpx3 detectors, b) an alignment for measurement under angle of 30 degries with respect to the axis of the beam. Reprinted from [34].

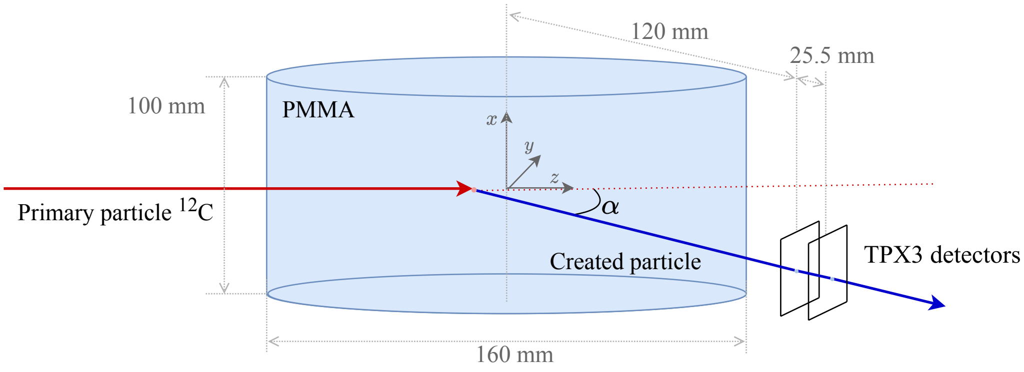


Figure 3.5: Schematic figure of the simulation setup and all important distances are included.

additional analysis (e.g. presence of cavities within a beam trajectory, the subsection 3.4).

The current proposition for the range beam monitoring can be divided into two parts: the lateral (position on the x and y axis) and the horizontal (position on the z axis) determination.

The **horizontal monitoring** methodology is based on the work Gwosch et al. [33]. A back-projection of detected particles is made to a given position of the beam axis (for detailed formula, see the next subsection 3.2.2) which is an input parameter of this analysis. The simulated secondary distributions in the z axis can be seen in the figures 3.7 for different detectors angles (blue distributions). The red distributions correspond to an original z-position of the primary vertexes. A correlation between position of a stopping beam position (around z = 0) and the shape of the distribution can be spotted from this figure. The method exploits this fact and performs linear fits of the part of the distribution closely to the stopping position. Two method were established for those purposes:

- **Distal edge of z back-projection distribution:**
  A position of a distal edge is defined as z value where linear fit reaches half of the distribution maximum and the fit is performed on the distal fall-off [6]. A linear correlation between the position of the beam Bragg peak and the distal edge can be spotted and utilized for the final beam monitoring.

- **Change of slope of distal fall-off:**
  As the distal fall-off transfers to remarkably slower decrease of the z back-projection distribution, a correlation can be investigated between this "transfer" point and the beam Bragg peak position. Two linear fit are exploited for this method: one covers mostly the distal fall-off and the other the end of the distribution (less decreasing part). The final correlation between intersection of these fits and the beam stopping position were observed and reveal possibility for the beam monitoring.

Both methods are based on the statistical approaches and minimaly take advantage of the individual tracking of the secondary particles. The application of the uncertainty of the individual tracking proceeder can be highlighted in

---

[6] This can be clearly seen in the figure 3.7 on the blue distribution for angle 3O degrees. By the term distal fall-off is meant a decrease of the distribution approximately in the z interval (-5,0). For detailed definition see article [33].

the case of cavities presence. It is needed to recognize that a inhomogeneity, cavity of air, lies within the trajectory of the beam. This could cause a shift of the Bragg peak position and adjustment of the beam energy should be performed. A cavity should ideally cause a decrease in the z back-projection distribution. If the information from the individual tracking is too uncertain than this cavity may not be recognized. This indifference in the distribution will propagate into the horizontal monitoring methods which would be therefore ineffective. This topic is more discussed in the section 3.4 together with the results for the uncertainty of the secondary particle tracking. The problematic of the cavity detection can be found in the work of Gaa et al. [41] for the Timepix detectors which concludes that the above presented methodology is sensitive to reveal presence of cavities for the detectors position of 30 degrees.

The **lateral monitoring** approach is summarized in the article by Bautista et al. [34]. Detected secondary ions are also back-projected to a horizontal plane (perpendicular to the beam axis) at certain z position. It is not possible to derive the actual z position of the plane form this method and a connection with the horizontal analysis has to be made to supply this information. After the back-projection, the mean value of the beam position is calculated (also its deviation) which creates the final output of the analysis. The current precision was estimated with the help of the beam aiming system recorded in a beam position file and the detectors at the angle of 30 degrees. The most recent analysis exploits the angle of 0 degree. Eventually, a possible modification can originate from particle specie selection due to a interaction dependency on a incident particle charge (further discussion in the section 3.3).

To investigate the uncertainty of the secondary particle tracking it is needed to obtain an information about a primary vertex position. The most simple (maybe only one) way how to approach this problematic is trough a simulation. For this purposes, it was decided to exploit a FLUKA Monte Carlo tool [42, 43]. Due to a partial invariance of the lateral monitoring method, it is focused on the horizontal monitoring, namely, on the z back-projection.

3.2.1 FLUKA Simulations and Distributions of Secondary Particles

The FLUKA code is one the basic simulation tools used in the particle and medical physics[7]. Its main purpose is a calculation of a particle transport and interaction with certain medium.

The FLUKA was adjusted to simulate setup 3.5. The main physics option in DEFAULTS card was set to HADROTHE which should accurately simulate conditions during the hadron therapy (all parameters of HADROTHE can be found in the publication [42]). Next to this default card, the CAOLESCE[8] and EVAPORATION[9] cards were included and adjusted in the simulation to maximize a real situation reproducibility.

The final simulation was also accompanied by routines written in a C++ and BASH language. They main benefit was parallel running of several simulations via changing the random seed to preserve independence of the output data and also to maximize the time efficiency by ongoing running of new simulations when the previous ones ended[10].

The simulation were repeated with different angles α and for given angle also several times to obtain sufficient statistics (the minimal level was 20 000 of detected particles). This very important as the angle α rises, see the table 3.1. It can be observed in the last column of the table that the number of detected particles quickly decreasing, exponentially (see figure 3.6 where exponential fit was performed).

The first interests of the investigation were basic distributions of the detected particles dependent on the angle α with a focus on the secondary particles, see the figures:

(a) **3.7 detected secondary particle distributions (red part)**:
The primary vertex z coordinate is exploited in this distribution to obtain information about the origin of the secondary particle.
One general feature can be observed in the distributions. The portion of the particles within a given z interval is almost independent on the

---

[7] The overall field of appliacation is much broader e.g.: proton and electron accelerators, shielding, dosimetry, cosmic rays etc. For more see article [42].

[8] This card serves to activate coalescence mechanism which serves for a creation of more complex nucleus from closely evaporated ones.

[9] The card setting was adjusted to also include this mechanism for the heavy fragments.

[10] Input parameters: starting seed number, number of simulations, the angle α and the number of simulated primaries.

α [°]	$N_{\text{prim}}*10^6[-]$	$N_{\text{det,mean}}[-]$	$N_{\text{det,tot}}*10^3[-]$
0	4	33521	134
10	8	8286	65
20	20	1768	34
30	40	567	22
40	116	207	24
50	270	79	21

Table 3.1: The total number of simulations N_{prim}, average number of detected particles $N_{\text{sec,tot}}$ and total number of detected particles $N_{\text{det,tot}}$ for different angles α.

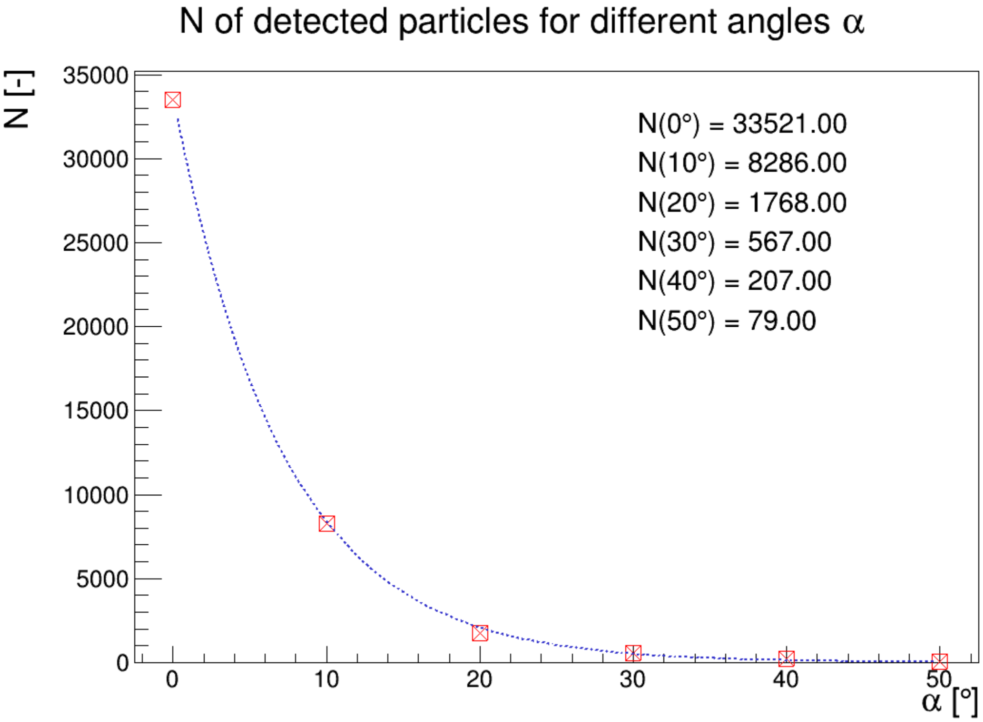


Figure 3.6: A graph of the average number of detected particles during the FLUKA simulations for different angles α. The blue line depicts a exponential function use to fit the data points.

angle α. In other words, despite the fact that total number of emitted particles is decreasing with the angle α which are of different nature (specie, energy - see the next figure description) it remains proportionally the same with respect to the z origin coordinate.
Most of the particles which are capable to reach detectors are created in the first half of the PMMA cylinder (z in interval from -8 to 0 cm). The cause can be found in the features of the fragmentation procedure. The secondary particles created in this region are capable to cross full medium due to their higher energy. As the energy of the primary particle decreases with a depth in the medium the probability of the secondary particle to escape from the target decreases.
This steep decrease of the distribution is continued by an almost constant part at a position of approximately 0 cm. In this region only second generation particles can create new particles of higher generations via the fragmentation. Increasing of α from the value 30 degrees causes that an original edge of the distribution at 8 cm for 0 degree gradually retreats to a value of 5 cm in the case of 50 degrees. This can be explained by the means of a geometry and the figure 3.6, especially the mutual position of the detectors and these particles almost at the edge of the PMMA cylinder. It can be seen that their angle is much higher than the angle α between the detectors and the beam axis and as it is approximately indicated in the figure 3.6 the amount of particles decreasing almost exponentially with the emitting angle.

(b) **3.8 generation of detected particles**:
A general correlation can be observed throughout all distributions. A increasing of the angle α causes that a higher generation particles are more detected. This can be attributed to their lower mass and charge nature and it is more discussed in the next subsection 3.2.2).

(c) **3.9 ID[11]/specie of detected particles**:
A gradual rise of proton domination can be seen from the distributions. The least contribution is for α equals 0 degree where the major component creates helium. This is the only case when also heavier fragments can be detected with sufficient statistics. As the angle increases the helium component diminishes and at the angle of 50 degrees it can

---

[11] A key for transfer between particle ID and its specie or category is written in the figure description.

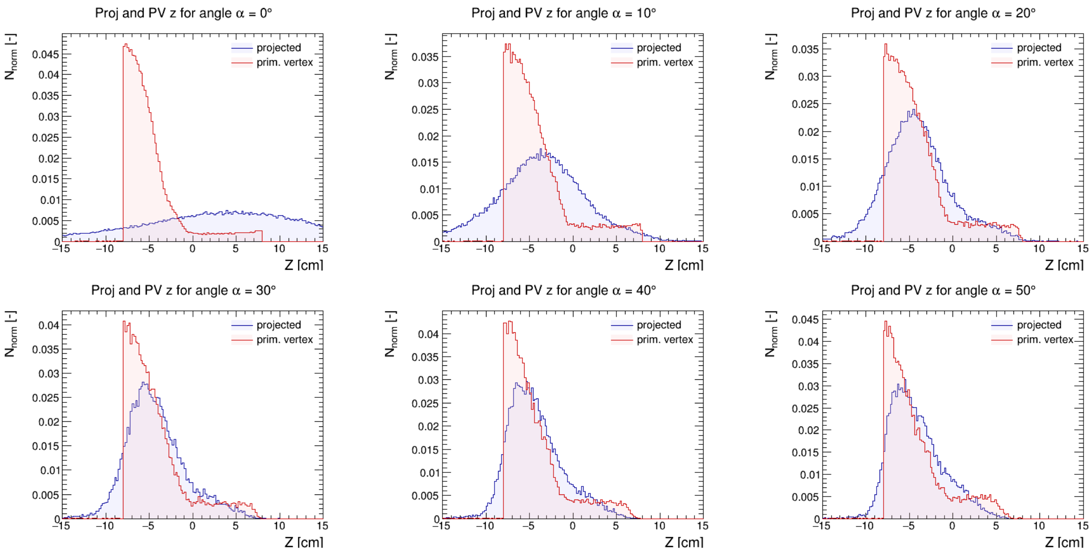


Figure 3.7: The detected particle horizontal distributions for different angles α. The red histogram is for primary vertex position of a secondary particle and the blue histogram depicts the reconstructed position based on the z back-projection.

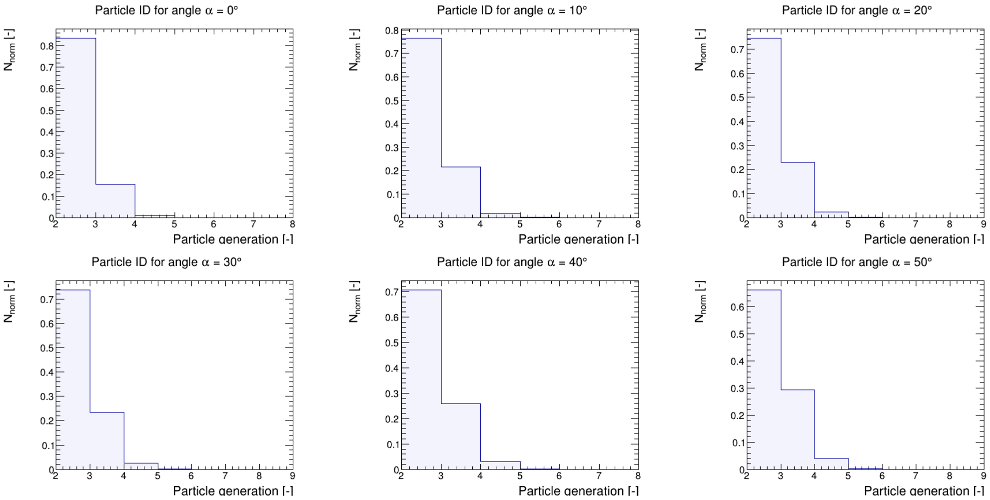


Figure 3.8: The generation distributions of the detected particle as a function of the angle α.

be approximately stated that only protons contributes to the detected secondary particles. A similar explanation can be applied as the one for the generation composition.

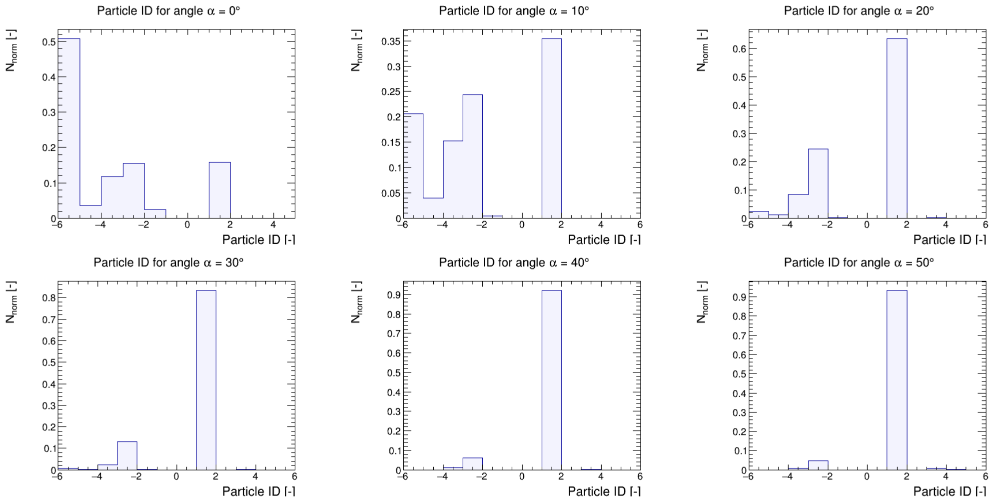


Figure 3.9: The ID distribution of detected particles for different angles α. A following key can be used to transfer ID of particle to their specie or a particle category: 1 - proton, -2 - heavier fragment than helium, -3 - deuteron, -4 - tritium, -5 - isotope of helium ^{3}He and -6 - helium.

3.2.2 Overall Uncertainty of Charged Particles Mapping

Before the investigation of the tracking uncertainty it is suitable to mention its assumed main contributions and the final dependencies which can be observed in the figure 3.10 as an illustration of secondary particle propagation through medium. The main contribution can be categorized in two classes as follows:

1. **Secondary particle propagation through medium:**
   The secondary particle interact with a medium which cause a deflection or a new particle are liberated (high-generation particles).

2. **Geometrical limitation of the back-projecting method:**
   As will be more discusses in the subsection 3.2.2, the back-projection exploits the nearest distance between the beam axis and the reconstructed direction/line of detected particles to obtain z-coordination of the reconstructed primary vertex. This is only simplification considering the beam width which is non-zero.

From the above most likely contributions, it is possible to deduce dependencies of the uncertainty (numbers label above mentioned list items):

- **[1] Medium composition:**
  This dependency can be easily included within a FLUKA simulation where different target materials can be set. Concerning the multiple Coulomb scattering, it can be observed from equation 3.2 (from the radiation length) that this dependency is:

  $$\sigma_R \approx Z^2 \tag{3.3}$$

  therefore material with higher Z is capable of higher deflection of a traversing particle. Eventually, this dependency was not tested and only one medium PMMA was used due to a time demanding nature of the simulations[12].

---

[12] A practical reason also arises because the real application will involve inhomogeneous target (body tissue) and therefore a simulation of this tissue as complex structure would be better solution than an effort to simulate its individual parts.

- **[1] Particle features (kinetic energy $E_{\mathbf{kin}}$, particles charge z):**

  As it is stated in the Highland formula 3.2, the deflection is disproportional to the momentum and velocity of the incident particle and proportional to its charge. The best resolution of tracking is assumed in the first half of the PMMA target before the stopping position of the beam. This idea is based on the z distributions originating from the primary vertex 3.7 and on the deposited energy to secondary particle because both of them reaches highest value in this region.

- **[1,2] Position of the detectors - angular dependency**:
  From the geometrical point of view, it can be proposed that the uncertainty should decrease with a rising of the angle α of the detectors and the beam axis. It should be also noted that the average total length L_{tot} of detected secondary particles in the medium is changing with α. The highest average should be most likely for 0 degree case when particles created in the entrance of cylinder have to cross almost its full width (see the figure 3.7 red part, most of the particles originate at beam entrance to the cylinder).

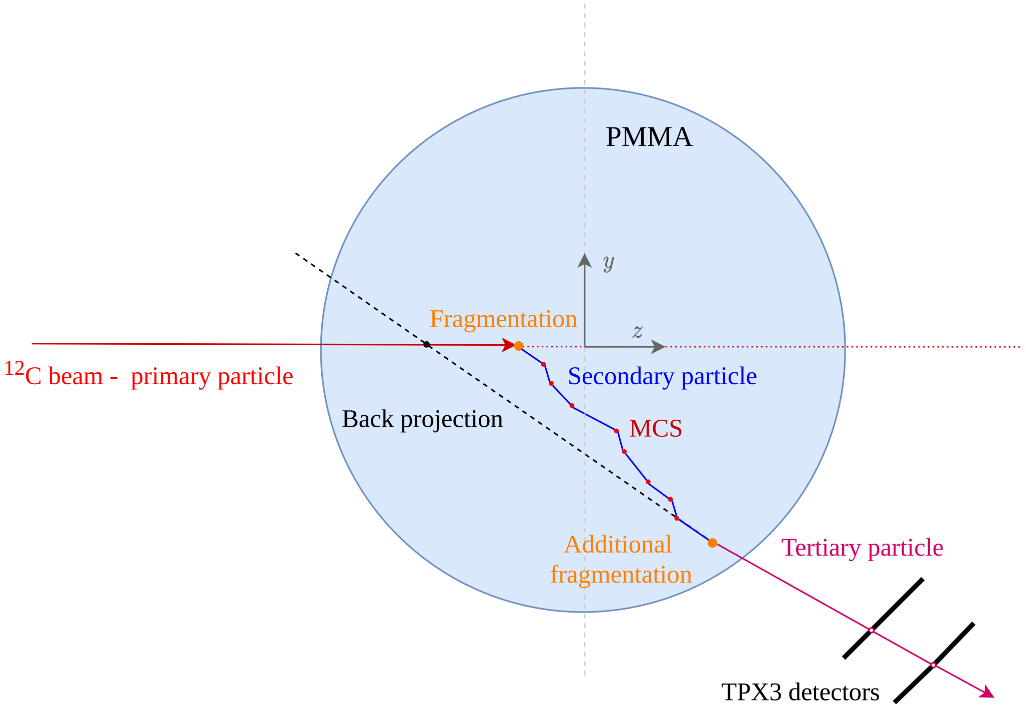


Figure 3.10: A basic scheme of a primary particle interaction in a target and a subsequent propagation of a secondary particle through this target. The main contributions caused by an additional secondary particle interactions are highlighted. The orange dots are positions of the fragmentation processes (it is assumed that also secondary particle can undergo the fragmentation) and the red dots represent the MCS. In the scheme is also depicted the reconstructed/back-projected line, black dashed one.

The first discussed result will be the overall uncertainty of z back-projection. The investigated variable was difference of z coordinates Δz between reconstructed/back-projected one z_{proj} and the originating form the primary vertex z_{PV}: $\Delta z = z_{\text{proj}} - z_{\text{PV}}$.
The simulation was performed and adjusted with respect to the main contributions and expected dependencies. All the results are expressed with the main dependency on the angle α which was discretely altered according to the following values: 0, 10, 20, 30, 40, 50 degrees. This choice was done with respect to the close proximity of the detectors and the PMMA cylinder.

The FLUKA simulations was accompanied with C++ programs to provide a fast data processing even for large input samples:[13]

1. [FLUKA] **Tracing individual particle:**
   A user routine MGDRAW was developed for this purposes by the FLUKA creators [42]. Its main function is to track currently transported particle and record particle events. This for example means that if particle traverses any boundary the MGDRAW routine is called (entry BXDRAW) and an user has possibility to approach this event and perform needed actions (e.g. store particle features). Together with additional conditions for crossing of both layers, it is feasible to save secondary particle features (crossed place at the detector, energy etc.).

2. [FLUKA] **Saving the primary vertex:**
   One of the necessary information is position of the primary vertex. The FLUKA provides possibility how to track origin of every new particle. This can be done with an usage of a routine STUPRF which handles all particles newly saved in the FLUKA storages (the main ones are called FLKSTK and EMFSTK) for their further transportation [14]. The primary vertex position was stored in a variables offered to users for free usage called SPAREK and called within the MGDRAW whenever particle fulfils the detectors crossing conditions.

---

[13] Only the most important parts are mentioned with discrimination of the FLUKA and C++ routine part in square brackets.

[14] It was also needed to skip all processes where the same particle is "created" as the incident one. This is for example radiation of photon when the FLUKA considers the incident particle in the output as new one. This can be distinguished in the STUPRF routine with a value of a variable NPPRMR.

3. [C++] **Calculation of z back-projection:**
   The z back-projection developed by Heidelberg team is based on the closest distance between the projected line and the beam axis. Because of accordance between this and z coordinate axis, it stays that a z position where the closest distance line interferes with the reconstructed line corresponds to a reconstructed primary vertex z coordinate. As was already mentioned, this is only approximation because the beam possesses Gaussian profile in the x a y direction. The formula for calculation of the projected/reconstructed z position is following:

$$z_{\text{proj}} = \frac{[(B_x - D_{1,x})(D_{1,x} - D_{2,x}) + (B_y - D_{1,y})(D_{1,y} - D_{2,y})](D_{1,z} - D_{2,z})}{[(D_{1,x} - D_{2,x})^2 + (D_{1,y} - D_{2,y})^2]^2} + D_{1,z}, \tag{3.4}$$

   where D_1, D_2 are the detected points and B is any point on the beam axis (as an illustration, see the figure 3.16).

4. [C++] **Investigation of Δz contributions and dependencies:**
   Additional conditions on the particle features were used to observed certain dependencies according to previously mentioned list. For further discussion see the next subsections about the individual contributions. To express a fluctuations in the results caused by the stochastic nature of the interactions a mean values and standard deviations are exploited:

$$\mu_x = \frac{\sum_i^N x_i}{N} \tag{3.5}$$

$$\sigma_x = \sqrt{\frac{\sum_i^N (x_i - \mu_x)^2}{N-1}} \tag{3.6}$$

   where investigated variable is x with values $x_1, ..., x_N$. In the case of the Δz, the mean value can be seen as a constant bias or shift of the method and the standard deviation as its error.

The overall uncertainty distributions as a function of the angle $\Delta z(\alpha)$ can be observed in the figure 3.11 and the resulting mean values with the standard deviations as error bars in the figure 3.12.

Two separate cases are shown for the mean and standard deviation values: with and without a condition on the z back-projected coordinate to lie within the cylinder (from 8 to -8 cm). The events which lies outside the cylinder can

$\alpha[^\circ]$	0	10	20	30	40	50
$N_{\text{outside}}[\%]$	43.8	22.1	13.8	10.5	8.5	6.4
$N_{\text{det,mean}}[-]$	33521	8286	1768	567	207	79

Table 3.2: The number of lost events N_{outside} which are outside the cylinder in a ratio to the total number of events for different angles α. The mean number of detected particles (total number of events) $N_{\text{det, total}}$ is also shown for 10^6 primary particles.

be taken as unimportant for the tracking due to their obvious impossibility. This decision improves the overall uncertainty as can be seen from the values in the figure 3.12. A lost rate can be defined with this additional condition which express number of projected events outside the cylinder N_{outside} to the total number of events. The results are in the table 3.2. They serve as a certain background uncertainty to the adjusted z back-projection method with the above additional condition. It can be assumed and stated that the lost rate follows behaviour of the overall uncertainty of $\Delta z(\alpha)$ with a decreasing tendency for larger angles α.

Concentrating on the bottom plot of the figure 3.12, the tracking bias and uncertainty possess a tendency to decrease with a increasing angle α which is in the accordance with the expected dependency mentioned in the previous text. It is worth noting that between the cases of 30, 40 and 50 degrees a rise of the uncertainty is manifesting with a slight change of the bias. This can play a crucial role for the decision which angle α should be used for the final monitoring. If the number of detected secondary particles is also considered, the 50 degrees position should not be utilized for this purposes instead of the 40 or 30 degrees one.

It the top figure, the differences between all cases are much more profound and they could be taken with greater importance due to their generality. An outstanding is conversion between 20 and 30 degrees cases when the shift and also the standard deviation decrease rapidly (more discussed in the section 3.4). This leads to a decision for the next uncertainty evaluation whether this condition should be included.

The method itself produces the projections outside the cylinder therefore it is its natural behaviour. Regarding the methods proposed for the beam monitoring, the horizontal distributions of the detected secondary particles are considered without this condition. Therefore other processing of $\Delta z(\alpha)$ (the individual contributions investigation, dependence on the ID and generation of secondary particles) has been performed without the condition but the $\Delta z(\alpha, z_{\text{proj}})$ (also a function of the z back-projected coordinate) is focused on this interval because the information importance is much higher.

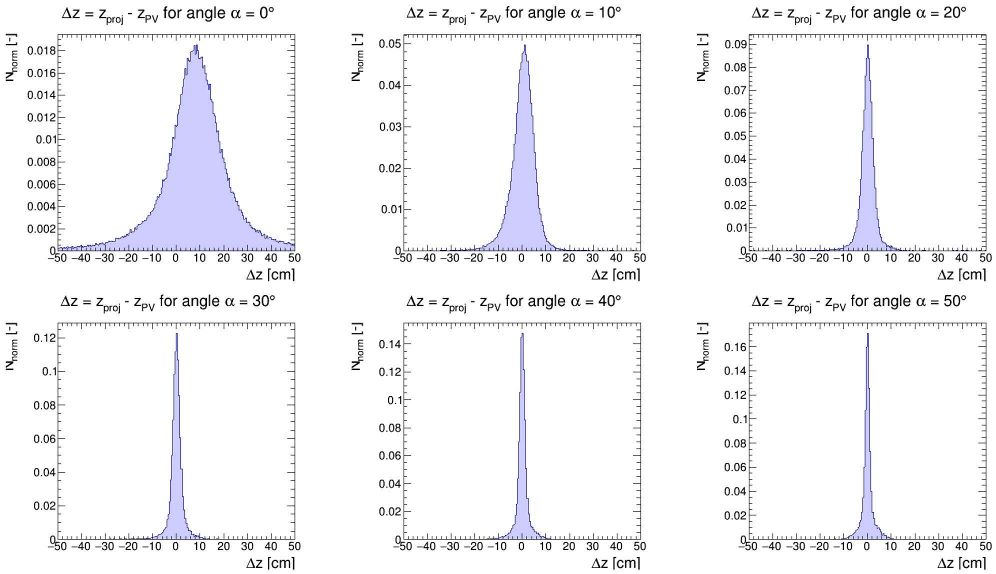


Figure 3.11: The distributions of Δz for different angles α (now additional condition was used).

The dependence of the uncertainty on the angle and also on the projected z coordinate z_{proj} can be seen in the figure 3.13. The z intervals of the graphs were set to 1 cm and it was constrained that 100 events has to be at least within a z interval to be considered in the graph. A general behaviour can be observed for all angles α, the z back-projection is more accurate for the region before the beam stopping position with raising tendency of the bias as approaching the end of cylinder at 8 cm. This could be ascribed to several effects, one of them is the higher energy nature of the secondary particles originating from the region which is decreasing with z coordinate. The MCS is more dominant for particle with lower momentum and causes their higher deflection. This will be more discussed in the subsection about individual

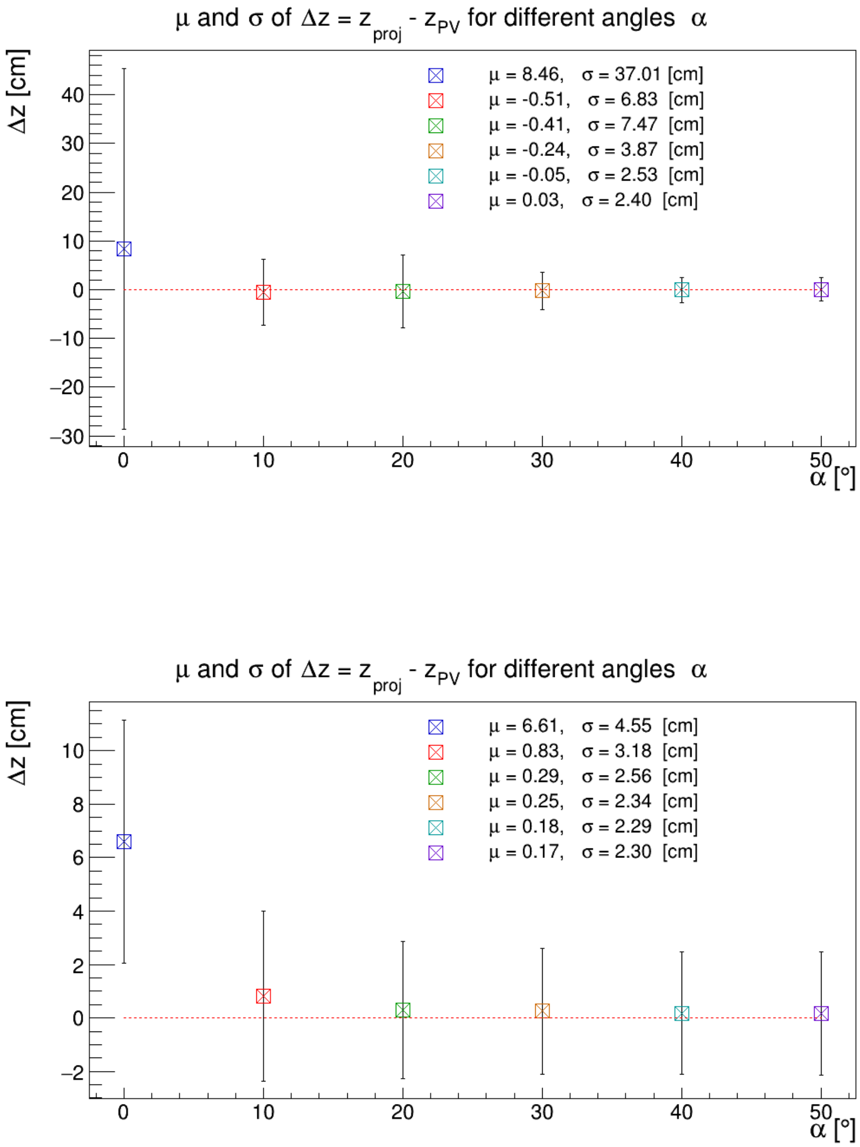


Figure 3.12: The mean and standard deviation values of Δz for different angles α. The **top** figure is without the additional condition on the z back-projection coordinate to lie within the cylinder. The **bottom** one is with this condition.

contributions where this effect is clearly manifesting.

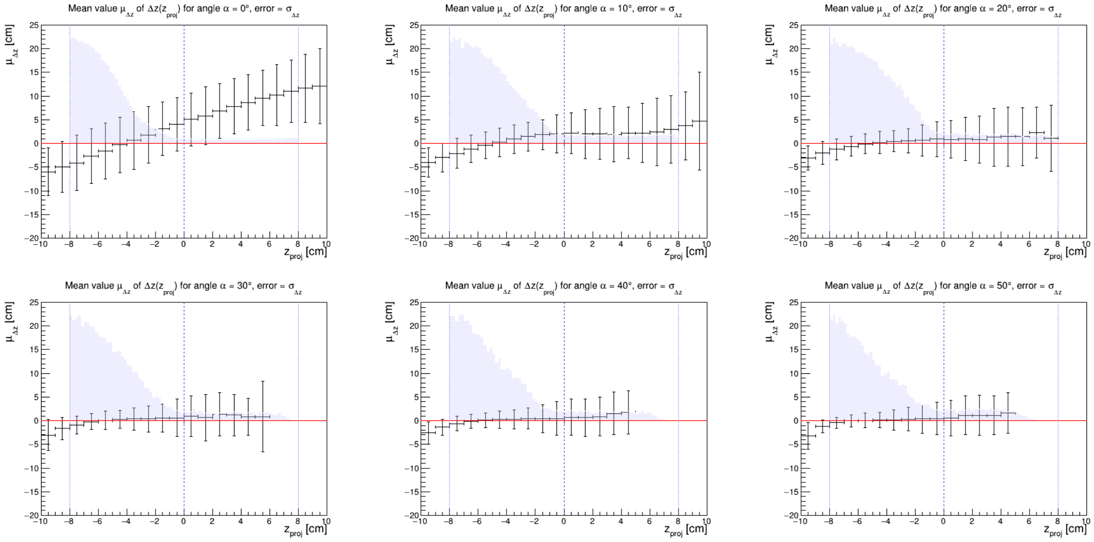


Figure 3.13: The mean and standard deviation values of Δz as a function of z back-projected coordinate z_{proj} for different angles α. In the background of each plot is an original distribution of secondary particles as a function of z coordinate of primary vertex.

The particle generation dependency of uncertainty $\Delta z(\alpha)$ is displayed in the figure 3.15.. It can be observed that till the angle of 30 degrees the uncertainty of 4th generation are smaller that those of 3rd one. In the case of 0 degree angle, the uncertainty is decreasing with higher generations. This can be viewed as a partially unexpected result if an idea is accepted that any other interaction of the secondary particle increases its deflection[15]. A possible explanation can be find in the features of the fragmentation process and MCS. It is known that projectile fragment has similar velocity as the projectile itself but possibly lower charge and mass. In the figure 3.9 can be seen that till the angle of 30 degrees, proton isotopes (deuteron and tritium) are present as a significant component. It was investigated that

---

[15] When the projectile fragment is created it preserves the direction of the primary particle which can be before the fragmentation more deflected than its the average deflection.

their significance is higher for the 4th generation than for the 3rd one[16]. This higher mass particles due to the additional neutrons but with the same charge should experience a smaller deflection by the MCS. It is also known from the fragmentation that the distribution of the secondary particles is peaked in the direction of the primary particle. Therefore together with the MCS features, it can be assumed that for larger angles (approximately more than 30 degrees) an occurrence of heavier nuclei and also proton isotopes is suppressed (this is in accordance with the figure 3.9) which is also manifested in the composition of generations. The original idea of disturbance by additional interactions becomes valid and particles of lower generations are more precisely tracked.

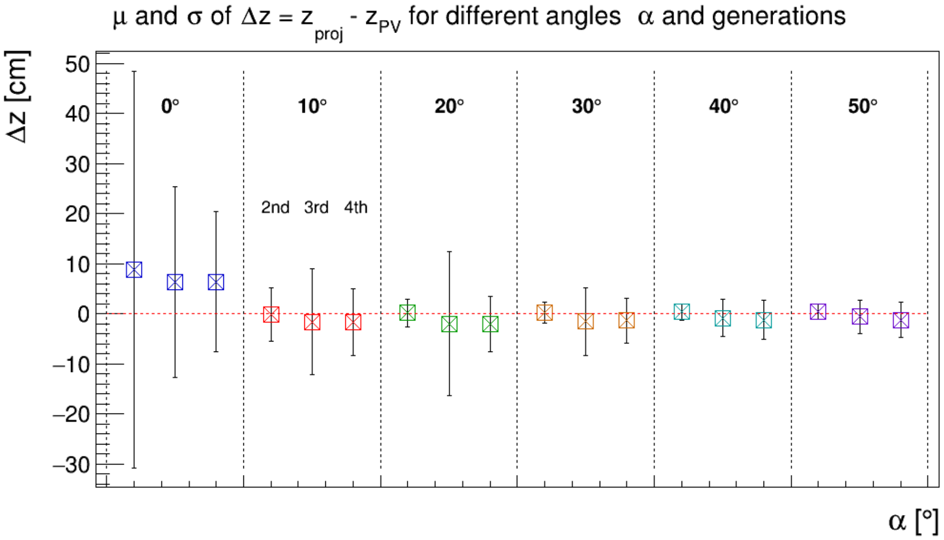


Figure 3.14: The mean and standard deviation values of Δz for different angles α and the particle generations.

One of the next examined correlations was the particle specie/type. The results are shown in the figure 3.15. An additional condition was used for this case to exclude those particle species which are not statistically significant for the monitoring. This was implemented as a rule of at least 5 percent occurrence in the sample. Particle specie/type remaining after this filtering

---

[16] The heavier particles than isotopes of proton are vanishing and therefore their portion is higher.

are: proton p, deuteron ^{2}H, tritium ^{3}H and helium ^{4}He.
Taking into consideration results from the figure 3.8, only angles 0 and 10 has significant contributions from helium (there is also very small amount of heavier nuclei) but the 0 angle case is unsuitable for the horizontal monitoring therefore it is not discussed individually (a partially similar behaviour can observed in between those two angles). The remaining case of interest is only the angle of 10 degrees when helium contributes with a portion of 20 percent to the overall number of detected secondary particles. The unexpected higher uncertainty of helium sample can be attributed to statistically fluctuations and their small amount.

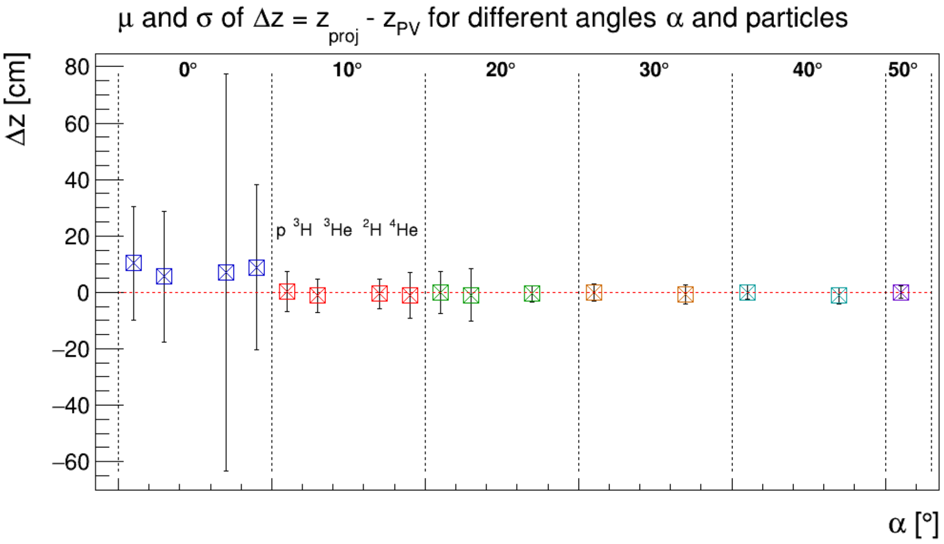


Figure 3.15: The mean and standard deviation values of Δz for different angles α and the particle specie/type.

3.2.3 Individual Contributions

This subsection should serves as an overview of the individual contributions to the z back-projection. The evaluation the uncertainty is the same as for the overall uncertainty with several exceptions in the FLUKA simulation settings. The Heidelberg team performed an additional analysis to investigate an optimal position/angle with respect to the monitoring methodology and sufficient number detected secondary particles. It was concluded that the most optimal angles α should be 20 or 30 degrees therefore more more focus lies on those angles. This decision was also made due to a time consuming nature of the simulations to obtain statistically sufficient data.

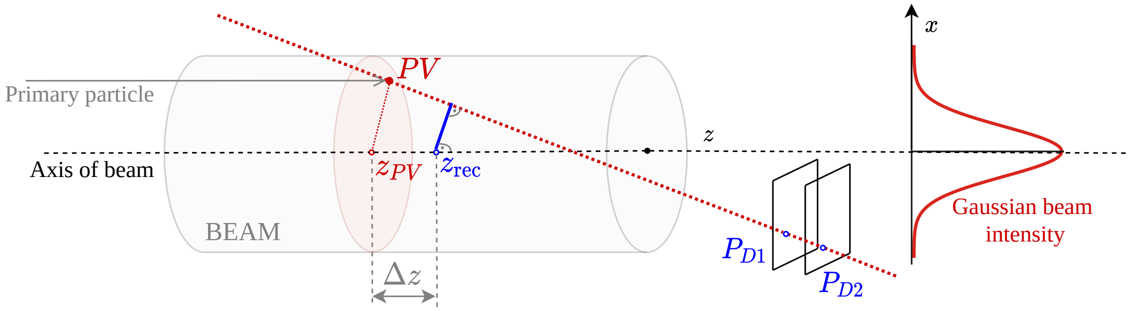


Figure 3.16: The schematic figure of the beam width contribution. The scale of all objects is adjusted to highlight the contribution. The beam intensity in x axis direction is also shown in the right part of the figure. The PV stands for a primary vertex.

1. **Beam width:**
   This contributions does not require an additional adjustments of the simulation and the final effect can be deduced with the equation 3.4. The full analytical can not be provided due to the missing knowledge about the secondary particles directions after the fragmentation process (an appropriate distribution would be combined together with the origin distributions and Δz calculation) and for this purposes the simulation is used.
   One replacement is needed in the equation 3.4 when the detected points are replaced with the primary vertex position P_{PV} and the direction of the secondary particle $n\vec{_{\text{PV}}}$ after its creation. The final formula for the difference Δz_{BW} states as:

$$\Delta z_{\text{BW}} = \frac{[n_{\text{PV},x}(B_x - P_{\text{PV},x}) + n_{\text{PV},y}(B_y - P_{\text{PV},y})]n_{\text{PV},z}}{[n_{\text{PV},x}^2 + n_{\text{PV},y}^2]^2}. \tag{3.7}$$

   The processing of the simulation had to be also adjusted due to the possibility that the secondary particles would continue after their creation outside the detectors. Those events do not reflect the needed directional distributions therefore only particles which crossed both detector sensors are considered.

The illustration of the beam width contribution due to the z back-projection is depicted in the figure 3.16. A approximate analytical expression for the dependence of the Δz_{bw} on the incident particle angle β can be derived from the equation 3.7 with an additional condition on the x coordinate to be 0:

$$\Delta z_{\mathrm{BW}} \approx \frac{1}{\tan(\beta)}, \qquad \beta = \arctan \frac{n_{\mathrm{PV},y}}{n_{\mathrm{PV},z}} \tag{3.8}$$

which follows natural conception that with rising angle of incidence particle the difference should be decreased. The resulting Δz_{BW} can be observed in the figure 3.17. The uncertainty is decreasing with the raising of the detectors angle α which also coincides with the raising of the incidence particle angle β.

The dependence on the back-projected z coordinate is in the figure 3.18. A similar behaviour as for he overall uncertainty can be seen for small angles, till 20 degrees. Higher angles behave almost constantly. High standard deviations in the 0 degree case are due to the emphasised geometrical contributions when slight deflection causes high differences in back-projections. Very important fact can be deduced from the missing points[17] after the center of cylinder for higher angle cases, more than 10 degrees. The absence insinuates that without the additional interactions of the secondary particle with the target there would be no detected particle from this region[18].

2. **Multiple Coulomb scattering:**
   The simulation had to be adjusted to reveal this individual contribution. For this purpose, the point of the beam in the equation was replaced with the primary vertex position which masked the added difference in the z position caused by the beam width[19]. To suppress

---

[17] The lowest number of events is set to 100.

[18] This also reflects the nature of the fragmentation process when created secondary particle capable to cross the target are mostly in the original direction of the primary particle.

[19] This option was chosen to avoid unnecessary simulations with 0 beam width and also to include the Gaussian characteristic of the beam as a possible enlargement of the particle trajectory in the medium.

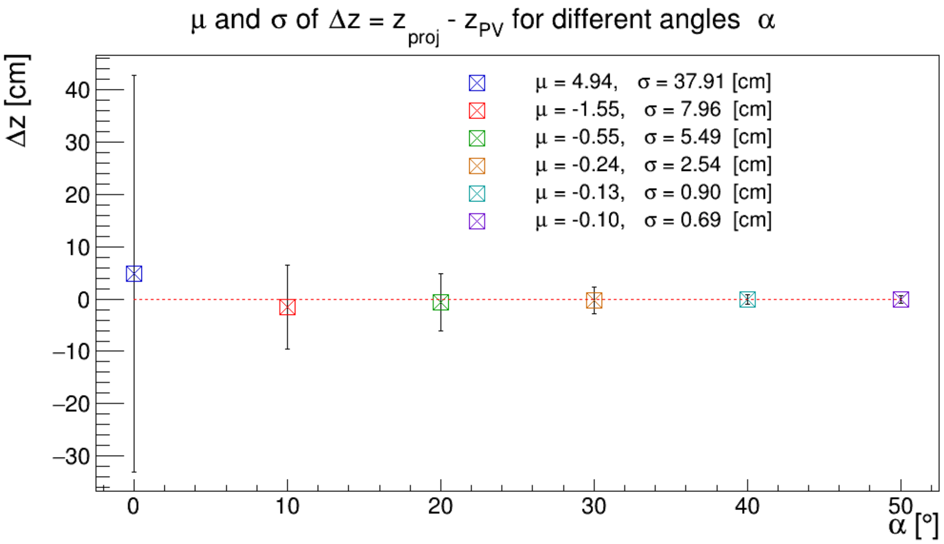


Figure 3.17: The mean and standard deviation values of Δz with only beam width contribution for different angles α.

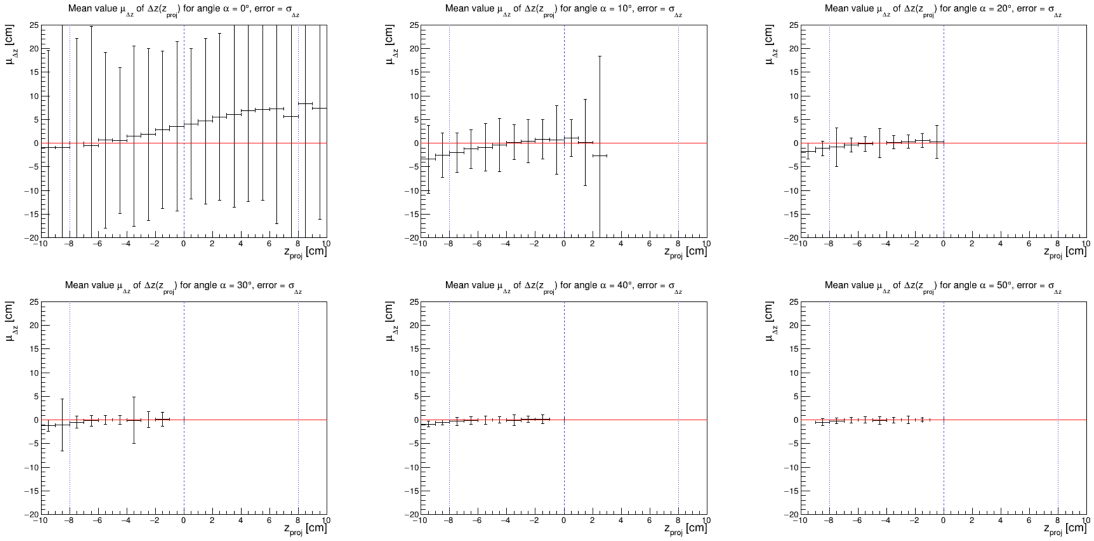


Figure 3.18: The mean and standard deviation values of Δz with only beam width contribution as a function of z back-projected coordinate z_{proj} for different angles α.

the contribution from an additional fragmentation, only particles of the second generation were considered which also means that the MCS contribution from 3rd and 4th generation is not included. This approximation can be partially justified with very small portion of higher generations in the detected particles[20].
The results are shown in the figure 3.19 and 3.20. The shift values observed in the figure 3.19 possess a slightly different behaviour than the overall uncertainty. They are decreasing till the angle of 20 degrees which is continued by an increase fo the shift.

Focusing on the Δz_{MCS} as a function of z_{proj}, there are no results after the z coordinate of 1 cm for angles above 20 degrees. This is caused by the filtering of higher generations which dominantly originates from this region. The option is mainly manifested in this result because in the overall results the higher generations are suppressed by their low number in the comparison with 2nd generation. A general tendency can be seen in all figures similar to the overall uncertainty behaviour, increasing of the bias and deviation with the z back-projected coordinate. Remarkable is a steeper slope of the increase in a comparison to other contributions.

3. **Fragmentation:**
The same approach in the alignment of the beam position with the primary vertex position was utilized to suppress the beam width contribution. The main adjustment of the FLUKA simulation was suppressing the MCS in all the materials of the simulation setup. This can be done by the means of the physics card MULSOPT with an option to include or exclude specific materials where MCS should not be performed. The result for angles 10, 20 and 30 degrees can be seen in the figure 3.21 and 3.22.
A raise can be observed in the shift values oppositely to the standard deviations which decreases with the angle α in the figure 3.21. A similar behaviour is manifesting in the case of the MCS which could be seen as a certain more fundamental contributor.

---

[20] In ideal case, simulations with one suppressed should be performed but this has not been done due to the time efficiency reasons.

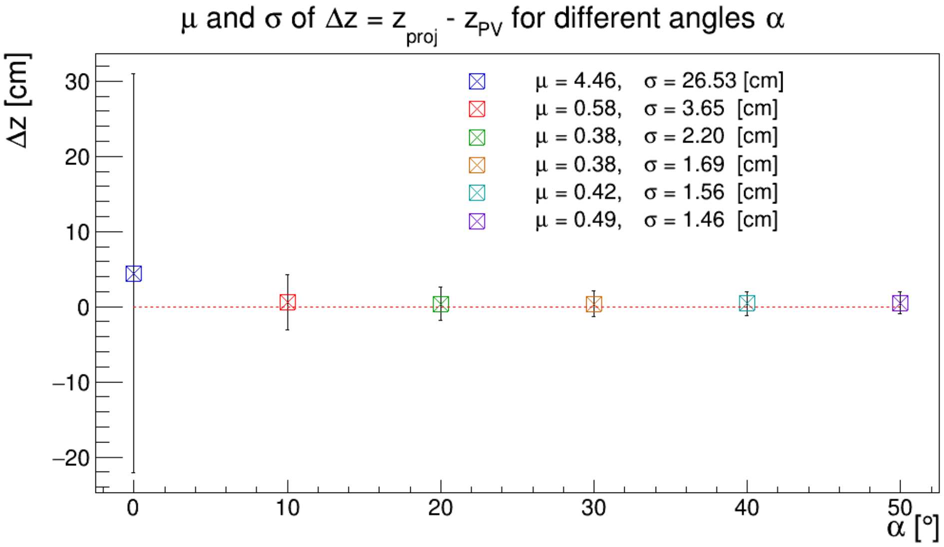


Figure 3.19: The mean and standard deviation values of Δz with mainly the MCS contribution for different angles α.

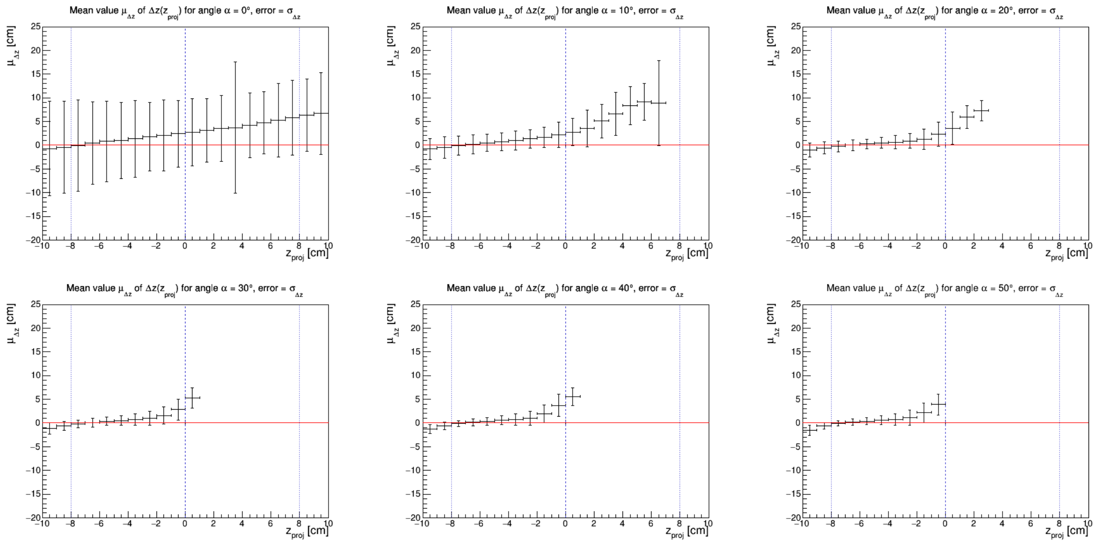


Figure 3.20: The mean and standard deviation values of Δz with mainly the MCS contribution as a function of z back-projected coordinate z_{proj} for different angles α.

A typical trend is manifesting in Δz as a function of the z back-projected coordinate.

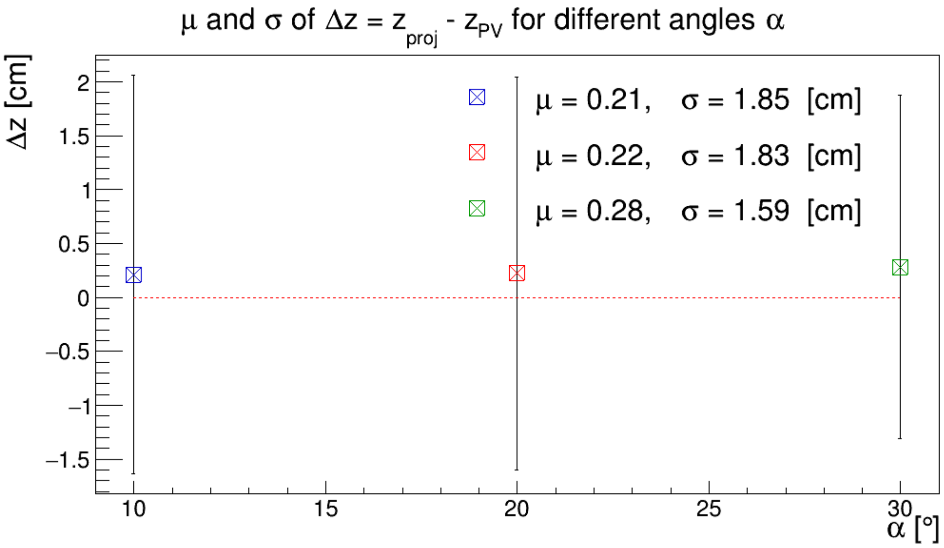


Figure 3.21: The mean and standard deviation values of Δz with mainly the fragmentation contribution for different angles α.

3.3 Particle Identification for the Lateral Monitoring

The deflection caused by the particle propagation through medium should be much lower for heavier nuclei. This can be utilized to more accurate tracking of particle for the monitoring purposes. It can be assumed that an additional filtering could improve its performance.

The methodology proposed in the section 2.2 can be utilized for this purposes. The measured data with the set-up described above were used. The position of detectors is at 0 degree of the angle α which was decided as the most optimal position for the lateral monitoring. The analysis of the data was done in the following way:

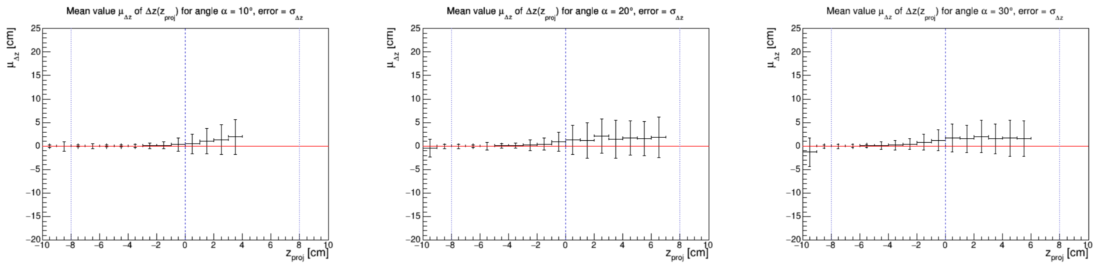


Figure 3.22: The mean and standard deviation values of Δz with mainly the fragmentation contribution as a function of z back-projected coordinate z_{proj} for different angles α (10 to 30 degrees from the right to the left part of the figure).

- Clusterisation of the data with time free parameter equals to 500 ns and all proposed cluster variables were calculated
- Coincidence process with the similarity algorithm for the most significant cluster features computed in the clusterisation were exploited (see the section 2.2)
- Only the information from the first detector was utilized for the final particle classification

The resulting fluxes can be seen in the figure 3.23. The simplified set of classes was exploited in this case: **protons**, **electrons** and **others**. The main purpose is to separate the proton component of the detected secondary particles form the heavier nuclei part. This task can be approached with the selected set of classes if every event labelled as a proton is excluded.

After the examination of the result in the figure 3.23, it can be stated that majority of the secondary particle should be created from protons with portion equals to 63.9 percent. Despite the higher efficiency in the energy region of protons detected in this measurement there is a large disagreement between the composition supplied by the FLUKA simulation. It is estimated to be around 15.9 percent. An improvement can be achieved if also proton isotopes are included which rises the final portion to approximately 43.1 percent. The total difference is estimated to be 20.8 percent.

An additional test was also performed with the detector position at the angle 30 degrees. This measurement should mainly serve for the horizontal beam monitoring but it can be exploited also for this purposes. The results can be observed in the figure 3.24. Individual values are following: protons 72.9 percent, electrons 25.7 percent and others 1.4 percent. If a comparison with the FLUKA predicted composition is repeated, protons 83.4 percent, the difference is 10.5 percent.

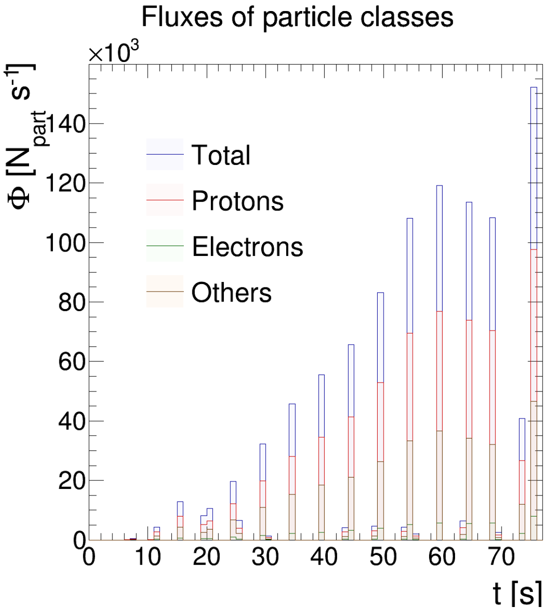


Figure 3.23: Fluxes of the particle classes as a fuction of time for the set-up with detectors at 30 degrees.

The relatively high differences between the simulation and estimated portion by the classification algorithm could be expected. The main cause can be found in the database which is mainly for a sensor with thickness 500 μm an applied bias of approximately 200V which is different from the above used settings. Additional source of uncertainty can originate from the simplified FLUKA simulation set-up but its significance should be low.

Therefore, it is proposed to perform additional referential measurements with this type of detectors settings or the more convenient way should be a development of the conversion model which was already mentioned in the section 2.2.

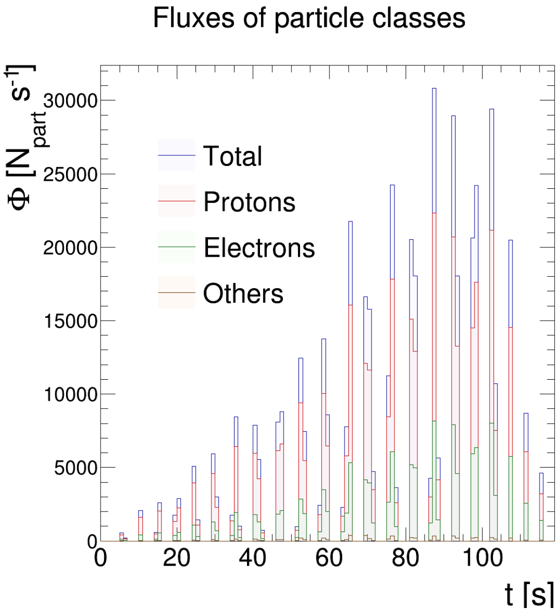


Figure 3.24: Fluxes of the particle classes as a function of time for the set-up with detectors at 0 degrees.

3.4 Discussion and Results Overview

The first part of this chapter is focused on the uncertainty of the secondary particle tracking. The work is mainly focused on the investigation of the horizontal back-projection method. In the case of the lateral monitoring, the used method would not benefit from the tracking uncertainty due to its more statistical approach. The difference between z coordinate obtained from the back-projection method and of the original primary vertex (where is secondary particle created) were examined for different detector angles, particle generation and specie with further focus on possible elemental contributors. The difference is expressed from a statical point of view with the mean value[21] and the standard deviation[22]. The back-projection method for some events reconstructs the z coordinate outside the cylinder which can be seen as impossible and therefore it was for some cases not considered.

The overall uncertainty for different angles shows that the bias but also the standard deviations are decreasing with higher angles where between 0 and 10 degrees case is main change mainly from the geometrical reasons. An important decrease is between the 20 and 30 degrees angle when both variables decline by almost 50 percent. The difference with including the additional condition (only projected PMMA events) causes large improvement of the tracking which is expected due to their obvious illogicality.

Dependencies on the on the particles generation and also their specie/type was also observed. In the case of the generations, the second generation exhibits the best tracking qualities except the case of 0 degree. An unexpected difference between third and forth generation emerges. An explained was offered based on the different particle composition with focus on the proton isotopes. The particle specie dependence revealed that heavier particle disturbs the tracking less with some exceptions. An extraordinary behaviour was observed comparing tritium, proton and helium samples. The heavier samples cause larger uncertainty than protons despite the features of the most likely interactions. This could be explained by lack of a sufficient statistics of the data samples for higher values of the angle α.

The dependence of the ΔZ (shift and the standard deviation) on the z back-projected coordinate reveals that the particle are most accurately tracked in the region before the beam Bragg peak. This can be attributed to their

---

[21] So-called shift or bias.

[22] In some cases when the bias is almost 0 than the standard deviation determines the full uncertainty

higher kinetic energy which mostly originates from the primary particle energy.

It can be concluded that the individual tracking based on the z back-projection can not be performed with a satisfying precision for all of the exploited angles. One of the possible contribution can be in the decision which angle should be used for the horizontal beam monitoring if for example cavities are also included. In this case the more favourable should be the 30 degrees due to their large tracking improvement in comparison to the 20 degrees case.

Following contributions were proposed: the beam width, the Multiple Coulomb scattering and the fragmentation process. The overall uncertainty can not be fully reproduced from the individual contributions. A reason can be found in the possible important correlations between the uncertainty elements or due to some more fundamental contributor which is therefore included in others (an overcounting is committed). This is substantiated by a similar behaviour of the shift and standard deviation in the case of the MCS and the fragmentation. Excepting these idea, the quantitative comparison can be made between individual contributions. The fragmentation process should cause the least increase of the uncertainty. The larger effect is generated by the MCS for larger angles higher than 20 degrees. The beam width is dominant for almost perpendicular positions of detectors to the beam which is mainly caused by geometrical reasons. It can be observed that the MCS possesses largest rise of the ΔZ as a function of the z back-projected coordinate in the region around the beam stopping position.

The second part comprehends the efforts towards the particle specie identification and filtering during the method for the lateral beam monitoring. The procedures exploited for this purpose are those from the chapter 2.2. In this part of the work, it was demonstrated that this uneasy task is feasible with proposed methodology. A limitation in the used settings of detectors lies constraints a which is not possible to overcome without a loss on the efficiency of the algorithms. A comparison between measurements and results supplied by the FLUKA simulation were done with following differences in the proton composition: for 0 degrees 20.8 percent including also proton isotopes and for 0 degrees 10.5 percent considering only protons. It is highly suggested to either develop a models to describe set-up conversions or to make additional measurements with focus on the desirable particle species with the currently used set-up.

Chapter 4

Conclusion

The goal of this thesis were to describe and develop a methodology for a measurement of mixed radiations field with semiconductor pixel detectors of the Timepix family. The application of proposed methods should be focused on the hadron therapy and the detection of the created secondary particles. This tasks were mostly achieved with some possibilities for improvements and continuing of the work.

The first chapter of this thesis summarizes the motivation for the exploitation of the Timepix family detectors for wide range of applications. The beginning of the chapter concentrates on the general features of the semiconductors and introduces the concept of their usage for registration of ionizing particles. This can be performed by the means of depleted PN junction together with an optimal readout structure. For this purposes, the Timepix detectors were developed with advanced features by the Medipix2 collaboration. The end of the chapter is dedicated to descriptions of two members of this family: the Timepix and the Timepix3 chips and recapitulates their utilization for the radiation field measurements.

The next chapter comprehends the problematic of the data processing after measurement with the Timepix detectors and focuses on application in the particle recognition. The data evaluation process includes two necessary steps: the clusterisation and the coincidence procedure in the case of the telescope architecture. In the first stage a raw pixel data of coordinates, energy and for the Timepix3 also time is converted to clusters which should mostly include information from one particle. During the creation of clusters, required features are calculated with respect to the final purpose. A several modified and new variables were introduced in the text from which the most important are the energy pixel variables and tuned computation of the cluster length. If needed, the coincidence procedure can be utilized to match events originating from one interacting particle in several synchronized detectors stacked on each other. The unsimilarity algorithm has been proposed to solve undesirable high portions of the multi-coincidence events.

The conclusion of the chapter is dedicated to the particle classification method resembling the modern machine learning kNN algorithm. Simplification and generality were the motivations during the method development. It is possible to implement the resulting algorithm on the microprocessor placed on the detectors providing an on-line particle recognition after an application of certain additional interface. The generality criterion was accomplished by the means of the database element which can comprehend any samples according to a given purpose. Therefore, adjustment and possibly also enhancement of the algorithm performance is achievable for any usage. The current database includes mainly data for the Timepix3 detector with 500 μm sensor thickness for protons in an energy range from 500 keV to 400 MeV and electrons from 100 keV to 7 MeV with additional heavy ions samples (helium, silicon and iron ions with certain energies) with a sufficient coverage of the elevation angle. It was demonstrated that the particle classification to protons, electrons and others can be performed with reasonable accuracy. An overall dependency on the used classes was observed and a possible improvement by the separation and class merging was proposed.

The application of this work and developed methodology lies in the ion carbon radiotherapy. The last chapter includes a discussion about the most recent approaches to the monitoring of the beam stopping position. The core of the work is the quantification of the uncertainty of the horizontal monitoring method with the simulations in the FLUKA tool. It was found that due to the medium effects on the particle propagation it is not possible to perform an accurate individual tracking. Three main causes of the uncertainty were suggested: the Multiple Coulomb scattering, the fragmentation process an the beam width. It has been shown that probably more general contributor is present which is reflected in all other suggested elements of the uncertainty. Despite this fact a quantitative comparison was offered. The overall results can be used for a selection of the most optimal position of detectors and as background knowledge for further investigations. The particle recognition algorithm was utilized for an improvement of the lateral monitoring via filtering of the proton component in the detected secondary particles. Unfortunately, this approach shows differences in a comparison with the results obtained in the FLUKA simulations. The cause can be found in the different settings of the detectors with respect to the database. A solution with a conversion model or additional measurements was suggested.

The future work will be dedicated to the database expansion and to the development of the conversion models between different detector settings and chip architectures. This should be sufficient for more accurate particle recognition within the hadron therapy. The result of all past, current and future efforts should be a creation of a general multi-purpose tool providing particle recognition and tracking. The goal will be achieved by the means of the machine learning algorithms combined with an extensive database and required models to describe the detector additional effects (the volcano effect etc.). A crucial role will also play Monte Carlo simulation tools which should supplement not performed measurements.

Bibliography

[1] G. Lutz. *Semiconductor Radiation Detectors: Device Physics*. Accelerator Physics Series. Springer Berlin Heidelberg, 2001.

[2] S. Ramo. Currents induced by electron motion. *Proceedings of the IRE*, 1939. `doi:10.1109/JRPROC.1939.228757`.

[3] H. Spieler. *Semiconductor Detector Systems*. Series on Semiconductor Science and Technology. OUP Oxford, 2005.

[4] M. Tanabashi, K. Hagiwara, K. Hikasa, et al. Review of particle physics. *Phys. Rev. D*, 2018. URL: `https://link.aps.org/doi/10.1103/PhysRevD.98.030001`.

[5] X. Llopart, R. Ballabriga, M. Campbell, et al. Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements. *Nucl. Instrum. Meth. A*, 2007. `doi:10.1016/j.nima.2007.08.079`.

[6] T. Poikela, J. Plosila, T. Westerlund, et al. Timepix3: a 65k channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout. *Journal of Instrumentation*, 2014. `doi:10.1088/1748-0221/9/05/c05013`.

[7] C. Granja, P. Krist, D. Chvatil, et al. Energy loss and online directional track visualization of fast electrons with the pixel detector timepix. *Radiation Measurements*, 2013. `doi:https://doi.org/10.1016/j.radmeas.2013.07.006`.

[8] D. Turecek, J. Jakubek, P. Soukup, et al. Usb 3.0 readout and time-walk correction method for timepix3 detector. *Journal of Instrumentation*, 2016. `doi:10.1088/1748-0221/11/12/C12065`.

[9] R. Ballabriga, M. Campbell, and X. Llopart. Asic developments for radiation imaging applications: The medipix and timepix family. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2017. `doi:10.1016/j.nima.2017.07.029`.

[10] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth. A*, 1997. `doi:10.1016/S0168-9002(97)00048-X`.

[11] J. Jakubek. Precise energy calibration of pixel detector working in time-over-threshold mode. *Nucl. Instrum. Meth. A*, 2011. `doi:10.1016/j.nima.2010.06.183`.

[12] C. Granja, J. Jakubek, U. Koester, and otehrs. Response of the pixel detector timepix to heavy ions. *Nuclear Instruments and Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment - NUCL INSTRUM METH PHYS RES A*, 2011. `doi:10.1016/j.nima.2010.06.166`.

[13] B. Bergmann, M. Pichotka, S. Pospisil, et al. 3D track reconstruction capability of a silicon hybrid active pixel detector. *Eur. Phys. J. C*, 2017. `doi:10.1140/epjc/s10052-017-4993-4`.

[14] T. Holy, E. Heijne, J. Jakubek, et al. Pattern recognition of tracks induced by individual quanta of ionizing radiation in medipix2 silicon detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2008. URL: `http://www.sciencedirect.com/science/article/pii/S0168900208004592`, `doi:https://doi.org/10.1016/j.nima.2008.03.074`.

[15] S Hoang, R Vilalta, L Pinsky, et al. Data analysis of tracks of heavy ion particles in timepix detector. *Journal of Physics: Conference Series*, 2014. `doi:10.1088/1742-6596/523/1/012026`.

[16] R. Vilalta, S. Kuchibhotla, S. Hoang, et al. Classification of sources of ionizing radiation in space missions: A machine learning approach. *Journal of the European Space Agency, Acta Futura 5*, 2012. `doi:10.2420/AF05.2012.111`.

[17] J. Bouchami, A. Gutiérrez, T. Holy, et al. Measurement of pattern recognition efficiency of tracks generated by ionizing radiation in a medipix2 device. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2011. `doi:https://doi.org/10.1016/j.nima.2010.06.163`.

[18] S Hoang, L Pinsky, R Vilalta, et al. LET estimation of heavy ion particles based on a timepix-based si detector. *Journal of Physics: Conference Series*, 2012. `doi:10.1088/1742-6596/396/2/022023`.

[19] C. Granja, J. Jakubek, S. Polansky, et al. Resolving power of pixel detector timepix for wide-range electron, proton and ion detection. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2018. `doi:https://doi.org/10.1016/j.nima.2018.08.014`.

[20] P. Mánek, B. Bergmann, P. Burian, et al. Randomized computer vision approaches for pattern recognition in timepix and timepix3 detectors, 2019. `arXiv:1911.02367`.

[21] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 1967.

[22] Miroslav Kubat. *An Introduction to Machine Learning*. Springer Publishing Company, Incorporated, 2015.

[23] J. Jakubek D. Turecek. Pixet software package tool for control, readout and online display of pixel detectors medipix/timepix. 2015.

[24] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 2003. `doi:10.1016/S0168-9002(03)01368-8`.

[25] R. R. Wilson. Radiological use of fast protons. *Radiology*, 1946. `doi:10.1148/47.5.487`.

[26] JERRY B. MARION. Chapter 6 - the effects of nuclear radiations. In *Energy in Perspective*, pages 140 – 169. Academic Press, 1974. `doi:https://doi.org/10.1016/B978-0-12-472275-0.50009-9`.

[27] D. Schardt, T. Elsässer, and D. Schulz-Ertner. Heavy-ion tumor therapy: Physical and radiobiological benefits. *Rev. Mod. Phys.*, 2010. `doi:10.1103/RevModPhys.82.383`.

[28] Dauvergne D., Battaglia M., Montarou G., et al. New methods of real-time control imaging for ion therapy. 2009. URL: `http://hal.in2p3.fr/in2p3-00363382`.

[29] Amaldi U., Hajdas W., Iliescu S., et al. Advanced quality assurance for cnao. *Nuclear Instruments and Methods in Physics Research*, 2010. `doi:10.1016/j.nima.2009.06.087`.

[30] M. Martišíková, J. Jakubek, C. Granja, et al. Measurement of secondary radiation during ion beam therapy with the pixel detector timepix. *Journal of Instrumentation*, 2011. `doi:10.1088/1748-0221/6/11/c11014`.

[31] J. Jakubek, C. Granja, B. Hartmann, et al. Imaging with secondary radiation in hadron therapy beams with the 3d sensitive voxel detector. 2011. `doi:10.1109/NSSMIC.2011.6153862`.

[32] P Henriquet, E. Testa, M Chevallier, et al. Interaction vertex imaging (ivi) for carbon ion therapy monitoring: A feasibility study. *Physics in medicine and biology*, 2012. `doi:10.1088/0031-9155/57/14/4655`.

[33] K Gwosch, B Hartmann, J. Jakubek, et al. Non-invasive monitoring of therapeutic carbon ion beams in a homogeneous phantom by tracking of secondary ions. *Physics in medicine and biology*, 2013. `doi:10.1088/0031-9155/58/11/3755`.

[34] R Félix-Bautista, T Gehrke, L Ghesquière-Diérickx, et al. Experimental verification of a non-invasive method to monitor the lateral pencil beam position in an anthropomorphic phantom for carbon-ion radiotherapy. *Physics in Medicine & Biology*, 2019. `doi:10.1088/1361-6560/ab2ca3`.

[35] K. C. Gwosch. Non-invasive monitoring of carbon ion beam therapy by track- ing of secondary ions – an initial study. Master's thesis, University of Heidelberg, 2012.

[36] H. A. Bethe. Molière's theory of multiple scattering. *Phys. Rev.*, 1953. `doi:10.1103/PhysRev.89.1256`.

[37] V. L. Highland. Some practical remarks on multiple scattering. *Nuclear Instruments and Methods*, 1975. `doi:https://doi.org/10.1016/0029-554X(75)90743-0`.

[38] R. Serber. Nuclear reactions at high energies. *Phys. Rev.*, Dec 1947. `doi:10.1103/PhysRev.72.1114`.

[39] Gaimard J. J. and K. H. Schmidt. A reexamination of the abrasion-ablation model for the description of the nuclear fragmentation reaction. *Nuclear Physics A*, 1991. `doi:https://doi.org/10.1016/0375-9474(91)90748-U`.

[40] K. Gunzert-Marx, H. Iwase, D. Schardt, et al. Secondary beam fragments produced by 200 MeV u-112c ions in water and their dose contributions in carbon ion radiotherapy. *New Journal of Physics*, 2008. `doi:10.1088/1367-2630/10/7/075003`.

[41] T. Gaa, M. Reinhart, B. Hartmann, et al. Visualization of air and metal inhomogeneities in phantoms irradiated by carbon ion beams using prompt secondary ions. *Physica Medica*, 2017. `doi:https://doi.org/10.1016/j.ejmp.2017.05.055`.

[42] A. Ferrari, P. Sala, A. Fasso, et al. Fluka: a multi-particle transport code. *CERN Yellow report*, 2005. `doi:10.2172/877507`.

[43] T.T. Böhlen, F. Cerutti, M.P.W. Chin, et al. The fluka code: Developments and challenges for high energy and medical applications. *Nuclear Data Sheets*, 2014. `doi:https://doi.org/10.1016/j.nds.2014.07.049`.