UNIVERSITÀ DEGLI STUDI DI PADOVA

Physics and Astronomy Department "Galileo Galilei"

Master's Degree in Physics of Data

Master's Thesis

Advanced tracking techniques for Active Target Time Projection Chamber detectors in Nuclear Physics experiments

Supervisor

Dr. Tommaso Marchi

Co-supervisor

Dr. Thomas Roger

Candidate

Lorenzo Domenichetti

Academic Year 2021/2022

Advanced tracking techniques for Active Target Time Projection Chamber detectors in Nuclear Physics experiments

Lorenzo Domenichetti

Abstract

Tracking detectors for low-energy nuclear physics experiments, acting at the same time as reaction targets, are very promising devices in a wide range of research topics. The ability to measure and reconstruct the trajectory of all of the reaction products with high efficiency and good geometrical resolution allows particle spectroscopy studies to be performed under experimental conditions below the sensitivity threshold of standard techniques. This is possible only if solid, efficient, and computationally fast reconstruction codes are implemented and validated. In the presented thesis project, reconstruction codes and classification techniques have been developed aimed at processing experimental data from the ACTAR Active Target. The purpose of such codes was to provide accurate information about track geometry, particle energy, and identification. In particular, the reaction ^{20}O(d,^{3}He)^{19}N* has been first analyzed using the Hough transform and RANSAC algorithms, comparing their efficiencies. In the second part, the results obtained applying machine learning techniques on the same data will be presented, with the aim of achieving a fast event classification employing for the first time on ACTAR these cutting-edge techniques.

Contents

Chapter 1

Studying nuclear physics extremes with an Active Target

1.1 Nuclear Physics basics

The atomic nucleus is composed of protons and neutrons, called nucleons. Protons have a positive charge, neutrons are neutral particles.

In the description of a nucleus, the mass number (A), defined as the sum of the number of protons (Z) and neutrons (N), is of special relevance. This number provides a first approximation of the binding energy of the nucleus and of its radius. Nucleons are bound to each other by the nuclear strong force, a short-distance attractive force, which, however, exhibits repulsive behavior when the nucleon-nucleon distance is further decreased. In addition, the repulsive Coulomb force is also present between the protons. The balance between these two forces is the pillar on which the stability of the nuclei is based. Unbalanced forces, on the other hand, give rise to unstable nuclei. Those have finite half-lives and might decay in many different ways until the newly formed configuration reaches a stable one. Each chemical element is characterized by a specific number of protons, while its number of neutrons can vary. Nuclei having the same charge and different masses are called isotopes.

The origin of the study of nuclear physics as we understand it today goes back to Rutherford's famous scattering experiment in 1911 [1]. In this experiment, the unexpected angular distribution of the scattered particles led to a new conception of the structure of atoms, consisting of a central nucleus and orbiting electrons, on which theories are still based today. Since then, nuclear physicists have studied the structures and properties of different nuclei. A useful tool to visualize the landscape of nuclei is the table of nuclides, which, showing on the vertical axis the number of protons and on the horizontal axis the number of neutrons, represents the whole landscape of nuclear physics, from stable to exotic nuclei. In the table of nuclides in Figure 1.1, the stable isotopes are represented in black; the observed unstable isotopes are colored in yellow, while in green are the predicted nuclei that have not yet been observed.

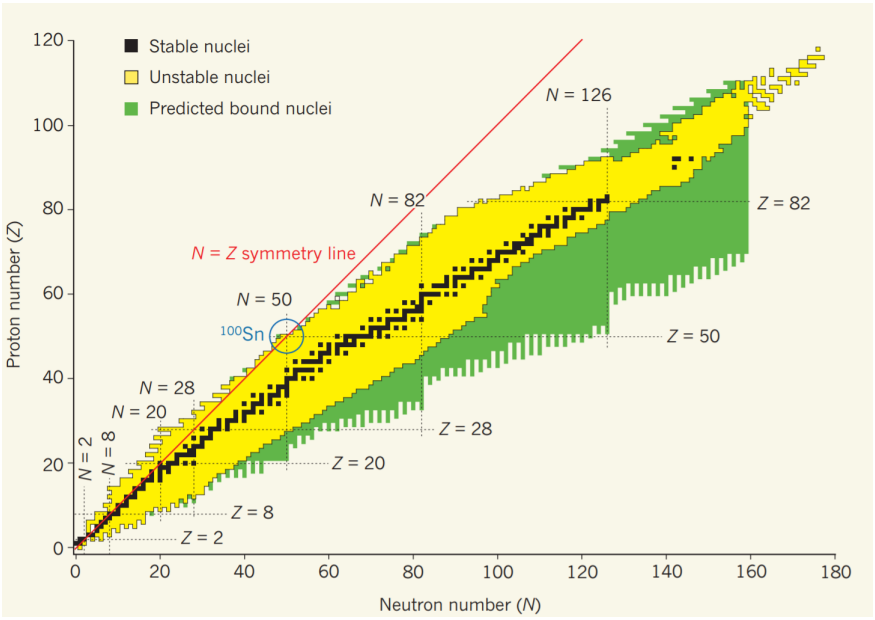


Figure 1.1: Nuclear Chart. Figure from [2].

The isotopes are represented on the same horizontal line. What defines the stability of an isotope is the number of neutrons. Indeed, protons tend to repulse themselves through the Coulomb force as they are positively charged. However, the presence of neutrons brings only an attractive nuclear force that balances electrostatic repulsion. When these two forces are balanced, the nucleus is said to be stable. A peculiarity of nuclides tables is the deviation of the so-called *stability valley*, the imaginary line where the stable nuclei lie, from the N=Z line. This deviation is related to the fact that a higher number of neutrons is needed to balance the larger Coulomb contribution related to an increase in the number of protons. Unstable nuclei have an N/Z ratio that deviates from the stable nuclei one. Such nuclei might decay in different ways, for example, by emitting an α particle (α decay), an electron (or a positron) and an antineutrino (or a neutrino) in a process called β^- (β^+) decay, or by emitting photons (γ decay). Nuclei far from stability might undergo different decay steps on their way to the stability valley. As a general rule, it can be said that the farther the nucleus is from the stability valley, the shorter its lifetime. The boundaries of the nuclear chart are defined by the proton and neutron drip lines, namely the imaginary lines where the energy needed to separate a proton or a neutron becomes negative. As shown in Figure 1.1, the proton drip line is partially observed, while the neutron-rich one is largely unknown. [3]

To study nuclei far from stability, complex apparata have been and continue to be developed to produce and accelerate beams of unstable ions. Radioactive ion beams (RIBs) are usually produced by exploiting a primary reaction and selecting a portion of the ejected particles. Many laboratories are already capable of producing these species, such as ISOLDE at CERN, GSI in Germany, GANIL in France, RIKEN in Japan or FRIB in the United States. Moreover, new-generation accelerators are also being built in Italy, such as the SPES at the Legnaro National Laboratories (LNL). The nuclear reactions induced by such ion beams on selected targets are then studied using state-of-the-art detectors.

1.2 Nuclear Structure

"Like many systems governed by the laws of quantum mechanics, the nucleus is a somewhat enigmatic and mysterious object whose properties are much more difficult to characterize than are those of macroscopic objects" [3]. The theoretical description of a nucleus can be performed with different degrees of accuracy. To some extent, the description can involve just a few parameters: electric charge, radius, mass, binding energy, angular momentum and parity, and energy of the excited states. These are the *static* properties of the nuclei. A completely different chapter in the description of nuclei regards their *dynamical* properties, including the decay and reaction probability of different species.

An appropriate description of the nuclei provides a good understanding of the nuclear force, the binding agent that gives them stability. As previously mentioned, at a short distance the nuclear force overcomes the electrostatic repulsion of protons and is stronger. However, for larger distances, this force is negligible. At a deeper level, finer characteristics arise: the nuclear strong force is almost independent of the type of nucleon (whether proton or neutron); the force depends on whether the spins of the nucleons are parallel or antiparallel; the nuclear force includes a very short-range repulsive force that keeps the nucleons apart.

The use of effective models allows a fair description of the observables in several regions of the nuclear chart. At the same time, the improvement in computing power has allowed the simulation of many system properties starting from basic calculations, thus starting from the nucleon-nucleon interaction itself (*ab-initio* theories). Today, the frontier of nuclear physics models is the study of nuclei properties based on theories well grounded in the fundamentals of quantum chromodynamics [4].

However, in trying to extend this approach to more complicated systems, difficulties arise from the technical mathematical challenges of solving many-body problems. Even using simple interaction potentials, the sets of equations that describe the interaction of A nucleons do not have an analytical solution. The second difficulty arises from the evidence that the nuclear force also includes three-body components, characteristics that do not have a classical analogue. The problem of describing the nuclei has to be simplified, and theories have to be designed balancing the physical details and the mathematical complications. Such simplifications aim at building a *nuclear model*, a simplified view of the nuclear structure that contains essential information on nuclear physics but at the same time reproduces some basic characteristics over a considerable number of nuclei. Such models are generally improved by adding second-order terms, which allow predictions to be closer to the experimentally measured data.

1.2.1 Liquid Drop Model

Historically, one of the first approaches that has tried to describe the nuclear structure is the liquid drop model (LDM), born from the analogies between the laws of physics that govern liquid drops and nuclei. In fact, as molecules in fluids, the nucleons feel a short-range force, have almost a constant *density* independent of the nucleus dimension, and show a binding energy per nucleon that is about to be constant varying the mass number. The LDM is

a *collective* model aiming at describing the binding energy (B) of a nucleus starting from collective properties. The formula that describes this quantity is called *semi-empirical mass formula* and is based on empirical parameters.

$$B = a_V A - a_S A^{2/3} - a_C \frac{Z^2}{A^{1/3}} - a_A \frac{(N-Z)^2}{A} \pm \frac{a_P}{A^{1/2}} \tag{1.1}$$

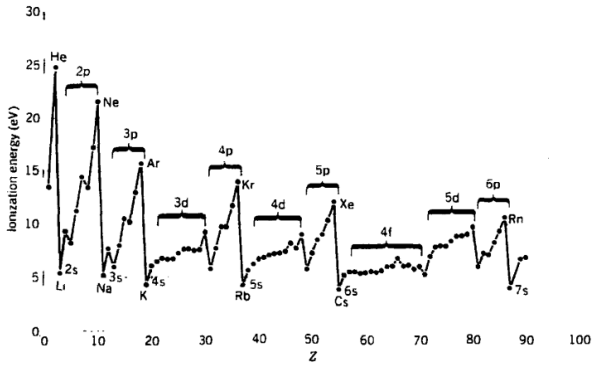
The formula depends on five parameters a_i, which are the empirical parameters tuned to resemble the experimental data. The first term is the only positive one and is proportional to the number of nucleons present in the nucleus. The second term accounts for the missing interaction felt by the nuclei at the surface with respect to those in the middle of the nucleus. The third term accounts for the Coulomb repulsion of the protons, while the fourth is the asymmetry term, whose explanation is not properly classic and is related to the asymmetry between neutrons and protons. The last term is the pairing one and is due to the fact that nuclei having 'coupled' nucleons tend to be more stable. This term is positive for even-even nuclei, negative for odd-odd nuclei, and null otherwise.

This model has had great success in describing global properties of the nucleus, such as, for example, the β decay instability of nuclear species. However, the lack of any microscopic contribution prevents the model from having a precise description and from reproducing nuclear excited states. Such properties can only be predicted considering internal degrees of freedom, as was tried in microscopic models. Among the first microscopic models that have been formulated is the *nuclear shell model*.

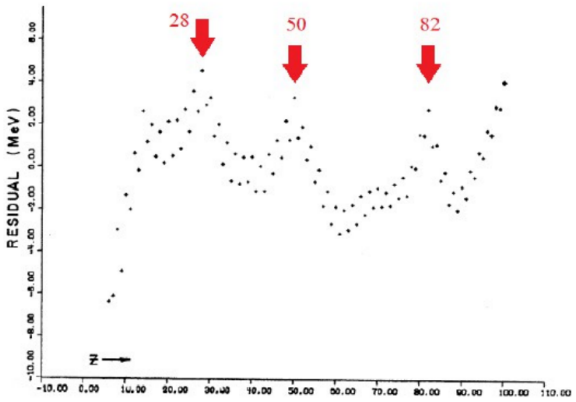
1.2.2 Nuclear shell model

Since the atomic theory based on the shell model gave remarkable insight into the structure of atoms, nuclear physicists tried to use a similar theory to tackle the problem of nuclear structure. However, one of the main differences between the nuclear shell model and the atomic one is the fact that in the former case the potential in which nucleons move is produced by the nucleons themselves. The Coulomb potential in which electrons move is generated by the nuclear core, and the interaction between electrons only accounts for second-order interactions. However, the hypothesis of a nuclear shell model is based on some evidence linking these two different systems. Indeed, considering the atomic shell model, it is clear how the properties of an atom, such as radius and ionization energy, change when a subshell is filled and an electron is added to the next shell. The existence of the nuclear shells can then be motivated by analogy considering the deviation of the proton and neutron separation energies compared to the semi-empirical mass formula, which considers no shell effect.

The similarities between the two figures in Fig. 1.2 is clear. Additional evidence of the existence of nuclear shells is provided by neutron capture cross sections and the scheme of excitation energies of the first excited state in even-even nuclei. The change in the trends of these observables occurs at the same proton or neutron numbers (2, 8, 20, 28, 50, 82, and 126). Nuclei that have a number of neutrons or protons belonging to this set generally show increased stability with respect to their neighbors. Thus, such numbers are called *magic numbers*, because of such effects [3].



(a) Ionization energy varying atomic number [3].



(b) Difference in binding energy varying proton number.

Figure 1.2: Comparison between the effects of nuclear and atomic shells [5].

The idea of basing the description of nuclei on a mean-field potential is due to Goeppert-Mayer, Wigner, Jensen, and Haxel [6,7]. Their theory is based on quantum mechanics, where systems are described by their wave function ϕ. The wave functions are calculated by solving the Schrödinger equation $H\phi = E\phi$, where E is the eigenvalue associated with ϕ and can be seen as the energy of the system, and H is the Hamiltonian of the system, including both the kinetic term (T) and the potential term (U). A non-relativistic form of the Hamiltonian can be written as:

$$H = T + U = \sum_{i=1}^{A} \frac{\boldsymbol{p}_i^2}{2m_i} + \sum_{i>j=1}^{A} V_{ij}\left(\boldsymbol{r}_i - \boldsymbol{r}_j\right) \tag{1.2}$$

In Eq. 1.2, the first term describes the kinetic energy, starting from the momentum and mass of the A nucleons, while the second term represents the interaction between nucleons i and j. Higher-order contributions, such as three-body forces, can also be included in the potential term. There is no fully explanatory potential V_{ij} and therefore it is not possible to calculate ϕ or E. The assumption that can be made to simplify the problem is that every nucleon moves in a potential generated by all other nucleons. Such types of approach are generally referred to as *mean field* approaches. In such frames, the Hamiltonian can be rewritten as follows.

$$H = \sum_{i=1}^{A}\left(\frac{\boldsymbol{p}_i^2}{2m_i} + U_i\left(\boldsymbol{r}\right)\right) + \sum_{i>j=1}^{A} V_{ij}\left(\boldsymbol{r}_i - \boldsymbol{r}_j\right) - \sum_{i=1}^{A} U_i\left(\boldsymbol{r}\right) = H_0 + H_{res} \tag{1.3}$$

The first term, H_0, is the mean field Hamiltonian and $U(r)$ the mean field potential, dependent only on the coordinates of a single nucleon. H_{res}, on the other hand, can be considered a correction. The smaller this correction, the more accurate the mean-field approximation.

Using mean-field approaches, such a term is often neglected, as will be done in the following. The main contribution of Goeppert-Meyer and her collaborators was the writing of the central potential based on the harmonic oscillator one.

$$U(r) = \frac{m\omega^2 \boldsymbol{r}^2}{2} + D(\boldsymbol{l}^2 - \langle N|\boldsymbol{l}^2|N\rangle) + C\boldsymbol{l}\cdot\boldsymbol{s} \tag{1.4}$$

The first term represents indeed the harmonic oscillator potential, with a characteristic frequency ω, the second term provides a correction to the potential related to the angular momentum ($\boldsymbol{l}$), and the last term represents the spin-orbit interaction term, considering the interaction between the spin and orbital angular momenta of the nucleons. Using such potential, the Schrödinger equation $H_0 E = \phi E$ can be solved. The allowed energy values turn out to be discrete. An example of the energy scheme is presented in Figure 1.3.

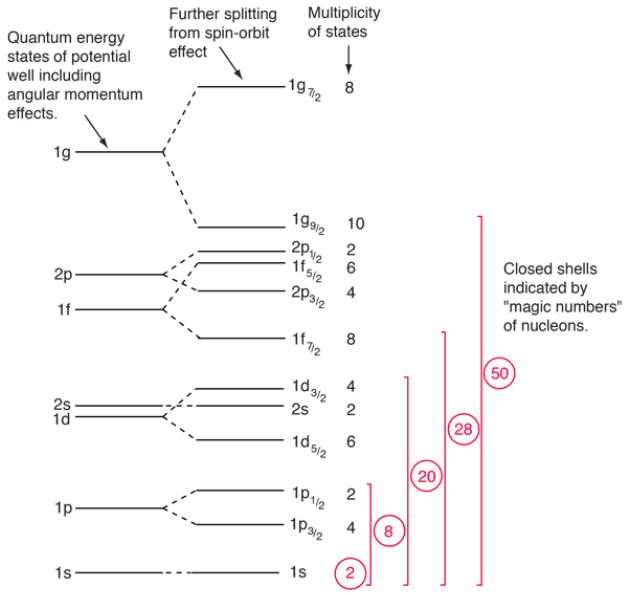


Figure 1.3: Shell model energy scheme.

Energy levels are usually labeled using the Nl_j notation, which describes the state energy, its orbital angular momentum $\boldsymbol{l}$ and the total angular momentum $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s}$. The degeneracy of the energy levels is broken, as shown in figure 1.3, by both the orbital angular momentum and the spin orbit term in the potential. However, each of these levels presents a degeneration of $2j+1$, due to possible orientations of the total angular momentum. Figure 1.3 shows that energetic levels are placed in a configuration such that some levels are grouped and others are divided by large energy gaps. These denser areas are called *shells*, and the nuclei that fill all these substructures show a more stable behavior; shell closures are indeed associated with the *magic* numbers seen experimentally. It is also worth noticing that the splitting of the levels due to the spin orbit coupling term is fundamental to achieve an agreement between theory and experiment. Even if accurate, the description using the harmonic oscillator potential might not be the most precise. Another often used potential is the Woods-Saxon potential,

which has the form [8]:

$$U_{WS}(r) = \frac{V_0}{1 + exp\left[(r - R)/a\right)]} + C\frac{1}{r}\frac{\partial V_{SO}(r)}{\partial r}\boldsymbol{l}\cdot\boldsymbol{s} \tag{1.5}$$

In this equation, V_0 represents the potential depth, R the nuclear radius and a the diffuseness parameter. Moreover, the first radial term already includes the effects of the $\boldsymbol{l}^2$ term, which in this case does not need to be added.

The contribution of H_{res} might not be negligible in cases where the shape of the nucleus deviates from the spherical assumption. In general, nuclei that are close to the shell closures present a spherical shape, while nuclei in the middle of a shell, having both proton and neutron number far away from a closed-shell configuration, show regular excitation modes, but different from the ones predicted by the shell model. These effects are often related to the *collective* behavior of the nucleons. A sign of the presence of collective features is the reduction in energy of the first 2^+ excited state, far below the energy needed to break a 0^+ coupled pair, which is particularly evident in the mass region $Z > 50$, $N < 82$. In describing midshell nuclei, the shell model, even considering the residual interactions, might not be an appropriate starting point for describing the coherent motion of valence nucleons. Configurations of nucleons that fill only partially the shells are tackled more effectively by models dedicated to predicting the *vibrational* and *rotational* properties of the nuclei [9]. However, the shell model is still a useful tool to describe the influence of the deformed shape on nuclear single-particle motion.

1.2.3 Nuclear models for deformed nuclei

A clear advance in understanding the motion of a single particle in deformed nuclei was made using the Nilsson potential [10]. Nilsson's contribution was the writing of a shell model potential that took into account the possible spatial deformations of nuclei. In the laboratory and in Cartesian coordinates, the Nilsson potential can be described as in Equation 1.6.

$$U(x, y, z) = \frac{m}{2}\left[\omega_\perp^2(x^2 + y^2) + \omega_z^2 z^2\right] + D\boldsymbol{l}^2 + C\boldsymbol{l}\cdot\boldsymbol{s} \tag{1.6}$$

Here, the parameters ω_i represent the stiffness in two different directions, the XY plane and the Z axis, chosen as the direction in which the nucleus shows deformation. Both parameters depend on the deformation parameter β, defined as follows in Eq. 1.7 [3]:

$$\beta = \frac{4}{3}\sqrt{\frac{\pi}{5}}\frac{\Delta R}{R_{av}} \tag{1.7}$$

Here ΔR is the difference in the semimajor and semiminor axes obtained when considering the elliptical section of a deformed nucleus. Another parameter often used to describe deformed nuclei is *elongation* ϵ, related to β by Eq. 1.8.

$$.\epsilon = 3\left(\frac{4\pi}{5}\right)^{-1/2}\beta \tag{1.8}$$

When the parameter ϵ is positive, the nucleus is said to be *prolate*, while when it is negative, the nucleus is said to be *oblate*. Prolate shapes are those where the deformed axis is stretched with respect to the non-deformed ones (as a rugby ball), while oblate-shaped nuclei have a deformed axis that is shrunk with respect to the other two (as a disk). The choice of such potential leads to a deformation of the level scheme, as reported in Figure 1.4

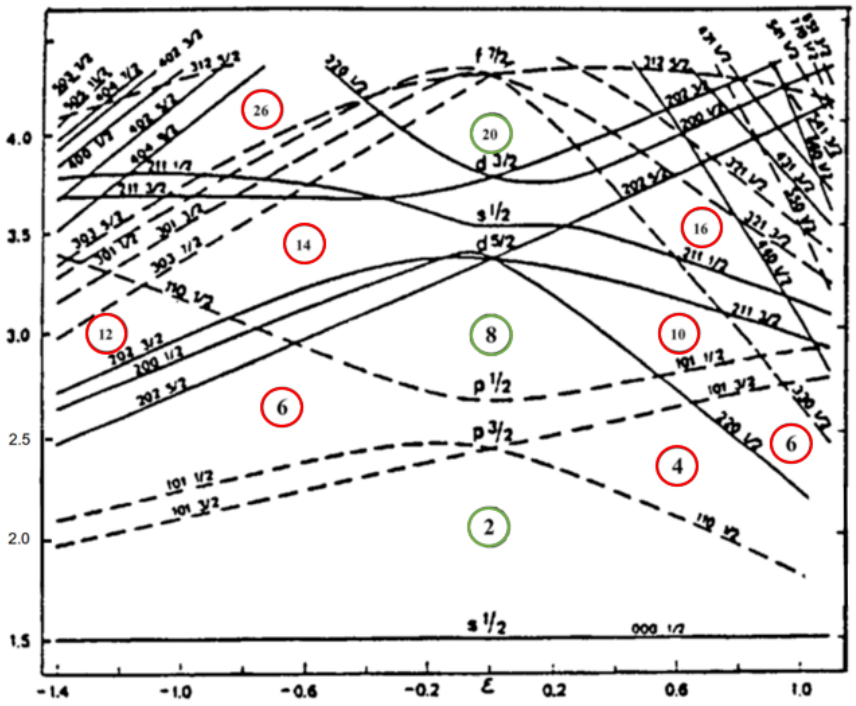


Figure 1.4: Level scheme for deformed nuclei, varying the parameter ϵ. Positive-parity states are shown in full lines, and dashed ones represent negative-parity states. The circles indicate the predicted magic numbers, since there are large gaps between the energy levels. Adapted from [11].

To achieve an appropriate description of the energy scheme under deformation, states are labeled using additional quantum numbers. These numbers are Ω and Λ. The former can be interpreted as the projection of the angular momentum $\boldsymbol{j}$ on the symmetry axis, while Λ has the same interpretation, but for the orbital angular momentum $\boldsymbol{l}$. Finally, the states of the Nilsson energy scheme are represented as $Nn_z\Lambda\Omega$. One of the most interesting consequences of the deviation of the energy levels is shown in Figure 1.4 is the appearance of new magic numbers for deformed nuclei. In fact, introducing the effects of elongation ϵ leads to the creation of new shell gaps, as marked in the red circles.

The prediction of enhanced gaps for nuclei with neutron or proton numbers equal to 6, 14, or 26 leads to the possible presence of new magic numbers. Therefore, studying the evolution of the characteristics of the nuclei along isotopic or isotonic chains might shed light on the underlying structure of these predicted new magic configurations.

1.3 Nuclear observables

1.3.1 Nuclear reactions

One of the objectives of nuclear theories is to provide predictions on quantities that can be measured, providing a comparison between calculated and experimental data. One of the most widely used tools to extract information about nuclei and their reciprocal interaction is the study of nuclear reactions. Theories try to predict observable quantities such as the energy, spin, and parity of populated states states or the reaction cross section. Nuclear reactions can occur when an energetic flux of particles collides onto a target, which can be solid, liquid, or gaseous. Such interactions are usually labeled through the formulas in Eq. 1.9.

$$a + X \rightarrow b + Y \quad or \quad X(a,b)Y \tag{1.9}$$

where a is the projectile, X is the target, usually fixed in the laboratory frame, b is the light ejectile, and Y is the heavy recoil. The right-hand way of indicating nuclear reactions is particularly convenient because it allows one to refer to general classes of reactions that share common properties, such as, for example, (α, n) or (n, γ) reactions. Nuclear reactions show different phenomena when the parameters of the reaction are changed. The two parameters that mostly affect the outcome of the reactions are the energy of the incoming beam and the *impact parameter*, the distance from the beam direction and the center of the scattering nucleus. Reactions with energies lower than 20 MeV/A can be considered low-energy reactions, those in the range between 20 and 100 MeV/A are intermediate-energy reactions, while higher energies categorize reactions as high-energy ones. In the present work, intermediate-energy reactions are of interest. Another categorization of nuclear reactions can be made considering the relative weight of the projectile and the target: reactions in which the projectile is lighter are said to be *direct kinematics* reactions, whereas reactions in which the projectile is heavier are said to be in *inverse kinematics*. Although the description of the interaction in the center-of-mass (CoM) frame does not change, these different types of reactions have large effects in the laboratory frame, where momentum and energy conservation laws often impose rigid restrictions on the trajectories of the ejectiles. Even if a theoretical description of the reaction might be simpler in the CoM frame, the implication of measuring particles in the laboratory frame has always to be taken into account. At intermediate energies, different processes can occur when two nuclei collide. If the incident and outgoing particles are the same, the process is called *scattering*. Depending on the energy and impact parameter, the nuclei can interact through *elastic scattering*, when both Y and b are in their ground state, or through *inelastic scattering*, where, for example, a nucleon can be excited to a specific state. Another interesting reaction mechanism is *transfer*, where one or more nucleons are transferred from the projectile to the target or vice versa. This mechanism is of interest for the set of analyzed data. The distinguishing feature of this type of reaction is that it can populate well-defined low excitation energies quantum states with a high degree of selectivity. Simpler transfer reactions are generally termed *quasi elastic*, whereas *inelastic transfer* reactions generally proceed through a sequence of direct transfers and internal rearrangements [12] [3].

1.3.2 Observables

Among the most widely used and measured observables in nuclear physics is the energy of the particles, which can be measured with great resolution. In nuclear reactions also the emission angle of the emitted particles is of particular interest, as it is necessary to observe the production angular distribution (usually relative to the beam direction). The *differential cross section* can then be interpreted as the probability of obtaining a particle b at a specific energy at an angle (θ, ϕ). Integrating the differential cross section over all angles yields the total cross section, the total probability of measuring a particle b of a specific energy. Integrating once more over all possible energies yields the *absolute* total cross section, the probability of forming the nucleus Y in the reaction. Another interesting observable that can be measured is the γ radiation yield, which comes from the decay of the excited states of Y. Measurement of the lighter particle b helps select reactions in which Y has been produced, reducing the background. Comparing only events where Y is present in the output channel is very useful, as it often leads to a cleaner energy spectrum and overall better accuracy in the measurement of the populated states. Moreover, the angular distribution of the emitted γ rays helps in retrieving information on the properties of the populated states, especially in deducing their spin and parity.

When analyzing the outcome of nuclear reactions, conservation laws help relating measured quantities to properties of the involved nuclei. In particular, in inelastic scattering reactions, exploiting energy and momentum conservation allows one to measure excited states in the nucleus Y by measuring the sole particle b. The energy missing from the incoming beam defines the excitation energy. Moreover, measuring the angular distribution of the outgoing particle brings information on the angular momentum it carries. Thus, the assignment of the angular momentum of the states populated in the Y nucleus can follow. Finally, in a wide range of energies in nuclear physics reactions the conservation of both proton and neutron numbers can be assumed, as well as the conservation of parity, which once again helps assigning the unknown parities to the excited states [3]. Radiation measurement is at the core of experimental nuclear and particle physics. While designing a detector, one has to take into account the different ways in which different types of radiation interact with materials. Indeed, there is a large difference in the interactions with materials of charged particles (electrons and ions) and non-charged particles (photons and neutrons). Moreover, the same type of radiation may have different kinds of interaction, mainly depending on the energy of the radiation itself and on the material crossed. Nuclear physics is, in general, involved in the measurement of all kinds of radiation, although the measurement of heavy ions is key in most nuclear physics experiments.

1.4 Interaction of heavy ions with materials

1.4.1 Energy loss of charged particles

Charged particles, such as α particles, interact with matter primarily through Coulomb forces between their positive charge and the negative charge of electrons in the absorber material. Interactions between nuclei might also occur (as in Rutherford scattering), but

they play a minor role in the design of a detector given the low probability of these reactions. After entering any absorbing medium, the charged particle starts to interact at the same time with many electrons, given the long range of the Coulomb force. Depending on the relative distance between the ions and the electrons, the transferred energy can be enough to raise the electron to a higher state (*excitation*) or completely remove it from the atom (*ionization*). The energy transferred is taken from the kinetic energy of the incoming particle, and given the large difference in mass between the ion and the electrons, many interactions take place before the ion stops. Moreover, because of the mass difference, the trajectory of the ions tends to be quite straight, since the ion momentum does not change much after every encounter and interactions occur in any direction simultaneously. Charged particles are generally characterized by their *range* that is the distance they can travel within the material before being stopped. The products of the encounters in the absorber medium are excited atoms and ion-electron pairs, which have a natural tendency to recombine to form neutral atoms. In many types of detector their recombination wants to be suppressed as these free electrons are the basis of the detector response.

The quantity that defines the energy loss of a particle through the material is the *stopping power*, defined as the differential energy loss divided by the corresponding differential path length. In a semi-classical assumption, the stopping power can be approximated by the Bethe formula:

$$S = -\frac{dE}{dx} = \frac{4\pi e^4 z^2}{m_0 v^2} NZ \left[\ln\left(\frac{2m_0 v^2}{I}\right) - \ln\left(1 - \frac{v^2}{c^2}\right) - \frac{v^2}{c^2} \right] \tag{1.10}$$

where v and z are the velocity and charge of the primary particle, N and Z are the number density and the atomic number of absorber atoms, m_0 is the rest mass of the electrons and e is the electric charge. I represents the average ionization and excitation potential of the absorber medium and is often evaluated experimentally. For non-relativistic particles, only the first term in squared brackets is relevant, and it varies slowly with particle energy. Therefore, the general trend of dE/dx follows the proportionality of $1/v^2$ [13]. The trend of the energy loss profile for different charged particles is shown in Figure 1.5.

Plotting the specific energy loss along the track of a charged particle yields the so-called *Bragg curve*, an example of which is shown in Figure 1.6.

As shown in the Bragg curve plot, the energy loss of the α particle is constant for most of the track and increases when the energy decreases. At the end of the curve, when electron pickup starts, the energy loss sharply decreases, and the curve falls off. The plot shows the Bragg curve for both a single α particle and a beam of particles, where the effects of *straggling* are taken into account. The energy loss process takes place through microscopic interactions that the incoming particle undergoes in a somehow random way. Such processes then introduce stochasticity in the energy loss and, therefore, in the path of the particles. In general, every beam of monoenergetic particles after traversing a certain thickness of material will have an energy distribution whose width is related to the stochasticity of the interactions.

1.4.2 Target thickness in nuclear reactions

In nuclear physics experiments that involve reactions, especially at low and intermediate

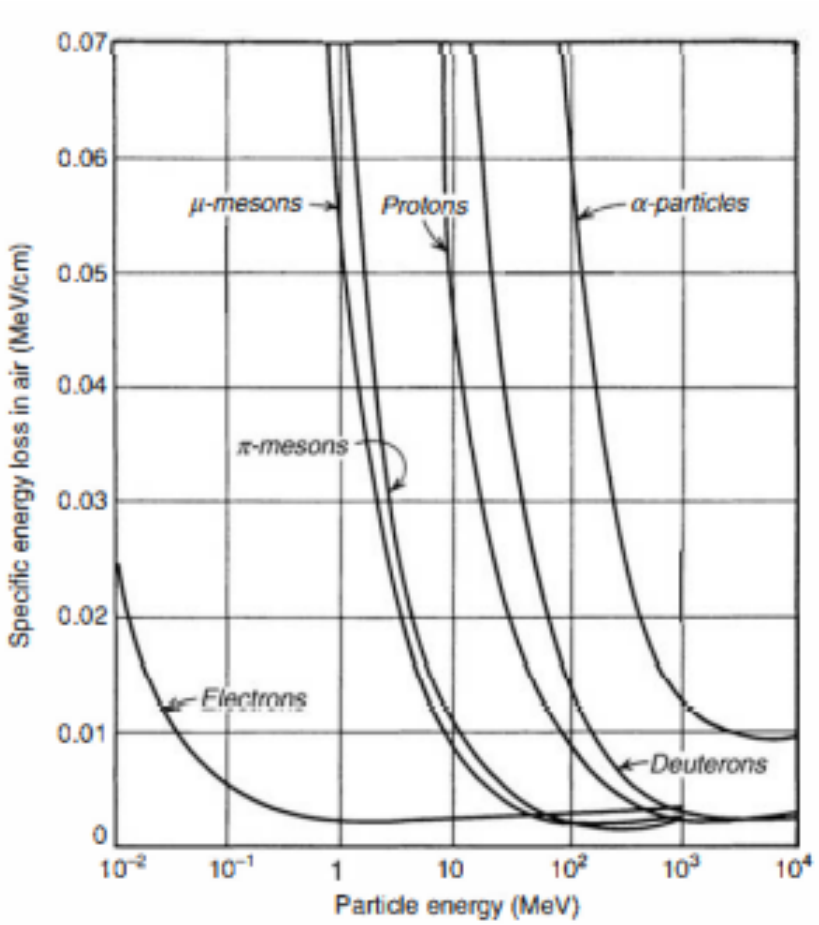


Figure 1.5: Energy loss trends of different particles in air [13].

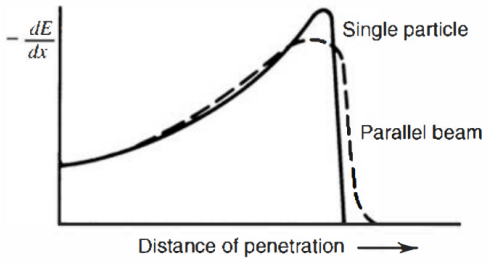


Figure 1.6: Specific energy loss along an α particle track [13].

energies, the thickness of the target is of paramount importance. On the one hand, a thick target guarantees a large presence of scattering centers, which increases the reaction rate. However, such types of targets reduce the energy resolution of the ejectiles. In fact, both the projectile and the ejectiles continuously release energy in the target material; the error caused by the energy straggling can then compromise the resolution needed for spectroscopic measurements. Moreover, a too thick target would also impede the coincident measurement of all the ejectiles, a technique that usually helps reducing the background. The use of thin

targets may be limiting in experiments in which the particle beam rate is not high enough. This case is often faced in experiments involving radioactive ion beams, which typically have rates of orders of magnitude lower than stable ones. Among the tools developed to increase the reaction rate while avoiding loss of resolution are detectors capable of measuring the interaction point within the target material [14]. Another solution is provided by the *active target* detectors [15], the detection principle of which will be explained in Section 1.6.

1.5 Example of physics cases with stable and unstable beams

The past decade has seen a rapid increase in the number of available radioactive beams. Such beams provide opportunities for studying properties of nuclei far from the stability and probing structure models in regions where unexpected features show up.

1.5.1 Clustering in nuclei

Nuclei have been known to exhibit a cluster structure [16], especially in the light-mass region. Strong efforts have been made to progress in understanding the nature of clustering both theoretically and experimentally [17]. Although much experimental evidence has been accumulated over the past decades, there is still a lack of data for unstable nuclei, both close and far from the stability valley. Thus, this information can be collected using radioactive ion beams. Reaction channels that have been probed for stable nuclei can also be considered for unstable ones. As an example, reactions involving α particles in the entrance or exit channel can be used to investigate the structure of the α cluster.
It is well known that the nucleus ^{8}Be is unbound for breakup into two α particles but it has a very long lifetime. As a consequence it has a well defined cluster structure. Other nuclei might contain three-cluster structures in which the subgroups can rearrange in a triangular or a linear combination [11]. These structures have been predicted, for example, in ^{14}C and ^{14}O [18]. Experiments were carried out to determine resonances above the α emission threshold in the ^{14}C chain looking for the α-chain states. Higher-mass nuclei, such as ^{12}C, could also be seen as consisting of integer numbers of clusters of α, given the great binding energy of α particles. In particular, the ^{12}C $+\alpha$ cluster structure has long been a topic of interest because of the possibility that low-energy resonance with such structures can dramatically enhance the ^{12}C$(\alpha,\gamma)^{16}$O reaction, which is considered one of the most important reactions of nuclear astrophysics as it determines the carbon-oxygen ratio of the universe [19]. Increasing the number of αs leads to ^{20}Ne, but the presence of more nucleons allows many more degrees of freedom and the situation is more complicated overall [20].

1.5.2 Z = 6 magicity

The nuclear shell structure, based on the independent motion of nucleons in an average field, became famous for explaining the existence of magic numbers. Although the introduction of the spin-orbit coupling term is necessary to explain the presence of such numbers, its origins are still an open question. As stated in her Nobel Prize lecture, Goeppert-Mayer clarified

that magic numbers arise from two different series of numbers. The first one - 2,8,20 - is attributed to the harmonic oscillator, while the second one - 28,50,82, and 126 - is due to the spin orbit (SO) coupling force [21]. However, the introduction of the spin-orbit interaction is still not fully justified from a microscopical point of view. The difficulties of explaining the structure of atomic nuclei from ab initio nuclear theories with two- or three-nucleon forces got in the way of an accurate description. Theoretical studies of SO splitting of single-particle states $1p_{1/2}$ and $1p_{3/2}$ in ^{15}N have suggested a possible influence of two-body SO and tensor forces, and the discovery of a prevalent SO-type magic number 6 is expected to offer new opportunities to understand the origin of such a force [22]. Already in her Nobel lecture, Goeppert-Mayer cited 6 and 14 as the two lowest numbers belonging to the set of magic numbers generated by the SO, admitting, nevertheless, that the energy gap between the states $1p_{1/2}$ and $1p_{3/2}$ due to the SO interaction is very small. Experimental results have also pointed out the possible presence of a magic number 6 in some semi-magic nuclei, such as ^{8}He, ^{14}O, and ^{14}C. Among these elements, it is of paramount importance to understand whether the closure of the subshell in the $Z = 6$ isotopes is driven by the closure of the shell $N = 8$ or whether it is also present in other carbon isotopes [23]. Many factors corroborate the presence of a $Z = 6$ shell closure, starting from the high excitation energy of the first 2^+ excited state to the small values of $B(E2)$ found for neutron-rich carbon isotopes (^{16}C and ^{18}C) [23, 24]. However, considering ^{14}C as a starting point, the reduction in spin-orbit splitting could also be due to a modified proton-neutron interaction as the filling of the $1d_{5/2}$ state begins. In ^{14}C, five $\boldsymbol{l} = 1$ transitions have been measured [24]. The ground state configuration contains the $1p_{1/2}$ state, and the first 7MeV transition has been linked to the spin orbit partner $1p_{3/2}$. However, no measurement of the spin orbit gap is present in the $Z = 6$ region after ^{14}C. A widening or reduction of the gap between $1p_{1/2}$ and $1p_{3/2}$ as neutrons are added to $1d_{5/2}$ could have crucial implications in the structure of neutron-rich carbon isotopes. Moreover, to probe the magicity of $Z = 6$, single-particle proton states in carbon neutron-rich isotopes have to be precisely measured. Such measurements will shed light on the strength of the spin orbit interaction in low-mass nuclei, helping to understand the deep motivations that lie behind it.

The study of such topics often requires the use of radioactive ion beams. To be effective, the experiments require the use of detectors specifically designed for the task. A very versatile type of detectors, which is being widely employed with radioactive ion beams, are active targets. Those detectors show very flexible properties and can be used to measure a wide range of processes with high efficiency.

1.6 Active target detectors

Active targets are detectors in which the target itself acts as a tracking device for the reaction products. The energy released in the material traversed by the projectiles and ejectiles is used in these type of detectors to measure and track them. This detection technique provides several advantages, such as increased angular coverage and increased effective target thickness. Such characteristics are of particular interest when working with low-intensity beams, such as those of ions far from the stability valley. Another advantage of active targets is the possibility of using gaseous or liquid targets, which are very helpful in inverse kinematic

reactions. Indeed, as previously stated, an extremely useful tool for precisely measuring nuclear states through particle spectroscopy is transfer reactions. However, when studying exotic species, such reactions must take place in *inverse kinematics*, using the heavy fragment as a projectile and the light elements (H, He) as targets, which under standard conditions are not solid. Recent studies have led to the use of gaseous *time projection chambers (TPC)* as active targets, allowing a three-dimensional reconstruction of reactions. An example of a gaseous active target detector is the ACTAR active target, whose data are the basis of this thesis work [25].

1.6.1 Working principle of gaseous detectors

As described in Section 1.4.1, the energy loss process of heavy ions leads to the creation of ion-electron pairs. In a gaseous detector, such free charges are those that create an electronic signal. To avoid recombination, the presence of an electric field sends the free charges to the collection electrodes. Usually, the choice made in gaseous detectors is to read the signal produced by the electrons, which have a higher drift velocity in the gas itself [13]. Moreover, in gas detectors, the number of electrons produced by ionization is typically increased by orders of magnitude in a region appropriately designed. The amplification of charges produced by primary ionization can be achieved by using strong electric fields in a gas volume, in a process called *avalanche multiplication*.

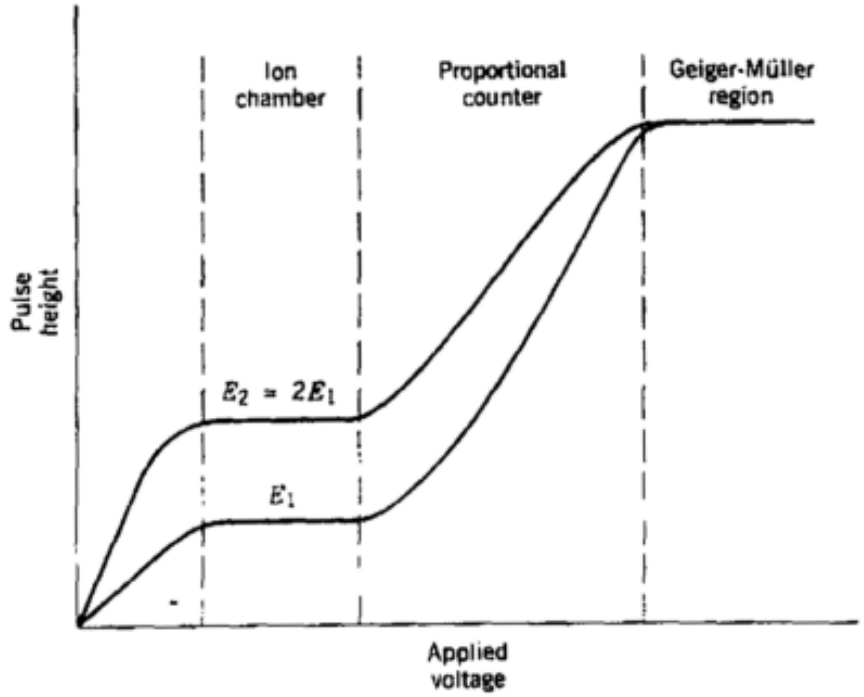


Figure 1.7: Electrons behaviors in gas [3].

Figure 1.7 shows that electrons coming from primary ionization can be amplified by different factors that vary with the applied electric field. In the *ionization chamber* regime, the gain is unitary. For stronger electric fields, one obtains the *proportional* regime, in which

the amplitude of the output signal is still proportional to the initial energy released but increases exponentially with the applied voltage (the Y-axis in Figure 1.7 is in log-scale). This region allows for charge amplification, as the ratio between the freed charges and the signals amplitude is constant. For higher electric fields, one obtains the *discharge* region, where the pulse height is constant regardless of the initial charge released. Usually, in a drift chamber, two zones can be distinguished: a first zone sensitive to incoming radiation, which operates in the ionization chamber regime, and a second zone operating in the proportional regime dedicated to charge amplification. The first region aims to collect the ions produced toward the amplification zone without modifying the number of charge carriers during the drift. The mixture and conditions of the gas are usually chosen based on their density, stopping power properties, dielectric properties, and finally taking into account the drift velocity of electrons and ions [26] [27]. The movement of charges (both electrons and ions) within the drift region induces a signal on the collection electrodes, which can be estimated from the Shockley-Ramo theorem [28]. The temporal duration of this signal is usually dominated by ion drift, since its drift velocity is about 100 times slower than that of electron drift [27]. Therefore, it is useful to introduce a tool that excludes the long tail of the signals induced by the ions. This instrument is called *Frisch grid*, and is a grid placed in close proximity to the collection anode. This grid screens the signal produced by every charge that is not within the region between the grid and the collection anode. In Figure 1.7, one can also notice a region where the output signal is weak, since in this region electrons undergo different effects that prevent charge collection, such as recombination.

Chapter 2

The ACTAR project

The detector of interest for this thesis work is an active target-time projection chamber, a detector that allows for three-dimensional reconstruction of the trajectories of the incoming and outgoing particles. The working principle of such detectors is shown in Figure 2.1. The segmented pad plane allows for a two-dimensional reconstruction of the tracks, exploiting the geometric information resulting from the independent reading of the pixels. Furthermore, it is possible to reconstruct the distance of the track from the collection plane by measuring the arrival time of the signal.

2.1 Detector

The ACtive TARget Time Projection Chamber (*ACTAR TPC*) is a detector developed in a collaborative European project whose goal was to construct and perform experiments with this new TPC detector, taking advantage of cutting-edge technologies. ACTAR is a gas-filled detection system designed to study the nuclear properties of exotic nuclei using rare isotope beams. The main feature of this detector are his large volume and high 2D-reconstruction granularity, which makes it well-suited for the study of nuclear reactions using low-intensity beams. The ACTAR TPC detector, shown schematically in Figure 2.2, consists of a chamber whose internal skeleton measures $606 \times 606 \times 335\, mm^3$. The gas input and output are placed at opposite corners of the chamber to ensure a good homogeneity of the gas within the whole chamber. All flanges around the structure are made of a thick $1.5\, cm$ aluminum layer, to support up to $1\, bar$ pressure difference with the environment, avoiding deformation.

The front flange is equipped with a cylindrical nose with a diameter of 80 mm to place the $6\, \mu m$ Mylar window that separates the gas volume from the beam line as close to the wire field cage as possible. This design reduces the distance between the entrance window and the active volume to only 60 mm, with the aim of reducing energy loss and straggling before the active region. The bottom and top flanges are used to support the field cage and provide the appropriate connections to the segmented pad plane. Side flanges can be used, depending on the physical case of interest, to hold ancillary detectors. The active volume is surrounded by a $295 \times 295 \times 255\, mm^2$ double wire field cage [30], which provides

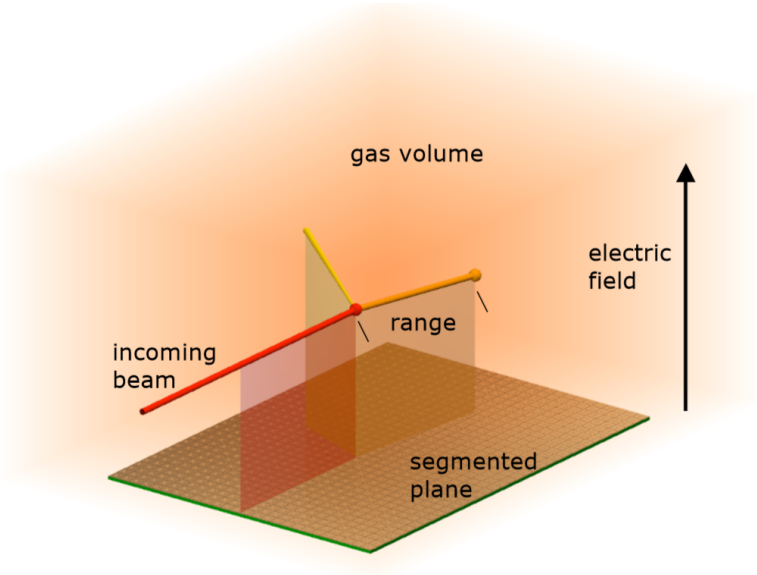


Figure 2.1: Working principle of an Active Target TPC [15].

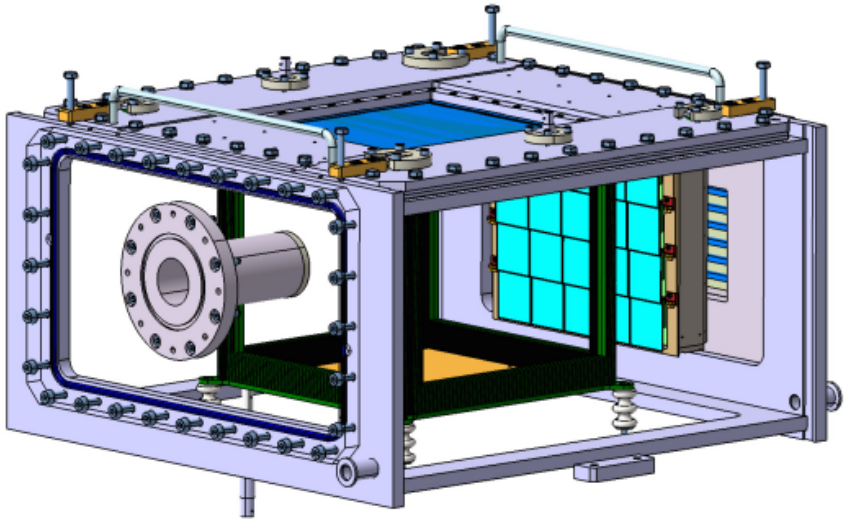


Figure 2.2: 3D scheme of the ACTAR TPC [29].

the electric field that transports the electrons produced by ionization to the collection plane. The field cage is attached to the cathode and is fed through a $20\,kV$ high-voltage supply. The field cage consists of $20\,\mu m$ diameter wires with a $1\,mm$ spacing in the inner plane and a $2\,mm$ spacing in the outer plane, connected through $4.7\,M\Omega$ resistors. As tested in the development phase, the use of a two-plane field cage increases the homogeneity of the electric field, preventing interference from external fields. The segmentation of the anode into 128×128 $2 \times 2\,mm^2$ pads provides excellent geometric resolution. However, the high density of the pads (25 channels/cm^2, 16384 total pads) represented a challenge in the design

of the mechanics and of the pad connections. Furthermore, such a plane had to be designed to withstand up to 1 *bar* pressure difference with respect to the environment with minimal deformations. Typical experiments are carried out with pressures ranging from a few tens of *mbar* to one *bar*. The anode construction started from a metal core printed circuit board (PCB), with a direct connection through the circuit from the pads to a connector with a pitch of $2\,mm$ [31]. Charge amplification is obtained directly in the active volume of the gas, placing a *MICROMEGAS* (described in Section 2.2.2) amplification system on top of the pad plane [29]. The 16384 pads are connected to the readout electronics through Kapton flex spark protection circuits (ZAP) to the GET system, described in Section 2.3 [29]. A first step for the final construction of the ACTAR TPC was the construction of two demonstrators [30], with the explicit objective of testing the critical points of the design, evaluating the feasibility of the mechanical construction, and characterizing the response of the detector. These detectors were on a 1:8 scale with respect to the final TPC, having a pad plane segmented into 2048 (64×32) channels exploiting the same geometry and dimension as those installed in the final detector ($2 \times 2\,mm^2$). The demonstrator used the same GET electronics and the same charge amplification system as was implemented in the final detector.

2.2 Charge amplification

To measure the energy deposited, the charge produced by the ions inside the active volume is collected in the pad plane. However, to have a proper electronic signal as the output, one needs to amplify the charge collected by orders of magnitude to have a relevant electronic signal.

2.2.1 Avalanche multiplication

In gaseous detectors, the amplification of the collected charge is often obtained by a process called *avalanche multiplication*, where the number of charge carriers increases by orders of magnitude. In the region dedicated to amplification, the electric field is tuned to obtain a proportional regime, where the amplitude of the signal is proportional to the number of primary ion-electron pairs produced. A simulated avalanche produced by a single electron is represented in Figure 2.3.

The avalanche process starts when an electron feels an electric field such that its energy overcomes the ionization threshold of the gas constituents. In figure 2.3, one can notice the increase in the number of free charge carriers once the primary electron reaches the avalanche area; the large black points in the scheme represent the grid that separated the two areas. After traversing the grid, the number of electrons grows rapidly: every free electron in the avalanche area is quickly accelerated until it reaches the ionization threshold. Once the process starts, the number of electrons (and ions) produced follows an exponential increase curve. Having a sufficient number of free charges allows the signal induced on the collection anode to be distinguished from the electronic noise.

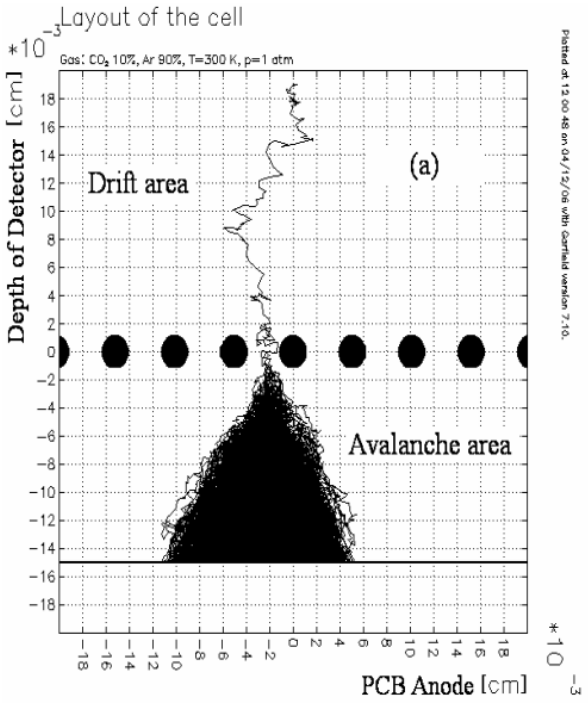


Figure 2.3: Bi-dimensional trajectory of an electron in a drift and in an amplification zone, where the avalanche takes place [26]. Simulated through [32].

2.2.2 MICROMEGAS

In the active target TPC, the charges produced in the active volume are amplified through a *MICROMEGAS* (MICROMesh Gaseous Structures), whose working principle is shown in Figure 2.4.

This system consists of a thin grid (few μm), with a spacing between the holes of a few tens of μm. Given its dimensions, such a grid is often referred to as *micromesh*. This grid lies on some insulating pillars, which hold it about a hundred micrometers on top of the collection anode. The potential between the cathode and the mesh is set to cause the electrons to drift, and the reduced electric field is often on the order of $1kV/(cm\,torr)$, while the anode voltage is tuned so that the region between the anode itself and the mesh operates in a proportional regime. The difference in the reduced electric field between the drift and avalanche areas allows for a transmission of electrons almost free from losses. In ACTAR, the distance between the anode and the grid is chosen to be $220\,\mu m$ [34]. Although such a distance is usually on the order of $100\,\mu m$, this distance has been chosen since low-pressure conditions would lead to a considerably small gain. Increasing the available area for amplification reduces this effect. However, since amplification is highly dependent on the distance between the mesh and the anode, this distance must be as constant as possible over the entire collection area. Moreover, the exponential nature of the avalanche process makes a few percent change in the distance impact for tens of percents in terms of amplification produced. Thus, the homogeneity of the distance, gas conditions, and gas constituents must be kept as constant

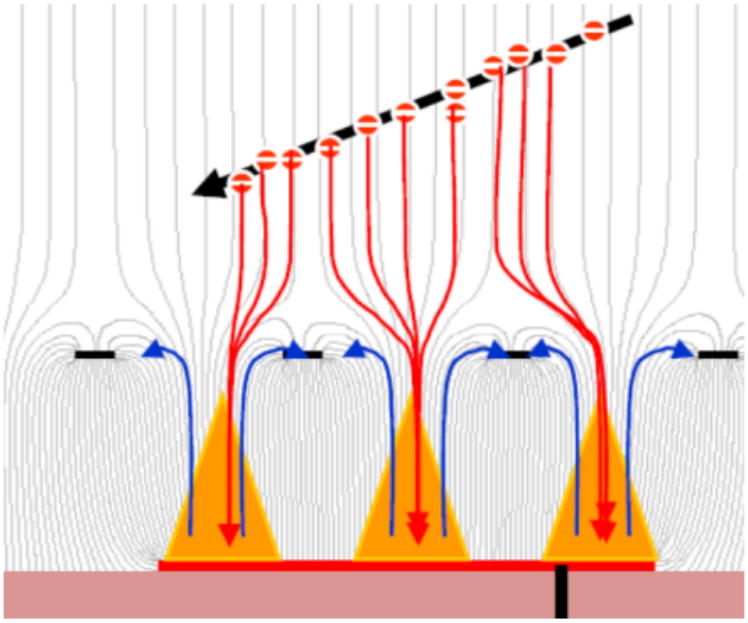


Figure 2.4: Charge carrier path into a MICROMEGAS [33].

and stable as possible to achieve a reliable amplification. In addition, the mesh acts as a Frisch grid, shielding the anode from the signals induced by the charges present in the gas volume, especially the slow ions produced in the ionization process. An other important feature of this amplification method is that, given the strong difference in the electric field generated in the drift and amplification regions, the ions produced in the avalanche process do not return to the active region, but are captured in the mesh itself, as shown in Figure 2.4. This feature is essential to avoid contamination by free charges inside the active region, which could lead to a modification of the ionization threshold, the homogeneity of the electric field and the overall electron collection efficiency.

2.3 GET system

The GET system (General Electronics for TPCs) [35], developed specifically for time projection chambers, is the system used by ACTAR for signal sampling and storage. The development of the project has been carried out by different institutions based in France (CEA, GANIL, CENBG) and the United States (NSCL). Even if the system has been designed for nuclear physics experimental setups, such an acquisition scheme can be applied to a broad spectrum of experiments and allows for the management of up to 30000 independent channels. The basic scheme of the GET system is reported in Figure 2.5 As shown in Figure 2.5, the front-end part consists of ASICs (Applied Specific Integrated Circuits), each containing 64 channels. These have a specific design and are called ASICs for GET (AGET). Each AGET has, for each channel, a preamplifier with variable gain, an analog filter, a threshold discriminator, and an SCA (Switch Capacitor Array), composed of a circular buffer array of 512 cells, where the signal coming from the shaping stage is continuously sampled and stored, with a frequency in the 1 to 100 MHz range. Different trigger settings can be set,

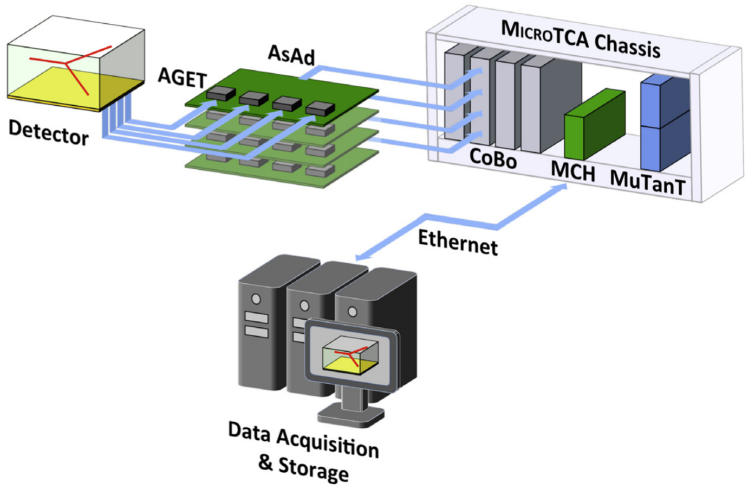


Figure 2.5: Scheme of the elements in the GET system [35].

and the acquisition phase starts once a trigger signal is recorded. Once reached, a threshold discriminator is used to test whether a channel is hit using its threshold. The outcome of such a procedure is binary, positive if the channel is hit, and null otherwise. In any case, when a trigger signal is launched, the acquisition of the signals from the AGETs is stopped and the content of the SCA is saved and digitized. The digitization of the signals can involve every channel (full readout) or only those that got hit (partial readout). Moreover, the AGETs host four more FPN (Fixed Pattern Noise) channels, which only sample noise and allow for better noise reduction. Four AGETs are hosted on an AsAd board (ASIC support and analog-to-digital converter), which also includes an FPGA and a 4-channel ADC converter. Thus, every AsAd manages 256 channels, plus 16 dedicated to noise, for a total of 272 channels. The ZAP connectors act as a link between the detector and the AsAds and consist of flexible Kapton connections that allow to manage the high channel density. Up to four AsAds can then be linked to a CoBo (Concentration Board), designed to be compatible with the μTCA standards (Micro Telecommunications Computing Architecture). Each CoBo can then read up to 1024 channels, and therefore 16 CoBos are needed for the pad plane channels of ACTAR. The main role of these boards is to act as a medium between the user and the electronics front-end, as they are both able to read and save the data coming from the AsAd, and receive signals from the user and set parameters on the AGET boards. An additional board (MuTanT) is mounted, together with the CoBos, on the μTCA chassis and takes care of the alignment of the timestamp, of the event number, and the CoBo sampling frequencies. Furthermore, the MuTanT board also sets the trigger, which can come from an external signal (L0), or it can be set as a threshold in the multiplicity of channels hit (L1), or it can be personalized (L3) [35].

2.4 Ancillary detectors

As shown in the right part of Figure 2.2, ACTAR is often coupled with auxiliary detectors, able to measure charged particles that escape the active volume or different types of radiation,

such as γ rays. In the former case, the presence of such detectors is particularly useful in achieving particle identification using the $\Delta E - E$ technique [25]. In fact, also particles escaping leave a track inside the active region, providing the partial energy loss. If the emitted fragment stops in the auxiliary detector, the total energy can be measured. Even if many ancillaries are hit at the same time, fitting the track and knowing the position of the side detector leads to a unique identification. Provided that tracks can be seen on the pad plane, the granularity of the auxiliary detectors is not a primary factor to consider, provided the area is small enough to limit the chance of double hits. However, an important factor to take into account is the energy resolution of the ancillary detectors, which should at least match that of the gas chamber.

2.5 Preprocessing algorithms

The acquisition of data is managed by the *NARVAL* system, which takes care of data collection and merging from all AGETs. The data are then stored in MFM format, a framework developed at GANIL, and acquired events are monitored online through a GRU application (Ganil ROOT Utilities). Figure 2.6 shows the shape of a typical signal recorded.

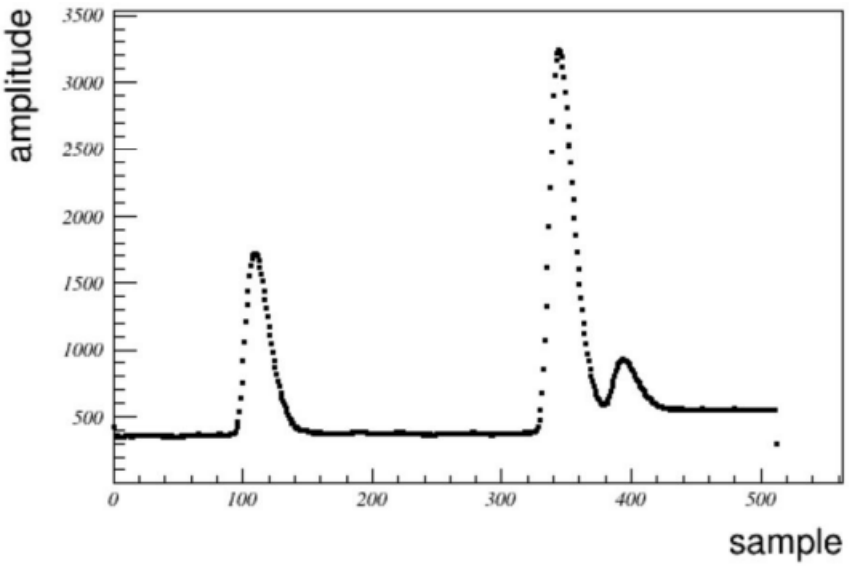


Figure 2.6: A signal obtained through the GET electronics. Every pulse corresponds at the arrival of charges in a given moment. Superposition of many pulses is related to spatial pile-up [36].

To retrieve the information from such signals, the height and timestamp of the maxima must be found. The processing of digitized signals is generally referred to as Pulse Shape Analysis (PSA). Since different detectors are related to the same GET acquisition, different PSAs must be used to achieve proper accuracy. In the case of a gas detector such as the ACTAR TPC, the charge collected by creating the input signal corresponds to the energy loss of the charged particles in the active volume. As shown in Figure 2.7, the shape of the

collected signal may vary depending on the angle of incidence of the particle with respect to the collection plane. Moreover, the time sampling of the signal should, in principle, allow for the distinction of the contributions of different tracks to the total signal, as shown once more in Figure 2.7 [37].

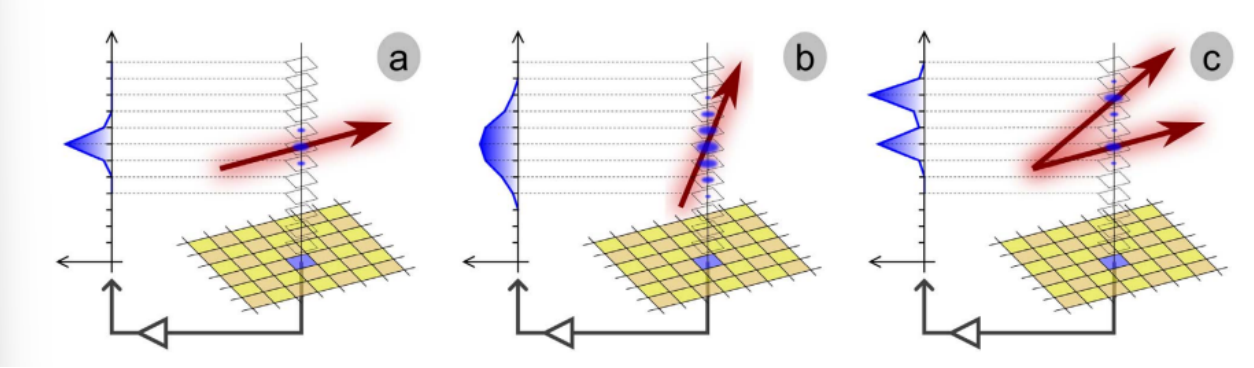


Figure 2.7: Variation of the shape of the signal depending on the impinging angle of the particle (*a and b)*) and on the presence of multiple tracks (*c*)

The measured output sample $s_{out}(t)$ results from the convolution of the input signal $i_{in}(t)$ and the electronic response function for the channel $h(t)$. If the response function is known, it is possible to deduce the shape of the input signal from the output sample. To deal with the convolution and deconvolution of the signals recovered from the GET electronics, an extensive study has been carried out in the Fourier space [37]. In particular, the analytic response function given by the design turned out to be only a rough approximation, and empirical response functions had to be derived from experimental measurements. Then, the signal could be reconstructed starting from the estimated response function. In order to improve the resolution, the signal is filtered with a low-pass filter, smoothing the signal and attenuating quick variations. In the end, the reconstructed input signal $i_{rec}(t)$ can be estimated from the output sample $s_{mes}(t)$. Applying the deconvolution on the Fourier transforms in the frequency space, one gets Eq. 2.1.

$$I_{rec}(f) = \frac{S_{mes}(f)}{H(f)} \cdot \Phi(f) \tag{2.1}$$

As shown in the paper, the method proposed in [37] allows reconstructing the effective input signal, which can then be used to analyze the shape of the signal. The intrinsic time resolution found is on the order of 5 sampling steps. Moreover, the event-by-event fit of the reconstructed signal showed a precision of about 1%, achieved with a simple filtering and fit function, composed of the convolution of a Gaussian and a square pulse for an input squared pulse.

Such a preprocessing procedure, including the baseline subtraction averaging the four FPN channel levels, the reconstruction of the signal through the Fast Fourier Transform, and the fit of the signals to retrieve the total charge and the timestamp of the maximum, is routinely applied to the recorded signal. The output of the analysis of the shape of a signal is called

a *hit*. A hit is an object that stores the three-dimensional position of the charge deposit and a value related to the charge deposit itself. Given the segmentation of the pad plane, the planar information can be retrieved by knowing which pad has been hit. The last spatial dimension can be estimated from the arrival time of the signal, knowing the electron drift velocity, and the charge deposited is estimated measuring the amplitude of the signal. The charge deposited on each pad could, in principle, be calibrated to achieve an energy deposit measurement. Given the finite resolution of the detector, the reconstructed points can be represented in space as *voxels*, which in this application can be considered boxes whose dimensions are defined by the spatial and time resolution of the device. For the sake of the following thesis work, calibrations have not been considered, as they would not have brought any advantage to the analysis run. The output of the analysis described previously is stored in *ReducedEvents*, `C++` objects that contain lists of hits related to a single acquisition.

Chapter 3

The experiment at GANIL

The analysis of the presented thesis work is based on the data collected during that experimental campaign held in March 2022 at GANIL [38]. As stated previously, the reasons for studying the fundamentals of nuclear force today are many, and the experiment presented aims at achieving a deeper understanding of the spin-orbit effect.

3.1 Experiment proposal

The aim of the proposed experiment was to study the evolution of the supposed $Z = 6$ shell gap through the carbon chain. To fully understand the structure of neutron-rich carbon isotopes, the experiment aimed to measure the spin orbit splitting between the $1p_{1/2} - 1p_{3/2}$ states. Such states were proposed to be measured using the d(^{19}N, ^{3}He)^{18}C transfer. This transfer reaction was chosen to selectively populate the proton single-particle states in ^{18}C. So far, the knowledge of the level structure of ^{18}C has been limited to only spectroscopic measurements [39], lifetime [40, 41], or neutron removal measurements [42], while more selective transfer channels have not yet been measured with sufficient precision. The measurement of a transfer reaction could also provide information on the parities and spins of the populated states through the measurement of the angular distribution of the ejectiles. The goal of the experiment run in March 2022 is to measure with unprecedented precision the spin-orbit splitting of the $1p$ states in a system where the populated states are close to the neutron separation energies. Such a challenging system might as well be a field of test for ab-initio calculations. Furthermore, to achieve a reliable normalization, measurement of neutron removal states in ^{18} N through the reaction channel (d, t) were planned.

The experiment was run and completed in March 2022. Previous tests had shown that the high-energy unstable beam produced a huge quantity of δ rays, leading to a large amount of free charge inside the ACTAR TPC. The problem was solved by measuring a similar reaction (^{20}O(d,^{3}He)^{19}N) allowing to study the same physics case by changing the secondary radioactive beam. The use of ^{20}O as a beam was a key factor in obtaining a positive result from the experiment. This beam could be produced at a higher rate (about $2 \cdot 10^4$ particles per second) and limited the *delta* electron problem encountered in previous attempts.

3.2 Experimental setup

3.2.1 ACtive TARget TPC

The apparatus used in the experiment was the ACtive TARget TPC, described in Chapter 2. To efficiently measure the proposed reaction, the usage of a deuterium target was foreseen. This could be achieved using ACTAR, filling the chamber with a mixture composed of D_2 (90%) and iC_4H_{10} (10%). The pressure in the chamber was about 950 $mbar$, and the gas flow inside the chamber was set at 155 sccm, to avoid any deterioration in the quality of the gas. In terms of electrical settings, the pad polarization voltage was set at $30\,V$, the mesh voltage at $560\,V$, and the voltages on the top and bottom of the field cage were, respectively, 6000 and $580\,V$. The gas chamber was surrounded on the front and lateral sides by silicon detectors, which were biased with a voltage of $40\,V$ and used to detect fragments that escaped the gas volume. Such detectors were also used to identify fragments, using the ΔE-E technique, where ΔE was measured as the energy loss in the gas and E as the energy released in the ancillary detectors. In this experiment, one of the main goals of the ancillary detectors was to discriminate helium isotopes, selecting events in which ^{3}He was produced in the transfer reaction. This goal has already been achieved, since identification in mass and charge has previously been reached up to lithium isotopes [43].

3.2.2 LISE spectrometer

The ^{20}O beam was produced from a primary reaction whose ejectiles were selected through the LISE separator [44], installed at GANIL. This apparatus allows for a clean selection of the fragments produced, in most cases achieving high purity and intensities. In the case of ^{20}O production, the primary reaction exploited a ^{22}Ne beam, accelerated through the two cyclotrons in GANIL [38] up to 60 MeV/A, impinging on a ^{nat}C target. The LISE spectrometer is a very broad-range and versatile instrument, capable of separating fragments by exploiting the different charge states and masses of the ejectiles of the primary reaction. Its design employs two dipoles and a Wien filter. The idea behind the usage of dipoles is to try to select the ions by exploiting their different magnetic rigidity. Combining different dipoles and using *wedges*, devices that can change the energy of ions between subsequent dipoles, allow for a clean selection. Furthermore, the LISE spectrometer employs a Wien filter which acts as a velocity filter. Such a device combines the presence of both an electric and a magnetic field and couples the electric and magnetic forces to achieve velocity selection. The particles that are selected are the ones that are not deflected and for which the electric and magnetic forces have the same magnitude. Considering that both forces depend linearly on the charge, it can be noticed that in this case neither the charge state nor the mass imply a direct role in the selection phase, as the only requirement for an ion to be selected is its velocity.

Chapter 4

Event reconstruction methods and implementation

The main part of the thesis work consisted of the analysis of the data collected. Those data are used as a test bench for different reconstruction algorithms. In particular, we focus on the development of efficient and reliable reconstruction techniques to retrieve information from the three-dimensional hits provided by the pads.

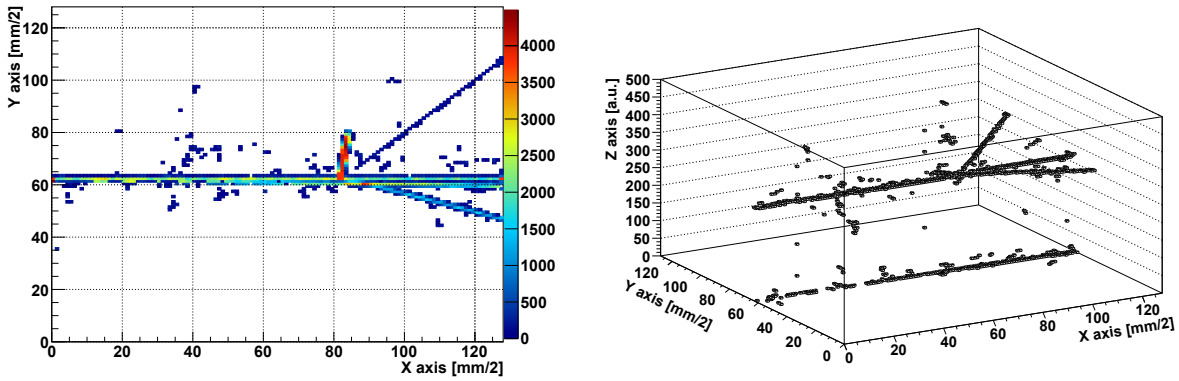
4.1 Translation to ARDA classes

In the last years, an extensive work has been done developing reconstruction algorithms for the ACTARdemo apparatus, the ACTAR demonstrator [45]. However, such algorithms are based on a different set of classes, inherited from the ARDA (ACTAR Raw Data Analysis) environment. Therefore, to facilitate the analysis process, one of the first objectives of this work was to link *ReducedEvents*, the signal processing output, to the ARDA classes, the input of the reconstruction codes.

4.1.1 ARDA

ARDA [36] is a C++/ROOT-based software package that converts signals recorded by GET electronics to retrieve physical information. The conversion of signals is achieved through the application of different PSA routines. The FPN channels are used to evaluate and subtract the baseline offset, and the analysis output is stored in C++ objects, named *cRawEvents*. Such objects contain raw hits that still have to be calibrated. In the calibration phase, through lookup tables, the hashed electronic channel related to the pad is converted into the 2D position, the arrival time of the signal is related, through the drift velocity, to the position along the axis perpendicular to the pad plane, and the amplitude of the signal to the charge deposited. One of the advantages of the ARDA framework is its simplicity and flexibility. Exploiting the ARDA environment allowed a simplification of the clustering and tracking task. In these classes, the RANSAC and Hough transform routines (Sect. 4.2.1 4.2.2) were developed. Among the tools developed are also visualization macros. Once the

hits were converted and calibrated, it was also possible to visualize the events as two- and three-dimensional figures. An example of this is shown in Figure 4.1.



(a) Two-dimensional representation.

(b) Three-dimensional representation.

Figure 4.1: Two- and three-dimensional representations of an event acquired with the ACTAR detector. The orthogonal dimension to the pad plane is reconstructed using the arrival time information. Figure 4.1b shows the presence of a non-interacting beam particle in the acquired event.

Figure 4.1 shows an issue that is typically present in many events stored, which is the presence of *pile-up*, due to the presence of tracks left by beam particles that do not undergo any nuclear interaction. Such tracks have to be correctly classified and excluded from the event analyzed, as they do not contribute to the physics event recorded. The presence of pile-up can also be visualized by looking at the three 2-dimensional projections of the data on the X, Y, and Z axes, as shown in Figure 4.2.

The projections show the event from three different perspectives and overall give a more accurate description of the event than the three-dimensional representation. Indeed, the XY projection gives an idea of the present tracks but does not show piled-up tracks, as their projections all belong to the same pads. On the other hand, this information can be retrieved from the XZ projection, where the pile-up is clearly spot and, at the same time, the tracks belonging to the event can be seen. The YZ projection represents the view one can have while looking at the beam spot from the end of the chamber. Though traces are not well spot nor very clear visually, this view is also important as it gives an idea on the momentum conservation. In fact, one can visually check whether all the traces are measured by checking the projected points. Such projections are also monitored during experiments, since detection thresholds can be modified to avoid information loss. In the reference experiment, light- and low-ionizing particles, such as high-energy protons, could not be measured because of a specific threshold setting. The information provided by the three projections is quite complete and can truly give a comprehensive view of the event stored, even without considering the information brought by the charge collection, as shown in Figure 4.2. At this point, it is necessary to reconstruct the tracks by clustering the hits, so to measure the particle's energy, emission angle, and reconstruct the interaction vertex. The clustering routines developed are described in the following section.

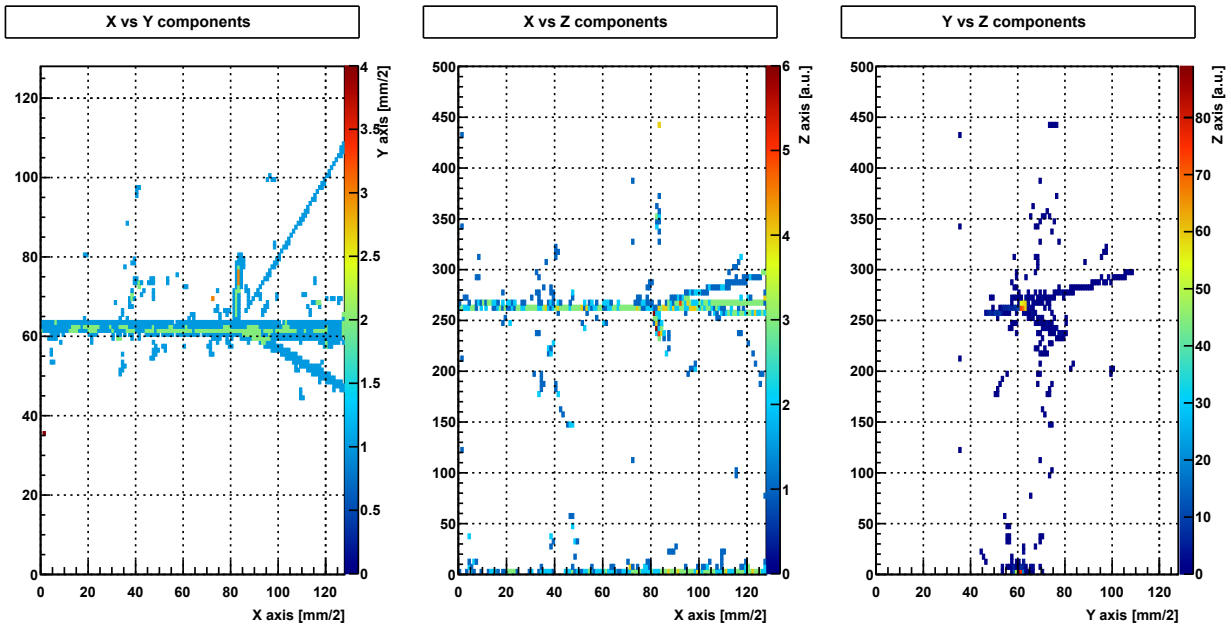


Figure 4.2: Three bidimensional components compared.

4.2 Clustering

To retrieve information from particle tracks, it is necessary to *cluster* together points belonging to the same track, avoiding at the same time including points belonging to other tracks. Such a procedure aims at a clean identification of the energy lost by the particles inside the active volume. In two previous works, two different clustering algorithms were built [36, 45]. However, both algorithms were designed for the ACTARdemo case. One of the goals of this thesis work was to write an implementation of such algorithms on the data coming from ACTAR, designing a common framework that could let the users decide intuitively which algorithms to use.

4.2.1 Ransac

One of the algorithms implemented in the ACTAR case is the RANSAC (RANdom Sample Consensus), an iterative routine that is usually used to estimate parameters on top of a mathematical model in data sets that present outliers, entries that must be excluded for a correct parameter estimation. The first publication on this algorithm dates back to 1981 by Fischler and Bolles [46]. Originally designed for image reconstruction and analysis, it is now widely used in computer vision tasks [47]. The algorithm aims at simplifying the research of patterns in a noisy environment. This is achieved by explicitly considering the least number of free parameters possible to define a model to be tested, and therefore simplicity is one of the main characteristics of the paradigm, originally stated as follows [46]:

- Given a model that requires a minimum of n data points to instantiate its free parameters and a set of data points P such that the number of points in P is greater than

n (#(P) > n), randomly select a subset *S1* of n data points from P and instantiate a model. use the instantiated model *M1* to determine the subset $S1^*$ of points in P that are within some error tolerance of *M1*. The set $S1^*$ is called the *consensus set* of *S1*.

- If $\#(S1^*)$ is greater than some threshold t, which is a function of the estimate of the number of gross errors in P, use $S1^*$ to compute (possibly using least squares) a new model $M1^*$.
- If $\#(S1^*)$ is less than t, randomly select a new subset *S2* and repeat the above process. If, after some predetermined number of trials, no consensus set with t or more members has been found, either solve the model with the largest consensus set found, or terminate in failure.

The RANSAC paradigm contains from the beginning three unspecified parameters, the error tolerance used to determine whether or not a point belongs to the model, the number of subsets to try (the number of iterations) and the threshold t, which defines when a good enough model has been found. Although an estimate of the first two parameters can be given by studying the dispersion and cardinality of the dataset, the third one is strictly dependent on the data set considered. In many applications, it was noticed that empirical tuning of the first two parameters can lead to accurate and precise results. A particular application of such an algorithm is the clustering of three-dimensional points, which directly translates into the reconstruction of the tracks left by charged particles inside the active volume of the ACTAR detector. In the ACTAR case, the outliers are points in space that do not belong to any trace, mainly related to electronic noise, or to tracks that come from particles not related to nuclear reactions, such as secondary *delta* electrons, or are related to other tracks. The goal of a clustering algorithm, other than discriminating points that belong to tracks or not, is to cluster points depending on the track to which they belong, allowing measurement of the total energy deposition and of the emission angle of single particles.

The steps of the RANSAC algorithm applied to the linear interpolation case can be split into five main steps, to be repeated for a user-defined number of iterations (n_{loops}).

Choice of a pair of points

The first steps consist of a random choice of a pair of points $\left(x_1^{rand}, x_2^{rand}\right)$ that belong to the data set. Then, the track linking these two points is drawn.

Inliers test

The distance of each point from the previously defined line is calculated. If for a certain point the distance from the line is less than a user-defined threshold d_{max}, the point will be considered an inlier and placed in a subset. Otherwise, the point will be discarded and considered an outlier.

Subset test

Once the inlier subset is built, the number of points, their total distance from the model line, and their total collected charge are calculated. If the number of inliers or the total energy does not exceed two user-defined thresholds (n_{min} and E_{min}), the algorithm restarts from the first step, increasing the iteration count by one.

Comparison of subsets

To understand whether the present subset is effectively an improvement over the previous best, the value of a *test function* is calculated. If the score obtained is better than the previous best score, then the information on the best track is updated. In any case, the iteration count is increased by one.

Cycle closure

The algorithm closes its cycle once a defined number of iterations is reached. If once the cycle is terminated, a track overcoming all thresholds is found, the algorithm restarts a new cycle looking for a new cluster on a data set on top of which the points already classified as belonging to a track are removed. If, on the other hand, no subset of inliers has crossed the thresholds, the algorithm stops.

Although the RANSAC paradigm only contains three free parameters, this specific application needs more in order to lead to accurate results: the number of iterations to be performed n_{loops}; the minimum energy (E_{min}) and the multiplicity of inliers (n_{min}) to accept a subset as a track; the maximum distance of inliers from the sample line (d_{max}). These parameters influence both the computational burden of the algorithm and its output, distinguishing tracks to be stored from discarded ones, events which are considered from ones not analyzed. If, on the one hand, the optimization of the parameters can shrink the clustering time, on the other hand, this might be related to information losses. While the application of the algorithm is broad-sighted and exploitable on various datasets, it is true that parameters have to be tuned specifically case by case, experiment by experiment. Another free parameter of this RANSAC application is, in the end, the test function used to compare different subsets. Notwithstanding the simple paradigm proposed by the authors, the subsets that overcome the thresholds are often still suboptimal. A test function needs to be defined to compare different clusters. Eq. 4.1 shows the one chosen for the following work: it depends explicitly on the randomly chosen pair $\left(x_1^{rand}, x_2^{rand}\right)$ and on the parameters chosen by the user, which determine the *inliers* set.

$$f_{test}\left(x_1^{rand}, x_2^{rand}\right) = \frac{\sum_{i=1}^{N_{inliers}} d\left(x_i\right) \cdot E\left(x_i\right)}{\left(\sum_{i=1}^{N_{inliers}} E\left(x_i\right)\right) \cdot N_{inliers}} \tag{4.1}$$

In Eq. 4.1, x_i is the i-th point within the tested subset, $E(x_i)$ its energy, $d(x_i)$ the distance from the line passing from the pair of points chosen randomly. This function aims at selecting the most energetic tracks with the largest number of pads. Therefore, it favors large subsets including points that carry a great amount of energy rather than those containing smaller amounts of collected charge, most likely related to noise.

Application example

To clarify the procedure undergone by the algorithm, Fig. 4.3 shows the 2D application of the algorithm in the case of ACTAR data in two dimensions. However, it is important to note that to reconstruct the events, all spatial dimensions are considered. In the upper left subfigure, the choice of two random points (in red) is shown. In the second one, the inliers are displayed in green: such a selection is not satisfactory, and the points do not belong to a real trace. If the number of inliers and their energy overcome two user-defined thresholds,

the test function is computed. The bottom left figure shows the best line found after different iterations. This line, which links the two randomly chosen red points, represents the direction of the track well, and due to the definition of the test function, clusters containing more points and being more energetic are favored with respect to smaller ones. At the end of a cycle, the points included in the cluster are removed from the data set and the algorithm restarts its analysis on the reduced set as shown in the last figure. This process is repeated until no more tracks are found and the algorithm stops.

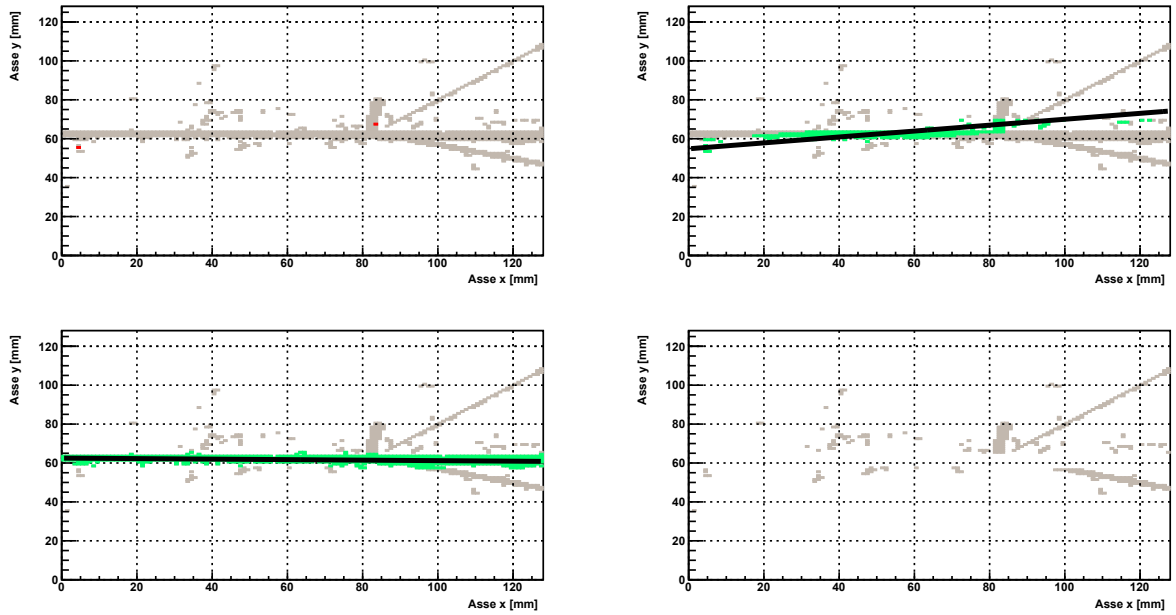


Figure 4.3: Track reconstruction through the RANSAC algorithm. In the upper left, the choice of a pair of points. Upper right is the construction of the line, and the inliers are represented in green. In the bottom left, the final line found after the number of selected iterations, and on the bottom left figure are the points left after the points of the first track are removed.

4.2.2 Hough Transform

The second algorithm implemented to reconstruct tracks in the ACTAR case is the Hough transform. The Hough transform was originally designed by Paul Hough to analyze the data of a bubble chamber using a calculator [48]. Today, the classes of problems that can be tackled through this algorithm are not only restricted to lines belonging to a plane, but include a broader range of curves. The algorithm aims to find the parameters that best describe a set of points, which in the ACTAR case are the parameters that describe the lines that interpolate the clusters in the best way. The general idea of the algorithm is the following: given a curve that needs to be fit and its parameterization, a histogram is created in the space of the parameters of the curve itself, discretizing the range of possible parameters in bins. If the analyzed curve has N parameters, the requirement to pass for a point defines a N-1-dimensional surface. The algorithm tests all configurations in the

parameter space, draws the corresponding curve, and adds counts to the corresponding bin according to the number of points that the tested model includes. The bin with the highest value in the histogram, which corresponds to the points where the most *N-1* dimensional curves intersect, determines the parameters that identify the model that crosses the highest number of points. The correct choice of the model parameterization is paramount. For example, if a redundant parameterization were defined, the investigation of the maximum could lead to ambiguities because many parameter sets could describe the same line.

Clustering in two dimensions

Among the simplest applications of the Hough transform is the detection of straight lines in two dimensions. As an example, a possible parameterization of a straight line is the following: $y = mx + b$, and can be represented as a point (b, m) in the parameter space. However, such a simple parameterization is unable to describe vertical lines. To avoid such a case, Duda and Hart proposed the use of *Hesse normal form*, which in the case of a line becomes:

$$r = x \cos\theta + y \sin\theta \tag{4.2}$$

In Eq. 4.2, r is the distance of the closest point to the origin and θ is the angle between the x axis and the line connecting the origin with that closest point. In general, using polar coordinates avoids the complications related to vertical lines. Using this parameterization, the *Hough space* becomes the (r, θ) space. Given a single point in the space, the set of all lines passing through that point corresponds to a sinusoidal curve in the (r, θ) space, which is uniquely related to that point.

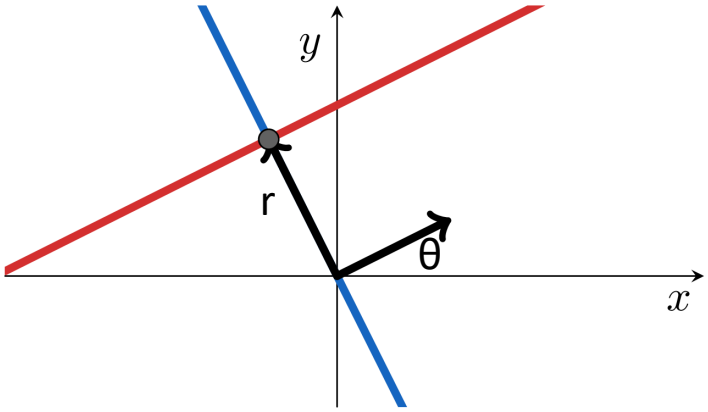
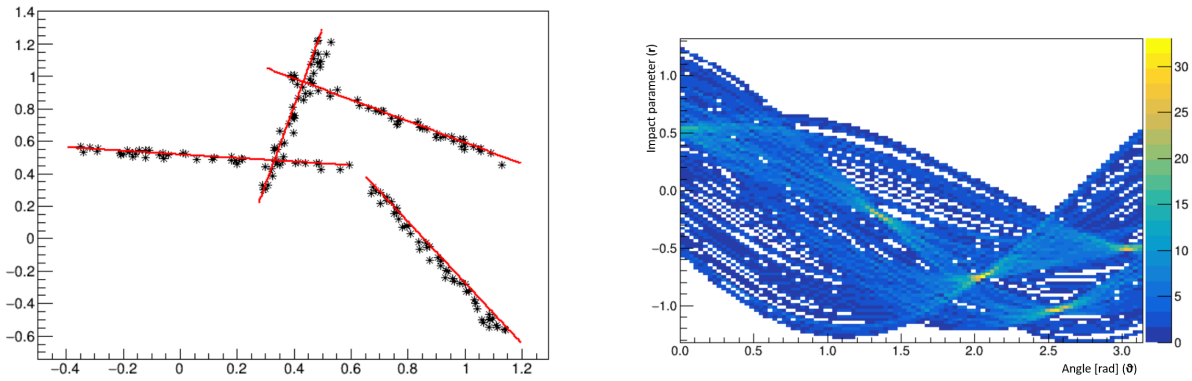


Figure 4.4: Parametrization of a straight line in two dimension in polar coordinates.

In the left subfigure 4.5a is represented the simulated dataset, composed of points belonging to four different lines that also overlap. Figure 4.5b is the Hough space representation of those data. The interval of possible impact parameters depends on the data set itself, while the possible angles range from 0 to π. In the Hough space, the maxima of the bidimensional histogram identify the parameters that accurately define the four red lines in Figure 4.5a.



(a) Set of simulated points.

(b) Representation in the (r, θ) space.

Figure 4.5: Hough transform applied on a 2D set of simulated data [36].

Accuracy is, in any case, limited by the binning of the Hough space. To improve such limits, the Hough transform can be implemented as a clustering algorithm, identifying the points belonging to different clusters and fitting them using a standard fitting technique.

The implementation of the Hough transform in the ACTAR case resembles the original paradigm. The maximum distance that points can have from the origin is defined by the geometry of the detector. Then, an iterative loop looks for the inliers of the models defined by every possible combination of parameters. In fact, once the parameters defining the line have been chosen, the points whose distance from the line is below a user-defined threshold are considered as inliers and their energy is added to the bin corresponding to the parameters in the Hough space. Once the loop is completed, the algorithm looks for the bin that holds the maximum counts. The coordinates of this bin represent the parameters of the line that contains the highest energy deposit. At this stage, the points considered inliers by the model including the highest energy are grouped into a set. Both the cardinality and the total energy of this set are compared with two thresholds that define whether the set can be considered as a track. If the comparisons yield a positive outcome, the points belonging to the set are removed from the total data set, and a new cycle is started over the parameter space, in analogy to the RANSAC case. This loop ends when the maximum in the histogram represents a line whose set of inliers is unable to overcome the thresholds to be considered a proper track.

Clustering in three dimensions

Even though it is among the simplest applications, recognizing lines can become a complicated task while increasing the dimension of the space considered. In three dimensions, for example, lines can be parametrized as in Eq. 4.3:

$$\vec{x} = \vec{b} + t\vec{n} \tag{4.3}$$

where $\vec{x}$ is a point on the line, $\vec{b}$ is the *basepoint*, $\vec{n}$ the versor that defines the direction of the line, and t is a parameter in $\mathbb{R}$. However, it is possible to note that there are many parameter choices that define the same line. For example, adding to $\vec{b}$ any parallel vector to $\vec{n}$ yields

the same line. The same works by flipping the versor: using $\vec{n}$ and $-\vec{n}$ also determines the same line. To achieve a single possible parameterization, two additional conditions can be included considering all $\vec{b}$ that satisfy Eq. 4.4:

$$\vec{b} \cdot \vec{n} = 0 \tag{4.4}$$

This condition can always be met and is equivalent to choosing the base points that belong to the plane that passes through the origin and is orthogonal to $\vec{n}$. Moreover, one can also arbitrarily fix the magnitude of $\vec{n}$: for simplicity, one can choose a hemisphere with a unit radius where the only possible versors defining the direction are the points in the hemisphere. Therefore, taking advantage of these additional conditions, the number of parameters needed to parameterize a line can be reduced to four. The Hough space can be built using as the first two parameters the angles in polar coordinates θ and ϕ defining the versor on the hemisphere and considering two more parameters for the definition of the basepoint. One of the main challenges faced during the implementation in many dimensions is the use of memory. The occupation of memory required to build the histogram scales as the fourth power of the number of bins per axis. Moreover, given the face on which the second parameters are computed starting from the angles, the number of times that the histogram has to be filled for every point scales as the number of bins squared. Therefore, increasing the number of bins to have better precision in the line identification might lead to a memory leak (the necessary memory would reach some GB per thread). To reduce memory usage, the Hough transform has been used with a reduced number of bins, solving the sole clustering task. Then, a second fit has been used to retrieve the precise fit parameters [36].

Double $\mathbb{R}^2$ Hough transform

To achieve acceptable rate performances, one trick that has been exploited consists in applying a *double* Hough transform on bi-dimensional data, using the projections of the data on planes. This application is based on the fact that the projection on a plane of a line is still a line, provided that they are not orthogonal. The Hough transform is then applied to the projections of the points on a plane, and a candidate to a possible line is found. Given the fact that there could be two parallel lines, whose projection is the same, the algorithm is run once more projecting the points on a plane orthogonal to the ones used for the first transform. The candidate line is then refined, and the data set is reduced accordingly to the points that belong to the line in both projections. Such a procedure is again iterated many times until a line is found or the algorithm yields a negative result. It has been found that the accuracy in track reconstruction is comparable to the direct $\mathbb{R}^3$ case, using a significantly shorter time. Indeed, even if the parameters to be analyzed are four in both cases, filling twice a two-dimensional histogram requires less time and memory than filling once a four-dimensional one. This is related to all the possible combinations of parameters that lead to a higher complexity while increasing the dimensions.

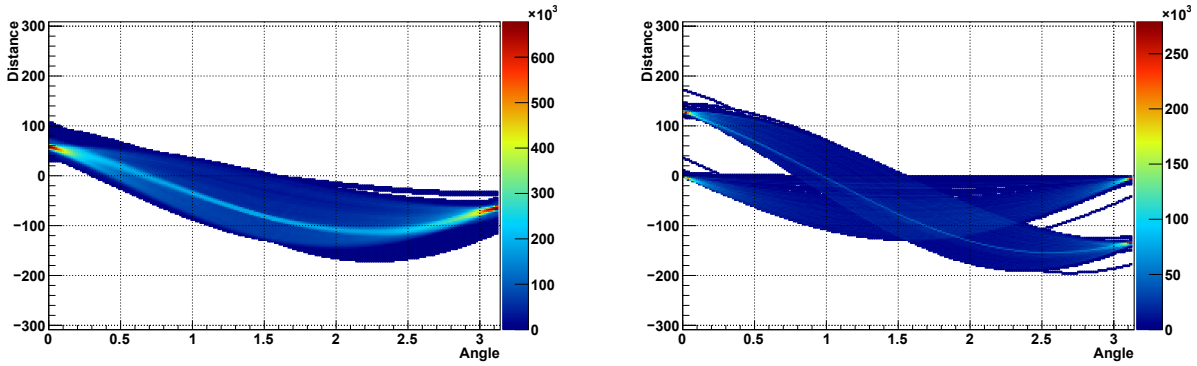
4.2.3 Beam clustering

As shown in Figures 4.1b and 4.2, and as visible in many other applications, the acquired data often carry along non-interacting beam tracks. These tracks do not belong to the physics

event of interest, but at the same time are considered by the algorithm. The distinction of such tracks is easier in the case where the Z position of the track is far from the event considered, as in the event in Figure 4.2, but it can become more difficult when the event and the pile-up almost overlap. Therefore, specific routines and thresholds have been developed and set to classify beam tracks. While the settings of the energetic and multiplicity thresholds can be shared between the RANSAC and Hough Transform algorithms, the routines differ. In the RANSAC case, the trick used to classify the beam tracks is the restriction of the initial couple of points to points that share the same Y and Z coordinate. Indeed, in the case of a homogeneous electric field, the electrons produced from ionization should arrive at the same time on the pad plane, and thus produce a horizontal track. Moreover, given the momentum of the incoming beam, the sole interactions with the electrons do not modify the path of the incoming particles. Restricting the possible choice of points leads to the identification of beam tracks in a few iterations. In the Hough case, on the other hand, there are no specific routines developed in the process. A reduction in computing time can be achieved while looking for beam tracks by reducing the range of possible angles. This feature is clear by looking at the XY projection in Figure 4.2, where it can be seen that the angle formed between the beam track and the X axis can only be close to 0 or π. On the other hand, it has been noticed that when looking for beam tracks through the $\mathbb{R}^2$ Hough transform, better performances are achieved considering as the first projection the one on the XY plane. This projection indeed allows for a first distinction between tracks that belong to the central area of the pad plane with respect to others that belong to the sides. Its then the second Hough transform that, once applied to the reduced set, easily manages to discriminate between points at different heights.

Application Example

The search of a beam track in the event in Figure 4.2 with the Hough transform is presented in Figure 4.6. Figure 4.6a represents the Hough space view of the XY projection of the event: even if two peaks are visible, on the right and left edges, they both represent the same line, which can be described by angles close to zero or π with the X axis. This is because the projection on the XY plane includes both the beam track that produces the event and the pile-up track underlying it. This overlap translates into the identification by the Hough transform of a very energetic track, which hides the other maxima related to the other tracks. At this point, the points that lie close enough to the line are stored in a subset, and this subset is analyzed applying a second Hough transform on a plane orthogonal to the previous one, in this case the XZ plane. The second transform on the reduced set, the result of which is shown in Figure 4.6b, results in two distinct maxima (and clusters). The one having a lower intercept is directly related to the cluster of points of the non-interacting beam track. On the other hand, the second cluster can be related to the beam particle that produces the interaction. It is important to note that both representations in the Hough space have a maximum that lies in the low-angle region. This effect can be noticed from the XY and XZ projections in Figure 4.2, and can be exploited, as stated previously, to reduce the possible angles on which to look for beam tracks. The clustering algorithm, also in this case, removes the points that contribute to the maximum in the Hough space and restarts the process on a reduced set. Once no more beam tracks are found, on both RANSAC and Hough cases, the



(a) First Hough transform on the XY plane. The horizontal track representing the beam tracks overlapped represents the maximum.

(b) Second Hough transform for the XY - inliers on the XZ plane. Two clusters define the two present beam tracks.

Figure 4.6: Hough spaces for the XY projection of the complete set and XZ projection of the reduced set.

algorithms restart changing thresholds and returning to default settings.

4.3 Tracks Cleaning

In application-specific cases, the tracks found by the clustering algorithms can be cleaned from possible outliers present in the subset. An application of this secondary processing technique is, for example, the exclusion of points that cannot physically belong to the same trace. If no magnetic field is applied, the tracks are expected to have a linear shape. Moreover, considering that the beam pipe is positioned at the center of the chamber, it is clear that points belonging to a trace emitted after a reaction can only belong to one half of the pad plane. Therefore, a track that contains points on both sides of the pad plane suggests the presence of outliers. The chance of including outliers is expected from the construction of both algorithms. In fact, if the tested line passes close enough to a point, it includes it, regardless of the half of the pad plane to which the point belongs. It is relevant to exclude such points if, for example, the length of a track has to be determined. Including points from the other side of the pad plane leads to an overestimation of the track length. Therefore, two different routines have been added for each clustering algorithm. The solution to this problem is based, in both cases, on the division of the pad plane into three zones. If, indeed, the beam is well centered inside the active area of the detector, it is possible to define *beam area* as the central area, whose width depends on the characteristics of the events analyzed. This zone separates two other areas on its sides. In the RANSAC case, a loop controls whether outliers are included at the end of the defined number of iterations. If a track has been found, a loop starts on the subset: the total energy is calculated, and it is checked whether the average position lies in the central area or on the side ones. In the first case, everything is left as is, as most likely the trace is coming from a beam particle. In the other case, the points belonging to the opposite sides (if present) are removed. After this removal, the line is required to pass once more the tests on the number of inliers and on the energy

threshold. If any of these tests are no longer passed, the algorithm stops. This condition is necessary to avoid the algorithm to fall on an infinite loop: if the algorithm were to be restarted when a suboptimal track containing points on both sides was found, it would be running over the whole dataset over and over. This requirement is anyway motivated by the fact that in the events, as far as the so far analyzed data are concerned, the main part of the energy deposit clearly belongs to a side, and often times having points on the other sides is only related to a few noisy points (so with a low energy associated) or to unlucky cases in which the model line on one side crosses part of another line on the other side. In the Hough transform case, the same pad plane division in three different areas is exploited, and the selection happens when a candidate line has been found. The average is computed on the Y-axis of the subset points. Then, if the highest fraction of the energy lies on one side, the data belonging to the other side are excluded. Given how the algorithm is structured, it is necessary in this case to move the deletion of the points from the total set of data after the removal of the outliers. After the outliers are removed, the candidate line is required to pass both tests on the number of inliers and on the total energy to be considered as a track. Otherwise, the algorithm stops and the points previously included in a subset and removed from the total set of data are placed back into the initial one.

At the end of the clustering process, it is possible to visualize the result of the clustering by plotting once more the three projections of the event, but distinguishing each cluster with a different color, as shown in Figure 4.7. There, one can notice that every track is correctly classified, and the sparse points corresponding mainly to electronic noise are recognized as not belonging to any track.

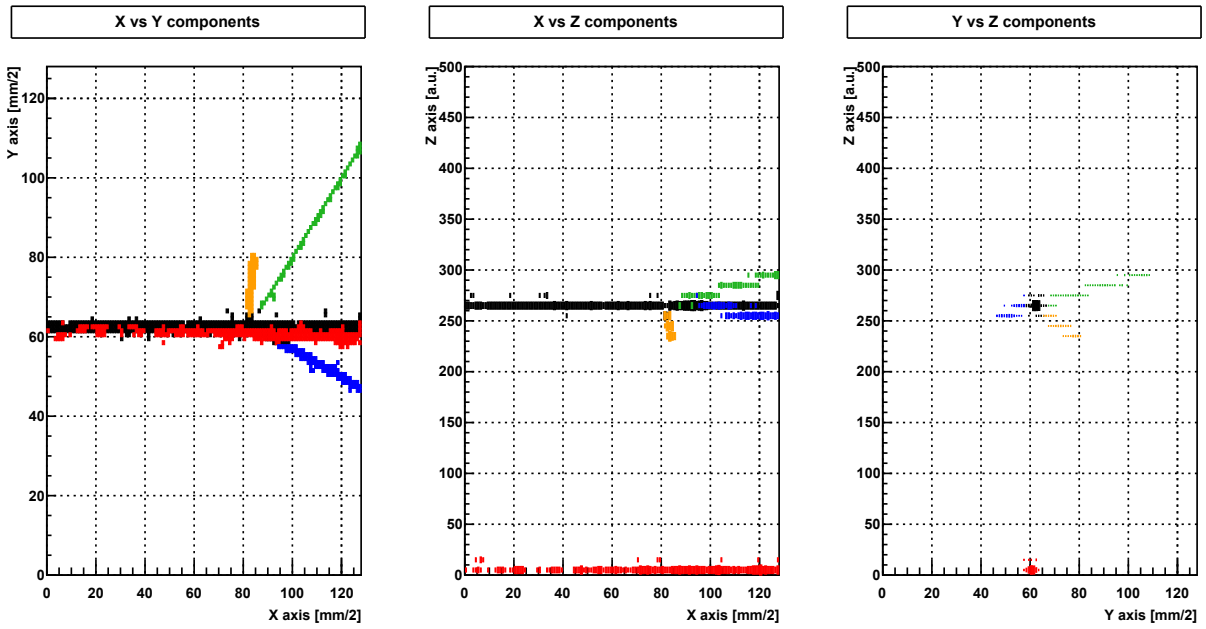


Figure 4.7: Projections of a clustered event. Different colours represent different clusters.

4.4 Tracking

As stated previously, both RANSAC and the Hough transform are used in this implementation only as clustering algorithms. Though the routines themselves have as an output a model line, in order to achieve a better resolution, a fit of the clusters is necessary. Such a fit has been implemented by exploiting the ROOT Fitter class, which is part of the ROOT environment. One of the possible ways to use this class is by defining a loss function and passing some of its parameters to a minimization routine (such as MINUIT, for example), which will then look for the parameters minimizing such a function. One of the challenges is then to find a minimization function that allows the algorithm to achieve a proper fit. Moreover, to achieve a reasonable tracking, it is necessary to consider the energies recorded in the different pads. One can use the energies themselves as a weight in the fitting procedure. This truly improves accuracy, as during the drift and avalanche the charge tends to be distributed on a region on the pad plane centered on the true path of the particle and where the maximum energy deposit lies. Two different parameterizations were implemented. The first one considers six parameters to be minimized, as no condition was posed on the lines. The function to be minimized in this case was simply the chi-square between the line described by the parameters and the points in the cluster. In the second case, the lines were parametrized as expressed in section 4.2.2, considering only four free parameters applying the two conditions described above. Thus, calling $\hat{n}(\theta, \phi)$ the versor defining the direction of the line, $\hat{v}(\theta, \phi)$ and $\hat{w}(\theta, \phi)$ the projections of the versor on the axis and $\tilde{x}$ and $\tilde{y}$ the coordinates of the basepoint, P_i the considered point and E_i its energy, one can define a loss function as in Eq. 4.5.

$$f(\theta, \phi, \tilde{x}, \tilde{y}) = \sum_i E_i \left[\left(\vec{P}_i \cdot \hat{v}(\theta, \phi) - \tilde{x} \right)^2 + \left(\vec{P}_i \cdot \hat{w}(\theta, \phi) - \tilde{y} \right)^2 \right] \tag{4.5}$$

Using such a function yields an accurate outcome. The computing time required for minimizing a function of four parameters is lower than that necessary for the first parameterization, and using six parameters does not show any improvement. Therefore, the method that will be considered in the following is related to the second parameterization.

4.5 Vertex Reconstruction

One of the main pieces of information that is necessary to retrieve from the acquired events is the position of the reaction point. The position of the interaction is indeed related to the energy at which the reaction takes place, and a precise reconstruction is necessary to justify the usage of an active target. The routine exploited is the following: around every line coming from the fit, a cylinder with a user-defined radius is defined. Two different tracks are considered to belong to the same reaction vertex if points coming from one cluster fall inside the cylinder built around the fitted line of the second cluster. Or, by introducing a more strict condition, the second implication can also be requested, meaning that at least one point must fall inside the cylinder of the other *and* vice versa. The single implication already yields a good outcome, and the double implication could be required in the case

of particularly noisy data sets. Once the traces involved in a reaction vertex have been identified, the position of the vertex is found through TMinuit. In this case, the parameters of the chi-square function to be minimized are the coordinates of the vertex. In case only two traces take part in a reaction vertex, an analytical solution could have been found. For the sake of generality, such procedure has not been implemented, and the algorithm always looks for the point minimizing the squared distance between itself and the fit lines of the traces inside the vertex. It is also worth noticing that such a procedure can be, in principle, applied in a iterative way in the case in which in the same acquisition there could be more than one interaction vertex.

Chapter 5

Analysis

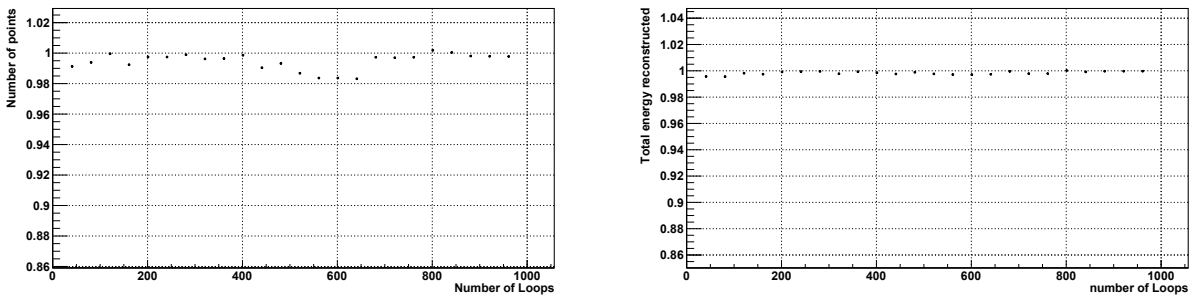
Although the data set coming from the GANIL experiment includes a large amount of physical information, the goal of the thesis was an in-depth study and comparison of different clustering algorithms. A great part of the work consisted, indeed, in repeating runs and tests over a selected set of events, trying to infer the implications of parameter changes for both RANSAC and the Hough transform. A reduced set of events has been constructed, manually selecting 100 events that contain different types of interaction. It must be noticed that picking randomly one hundred events from an acquisition run would lead to a much different set, as normally events in which no nuclear reaction is present are also included. However, a manual selection of the events was performed to test the performances on events of interest to avoid tuning the parameters over noise. Among the parameters analyzed in both algorithms, there is also an external one which has been shown to significantly improve clustering performance. This parameter is the Z rescaling factor, which is simply applied to the data and scales the Z axis accordingly. Such a scaling factor is useful to obtain a uniform distance between voxels, the small cubes that are defined by the resolution of the detector, in three dimensions. While this is guaranteed for X and Y, as the spatial information comes from the pad hit, it is not as such for the Z one, where the time information is converted to a spatial one by means of the drift velocity. In the following analysis, the Z rescaling factor has been set as a default value to a value that forced all axes to have the same range $[0, 128]$. In this case, where no calibration for the drift velocity has been considered, the rescaling factor is equal to four. Often, as in the events coming from the experiment there is no ground truth, the outputs are compared to the output of the analysis using the parameters that are supposed to perform best.

5.1 Parameters Analysis - RANSAC

The first algorithm presented is RANSAC. As stated previously, many parameters influence the performance and outcome of the algorithm. The ones that will be presented are the number of loops, the size of the tracks accepted, and the Z-rescaling factor.

Loops number

The first parameter analyzed was the number of iterations selected for the RANSAC loop over a single clustering step. Therefore, the reduced set of 100 events was analyzed, varying the number of iterations per cluster. It has been noticed that the computational effort depends linearly on the number of loops. To optimize such parameter, it is possible to observe how the number of points and of the energy reconstructed vary accordingly to the iterations considered. To achieve a clean estimate, the distinction between beam and non-beam tracks has not been taken into consideration. The maximum number of loops considered has been in this case 10^3. This upper limit is already an improvement over the previous analysis performed [45], as the number of iterations has been significantly reduced. This reduction is related to the fact that an additional requirement has been introduced for the couple of points chosen at random. In fact, to consider a couple of points as valid, the points must be at a distance greater than 4 mm. This requirement allows the algorithm to avoid choosing points that belong to the same track but are too close to each other, as the line that crosses both of them might not resemble the particle trace. This simple condition has allowed one to reduce the number of iterations by a factor of 10, without losing information. In the case of the parameter analyzed, the results obtained using the maximum number of loops are taken as a reference. Thus, the output obtained with the other iterations is normalized accordingly. to easily compare the outputs with each other. The averages of the scores obtained for the 100 events are plotted in Figure 5.1.

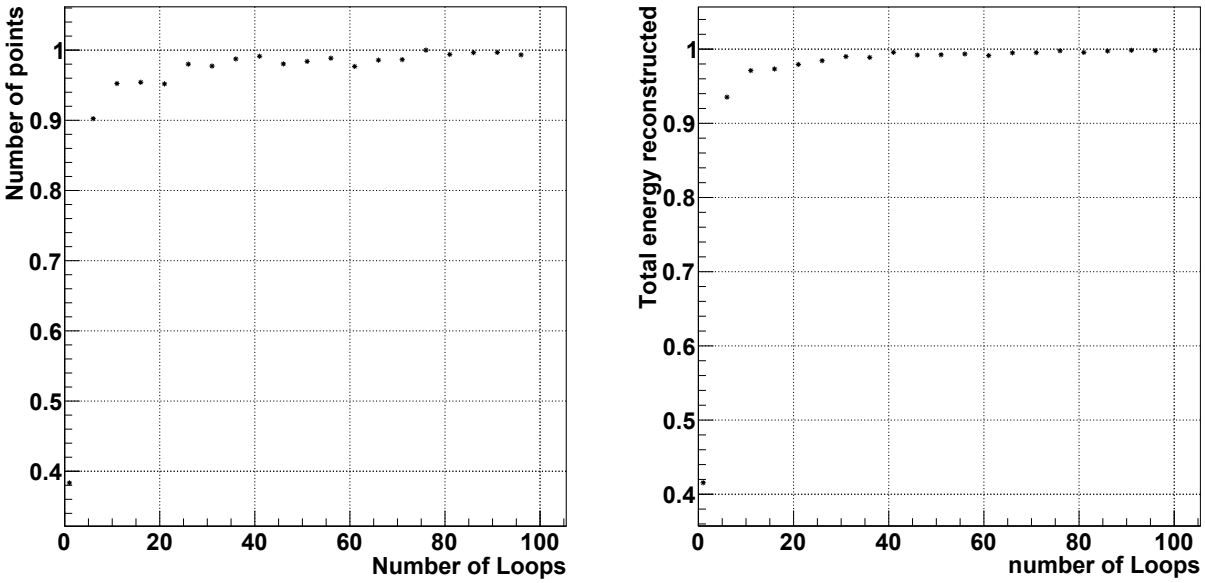


(a) Fraction of points clustered vs number of loops. (b) Total energy reconstructed vs number of loops.

Figure 5.1: Ratio of reconstructed points and total energy.

As Figure 5.1 shows, the number of points included in the reconstructed tracks and their total energy soon reach a plateau and are comparable to the values reconstructed in the maximum number of iterations. In principle, such graphs show that it could be possible to reduce the number of iterations and reconstruct the same amount of energy. Given the fact that most of the Figure 5.1 shows a flat trend, it becomes more interesting to reduce the number of iterations, studying in depth the region at a lower number of iterations. The result obtained by reducing the maximum number of iterations to 100 is presented in Figure 5.2.

Even by reducing the maximum number of iterations to 100, the points and energy reconstructed are comparable to the ones included with a much higher number of iterations. This result is interesting and probably depends on the choice of thresholds and on the definition of the test function. It should also be noticed that the graphs presented do not justify by



(a) Fraction of points below 100 iterations. (b) Energy reconstructed below 100 iterations.

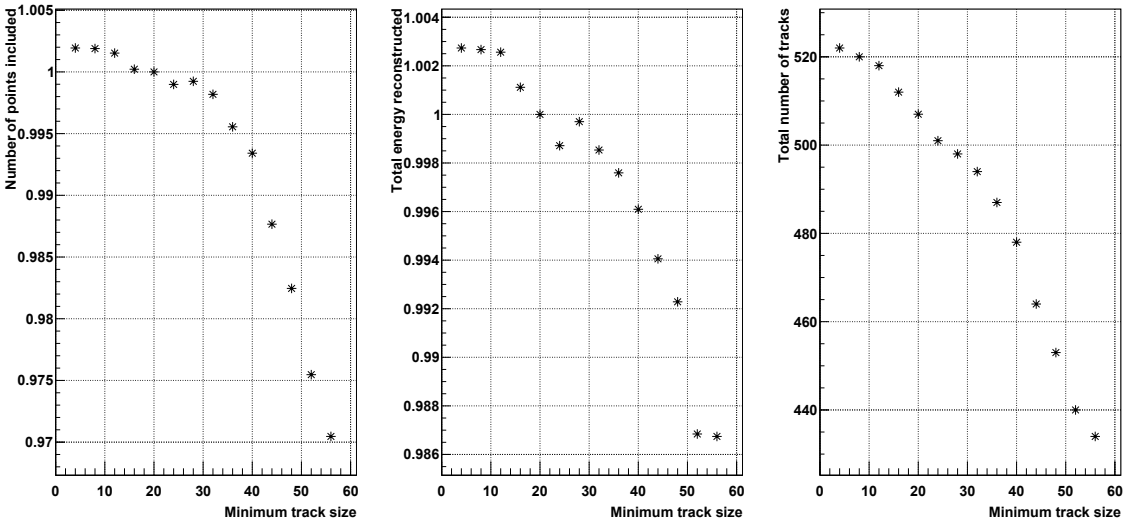
Figure 5.2: Ratio of reconstructed points and total energy.

themselves a reduction of the parameters. Indeed, even if the entire energy present in an event is reconstructed and fitted, there is no warranty on whether the data are correctly clustered. However, such a feature has been checked during the analysis phase and will be discussed further in Section 5.3. In conclusion, it has been shown that in this application the number of iterations can be strongly reduced with respect to the previous analysis, shrinking the computation time. More accurate metrics have to be tested to take into account the accuracy in the reconstruction of the single clusters.

Minimum tracks multiplicity

The second parameter analyzed was the minimum number of points required to consider a set of inliers as a proper track. This parameter represents one of the two thresholds that discriminate whether or not a cluster is worthy of being considered a track. In this case, the reference has been set to a minimum multiplicity of 20 pads, setting the minimum energy threshold to 5000 $[a.u.]$. The variables considered are once again the total number of points and energy included in the reconstructed event, normalized as for the previous analysis to the reference. Also, to also have an indication of how many tracks are actually identified during the analysis, the total number of clusters found has also been plotted. The results obtained on the reduced set of 100 events are shown in Fig. 5.3.

In this case, two different trends can be seen in the normalized Figures 5.3a and 5.3b. Indeed, considering multiplicities below the reference, in both cases the trend is quite flat, showing limited changes. Both ratios are above unity, as for lower thresholds it is easier to form clusters. However, fixing the energy threshold limits the acceptance of non physical clusters.



(a) Fraction of points included in tracks vs number of loops.

(b) Total energy reconstructed vs number of loops.

(c) Total number of tracks vs loops.

Figure 5.3: Comparison of RANSAC outcome varying the minimum clusters minimum multiplicity.

An increase in the threshold is related to a reduction in both points and energy. This decreasing trend is also seen in Fig. 5.3c, which represents the number of clusters found. Analyzing this last graph, one can see that the clusters increase even for multiplicities below 20, implying that the stability seen in the first two graphs is related to different cluster assignations, and to the splitting of the points belonging to the same particle in many smaller clusters. Moreover, the drop in the number of clusters is percentagewise higher than the other drops: higher multiplicities are also related to the exclusion of physical tracks coming from low-ionizing ions. Such traces generally show a limited energy depot and multiplicity, whereas traces left from heavier fragments, such as the beam traces, have higher multiplicities and depots. The comparison between these kinds of traces explains the sharper drop seen in the rightmost figure: even if low-ionization particles are not recognized, the most energetic clusters are still included and contain the largest fraction of the energy. In addition, the higher threshold hinders the chance for subgroups to form, as is the case with the lower thresholds. Depending on the physics case of interest, thresholds can be set to consider or neglect these low-energy tracks. In any case, the effect of increasing the thresholds has to be taken into account not to include biases in the analysis.

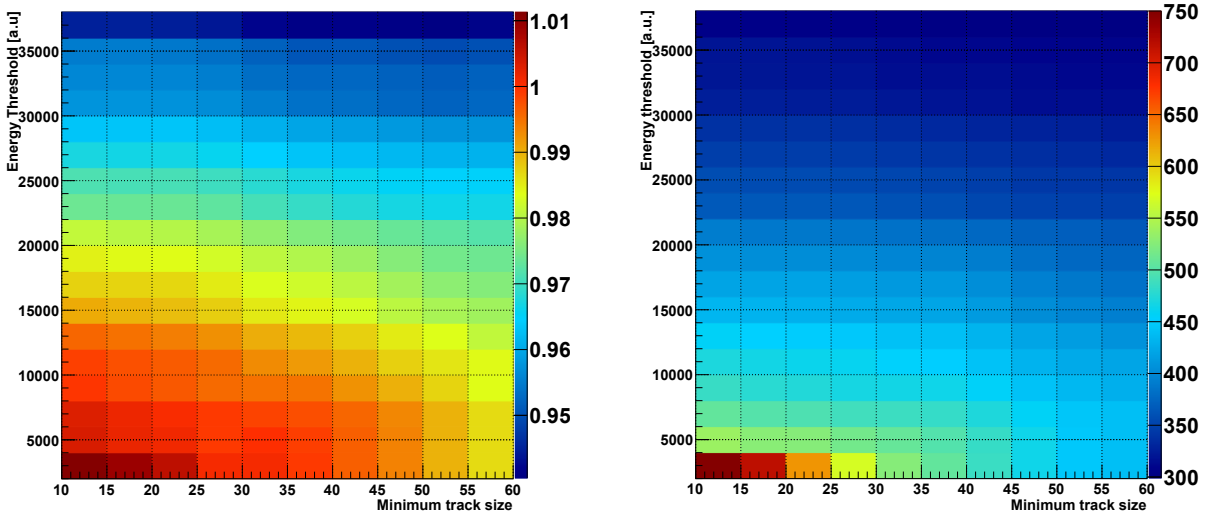
Energy threshold

An other parameter analyzed has been the energy threshold, which on its side has reported the same trend as shown by the analysis of the multiplicity threshold. The same variables as for the multiplicity have been analyzed, and the graph can be found in the repository. In this case, the reference has also been set using the same thresholds as in the previous case, 5000 $[a.u.]$ for the energy and 20 for the minimum multiplicity. Again, a characteristic of the

results is the steeper increase in the number of clusters found while excessively reducing the energy threshold, as it becomes easier for noise clusters to overcome both thresholds.

Energy vs multiplicity threshold

Of particular interest is the combined analysis of the energy and multiplicity thresholds, as the performance of the algorithm is more dependent on the combination of thresholds, rather than on a single one. Figure 5.4 represents how the fraction of energy and the total number of clusters change by varying the two main thresholds.



(a) Fraction of energy included.

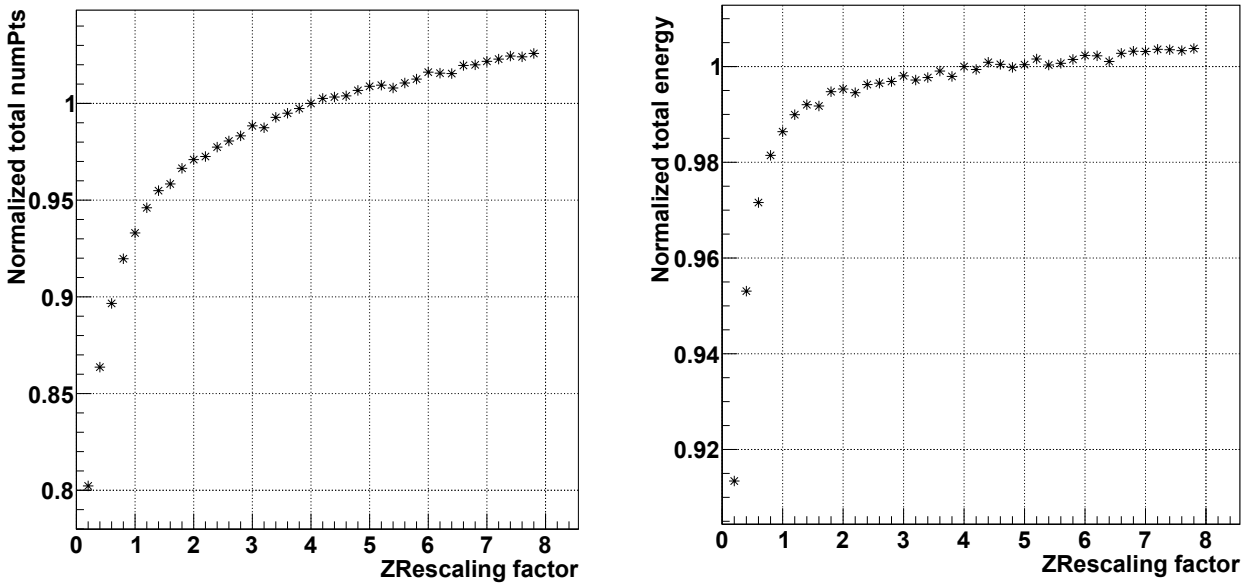
(b) Total number of tracks found in the events set.

Figure 5.4: Comparison of performances varying both the energy and the multiplicity threshold.

As can be seen in Figure 5.4a the variation in the percentage of energy collected is large, especially when comparing the reference value, which in this case is chosen as the combination (20,500). The output obtained using the higher thresholds excludes up to 5% of the energy with respect to the reference. This trend is related to the graph in Figure 5.4a, as the different analyzes are linked to each other. It is also worth comparing the graph 5.4a with the second subfigure, Fig. 5.4b: while the reduction in the collected energy reaches about 5%, the difference in the number of identified clusters quickly reaches 50%. The difference in these percentages is again related to the fact that, increasing thresholds, tracks coming from low-ionizing particles do not fulfill the requirements, while the most energetic ones, as the beam traces, always overcome the thresholds. Once more, the comparison of the graphs in Fig. 5.4 is important to achieve a first-order parameter tuning, understanding the behavior of the algorithm on the whole dataset and comparing global variables rather than single events.

5.1.1 Z rescaling

The last parameter analyzed was the scaling factor on the Z axis. As mentioned above, the reference rescaling factor was chosen so that all axes had the same range ([0,128]). In this specific case, where no calibration is included, such a factor corresponds to a factor 4, range of the SCA and the pad plane dimension. The result of the test procedure obtained by varying this factor is presented in Figure 5.5. Again, the number of points and the total energy reconstructed are compared, normalizing the result to the reference $Z_{resc} = 4$[1]. The number of iterations chosen was 1000, and the previous reference thresholds were used.



(a) Fraction of points included vs rescaling factor. (b) Total energy reconstructed vs rescaling factor.

Figure 5.5: RANSAC performance by varying the rescaling factor.

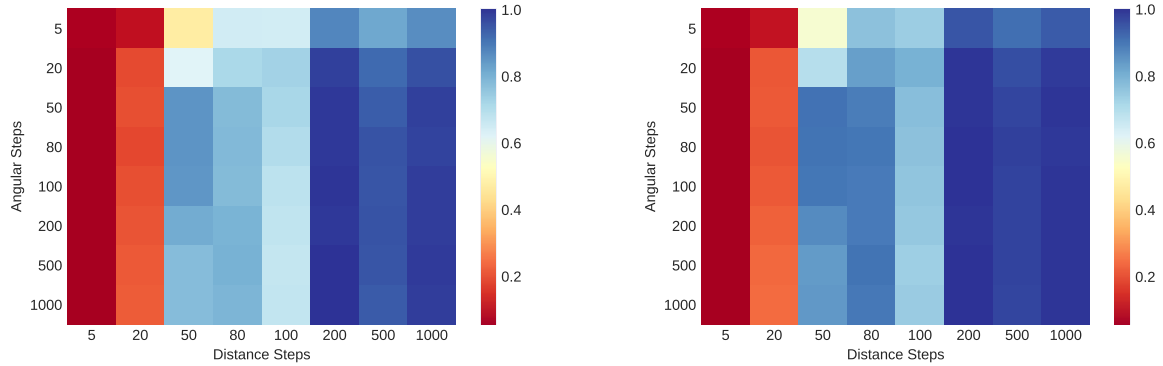
Both graphs in Figure 5.5 show a monotonically increasing trend. Figure 5.5a represents the fraction of points included in the events. The trend found means that shrinking the axis on itself leads more and more points to be grouped together as their vertical distance is reduced. This behavior of the algorithm is intuitive, since the distance considered in $\mathbb{R}^3$ is indeed homogeneous in all three dimensions. If on the Z-axis points are more distant from each other, as in the case of lower rescaling factors, the fraction of the ones included in a cluster may be reduced, not because they do not belong to the same track, but because their sole distance on the Z-axis makes the points being excluded from the inliers set. This feature has been visually observed while analyzing events, and a factor avoiding such an effect was the one that yielded the same interval on the three axis, in analogy with the homogeneous distance chosen. An alternative solution to the problem could consist of a definition of a distance that treats the X, Y, and Z components differently, but the solution of a rescaling

[1]For the sake of simplicity, the rescaling factor is applied on the denominator, so that a rescaling factor of four maps the SCA range [0,512] to the [0,128] range.

factor has been considered more efficient. Nevertheless, if the rescaling factor is increased too much and points are compressed on top of each other, the cluster may also include points which are not linked to the same charge deposit. The difference in multiplicity is again more evident than in the charge case, as the comparison of the two sub-figures shows. In conclusion, the rescaling factor happened to be a very important parameter for a precise clustering result. Tuning such parameters in a way that every axis keeps the same range is useful in case of a homogeneous distance as the one chosen. A wrong choice of the rescaling factor might lead to the exclusion of a significant amount of points and charge, which can even be spot visually while watching the clustered events.

5.2 Parameters Analysis - Hough

Given the analogies between the two algorithms, for the Hough case only the parameters specific to the clustering routine were considered, in particular the number of steps in the angular range and in the distance. The distance range has been taken as the interval [0,200], while the angular range was chosen as $[0, \pi]$. This choice is related to the fact that, after applying the rescaling factor, the maximum distance from the origin is $128\sqrt{3}$, and to the fact that in two dimensions every line direction can be identified from an angle in the chosen range. Once again, the results obtained for both multiplicity and total energy were normalized to the output obtained with 200 angular and distance steps. The results of the scan of the reconstructed energy and the number of points obtained by changing both parameters simultaneously is shown in Figure 5.6.



(a) Fraction of points included vs steps number. (b) Total energy reconstructed vs steps number.

Figure 5.6: Global performance varying number of steps.

The trends in Figure 5.6 show a behavior that is expected in most cases. The axis that appears to have the greatest influence on the outcome is the number of distance steps. As one can see in the two lowest bins, considering 50 and 100 distance steps yields quite a scarce point collection. At the same time, increasing the number of distance steps of angular steps above 200 does not bring any improvement, as the total energy and multiplicity of tracks remain constant. This choice has no basis in the underlying algorithm and is probably related to the data set analyzed. Both sub-figures show the same trend, as in this application the

energy and number of points are not correlated to a specific loss function, as in the RANSAC case. On the other hand, considering the results of the Hough transform, no fluctuation due to random number generation is present, in contrast to the RANSAC case. In this case, it is of particular interest to consider a graph that reports the total clustering time versus the number of steps chosen. Comparing the timestamps at the beginning and end of each iteration varying parameters, one gets the graph reported in Figure 5.7.

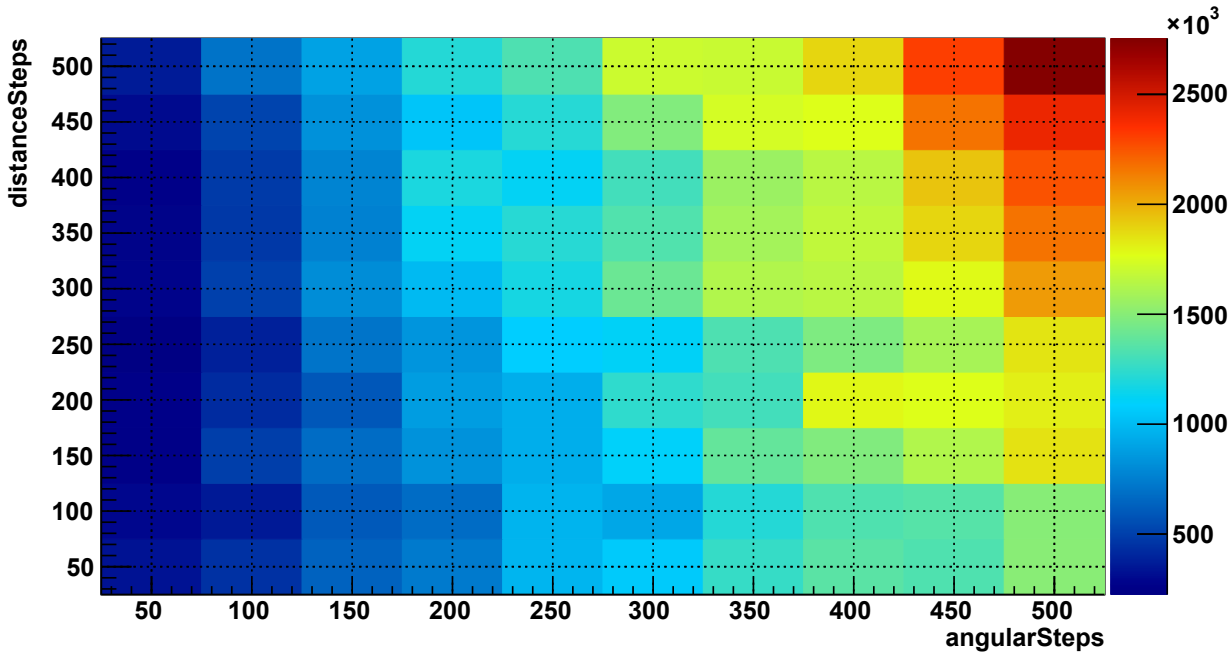


Figure 5.7: Clustering time for the 100-event dataset varying steps numbers.

The pattern of the graph is once more consistent with the expectations. Increasing the number of steps, the number of traces to be included changes, and therefore also the computing time required by the clustering stage. The comparison of the results of the two figures, 5.6 and 5.7, must be taken into account in any application. Although Figure 5.6 did not show any particular pattern once the reconstruction accuracies were considered, Figure 5.7 showed a very precise one. The computing time can then be minimized avoiding accuracy losses and finding a good balance between the number and distance steps.

5.3 Comparing clusters

Once the performances of the single algorithms have been considered, a question arises spontaneously: which one of the two is the best? Many characteristics and peculiarities of the data set have been detected and analyzed through the algorithms and their output. Different general variables, such as the total charge and the total number of points collected, give an idea of the overall performance of the algorithms. However, to achieve a detailed estimate of the clustering accuracy of the algorithms, the subgroups themselves need to be directly compared. Many studies have been conducted on the definition of clustering comparison

metrics [49]. In fact, clustering is a very generic task and can be found under different names in different contexts: partition (in graph theory), numerical taxonomy (in biology), and *unsupervised* learning (in pattern recognition). In particular, the last name refers to clustering as an unsupervised task, referring to the fact that no ground truth is present, and any label that can be trusted or taken as a reference. Thus, no predefined classes are available, nor any relations that should be avoided among the data are available. Such characteristics of the clustering process make the evaluation of the quality of the partitions one of the most important issues in cluster analysis.

5.3.1 Clustering validation metrics

There are mainly two classes of indices that can be used, external and internal, which, respectively, evaluate the goodness of a clustering structure with or without using external information. In the case considered, the information that can be retrieved from internal metrics is not very useful. Indeed, these metrics focus on the measurement of the proximity between instances and clusters, or on how much a cluster is separated from others. Therefore, in this application, external metrics will be considered: many have been studied, and two examples of those metrics are the Adjusted Rand Index (ARI) [50], the Fowlkes Mallows index (FM) [51], and the normalized mutual information (NMI) [52]. To define such indexes, the following notation is considered. Let $U = \{u_1, u_2...u_R\}$ represent the original partition of a dataset, where u_i denote a subset of objects associated with the cluster i. Equivalently, let $V = \{v_1, v_2...v_R\}$ represent the partition found by a cluster algorithm. Then, the following variables are defined as follows:

- Denote a as the number of pairs of objects placed in the same group in both U and V, it can be defined as $a = \sum_{i,j} \binom{n_{ij}}{k}$, where n_{ij} is the number of objects belonging to both subsets u_i and v_j;
- Let b indicate the number of pairs of objects belonging to the same group in U but different groups in V, defined as $b = \sum_i \binom{n_i}{2} - \sum_{i,j} \binom{n_{ij}}{2}$, where $n_i = \sum_j n_{ij}$;
- Finally, let c be the number of pairs of objects belonging to different groups in U and to the same group in V, which can be written as $c = \sum_j \binom{n_j}{2} - \sum_{i,j} \binom{n_{ij}}{2}$.

Then, the Adjusted Rand Index, can be defined as follows:

$$ARI = \frac{\sum_{ij} \binom{n_{ij}}{2} - \left[\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}\right] \Big/ \binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}\right] - \left[\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}\right] \Big/ \binom{n}{2}} \tag{5.1}$$

and the Fowlkes Mallow index can be defined as:

$$FM = a\frac{\sqrt{a+b}\sqrt{a+c}}{(a+b)(a+c)} \tag{5.2}$$

It can be noted that the ARI is the corrected-for-chance version of the Rand Index [53], a simpler index that can be interpreted as a measure of the percentage of correct decisions

made by algorithms.

The NMI, on the other hand, is a symmetric metric that quantifies the shared information between two distributions and provides also a clear indication of the shared information between a pair of clusters. The NMI is defined as:

$$NMI(C,T) = \frac{I(C,T)}{\sqrt{H(C)H(T)}} \tag{5.3}$$

where C represents the random variable that denotes the assignment of groups of points and T is the random variable that denotes the reference class labels on the points. $I(C,T) = H(C) - H(C|T)$ is the mutual information between the random variables C and T. $H(C)$ is the Shannon entropy of C, and $H(C|T)$ is the conditional entropy of C given T. We can observe that $NMI(C,C) = 1$, so that when two sets have a perfect one-to-one correspondence, the metric is equal to unity.

5.3.2 Implementation

As stated previously, external metrics will be considered, even if no true labels are present. To use such metrics, a reference clustering assignation will be chosen using a defined set of parameters for the clustering algorithm. The metrics considered are among the most common ones for comparing clusters and are already included in programming language libraries such as `R` or `Python`. The implementation is based on a `Python` environment and the `scikit-learn` library [54]. Practically, passing to this environment requires data to be reorganized for proper comparison. Indeed, the output of the clustering algorithm is constituted of objects containing the different lines (i.e. clusters), but ideally the input of the scikit-learn metrics would be lists of indexes, defining the cluster number to which the points are associated. After the data have been classified, standard `C++` maps are used to obtain a unique identification of the points. Through maps, single entries are identified using an integer number that hashes the physical coordinates. Given a point $P(x,y,z)$ and the physical boundaries of the pad plane, an identification can be found as:

$$n = x + 128 * y + 128 * 128 * z \tag{5.4}$$

This choice has been made to avoid any complication with the Z rescaling factor, which has to be stored to correctly retrieve the position. In this case, provided that the factor is not changed, the identification is unique. Moreover, the solution of the encoding points with an integer number is particularly useful when using `std::maps`, where the entries are automatically ordered depending on their keys, and also allows one to compare different maps referring to the same point simply by comparing their keys. After such preprocessing, points belonging to a cluster are associated with the running number of the cluster, while points which are not clustered are associated with a null label.

As an example, one can use the metrics to compare how the same algorithm's cluster vary using different parameters. The three metrics are compared by varying the number of loops of the RANSAC, considering the average score over the reduced data set of 100 events. The

reference needed for defining an external metric has been defined in this case as the output of the algorithm with the highest number of iterations, in this case 500. Figure 5.8 shows that the metrics have the same trend and, as expected, are normalized to one. Once compared to the same classification, all metrics yield, as a result, unity. Moreover, the trend of the lines resembles the one shown in Figure 5.2, though in this case the trend extends for a larger number of iterations.

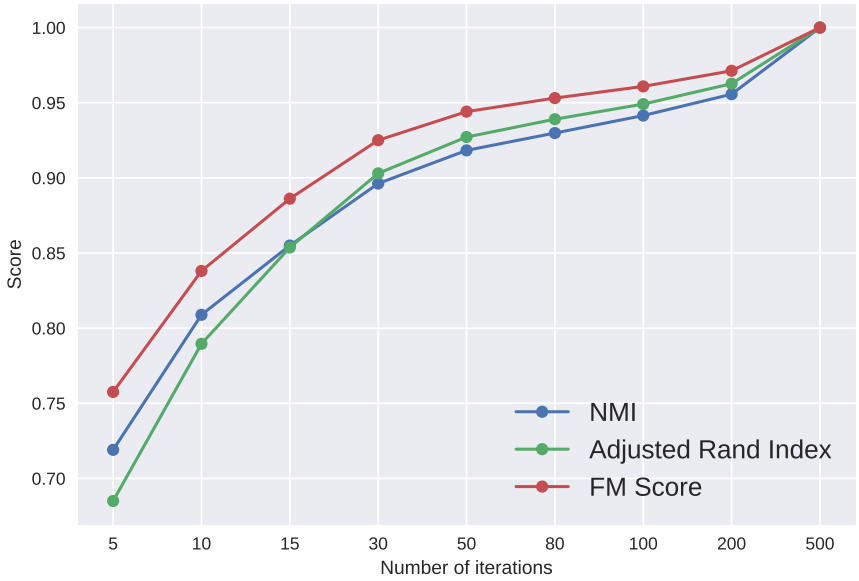


Figure 5.8: Comparison of the metrics varying RANSAC maximum loops number.

To better visualize what the metrics used represent, we discuss and show one event after clustering considering different parameters. The event analyzed has been chosen on purpose to underline one of the most critical cases: the interaction vertex is at the end of the chamber, some traces only consist of a handful of points, and there is a pile-up track next to the one producing the interaction. Figure 5.9 represents the result taken as a reference for the computation of the metrics values. As one can guess also visually, the clustering output is accurate enough, though some points are included in the wrong trace (as in the beginning of the orange one). Figure 5.10, on the other hand, represents the sole XY components in the case in which the RANSAC algorithm is let loop for 15, 50 and 80 times, respectively. As this figure shows, the clustering accuracy improves as the number of iterations increases. For a more quantitative comparison, the results obtained for the different runs are reported in Table 5.1 and Table 5.2. In the second case, for the sake of simplicity, the ratio to total energy is reported for each track, as the total energy collected in the different runs is comparable and only changes of fractions of percentage. The three subfigures in Fig. 5.10 show the change in the clustering accuracy. After only 15 iterations, the tracks are not well identified, as the subgroups clearly include points related to different ions. Already at 50 iterations, the identification of clusters seems to better suit the experimental data (Figure

5.10b). However, looking at tables 5.1 and 5.2, one of the problems that can be spot is the inclusion of points belonging to the pile-up track, which has not interacted. Such points are very energetic and quickly raise the energy of the horizontal track. In case of runs with more iterations, the two beam traces are well distinguished, and at 500 iterations the points that are mistakenly associated to a different trace are just a few. These points are indeed related to those close to the reaction point, and a clear identification of the emission angle is not trivial. However, the classification found after 500 iterations can be said to be satisfactory: also in three dimensions, notwithstanding the technical difficulties related to pile up and to the vertex position, most traces are well recognized and reconstructed.

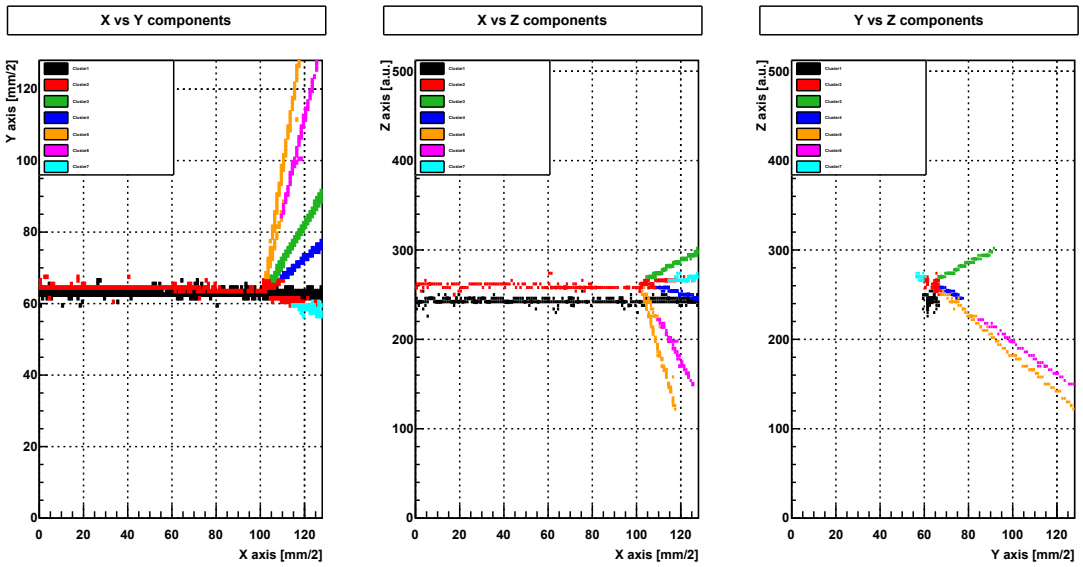


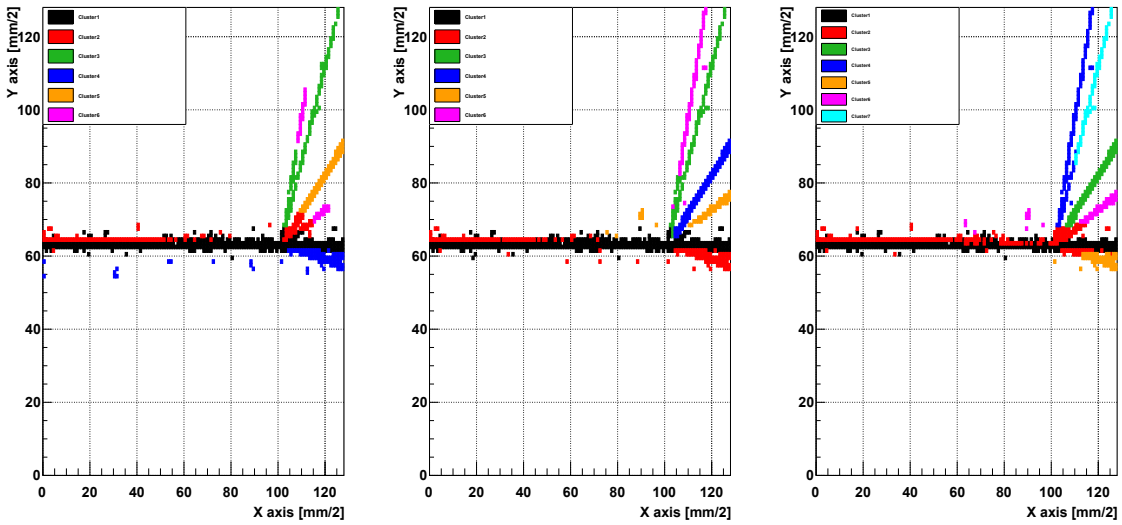
Figure 5.9: RANSAC output using 500 iterations - three projections.

Iterations	NMI	$N_{clusters}$	N_1	N_2	N_3	N_4	N_5	N_6	N_7
15	0.66	6	369	116	98	83	58	34	-
50	0.76	6	378	137	93	80	45	60	-
80	0.86	7	311	181	71	94	49	45	55
500	1	7	295	190	77	44	93	56	36

Table 5.1: Numerical comparison of the multiplicity included in each cluster, varying iterations.

Furthermore, it is particularly interesting to analyze the relationship between Fig. 5.10 and the associated tables, which allows one to perceive what the values of the different scores represent once applied to the ACTAR case. This comparison shows that once a correct reference classification is assumed, the NMIs[2] below 0.75 are not satisfactory, and

[2]NMI is taken here as a reference, but considering other indicators leads to the same conclusions.



(a) 15 iterations.

NMI	ARI	FM
0.66	0.63	0.71

(b) 50 iterations.

NMI	ARI	FM
0.76	0.71	0.77

(c) 80 iterations.

NMI	ARI	FM
0.86	0.87	0.89

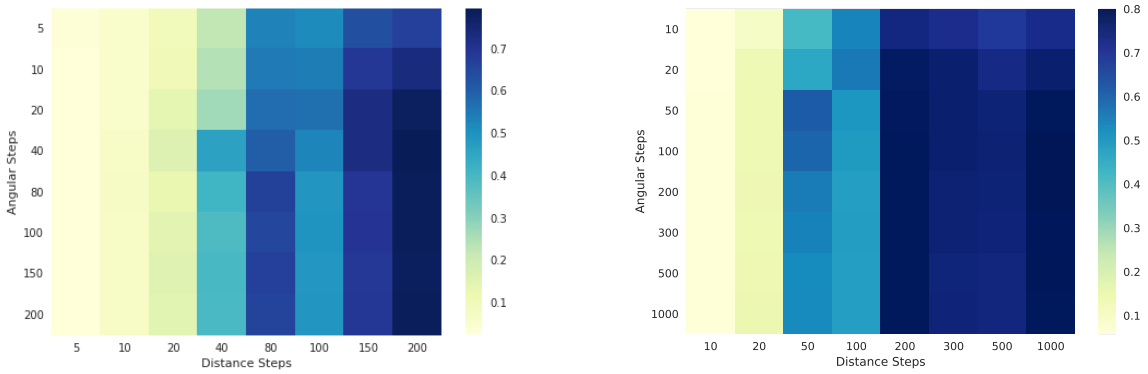
Figure 5.10: RANSAC clustering output using different number of iterations, and relative indicators with respect to the 500-iterations case.

Iterations	$N_{clusters}$	N_1	N_2	N_3	N_4	N_5	N_6	N_7
15	6	77.1	9.6	1.6	6.4	4.0	1.3	-
50	6	77.5	12.9	1.7	5.3	2.2	1.1	-
80	7	66.9	20.6	4.8	1.7	4.4	2.1	0.9
500	7	62.6	25.0	5.1	2.4	1.7	1.0	3.3

Table 5.2: Fraction of energy included in the different clusters.

the clustering of the event must be re-run and improved. On the other hand, the 0.86 value achieved after 80 iterations represents an accurate result. Higher values of NMI would lead to cluster assignations that are closer and closer to the 500-iteration one. Once a reliable configuration of parameters has been chosen, both for RANSAC and for the Hough transform, the performances of the algorithms can be related to the reference classification. Moreover, using these metrics, different algorithms can be compared to each other. In fact, it is enough to define a reference to compare it with the output of another algorithm. This is the case of the analysis shown in Fig. 5.11a.

In Fig. 5.11, the NMI is drawn in a 2D histogram for different choices of angular and distance steps, the main parameters of the Hough transform algorithm. The outcome of the Hough algorithm has been compared to the classification obtained using RANSAC (500 loops) and



(a) Lower number of steps.

(b) Higher number of steps.

Figure 5.11: Comparison of NMI between Hough and RANSAC outputs (500 loops), varying number of steps.

the standard set of parameters. The trend of the heatmap in Fig. 5.11a shows that it is more important to select appropriately the number of distance steps than the number of angular steps. The fact that the heatmap is colored in vertical bands is related to the improvement in clustering performance related to the increase in distance steps. However, such a trend stops for higher numbers of distance steps. As shown in Fig. 5.11b, after 200 distance steps the algorithm is already clustering the tracks at its best and is not able to improve anymore. On the other hand, increasing the number of angle steps above 80 or 100 becomes unnecessary, as in this application no major improvement is spot using a thinner binning, while more steps are always associated with a higher computing time. It is also interesting to compare 5.6b and 5.11, as both show the same trend but with slightly different Z-axis values.

5.4 Event classification through neural networks

5.4.1 Why machine learning

Both algorithms constructed for the clustering of events work fine on the data set. However, these traditional techniques remain extremely computationally expensive and require the simulation of thousands of tracks per event. The pace of the clustering and tracking algorithm makes this step the bottleneck of the analysis, especially compared to the event acquisition rate, that nominally counts up to few kilohertz [25]. Moreover, in case the reaction of interest in the acquired set represents only a small percentage of the data (as often happens), the direct application of the clustering algorithm on the dataset becomes highly inefficient. More efficient analysis tools need to be developed to reduce analysis time, especially when considering large acquisition times (from days to weeks). Having the ability to separate events *beforehand* would lead to a significant reduction in computing time, as clustering could be applied only to events of interest. An *unsupervised* ML algorithm capable of separating data without a priori knowledge of the different types of events would be among the best solutions, given the possibility of applying the technique to data from different ex-

periments [55]. However, such a technique might have limited performances in cases in which many reaction channels are present or in which the events of interest belong to two separate acquisitions. A task that can be more easily tackled is the identification of the presence of an interaction vertex. This task can be faced using a *supervised* approach, labeling events accordingly to whether they contain or not an interaction vertex, and training a model on top of such a set (called a training set).

5.4.2 Classification of events

To train a machine learning model in a supervised setting, it is necessary to associate a label with the input data. Such labeling can be achieved by reconstructing the events and then labeling them on the basis of their characteristics. For example, the multiplicity of particles emitted in the reaction or the depth at which the reaction takes place can be used as labels. Such values, combined with the information coming from the tracking, can provide important information on the underlying physics. However, it can also be considered that a large fraction of the acquired events does not include any reaction. The number of these events depends on the trigger settings and needs to be recorded for normalization. These are present in every data set, and typically include only beam tracks without any ejected particle. However, given the fact that no selection is made, these events also undergo the signal processing and clustering routine, which is slow and computationally demanding. The requirement of the labeling of the data can be seen in some applications, like the one we are discussing, as a downside. This process is indeed based on algorithms that may fail or introduce biases. However, the presence of labels has the advantage that the performance of the algorithm can be monitored and controlled.

5.4.3 Machine learning background

In a machine learning framework, events are called *examples*, and are in general composed of sequences of real numbers, where each entry is called *feature*. In a supervised learning setting, the examples are associated with a label. If the number of possible labels is limited, the learning task is referred to as a *classification* task, which is the case of the applications presented. The goal of the learning process is to reproduce the labels of a training set as best as possible. This is achieved by trying to minimize a *cost function*, defined to characterize the cumulative measure of the classification errors made by the model. Using standard techniques as the gradient descent [56], one can find the values of the model which best reproduce the labels.

5.4.4 Neural networks

Machine learning algorithms comprehend a diverse set of routines capable of handling the most different computing problems. The term *neural network* describes computing systems inspired by the connections between neurons. This class of models has been successfully applied to the most diverse domains, ranging from image recognition to autonomous driving, as well as experimental physics [57]. In these networks, each neuron is a computing unit that can perform operations of different kinds. Neurons are linked to each other to form

a graph. Each neuron accepts many inputs from other neurons, aggregates them in some non-linear manner, and produces an output that can be analyzed by the units in the next computing steps. Each unit contains parameters that are usually randomly initialized at the beginning and then modified during a training process. Neurons are often organized into layers, which are different computation stages. Such architectures enable neural networks to learn complicated patterns, and it has also been shown that a neural network with a *finite* number of units can approximate any continuous function with arbitrary precision [58]. Classic neural network architectures are made up of nodes computing a weighted sum of inputs, plus a bias term. An *activation function* is applied to the sum, and the non-linearity of this operation is the key to the success of neural networks. The output of each computation node is then sent to the next layer. If every neuron in a layer is connected to all neurons in the following one, the architecture is called *fully connected*. In such networks, the information flows from an *input layer*, which receives the input features, to the last layer, the *output* one. If any other layer is present between them, they would be called *hidden layers*; networks with hidden layers are called *deep networks*. The optimization of the network parameters starts with the calculation of the so-called *loss* function, which measures the correctness of the predicted labels with respect to the real ones. Once the performances are checked, a minimization algorithm is employed to find the lowest possible value of the loss. *Back-propagation* algorithms start at the end of a network and change the parameters layer by layer, adjusting the parameters according to the gradient of the loss function. In training complicated models, a single sweep over the training set (an epoch) is not enough to obtain the best results. Many epochs are necessary for the parameters to converge. Once the training is over, the performances are usually checked on a different set, which contains data that the algorithm has never analyzed before. In the end, the final goal of a network is to correctly predict the data belonging to this second set, as the overarching aim of a neural network is to classify events that were not labeled before.

5.4.5 Convolutional Networks

Convolutional Neural Networks are networks specialized in the recognition of patterns in grid-like data, such as time series (1D) or images (2D pixels grids). Starting from an n-dimensional matrix of inputs, CNNs are typically built by combining three types of layers. The heart of these kind of networks are the convolutional filters, which process a set of neighboring values that span the same convolution filter over the input data. This step represents a linear convolution of the input data, sending an n-dimensional input to an n-dimensional output. Many filters are applied to the same input, and the n-dimensional output of a single filter is called a feature map, which constitutes the output of a convolution layer. The second type of layer is the *activation layer*, which acts by applying a non-linear activation function to the feature maps. The computational burden associated with the many feature maps created during the convolution phase can easily diverge, and the application of pooling layers aims to tackle possible memory leaks. These layers typically calculate the maximum, or average over a certain interval size, to reduce the resolution of the matrix while preserving the main features of the input. This kind of network is usually made up of many convolutional layers that can be applied directly to the input data. After these layers, the intermediate results are converted into a vector and, more often than not, fed into a fully connected network,

which leads the flattened vector to the output classification. An example of a typical CNN architecture is shown in Figure 5.12. The advantage of this type of model with respect to

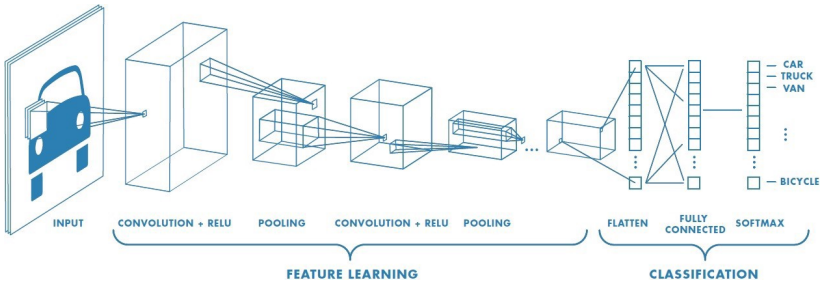


Figure 5.12: Scheme of a convolutional neural network architecture.

other types of network is that the recognition of the patterns is position independent, as the same filters span over the whole input data. This feature is particularly important for image classification tasks. The implementation of CNN on ACTAR data takes advantage of `Keras`, a `Python` library that offers tools for a simple construction of models [59].

Pretrained networks

Developing effective high-performance neural networks from scratch often requires large data sets and long computation times. However, it has been found that models that are trained on particular data sets learn general features and are effective in a variety of tasks. For example, large neural networks that are trained in the ImageNet data set [60], learn to identify lines, edges, and other common shapes that are useful to solve various problems. Thus, oftentimes CNNs are built starting from the architecture and weights of a network pre-trained on such type of data sets. Then, the training process has only to fine-tune few parameters for the specific task, or to tune the parameters of a part of the total network. In both cases, given the fact that the networks are built on top of previous knowledge, the learning and tuning of parameters becomes much faster, as they only need to be adapted to the specific application. An example of such networks is the VGG16 network, a sophisticated multi-layer CNN which has proven to achieve great performance on various data sets [61].

5.4.6 Construction of the data set

In many machine learning applications, more than the choice and tuning of the model, the most delicate step is the creation of the data set. Indeed, the weights of any chosen model will depend on the data that are fed in during the training process, which therefore have a strong influence on the prediction accuracy that a model can achieve. First, it must be said that the conversion to a *computer-vision* problem required the conversion of the data from *hits* to a grid-like image. Given the sampling range of the SCA, encoding an event in its original grid would lead to a matrix of dimension [128,128,512], which is clearly not manageable memory-wise on small clusters. Therefore, it has been chosen to train networks over the projections of the events on bidimensional planes. Two ways of building data sets have been considered: the first one only considered the [128x128] projection of the event on the XY plane, while for the second one entries were built stacking on top of each other points belonging to the

three XY, XZ, and YZ projections. Given the different range of the Z axis, the coordinates of the points in this dimension have been rescaled by a factor of four, so to have a dataset of dimension [128x128x3]. The first type of data set will be labeled as $1x2D$, while the second will be called $3x2D$. This non-efficient storing of the data leads to a limit in the number of considered entries. The data sets were built starting from events coming from the same run (E796, run number 140), consisting of a total of 38787 entries. Given memory limitations, the 1x2D set contained up to 30000 entries and the 3x2D set up to 10000. Once the features are constructed, labels have to be assigned to the events. The procedure chosen exploits the RANSAC algorithm for clustering. Different labels can be associated with events. As a first example, the events were divided into two classes, discriminating on whether an interaction vertex was found in the stored event. RANSAC settings have also been changed during the process to understand how strong the biases introduced can be and how the networks would cope with different labels. Given the same clustering parameters for the RANSAC algorithm, three different labeling principles have been considered to discriminate events in two different classes.

- **All events.** In the first case, the events were labeled using the default routine settings, placing all events in which an interaction vertex was found in the same class regardless of the multiplicity of the ejectiles. Such labeling lead to an unbalanced dataset, as in the 87% of the events an interaction vertex was found.

- **Balanced classes.** This second data set has been built considering the same labeling principle, but the number of entries has been reduced to 9000, with the aim of building a set with the same number of entries in both classes. This choice was made to study how networks performed in a more challenging environment, where looser trigger settings might not lead to such unbalanced sets.

- **Multi-fragmentation.** This labeling principle aimed to define a more complicated classification task. Here, the tracks that were considered beam traces (see Section 4.2.3) were considered as *non-fittable* tracks. This led to the exclusion of such traces from the identification of ones belonging to the interaction vertex. Thus, scattering-like events could or could not be considered as interaction points depending on the scattering angle between the beam and the ejectiles, and on the interaction depth. This labeling choice led to a slightly unbalanced data set, where in 57% of the events no vertex was found. The name of this dataset comes from the fact that the events containing many ejected traces (as happens in multifragmentation events) were labeled as containing an interaction and were not affected by avoiding the fitting in the beam traces.

- **Normalized data set.** This last data set was constructed after considering the performance of the networks over the previous data sets. The purpose of this data set is to simplify the learning task. Here, the only projection on the XY plane was considered, and the pads hit were associated with a value of one, regardless of the associated energy deposit, whereas those not hit were associated with a zero. Such a case is particularly interesting because it is not considering the energy parameter, emulating a classification that could be directly placed on the electronic chain.

It must be noticed that the choice of events and of the labeling criterion is probably the most delicate step in the whole process, as the different choices might introduce biases in the final predictions. Anyway, once this has been taken into account, one can start training machine learning models on the data. The models have been trained over all data sets that have been divided between a training set and a validation one with a 85/15 ratio. In the end, no test set was considered.

5.4.7 Application to ACTAR events

Model definition

Two different types of network were applied to the ACTAR data, including both models trained from scratch and pre-trained ones. A typical network architecture is shown in Fig. 5.13. Using a network employing three convolutional layers and no dense layers allowed in most combinations of kernel and pool sizes to achieve the same results on the binary classification task. In all networks tested, the convolutional layers were followed by an activation layer whose associated function was the ReLU; max and average pooling were tried, but no significant performance difference was found. For simplicity, max-pooling was in the end preferred. Padding, a technique that lets the filters span also on the edges of the input matrix, was employed, and to try to avoid overfitting, a regularization of the weights of the kernels was imposed through an L2 regularization. In the training phase, the loss function used was the *categorical crossentropy*, defined in Equation 5.5.

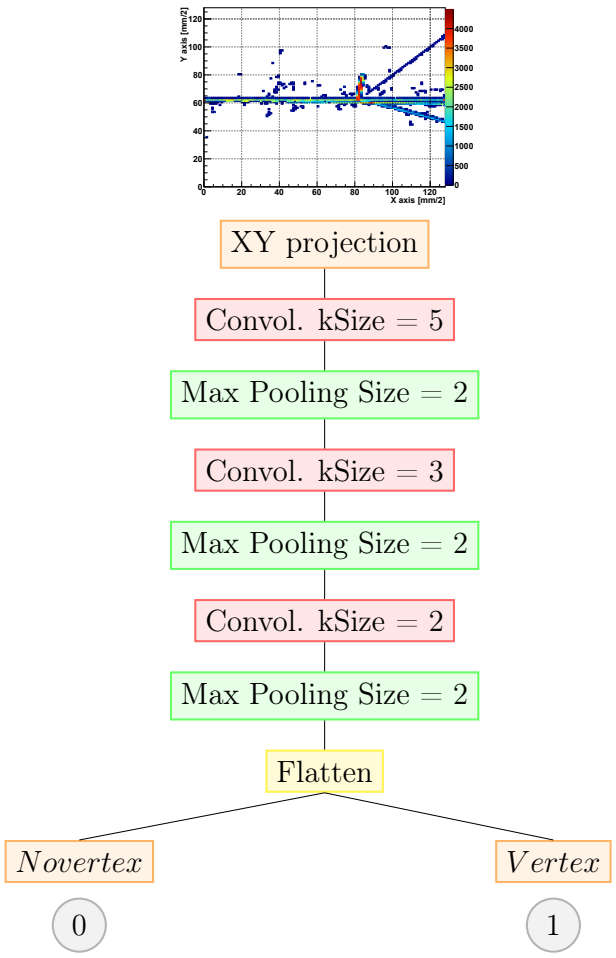


Figure 5.13: Structure of a convolutional network.

$$C(p,q) = -\sum_i p_i \cdot \log(q_i) \tag{5.5}$$

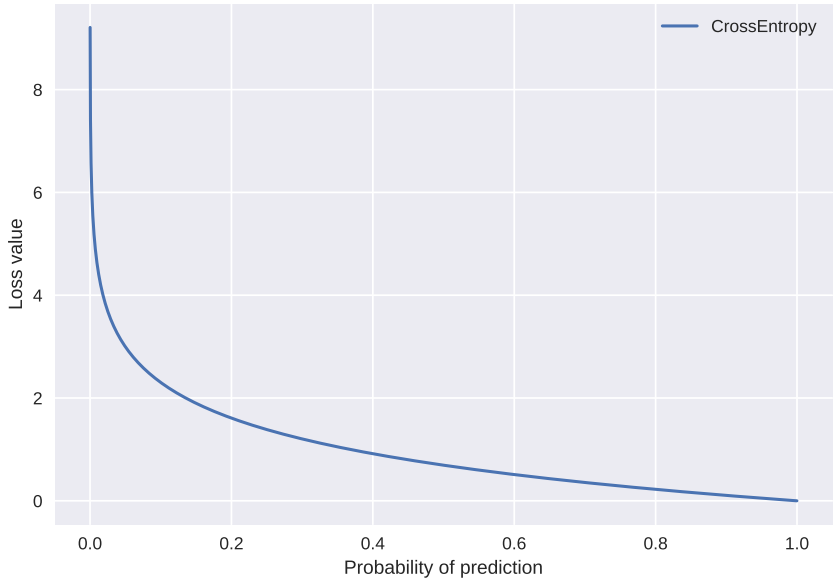


Figure 5.14: Categorical crossentropy varying the probability of correct prediction.

This function is particularly useful in classification tasks, where a single p_i is not null and the loss produces a divergence value when the prediction probability is very far from the actual label. The algorithm used to minimize such a loss function is *Adam*, included among the default ones in the `Keras` library. Learning rates, the parameters determining how much the value of the weights changes after each iteration, of both 0.01 and 0.001 acted fine on the data sets, leading to reasonable training times and producing smooth training curves.

In the case of the figure in 5.13, the number of filters per convolutional layer was set to [128,128,64], respectively, but many different configurations of feature maps allowed to achieve comparable accuracy. As often done in classification tasks, the last layer before the final classification used a *softmax* activation function, which guarantees a classification-like output, where a single neuron in the output layer indicates the outcome of the classification. An other approach used for model construction considered pre-trained models, an example of which is the VGG16 network. The weights applied to this network were those from the training over the ImageNet set, and a dense network was applied at the end to retrieve a binary classification output. It is also interesting to compare the complexities of the two types of model. An architecture like the one in 5.13 brings $2 \cdot 10^5$ parameters that were effectively trained on the dataset. On the other hand, the sole VGG model has $14 \cdot 10^6$ parameters. It is then clear that the training of such networks takes a different computational effort, and this is the reason why networks such as VGG are used as pre-trained ones, training only a handful of parameters that link the VGG16 latent space to the output.

5.4.8 Results

The first results presented will be those from the *multi-fragmentation* events. This data set, which labeling principle was the most challenging one, led to lower accuracies compared to the others. However, the influence of the definition of the dataset is, in this case, very clear and is worth being studied.

Multi-fragmentation

The application of both types of network, pre-trained or not, led to comparable results in terms of precision. The training has been performed for a number of epochs between 30 and 80. In most cases, overfitting was avoided by including early callbacks and fine-tuning network parameters. Moreover, it was always monitored by looking at the accuracy and loss curves. As an example of the trend of these two curves throughout the training process of the network presented in Figure 5.13, while the one applied to the 1x2D data set is shown in Figure 5.15. As usually happens, the accuracy reached on the training set is higher than that of the validation set, but the fact that the two curves have the same overall trend hints that no overfitting is present. In this application, the accuracy tends to flatten already after 20 epochs. From this point on, the iterations do not bring about major advantages in the validation set, while only the training accuracy increases. In Table 5.3 the different scores

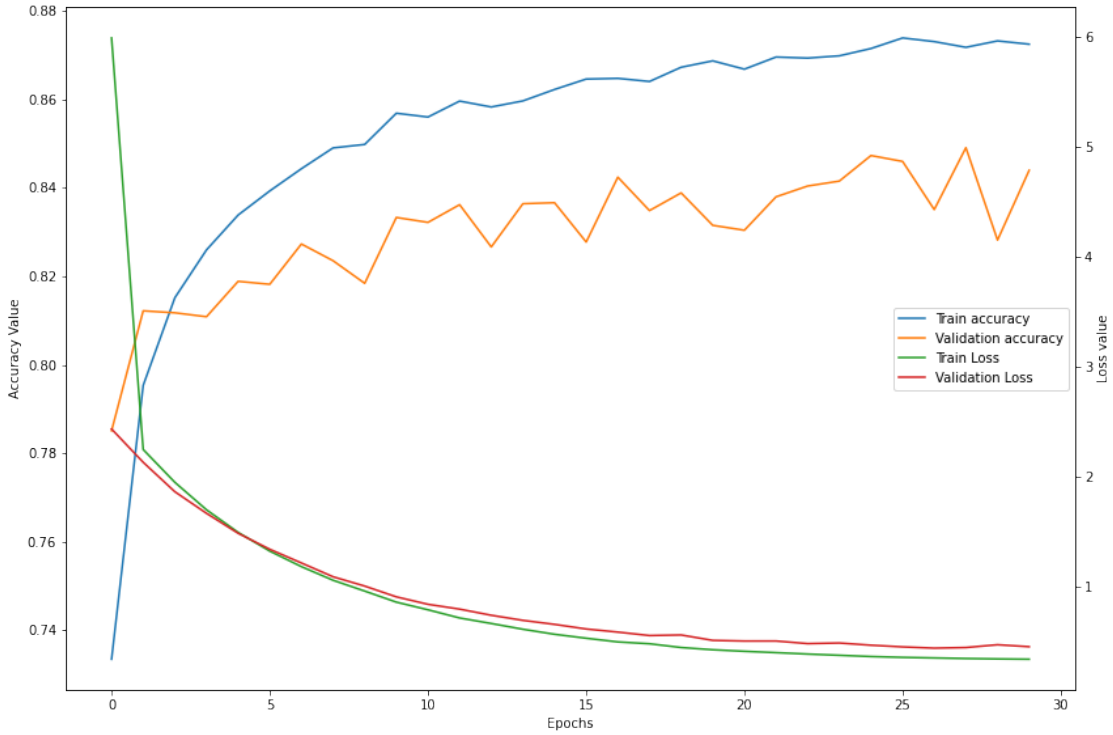


Figure 5.15: Accuracy and loss training curves for the cNN in figure 5.13, trained on the 1x2D multi-fragmentation data set.

achieved by applying different networks to the multi-fragmentation data set are summarized. In the model column, *cNN* refers to the network in Fig. 5.13, *VGG* to a VGG16 model to

which only a dropout layer has been applied after flattening and *VGG Dense* to a VGG16 model in which both a dropout and an additional 25-large hidden dense layer were added before the output.

Model	Parameters	Data set	Accuracy	
			Train	Validation
cNN	2.16e5	1x2D (30000 entries)	91%	84%
cNN	2.16e5	1x2D (10000 entries)	97%	82%
cNN	2.23e5	3x2D (10000 entries)	100%	82%.
VGG	0.16e5 (15e5)	3x2D (10000 entries)	77%	76%.
$VGG\,Dense$	2.20e5 (15e5)	3x2D (1000 entries)	57%	57%.

Table 5.3: Summary of results of different cNN training.

The different networks, though trained on different data sets, show interesting and different behaviors. The first line in Table 5.3 shows the results obtained in Figure 5.15, but when the training is allowed to run for more epochs. In this case, the validation accuracy is above 84%, a value already obtained after a few epochs, while the training accuracy constantly increases to 91%. This effect is clearly related to overfitting, although even in the last iterations the validation accuracy remained around the previously reached value. Tuning the optimizer learning rate and the number of epochs also in this case is enough to avoid unnecessary long training phases. Training the same network over the reduced 3x2D data set led to a similar result in terms of validation accuracy. Indeed, the validation accuracy reached 82%, while in this case the training accuracy reached 100%. For the sake of a fair comparison, the same network has been trained over the 1x2D data, reducing the available number of samples (the 3x2D dataset contains 10000 entries). In this case, the validation accuracy did not show any drop, while the training reached 97%. Comparing these trainings is useful to understand the effects of both the dimension of the input features (1x2D or 3x2D) and of the data set size on the training of the networks (30000 or 10000 entries), given the fact that the same network is employed. Therefore, it can be noticed that using a larger amount of data still leads to a training accuracy far from perfection, though it leads to an overall more robust network. However, training the same network on a reduced number of data leads to a more overfitting model: the training accuracy reaches 97% while no gain in the validation metrics has been found. This effect is expected from the different ratio between the network parameters and the number of training samples, which also explains the fact that using 10000 3x2D entries the network fitted perfectly the training set: being the number of samples equal, the information contained in the latter data set is enriched by two more projections. If already in the 1x2D case the network was almost perfectly fitting the training data, such a result is expected for a more complete data set. However, it is interesting to notice that no validation accuracy chance is found between the two datasets. This, on the one hand, might be related to the overfitting of the latter network; in any case, it points to the fact that the sole XY projection is sufficient to predict whether a vertex is present in the event analyzed. This feature is mainly related to trigger settings. In fact, in the data from the analyzed experiment, most

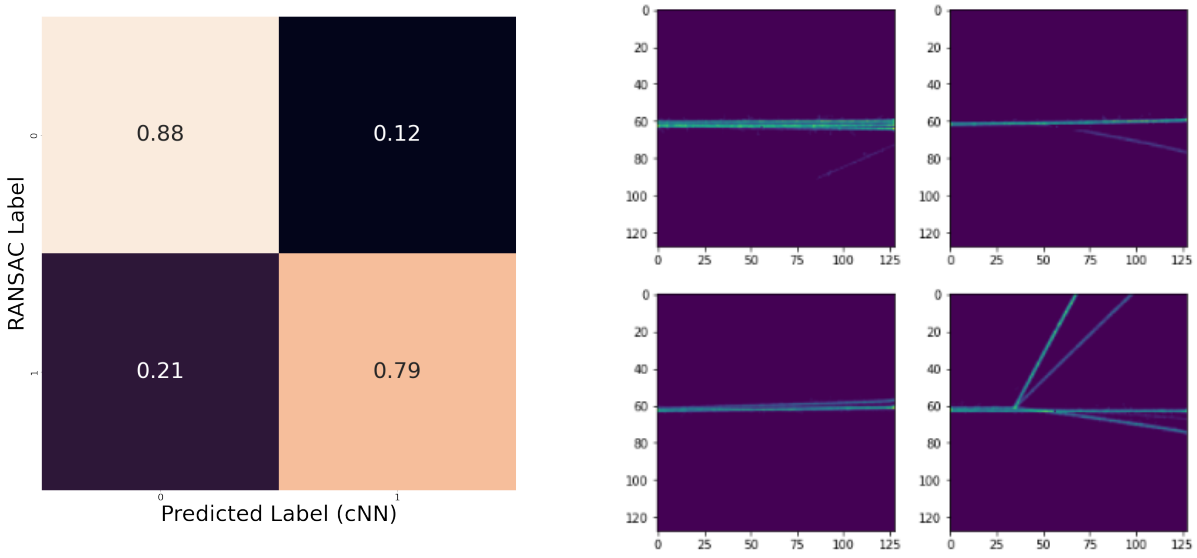
events containing an interaction vertex were stored when a trigger signal came from one of the silicon detectors placed on the sides of the detector. For example, such a trigger setting helps avoiding most events in which ejectiles have projections on the XY plane that overlap on the beam area. On the other hand, using a more complicated model such as the VGG one leads to shattered results. In Table 5.3, the scores obtained with two different VGG networks are presented. The number of trainable parameters is represented without brackets, and in both cases represents the parameters associated with the dense network attached to the model. The numbers in brackets represent, on the other hand, the non-trainable parameters associated with the VGG16 default model, taken from the training on the ImageNet set. The simpler network with a smaller number of parameters reached an accuracy of 77% on the validation set, which is not far from the case of the network trained from scratch, but it is a little behind. On the other hand, the VGG+Dense network failed the classification task. Adding a hidden dense layer got the network stuck in a minimum and no improvements were seen in the training phase: the 57% accuracy found reflects the ratio of the different classes in the data set. In this simple application, more complicated models did not lead to any improvement in the predictions. The simplest model, the better the results.

More insights on the network classification can be studied by splitting the predictions into classes and analyzing the confusion matrix that compares the different subsets. Figure 5.16a shows the confusion matrix of the predictions on the validation set of the *cNN* in Fig. 5.13 after being trained over 30000 1x2D entries. The confusion matrix shows that in most cases the labels are correctly classified, as already noticed considering the global accuracy. However, splitting the predictions into four classes showed that events that did not contain an interaction vertex (labeled with a 0) were better predicted than those containing an interaction vertex (labeled with a 1). This feature has been reported in all trained networks and could be related to the larger fraction of events present in the multi-fragmentation data set labeled as not containing an interaction vertex.

Even if the best accuracy found was about 85%, in principle better overall results could be expected after tuning the parameters of the networks. However, the change in network architectures and parameters did not lead to any improvement in validation accuracy [3]. At the same time, this limited accuracy found reflects the choice previously made to partially exclude elastic scattering events. Therefore, it has been decided to compare the predictions to the events themselves, evaluating, starting from the confusion matrix in Figure 5.16a, the projections of the events correctly and incorrectly classified by the network. Typical events that have been found are represented in Figure 5.16.

Figure 5.16b shows why in most cases the accuracy of the network did not exceed the 85% limit. This is mostly related to elastic-like events that are *misclassified* during the labeling phase. Indeed, the networks did not just copy the RANSAC classification, but also classified elastic-like events as containing an interaction vertex, as it actually should be. In the upper right subfigure in Fig. 5.16, the scattering event is recognized by the network, while the presence of a straight beam trace would not let the RANSAC algorithm reconstruct the interaction. On the other hand, the bottom-left subfigure shows an event that, from the XY projection, does not seem to contain a vertex. However, during the labeling process,

[3]For the sake of brevity, the different trials will not be displayed as they would not add to the discussion.



(a) Confusion matrix of the accuracy of network in Fig. 5.13 trained on 1x2D data.

(b) Examples of events associated to the confusion matrix in Fig. 5.16a.

Figure 5.16: Confusion matrix and classification example associated to the four different boxes.

a vertex was found. This might be related to the trace that can be seen slightly deviating from the beam axis, and might be related to an ejectile whose trace on the XY direction is only slightly deviating from the beam area, but which might have a higher scattering angle in three dimensions.

All events

In this section, the results of the networks trained on the data set labeled considering the most comprehensive criterion are presented. To compare the results with those presented in the previous paragraph, the network architecture will not be changed, in order to emphasize the effects of the labeling on the network performances. Once again, the results obtained are summarized in Table 5.4, using the same conventions as defined before. In this case, due to the fact that the labels are unbalanced, in the validation accuracy column the True-True and False-False accuracies will be specified (corresponding to the diagonal values of the confusion matrices as in Fig. 5.16a). The situation found in this case is quite different from the previous one. First, it can be noticed that using this looser labeling criterion, the cNN architecture taken as a reference achieves an optimal classification on the complete 1x2D data set. In fact, even if the training accuracy reaches almost 100%, the same result is achieved in the validation sets. Furthermore, after training over the entire data set, the predictions on both classes are well reproduced, having an accuracy of 95% for both subsets. Normalizing the accuracies to the actual number of entries is important, especially in the case of a very unbalanced data set. Adding to this, it can be seen that training the same network over a

Model	Parameters	Data set	Accuracy Train	Accuracy Validation (00/11)
cNN	2.16e5	1x2D (30000 entries)	98%	95% / 95%
cNN	2.16e5	1x2D (10000 entries)	98%	68% / 99%
cNN	2.23e5	3x2D (10000 entries)	79%	0% / 100%.
VGG	0.16e5 (15e5)	3x2D (10000 entries)	80%	0% / 100% .
VGG Dense	2.20e5 (15e5)	3x2D (10000 entries)	87%	0% / 100%.

Table 5.4: Summary of results of different cNN training (*All Events* dataset).

reduced data set with 1/3 of the entries leads to a reduction of accuracy in the prediction of the less represented class. This effect can be pictured thinking that using an unbalanced dataset, the network did not have enough '0' on top of which it could train its predictions. However, the network concentrated much more on the 1/1 predictions, reaching a higher accuracy than in the previous case. Finally, it can be seen that in this case all networks trained on the 3x2D dataset were only able to classify points as having an interaction vertex. This classification is clearly wrong, as all events are classified with the same outcome. In conclusion, it can be once again stated that, given the actual trigger settings, the sole XY projection is, in most cases, enough to predict whether an interaction vertex is present or not. Moreover, it has been shown that the size of the data set is important during the training phase, especially when the labels are strongly unbalanced. This feature is demonstrated by the first cNN trained on the 30000-entries dataset, whose training takes less than 30 epochs, for a training time of around 10 minutes on a regular `Kaggle` [62] notebook.

Balanced classes

Given the trigger conditions set in the experimental setup, the ratio between events that contained or not contain an interaction vertex was quite high, which ultimately led to the construction of an unbalanced data set. Even if on such a dataset an excellent accuracy was found for both classes, the trigger conditions might be changed, and so the ratio of analyzed events containing or not an interaction vertex. Therefore, a reduced data set was built to check the performance of the network over data sets with a different true/false ratio. Due to the results achieved in the complete data set, the 3x2D data set has not been considered. For the sake of simplicity, 9000 entries have been selected starting from the whole set, considering every acquired event not containing a vertex and adding up events until reaching an equal number of entries with and without a vertex. Again, only the results coming from the usual network will be presented. The results are shown in Figure 5.13 and summarized in Table 5.5.

Once again, even in the balanced case, despite the reduced number of entries, an almost perfect classification is achieved. Such a result might be expected from the fact that in the unbalanced case the same result was reached, but it is interesting to see that this also happens in the case of a balanced dataset having a lower number of entries. Moreover, the

Model	Parameters	Data set	Accuracy Train	Accuracy Validation (00/11)
cNN	2.16e5	1x2D (9000 entries)	98%	96% / 97%

Table 5.5: Single cNN training result (*Balanced* dataset).

sole XY projection turned out to be sufficient to achieve a reliable estimate of the presence of an interaction vertex.

Normalized data set

The normalization of the data set was performed for the 1x2D projection, as the idea of avoiding signal processing does not require the consideration of the third component. Thus, both the unbalanced data set and the *All Events* data sets were normalized. The results found, applying once more the network described in Fig. 5.13, are presented in Table 5.6.

Model	Parameters	Data set	Type	Accuracy Train	Accuracy Validation (00/11)
cNN	2.16e5	1x2D (30000 entries)	Unbalanced	96%	84% / 97%
cNN	2.16e5	1x2D (10000 entries)	Unbalanced	95%	94% / 95%
cNN	2.16e5	1x2D (9000 entries)	Balanced	96%	98% / 92%

Table 5.6: Single cNN training result (*Normalized* dataset).

As Table 5.6 shows, the cNN manages to achieve a good classification over the three sets. In the first case, using the 30000-entry dataset, the unbalanced labeling also leads to an unbalanced accuracy on the labels. However, reducing the entries improved the null-null prediction. Such a case is not expected, but at the same time, it is seen that the balanced set also leads to accurate results. Training using this data set proved to be successful, but pointed out the need to take special care in the case of very unbalanced data sets. This result is probably the most interesting among the ones presented in this thesis work. Indeed, these trainings proved that under simple requirements, it is possible to predict the presence of an interaction vertex without needing to go through signal processing and cluster analysis. Therefore, in principle, an NN could be included in the electronic chain to help acquire the only events that included an interaction point. Indeed, the prediction velocity of such a network is on the order of kilohertz, which is the acquisition rate that a detector like ACTAR can withstand.

5.4.9 Vertex position studies

The analysis of the misclassified events in Figure 5.16 opens many other questions about the possible correlation between the classifications and physics variables, as the ejectile multiplicity or the interaction depth. First, the possible presence of bias in classification has been studied, comparing the positions of events that were correctly and incorrectly classified. The result of such an analysis is shown in Figure 5.17. However, the graph shows that there is no

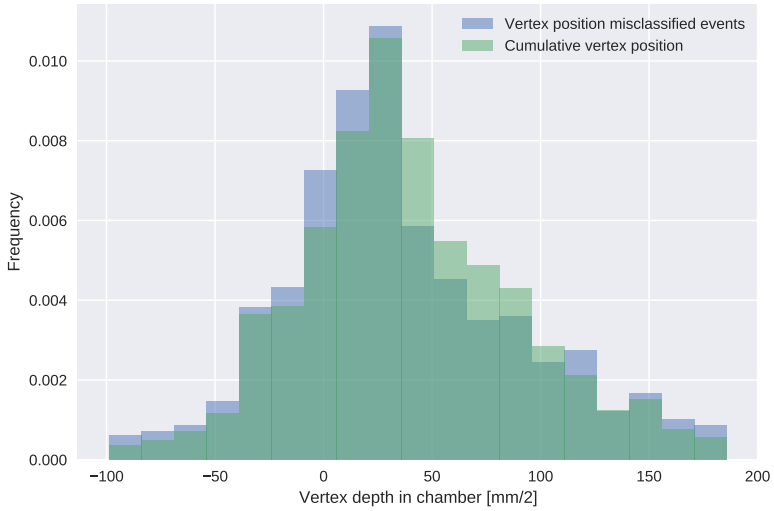


Figure 5.17: Comparison between the depth of the interactions of the whole set and on the set of unclassified events.

relevant difference in the distributions between the total distribution and the misclassified events one. The slight difference in the distributions might be related to the different cross sections (physical probabilities) of the processes involved. Trigger settings and the efficiency of the reconstruction algorithm itself might also play a role. Indeed, as shown previously, the position of the vertex might be an important variable in the definition of the efficiency reconstruction, since traces having the same emission angle but different interaction depths face different challenges in the reconstruction phase. This becomes interesting as a general study, as in every experiment the efficiency of the reconstruction techniques has an impact on the measured yields of events. To obtain an estimate of this efficiency, the Geant4-based ACTARSIM code was run [63]. This tool allows the simulation of the ACTAR TPC detector, but it has some limitations. In particular, noise generation is not included by default in the simulation, and the only reaction channels that can be simulated are elastic and inelastic scattering. A reference data set composed of 10000 events has been built, simulating inelastic scattering events trying to emulate the experimental settings described in Chapter 3. Figure 5.18 shows the three projections of a simulated event.

Although similar, the simulated events present substantial differences with the experimental ones. The absence of noise and pile-up are both factors that dramatically simplify the clustering task. In any case, these simpler events can be used to study the accuracy of the

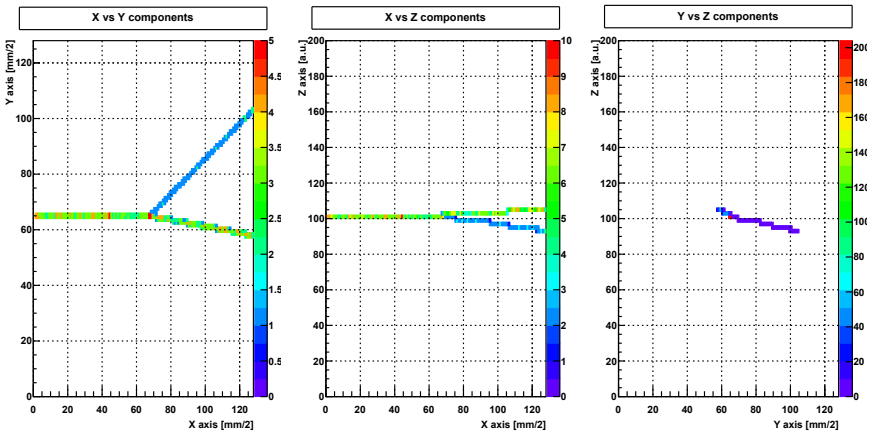


Figure 5.18: Projections of a simulated scattering event through ACTARSIM.

vertex reconstruction, since the interaction depth is known in simulations. To study the vertex reconstruction efficiency, RANSAC was run over the whole dataset, comparing the depth of the reconstructed vertices with the position of the true (simulated) ones. The latter distribution was chosen to be flat, to achieve an easier comparison with the reconstructed one. Figure 5.19 shows the comparison between the interaction depth distributions of the simulated, reconstructed, and unreconstructed events. The last case is represented by the events in which RANSAC (the algorithm used) was unable to find a vertex.

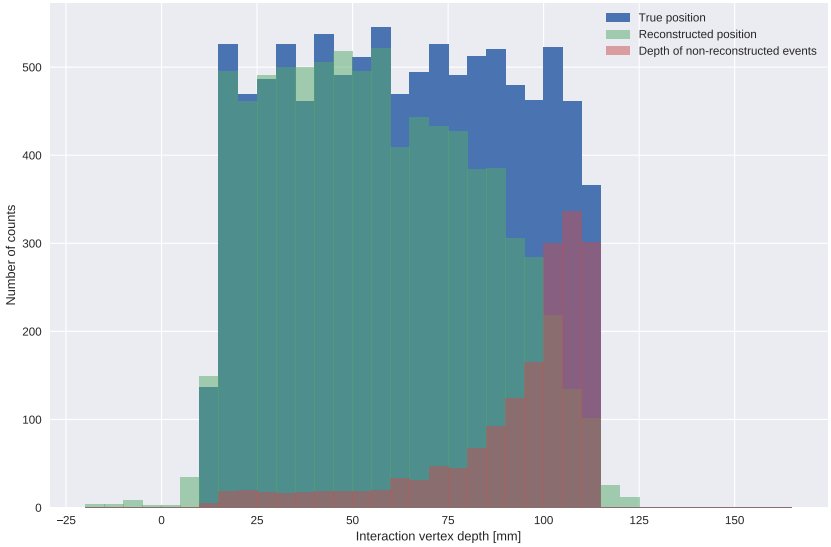


Figure 5.19: Comparison between the distributions of the simulated and reconstructed vertices.

The two distributions are similar. However, it should be noted that the number of entries in

the reconstructed graph is about 8700, and therefore the overall reconstruction efficiency is approximately 87%. However, events that are not correctly classified belong in most cases to interactions that take place at the end of the active area. Such a feature is expected, as once the interaction vertex lies at the end of the chamber, it might become more complicated to reconstruct tracks with low emission angles. In conclusion, even if such a study did not bring any further improvement to the interpretation of the NN classification, it has shed some light on the efficiency of the whole reconstruction algorithm in the case of elastic scattering events. In future experiments designs, the efficiency of the reconstruction algorithms must be taken into account to not lose statistics in the analysis phase. Moreover, such analysis has to be repeated for each reconstruction codes, and the results may vary depending on the physics process considered: one might expect different outcomes related to the different ejectiles distribution, the detector condition (influencing the range of the ejectiles), or the reaction mechanism that wants to be measured, as the multiplicity of the ejectiles might also play a role.

Chapter 6

Conclusion and outlook

Nuclear physics is still today a very active research area in basic physics. The challenges faced in the quest to study nuclei far from stability need to be tackled with the best technological tools, both on the hardware and software side. As presented in the thesis, active target detectors represent one of the means used to undertake new experimental endeavours. Thanks to their high dynamic range, granularity and resolution, such devices are very versatile and can be used in different experiments, such as scattering and implantation-decay experiments. ACTAR is an active target detector that has already been used successfully to perform experiments to study nuclei from different regions of the nuclear chart. However, the challenges that are faced from the software side have never been tackled with a systematic approach and no library is currently present for a definition of a standard routine analysis. Moreover, given the complexity of the acquired events, the amount of data collected during experiments can overcome the possible storing capacity of the present infrastructure that could cause an experiment to stop before the planned beam-time goal is reached. This thesis work wanted to address these issues, creating standards for the clustering and reconstruction of events, validating metrics for comparing clustering algorithms, and finally applying neural networks to classify events and reach a fast event selection that could be used online.

In particular, this thesis work started with the participation run in March 2022 at GANIL with ACTAR, taking part in the preparation of the apparatus and the shifts. Then, the analysis part has started, focusing on merging two already present clustering algorithms in a single framework. The goal of the designed framework was to possibly include any clustering or event reconstruction routine, creating classes inheriting from a parent that had minimum requirements, taking the singleton framework as a reference. Once this framework was built, both the Hough transform and the RANSAC algorithms were tuned to the data coming from the experiment in which I participated. The beam clustering and exclusion of the points were designed for both algorithms and implemented directly in the routines, and their implementation needed some technicalities. Technical challenges were faced when modifying the code, and a deep understanding of the action of every part of it was necessary to avoid any side effect. Then, the influences of the parameters of both algorithms and of the rescaling factor have been analyzed in both cases, comparing the times. Furthermore, to achieve an in-depth comparison of clustering performance, some metrics were implemented, such as the

NMI, the ARI, and the FM index. The values of such metrics were compared with those of the clusters themselves to achieve a proper interpretation of the values. However, one can conclude that, in the analyzed case, more accurate metrics could be defined. Indeed, the metrics used were only considering the assignation of the entries, while the energy of the single points or of the reconstructed traces, far more important observables, was not considered in the clustering performance evaluation. Thus, different metrics and validation techniques should be built, as not many are present, and validated on ACTAR-like data. The last part of the work consisted in the application of a machine learning model to the data to speed up the classification and reconstruction times. The first step consisted of a mandatory simplification of the problem, and the sets were constructed from a reduced set of data. Such reduction was anyways not enough to address the problem with reasonable but limited computing resources. Thus, the projections of acquired events have been considered as entries in the model, building two different datasets. The first one considered a single projection, while the second one considered three bidimensional ones. These entries were then labeled accordingly to whether an interaction vertex was found in the reconstruction phase. The performance of both pre-trained and models defined from scratch was tested on the two datasets. The accuracies found using the logic labeling principles reach almost a perfect accuracy. Furthermore, it was also shown that given the experimental settings, a single projection of the event was enough in most cases to achieve an accurate classification. On top of that, networks were able to classify correctly the events in both the balanced and in the unbalanced set cases. In conclusion, the events were also correctly classified when the energy values were not considered, allowing in principle the implementation of neural networks in the acquisition chain to help discriminate whether an event is worth being acquired or not.

As shown, the work on the software side of ACTAR is far from end. The implementation of cutting-edge technologies can clearly bring an advantage to the detector operation itself, managing to select data before being stored in memory and selecting only events of interest. The definition of custom metrics can help users to achieve a better and better understanding of the outcome of the clustering analysis. The metrics analyzed so far give a quite accurate but general view on the performances, but considering energy can lead to more effective conclusions.

Bibliography

[1] E. Rutherford, "The structure of the atom," *Nature*, vol. 92, 1913. [Online]. Available: https://doi.org/10.1038/092423a0

[2] D. Bazin, "Nuclear physics: symmetrical tin," *Nature*, vol. 486, no. 7403, p. 330—331, June 2012. [Online]. Available: https://doi.org/10.1038/486330a

[3] K. S. Krane, *Introductory nuclear physics.* New York, NY: Wiley, 1988. [Online]. Available: https://cds.cern.ch/record/359790

[4] S. Bacca, "Structure models: From shell model to ab initio methods," *The European Physical Journal Plus*, vol. 131, 2016. [Online]. Available: https://doi.org/10.1140/epjp/i2016-16107-6

[5] G. T. Garvey, W. J. Gerace, R. L. Jaffe, I. Talmi, and I. Kelson, "Set of nuclear-mass relations and a resultant mass table," *Rev. Mod. Phys.*, vol. 41, pp. S1–S80, Oct 1969. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.41.S1

[6] M. G. Mayer, "On closed shells in nuclei. ii," *Phys. Rev.*, vol. 75, pp. 1969–1970, Jun 1949. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.75.1969

[7] O. Haxel, J. H. D. Jensen, and H. E. Suess, "On the "magic numbers" in nuclear structure," *Phys. Rev.*, vol. 75, pp. 1766–1766, Jun 1949. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.75.1766.2

[8] R. D. Woods and D. S. Saxon, "Diffuse surface optical model for nucleon-nuclei scattering," *Phys. Rev.*, vol. 95, pp. 577–578, Jul 1954. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.95.577

[9] K. Heyde, *Basic Ideas and Concepts in Nuclear Physics, An Introductory Approach*, ser. Fundamental and applied nuclear physics series. Taylor & Francis, 1999. [Online]. Available: https://books.google.fr/books?id=EAGXzQEACAAJ

[10] S. G. Nilsson, "Binding states of individual nucleons in strongly deformed nuclei," *Kong. Dan. Vid. Sel. Mat. Fys. Med.*, vol. 29N16, pp. 1–69, 1955.

[11] M. Freer, "The clustered nucleus—cluster structures in stable and unstable nuclei," *Reports on Progress in Physics*, vol. 70, no. 12, pp. 2149–2210, nov 2007. [Online]. Available: https://doi.org/10.1088/0034-4885/70/12/r03

[12] R. Bass, *Nuclear reactions with heavy ions.* Springer, 1980.

[13] G. Knoll, *Radiation Detection and Measurement*. Wiley, 2010. [Online]. Available: https://books.google.it/books?id=4vTJ7UDel5IC

[14] Obertelli, A. and D. et al., "Minos: A vertex tracker coupled to a thick liquid-hydrogen target for in-beam spectroscopy of exotic nuclei," *Eur. Phys. J. A*, vol. 50, no. 1, p. 8, 2014. [Online]. Available: https://doi.org/10.1140/epja/i2014-14008-y

[15] M. Milin, T. Niksic, S. Szilner, and D. Vretenar, Eds., *ACTAR: the New Generation of Active Targets*, vol. 1165, no. 1, 2009.

[16] H. Horiuchi and K. Ikeda, "A Molecule-like Structure in Atomic Nuclei of 16O* and 10Ne," *Progress of Theoretical Physics*, vol. 40, no. 2, pp. 277–287, 08 1968. [Online]. Available: https://doi.org/10.1143/PTP.40.277

[17] W. von Oertzen, M. Freer, and Y. Kanada-En'yo, "Nuclear clusters and nuclear molecules," *Physics Reports*, vol. 432, no. 2, pp. 43–113, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370157306002626

[18] T. Baba and M. Kimura, "Coulomb shift in the mirror pair of ^{14}C and ^{14}O as a signature of the linear-chain structure," *Phys. Rev. C*, vol. 99, p. 021303, Feb 2019. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevC.99.021303

[19] R. J. deBoer, J. Görres, M. Wiescher, R. E. Azuma, A. Best, C. R. Brune, C. E. Fields, S. Jones, M. Pignatari, D. Sayre, K. Smith, F. X. Timmes, and E. Uberseder, "The $^{12}\mathbf{C}(\alpha,\gamma)^{16}\mathbf{O}$ reaction and its implications for stellar helium burning," *Rev. Mod. Phys.*, vol. 89, p. 035007, Sep 2017. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.89.035007

[20] D. Bazin, T. Ahn, Y. Ayyad, S. Beceiro-Novo, A. O. Macchiavelli, W. Mittig, and J. S. Randhawa, "Low energy nuclear physics with active targets and time projection chambers," *Prog. Part. Nucl. Phys.*, vol. 114, p. 103790, 2020.

[21] M. Goeppert-Mayer, "The shell model," *Nobel Lectures, Physics*, 1963. [Online]. Available: https://www.nobelprize.org/prizes/physics/1963/mayer/lecture/

[22] S. C. Pieper and V. R. Pandharipande, "Origins of spin-orbit splitting in ^{15}N," *Phys. Rev. Lett.*, vol. 70, pp. 2541–2544, Apr 1993. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.70.2541

[23] H. G. e. a. Tran D.T., Ong H.J., "Evidence for prevalent z = 6 magic number in neutron-rich carbon isotopes," *Nature Communications 9*, 2018. [Online]. Available: https://doi.org/10.1038/s41467-018-04024-y

[24] F. Ajzenberg-Selove, "Energy levels of light nuclei a = 13–15," *Nuclear Physics A*, vol. 523, no. 1, pp. 1–196, 1991. [Online]. Available: https://www.sciencedirect.com/science/article/pii/037594749190446D

[25] ACTAR-Collaboration, "Actar: Conceptual design report," *tech. rep*, 2012.

[26] B. Mauss, "Réactions élastiques et inélastiques résonantes pour la caractérisation expérimentale de la cible active actar tpc." Ph.D. dissertation, These de

doctorat dirigee par Sorlin, Olivier Physique Normandie, 2018. [Online]. Available: http://www.theses.fr/2018NORMC226

[27] F. Sauli, *Gaseous Radiation Detectors: Fundamentals and Applications*, ser. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2014.

[28] W. Shockley, "Currents to conductors induced by a moving point charge," *Journal of Applied Physics*, vol. 9, no. 10, pp. 635–636, 1938. [Online]. Available: https://doi.org/10.1063/1.1710367

[29] B. Mauss, P. Morfouace, T. Roger, J. Pancin, G. Grinyer, J. Giovinazzo, and et al., "Commissioning of the active target and time projection chamber (actar tpc)," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 940, pp. 498–504, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016890021930912X

[30] T. Roger, J. Pancin, G. Grinyer, B. Mauss, and A. e. a. Laffoley, "Demonstrator detection system for the active target and time projection chamber (actar tpc) project," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 895, 07 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900218304777

[31] J. Giovinazzo, J. Pibernat, T. Goigoux, R. de Oliveira, G. Grinyer, C. Huss, B. Mauss, J. Pancin, J. Pedroza, A. Rebii, T. Roger, P. Rosier, F. Saillant, and G. Wittwer, "Metal-core pad-plane development for actar tpc," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 892, pp. 114–121, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900218303243

[32] "Garfield." [Online]. Available: https://garfieldpp.web.cern.ch/garfieldpp/

[33] G. Charpak, J. Derre, Y. Giomataris, and P. Rebourgeard, "MICROMEGAS, a multi-purpose gaseous detector," *Nucl. Instrum. Meth. A*, vol. 478, pp. 26–36, 2002.

[34] J. Pancin, S. Damoy, D. Perez Loureiro, V. Chambert, F. Dorangeville, F. Druillole, G. Grinyer, A. Lermitage, A. Maroni, G. Noël, C. Porte, T. Roger, P. Rosier, and L. Suen, "Tests of micro-pattern gaseous detectors for active target time projection chambers in nuclear physics," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 735, pp. 532–540, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900213013119

[35] E. Pollacco *et al.*, "GET: A generic electronics system for TPCs and nuclear physics instrumentation," *Nucl. Instrum. Meth. A*, vol. 887, pp. 81–93, 2018.

[36] L. Scomparin, "Studio sperimentale di una reazione nucleare con active target tpc," Bachelor's thesis, Università degli Studi di Padova, 2019. [Online]. Available: http://tesi.cab.unipd.it/62910/1/Tesi_LT_Scomparin_Luca.pdf

[37] J. Giovinazzo, T. Goigoux, S. Anvar, P. Baron, B. Blank, E. Delagnes, G. Grinyer, J. Pancin, J. Pedroza, J. Pibernat, E. Pollacco, A. Rebii, T. Roger, and P. Sizun, "Get electronics samples data analysis," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 840, pp. 15–27, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900216309408

[38] [Online]. Available: https://www.ganil-spiral2.eu/

[39] M. Stanoiu, D. Sohler, O. Sorlin, F. Azaiez, Z. Dombrádi, B. A. Brown, M. Belleguic, C. Borcea, C. Bourgeois, Z. Dlouhy, Z. Elekes, Z. Fülöp, S. Grévy, D. Guillemaud-Mueller, F. Ibrahim, A. Kerek, A. Krasznahorkay, M. Lewitowicz, S. M. Lukyanov, S. Mandal, J. Mrázek, F. Negoita, Y.-E. Penionzhkevich, Z. Podolyák, P. Roussel-Chomaz, M. G. Saint-Laurent, H. Savajols, G. Sletten, J. Timár, C. Timis, and A. Yamamoto, "Disappearance of the $n = 14$ shell gap in the carbon isotopic chain," *Phys. Rev. C*, vol. 78, p. 034315, Sep 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevC.78.034315

[40] H. J. Ong, N. Imai, D. Suzuki, H. Iwasaki, H. Sakurai, T. K. Onishi, M. K. Suzuki, S. Ota, S. Takeuchi, T. Nakao, Y. Togano, Y. Kondo, N. Aoi, H. Baba, S. Bishop, Y. Ichikawa, M. Ishihara, T. Kubo, K. Kurita, T. Motobayashi, T. Nakamura, T. Okumura, and Y. Yanagisawa, "Lifetime measurements of first excited states in ^{16,18}C," *Phys. Rev. C*, vol. 78, p. 014308, Jul 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevC.78.014308

[41] P. Voss, T. Baugher, D. Bazin, R. M. Clark, H. L. Crawford, A. Dewald, P. Fallon, A. Gade, G. F. Grinyer, H. Iwasaki, A. O. Macchiavelli, S. McDaniel, D. Miller, M. Petri, A. Ratkiewicz, W. Rother, K. Starosta, K. A. Walsh, D. Weisshaar, C. Forssén, R. Roth, and P. Navrátil, "Excited-state transition-rate measurements in ^{18}c," *Phys. Rev. C*, vol. 86, p. 011303, Jul 2012. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevC.86.011303

[42] Y. Kondo, T. Nakamura, Y. Satou, T. Matsumoto, N. Aoi, N. Endo, N. Fukuda, T. Gomi, Y. Hashimoto, M. Ishihara, S. Kawai, M. Kitayama, T. Kobayashi, Y. Matsuda, N. Matsui, T. Motobayashi, T. Nakabayashi, K. Ogata, T. Okumura, H. J. Ong, T. K. Onishi, H. Otsu, H. Sakurai, S. Shimoura, M. Shinohara, T. Sugimoto, S. Takeuchi, M. Tamaki, Y. Togano, and Y. Yanagisawa, "One-neutron removal reactions of ^{18}C and ^{19}C on a proton target," *Phys. Rev. C*, vol. 79, p. 014602, Jan 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevC.79.014602

[43] T. Roger, "Etude des réactions induites par le noyau à halo 11li avec la cible active maya," Ph.D. dissertation, These de doctorat, 2009. [Online]. Available: http://theses.fr/2009CAEN2030

[44] R. Anne and A. C. Mueller, "Lise 3: a magnetic spectrometer—wien filter combination for secondary radioactive beam production," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials*

and Atoms, vol. 70, no. 1, pp. 276–285, 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0168583X9295943L

[45] L. Domenichetti, "Algoritmi di tracciamento per reazioni nucleari misurate con active-target tpc," Bachelor's thesis, Università degli Studi di Padova, 2020. [Online]. Available: https://thesis.unipd.it/bitstream/20.500.12608/22705/1/Tesi_Lorenzo_Domenichetti.pdf

[46] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, p. 381–395, Jun. 1981. [Online]. Available: https://doi.org/10.1145/358669.358692

[47] A. Zanchettin, "Algoritmo di ransac: panoramica, confronti e applicazioni," Master's thesis, Università degli Studi di Padova, 2013. [Online]. Available: http://tesi.cab.unipd.it/44211/1/L'algoritmo_di_RANSAC_panoramica,_confronti_e_applicazioni.pdf

[48] P. V. C. Hough, "Machine Analysis of Bubble Chamber Pictures," *Conf. Proc. C*, vol. 590914, pp. 554–558, 1959.

[49] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "On clustering validation techniques," *Journal of Intelligent Information Systems*, vol. 17, no. 2, pp. 107–145, Dec 2001. [Online]. Available: https://doi.org/10.1023/A:1012801612483

[50] L. Hubert and P. Arabie, "Comparing partitions," *Journal of Classification*, vol. 2, no. 1, pp. 193–218, Dec 1985. [Online]. Available: https://doi.org/10.1007/BF01908075

[51] E. B. Fowlkes and C. L. Mallows, "A method for comparing two hierarchical clusterings," *Journal of the American Statistical Association*, vol. 78, no. 383, pp. 553–569, 1983. [Online]. Available: http://www.jstor.org/stable/2288117

[52] A. Strehl and J. Ghosh, "Cluster ensembles - a knowledge reuse framework for combining multiple partitions," *Journal of Machine Learning Research*, vol. 3, pp. 583–617, 01 2002.

[53] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *Journal of the American Statistical Association*, vol. 66, no. 336, pp. 846–850, 1971. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/01621459.1971.10482356

[54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[55] R. Solli, D. Bazin, M. Kuchera, R. Strauss, and M. Hjorth-Jensen, "Unsupervised learning for identifying events in active target experiments," Master's thesis, 08 2020.

[56] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, "A high-bias, low-variance introduction to machine learning for physicists," *Physics Reports*, vol. 810, pp. 1–124, 2019, a high-bias, low-variance introduction to Machine Learning for physicists. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370157319300766

[57] M. Kuchera, R. Ramanujan, J. Taylor, R. Strauss, D. Bazin, J. Bradt, and R. Chen, "Machine learning methods for track classification in the at-tpc," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 940, pp. 156–167, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900219308046

[58] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2, no. 4, pp. 303–314, Dec 1989. [Online]. Available: https://doi.org/10.1007/BF02551274

[59] F. Chollet *et al.* (2015) Keras. [Online]. Available: https://github.com/fchollet/keras

[60] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[61] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014. [Online]. Available: https://arxiv.org/abs/1409.1556

[62] Kaggle. [Online]. Available: https://www.kaggle.com/

[63] P. Konczykowski, B. Fernández-Dominguez, H. Alvarez-Pol, M. Caamaño, G. Grinyer, A. Laffoley, B. Mauss, J. Pancin, D. Pérez-Loureiro, and T. Roger, "Validation of the energy-loss response of α particles in ic4h10 with actarsim," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 927, pp. 125–132, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900219301883