Study on reconstruction efficiency

## I. Experimental setup [1]

The Micro Booster Neutrino Experiment (MicroBooNE) employes a large (~100 tonnes) Liquid Argon Time Projection Chamber (LArTPC) detector designed for precision neutrino physics measurements. MicroBooNE is the latest among a family of detectors that exploit the potential of liquified noble gases as the detection medium for neutrino interactions. These detectors combine the advantages of high spatial resolution and calorimetry for excellent particle identification with the potential to scale to very large volumes.

The MicroBooNE detector at Fermilab in Batavia, Illinois is sited in the Liquid Argon Test Facility (LArTF) on axis in the Booster Neutrino Beam (BNB), 470 m downstream from the neutrino production target. The BNB delivers a beam of predominantly muon neutrinos produced primarily from pion decays, with energies peaking at 700 MeV. MicroBooNE is also exposed to an off-axis component of the NuMI beam produced from pion and kaon decays with average neutrino energies of about 0.25 GeV and 2 GeV respectively. MicroBooNE is located about 600 m downstream from the NuMI neutrino production target.

Charged particles traversing a volume of liquid argon leave trails of ionization electrons in their wake and also create prompt vacuum ultraviolet (VUV) scintillation photons. In a LArTPC, the liquid argon is highly purified so that the ionization trails can be transported with minimal attenuation over distances of the order of meters under the influence of a uniform electric field in the detector volume, until they reach sense planes located along one side of the active volume. The electric field is created by introducing voltage onto a cathode plane and gradually stepping that voltage down in magnitude across a field cage, which is formed from a series of equipotential rings surrounding the drift volume. Non-uniformities in the electric field, diffusion, recombination, and space charge effects modify the tracks as they are transported. Calibration of these effects is critical to reconstruction of the initial ionization trails.

The anode planes are arranged parallel to the cathode plane, and in MicroBooNE, parallel to the beam direction. There are three planes comprised of sense wires with a characteristic pitch, held at a predetermined bias voltage, that continuously sense the signals induced by the ionization electrons drifting towards them. The electrostatic potentials of the sequence of anode planes allow ionization electrons to pass undisturbed by the first two planes before ultimately ending their trajectory on a wire in the last plane. The drifting ionization thus induces signals on the first planes (referred to as induction planes) and directly contributes to the signals in the final plane (referred to as the collection plane). Figure 1 depicts the arrangement of the MicroBooNE LArTPC and its operational principle.

The charged particle trajectory is reconstructed using the known positions of the anode plane wires and the recorded drift time of the ionization. The drift time is the difference between the arrival times of ionization signals on the wires and the time the interaction took place in the detector ($t_0$) which is provided by an accelerator clock synchronized to the beam (BNB or NuMI) or from a trigger provided by the light collection system. The characteristics of the waveforms observed by each wire provide a measure of the energy deposition of the traversing particles near that wire, and, when taken as a whole for each contained particle's trajectory, allow for determination of momentum and particle identity.

The scintillation photons are detected by a light collection system that is immersed in the liquid argon and faces into the detector volume. This system provides signals that can establish the event $t_0$ and supplies trigger information to an electronic readout system. The light collection system signals are vital in distinguishing detector activity that is in-time with the beam (and therefore possibly originating from beam interactions) from activity which is out-of-time (and therefore probably not associated with the beam), benefiting triggering and event reconstruction.

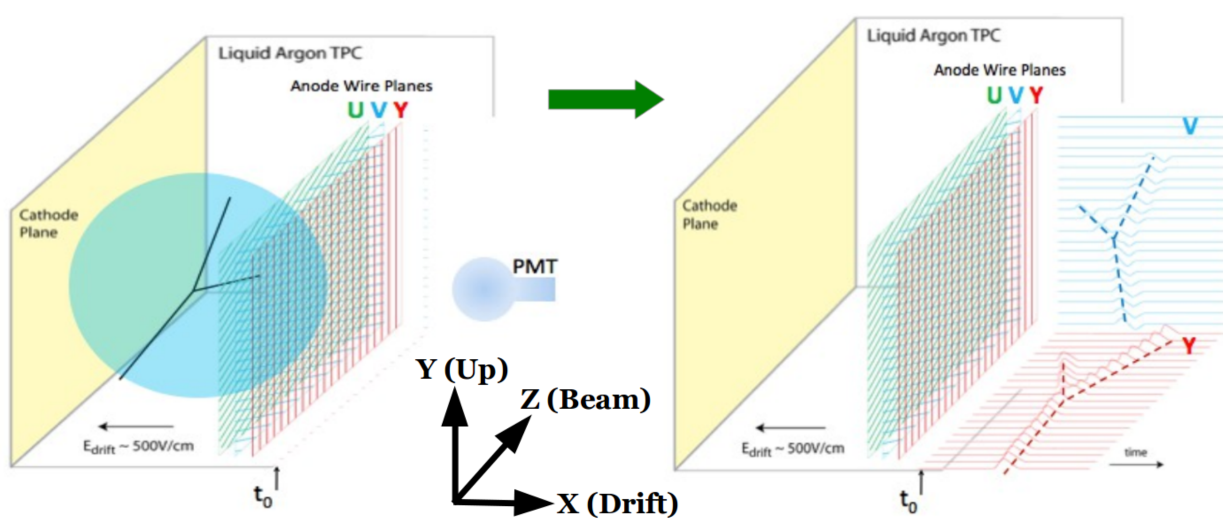


Figure 1. MicroBooNE operational principle.[1]

Liquid argon as a target for neutrinos is attractive due to its density, allowing a more compact detector with a substantial boost in event rate over a comparable detector using less dense media. A tradeoff to this aspect is the fact that the complicated structure of the argon nucleus (relative to hydrogen or helium, for example) will introduce nuclear effects that the data analysis must take into account. The cryogenic temperatures at which the noble elements are in the liquid phase also introduces the need for additional design considerations to ensure stable and safe operations.

## II. The problem [2]

A great challenge for neutrino based experiment is the fact that neither the energy of the incoming particle, nor the particles configuration and kinematics within the nucleus are known a priori, which means one as to rely on Monte Carlo (MC) simulations to generate probability-weighted maps that can connects observations in the detector to distributions of possible true kinematics. One way to partially get out of this situation, comes from the analysis and modeling of neutrino-nucleus scattering: given the correct interaction model in fact, neutrino's energy can be inferred from the final state of the system. Improving our understanding on how to model neutrino-nucleus interaction, is then for this reason directly reflected on the goodness of our knowledge on the neutrino beam energy spectrum.

The basic setup of a lepton-nucleus scattering experiment is shown in Fig. 2. A neutrino of unknown energy enters the detector made of heavier nuclei and interacts. In charged-current neutrino scattering, the final-state lepton is the charged partner of the incoming flavor while in neutral- current scattering the final state lepton is a neutrino of the same flavor as the incoming neutrino. Typically, the exchanged W or Z boson interacts with a bound nucleon, moving with Fermi momentum $P_F$ within the nucleus, producing an outgoing nucleon of four-momentum $P_1$ and, if the neutrino energy is high enough, additional hadrons, mostly pions. Occasionally the exchanged boson interacts with a pair of correlated nucleons and a second nucleon is released in the initial interaction: these "two-particle-two-hole" events are fascinating from the perspective of nuclear physics and, it turns out, of quantitative importance in measuring neutrino-oscillation parameters. These nuclear effects of the initial interaction; including the Fermi momentum of the bound nucleon and the existence of correlated multi-nucleon ensembles, affect the initial kinematic distribution of both the outgoing lepton and hadronic shower.

The final state lepton escapes the nucleus, however the initially produced hadronic shower undergoes significant further nuclear effects as it proceeds through the dense nuclear matter within the nucleus. These final state interactions (FSI) can change the energy, angle and even charge state of the originally produced hadrons with the pions having reasonable probability of even being totally absorbed within the nucleus and not emerging in the detector.

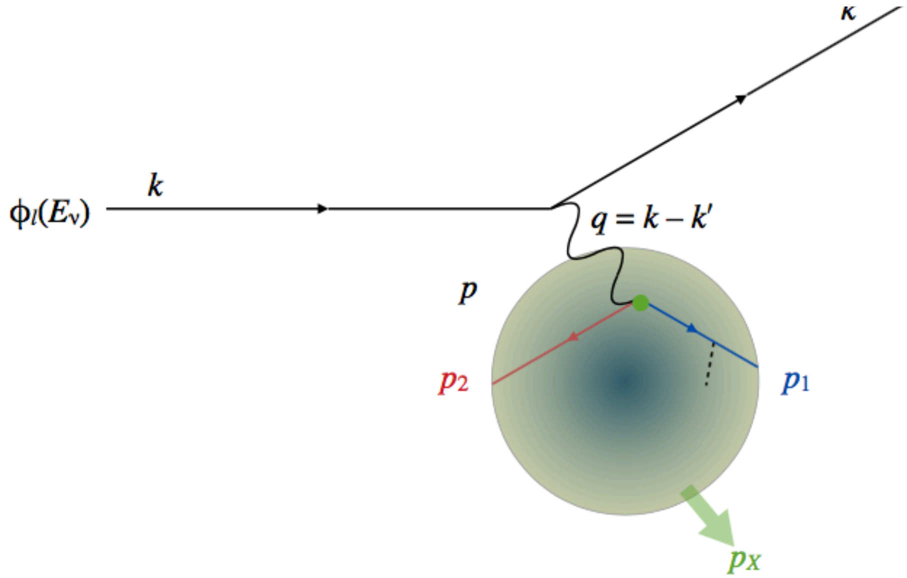


Figure 2. In neutrino-nucleus scattering a neutrino of energy $E_\nu$ and flavor l within a beam with energy spectrum $\phi_l(E_\nu)$, strikes a nucleus of atomic number A. In charged (neutral) current interaction the associated charged lepton l (neutrino of same flavor) emerges. Hadrons emerge from the initial interaction vertex as well that include one or more nucleons and, typically, pions (black dashed lines).[2]

A large and growing body of work over the past several years highlights how mis-modeling of the nucleus (the nuclear model) could lead to unacceptably large systematic uncertainties or, worse, biased measurements in current and future oscillation experiments. This suggests that since, for example, the discovery of CP violation at DUNE/LBNF will require as-yet unachieved percent-level control over the appearance signals, the understanding of the nuclear model has to be critically examined refined, and quantified

Given the importance of the physics of two-body two-hole (MEC) interactions in the reconstruction of neutrino beam energy spectrum, there is a great need to study it through measurements; and to being able to do that, is of primary importance to analyze and improve the performances of the detector, as well as the reconstruction efficiency, topic that, at least trying to understand how to, is the purpose of my work.

## III. My work: reconstruction efficiency study.

With this general idea, the first step of my work has been to write a script that could generate a first sample of two-tracks proton-muon events with defined kinematics and geometrical properties.

In the coordinate system chosen, the z axis is in line with the beam direction, the y axis is the one that point "upward", and the x axis perpendicular to this plane (see figure 1); **θ** is chosen to be the angle between the vector and the z axis, while **φ** is the angle between the projection of the vector in the x-y plane and the x axis.

These are the features of the first events generated:

Proton's kinematics:

- **Cos(θ)** distribution flat between -1 and 1.
- **φ** distribution flat between 0-2**π**.
- **Momentum** distribution flat between 0.2 GeV and 0.65 GeV.

Muon's kinematics:

- **Cos(θ)** distribution flat between 0 and 1 (muon travels only in the forward direction).
- **φ** distribution flat between 0-2**π**.
- **Energy** distribution flat between 0.4 GeV and 2 GeV.

The primary vertex is randomly generated so that it falls in the fiducial volume of the detector:

- **x position** ranges from 25 cm to 231.35 cm
- **y position** ranges from -91.5 cm to 91.5 cm
- **z position** ranges from 35 cm to 1011 cm

Then, to verify the propriety of these properties, I produced histograms of a 5000 events sample.

As can be seen by the plots (figures from 3 to 9) the distributions appear to be as wanted (bin fluctuactions expected≈10 around a mean value of 70).

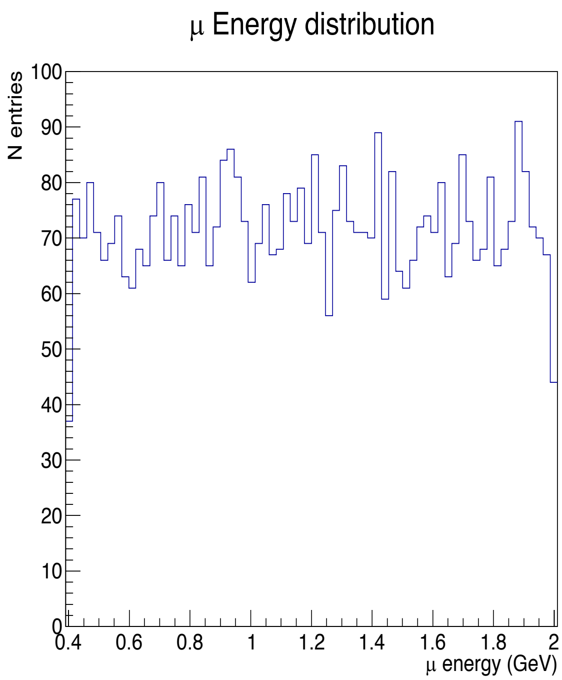


Figure 3. Muon energy distribution

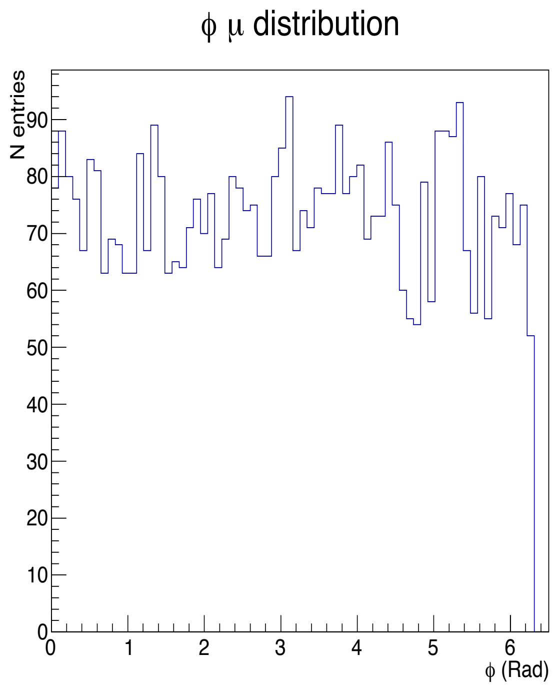


Figure 4. Muon **φ** distribution

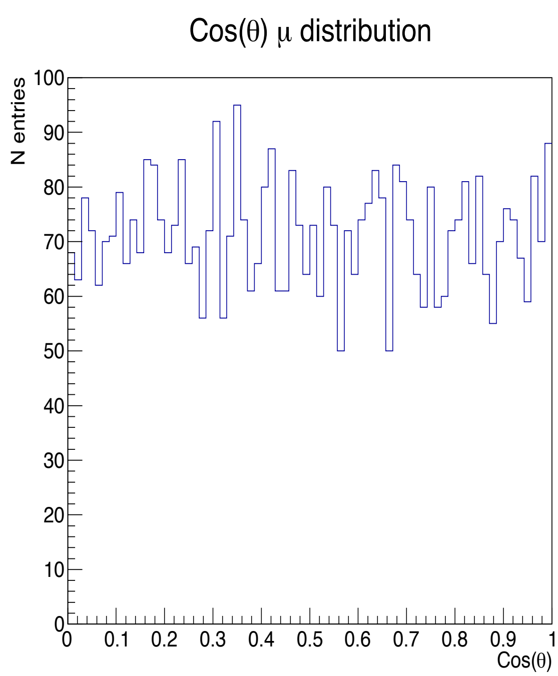


Figure 5. Muon cos(**θ**) distribution

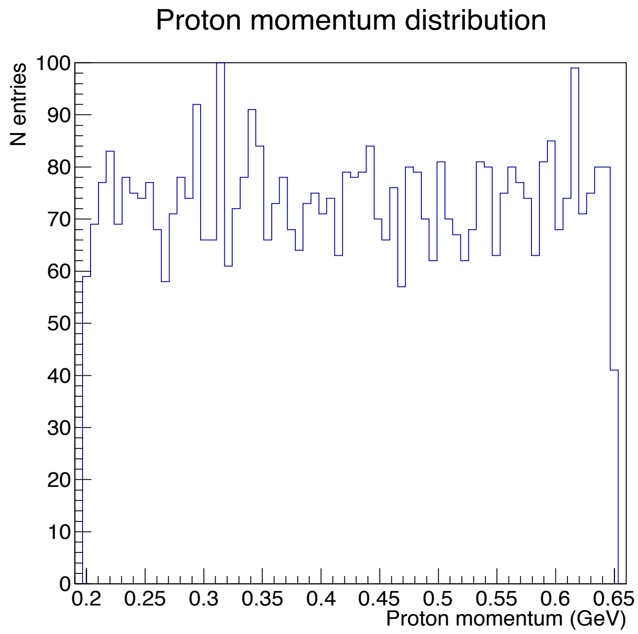


Figure 6.Proton momentum distribution

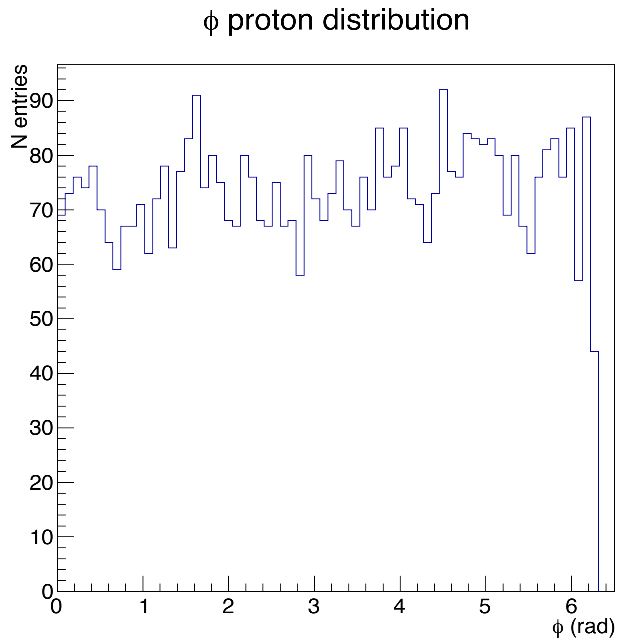


Figure 7. Proton **φ** distribution

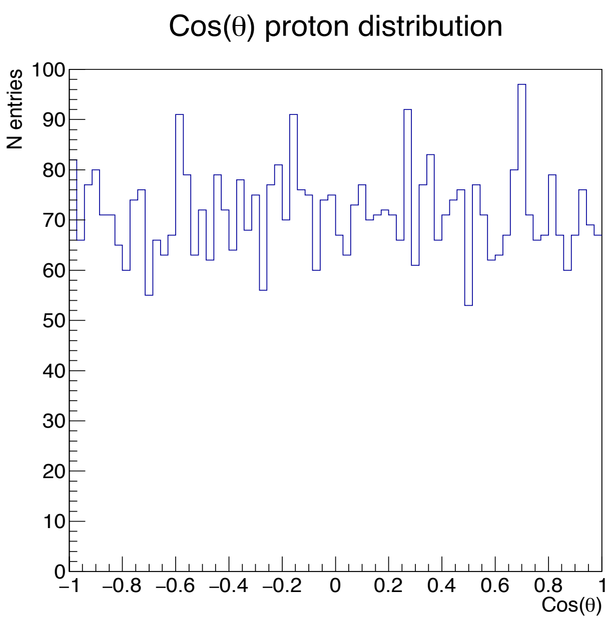


Figure 8. Proton cos(**θ**) distribution

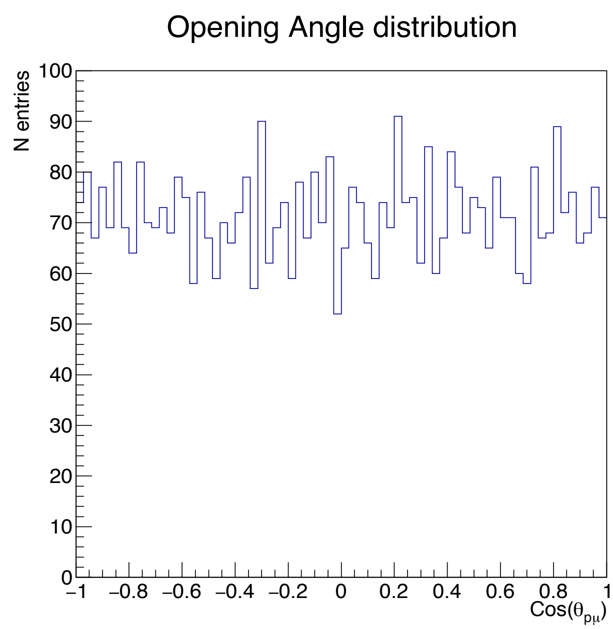


Figure 9. distribution of the opening angle between proton's and muon's directions

There are few things that need to be defined to understand how the reconstruction works

- **Hits**: Each represent a signal detected on a single wire of the LArTPC at a definite drift time.
- **MCParticles**: The Monte Carlo simulated particles, which represent the "true information".
- **PFParticles**: The main output in the LArSoft Event Data Model (LArSoft EDM, an ensemble of algorithms that reconstruct the event, matching Hits to the corresponding MCParticle and reconstructing tracks) is a list of reconstructed 3D Particles (termed "PFParticles") for the event. Each PFParticle corresponds to a distinct track or shower in the event, and has an associated collection of 2D hits
- **Efficiency:** Fraction of MCParticles with at least one matched reconstructed Particle
- **Completeness:** Fraction of Hits in a MCParticle shared with the reconstructed Particle
- **Purity:** Fraction of Hits in a reconstructed Particle shared with the MCParticle

Verified that the script works as it should, I proceeded generating a bigger sample of 10000 events, that processed by the reconstruction software; the reconstruction software's output is an NTuple containing all the information about how the event is tracked by the detector (i.e. vertex position, momentum direction, number of Hits associated to both MCParticles and PFParticles, number of Hits shared, etc.). I used this data in the ROOT's macros I wrote, to produce different plots and histograms, to check the quality of the reconstruction paying attention in particular to purity, completeness and efficiency.

To be a little bit clearer:

The purity gives a measure of how much of the PF particle is due to the matched MC Particle. It's given by the number of hits shared between the MC Particle and the reconstructed particle, over the number of hits associated to the PF Particle:

$$\text{Purity} = \frac{N_{MC \cap PF}}{N_{PF}}$$

The completeness gives a measure of how much of the true particle is on the matched track. It's given by number of hits shared between the reconstructed particle and the MC Particle, over the number of hits of the MC Particle:

$$\text{Completeness} = \frac{N_{MC \cap PF}}{N_{MC}}$$

**Muon analysis:**

Figures 10 (11) is the result of a profile on the z-axis made on a 3D histogram representing opening angle between proton and muon, against muon kinetic energy, against completeness (purity); the color scale on the right indicates the magnitude of completeness (purity), that as result of the profiling, is actually a bin averaged value. The muon appears to be always well reconstructed as expected, so I won't show any other always identical plot, or even talk about it.

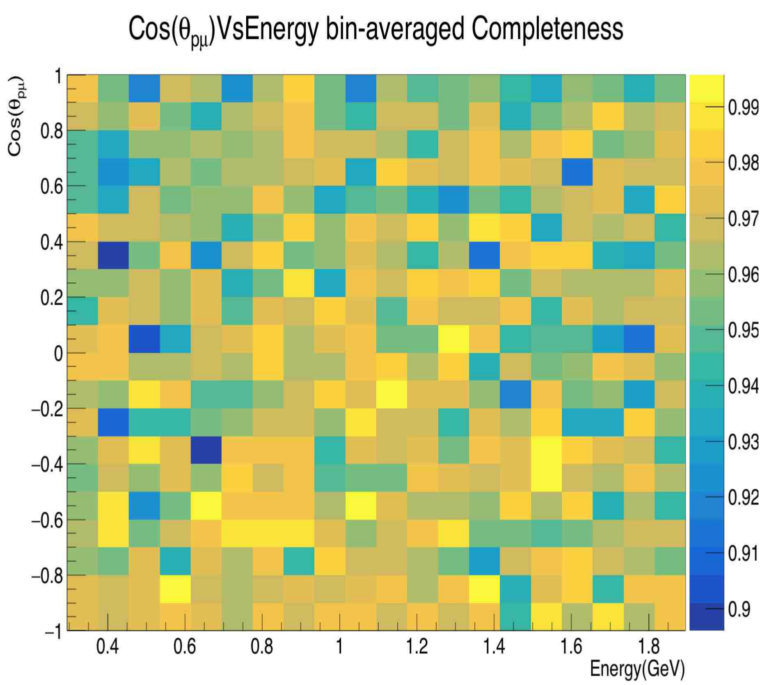


Figure 10

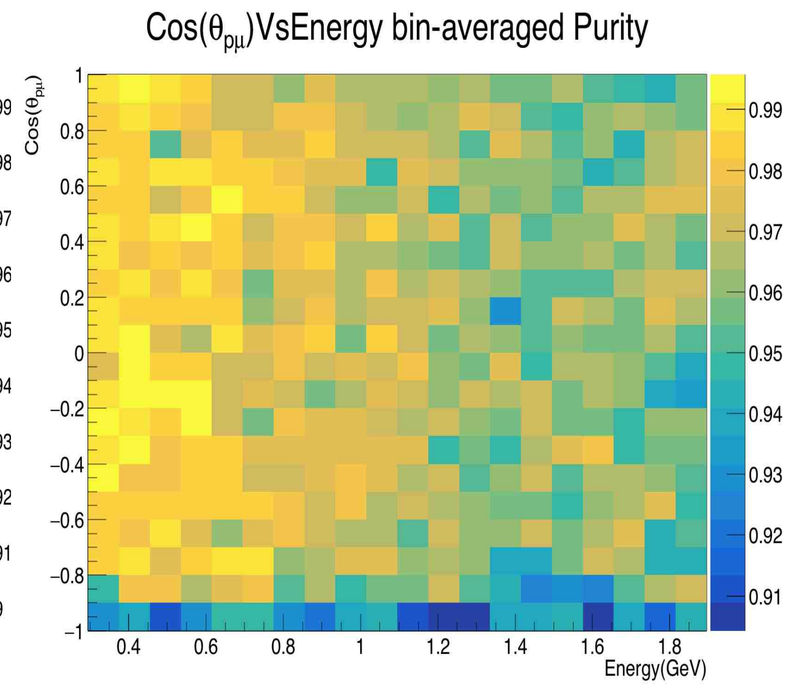


Figure 11

**Proton analysis:**

Figure 12 (14) is a 2D histogram that represent completeness (purity) against kinetic energy, with the color scale representing the number of entries for each two dimensional bin. On the right, figure 13 (15) is the product of the profiling on the y-axis, so that the result is bin-averaged completeness (purity) Vs kinetic energy.

Figure 16 (18) and 17 (19) are the result of the same operation, but with completeness (purity) against the cosine of the opening angle between muon and proton instead of kinetic energy.

Is important to point out that for each plot, only the events for which the software is able to track the proton are considered. The Ntuple in fact also contain a value both for muon and proton that indicates if they have a track associated, and if they share the same track. Each proton is counted only if it's associated to a track different from the muon's, or, if otherwise it's the same, if the number of Hits shared between the PF and MC proton is greater than the muon's one.

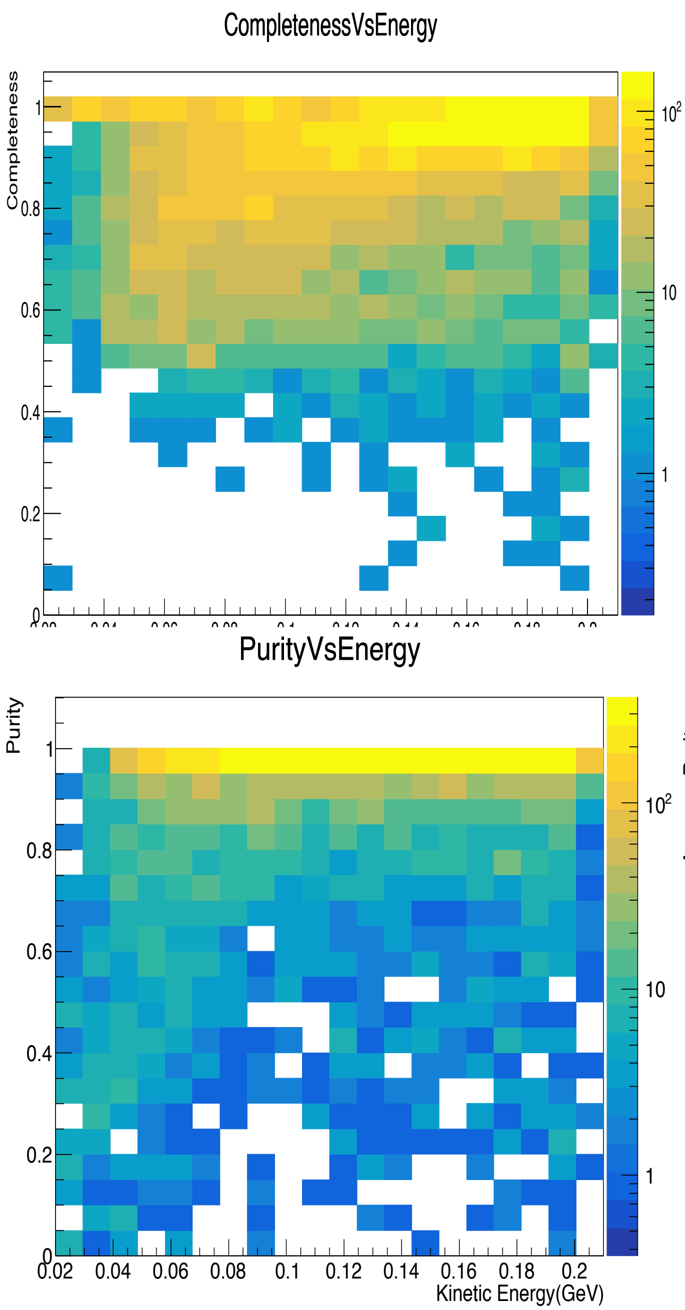

Figure 14

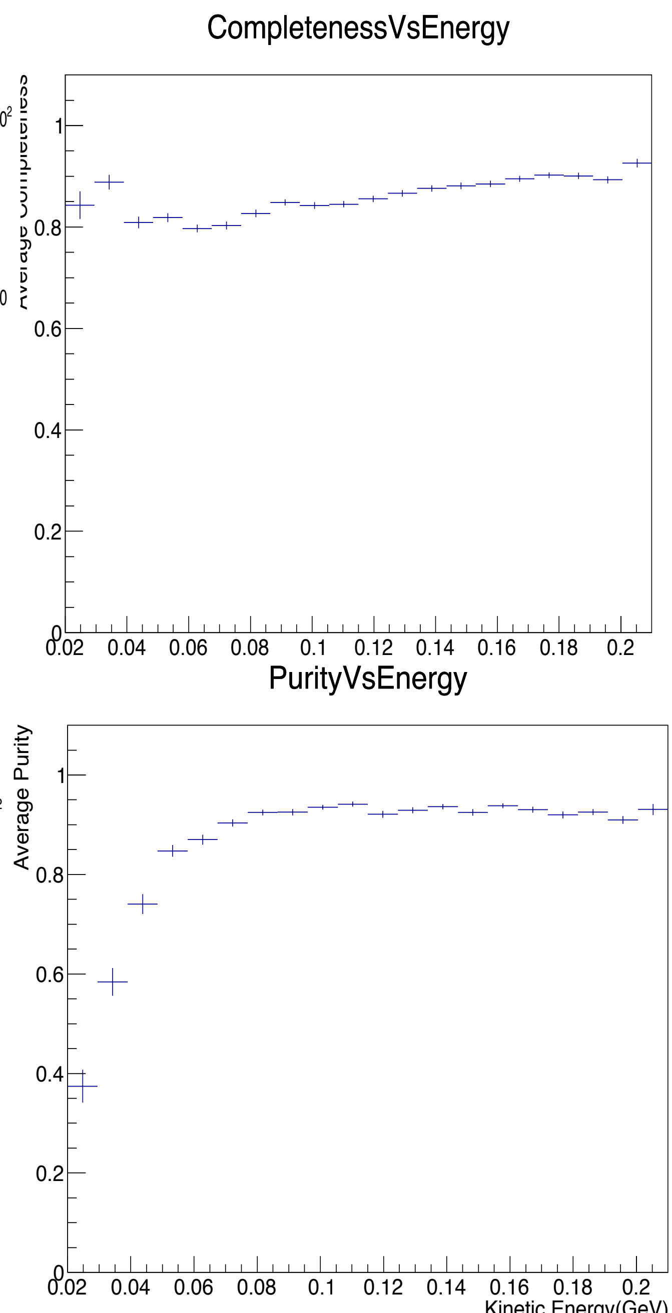

Figure 15

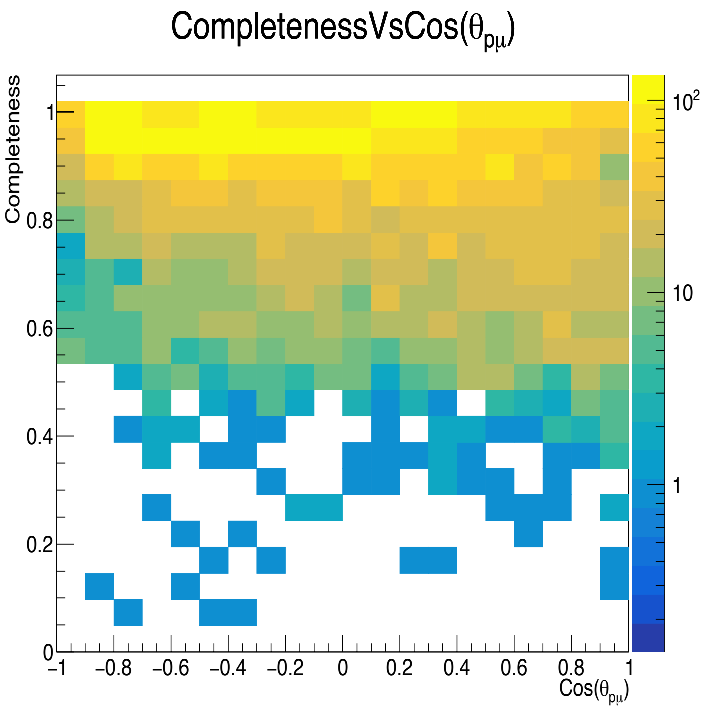


Figure 16

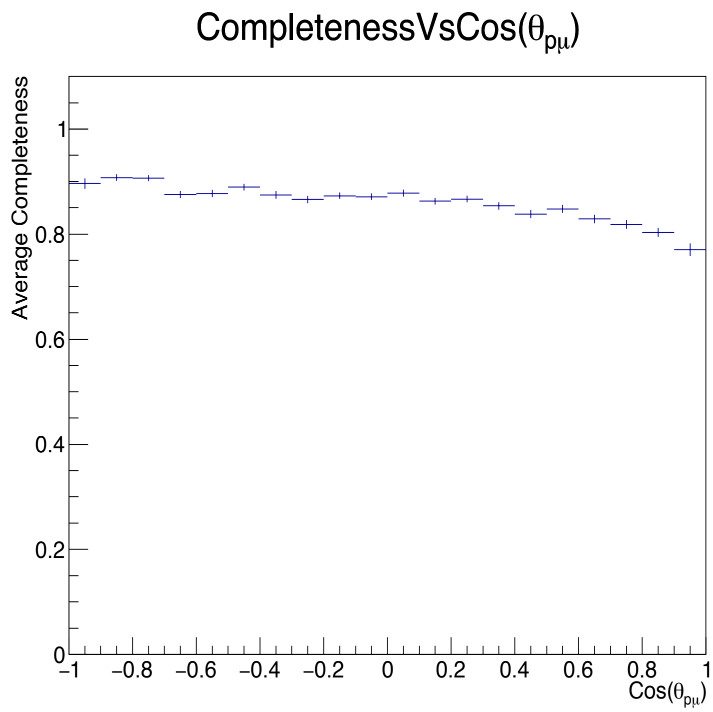


Figure 17

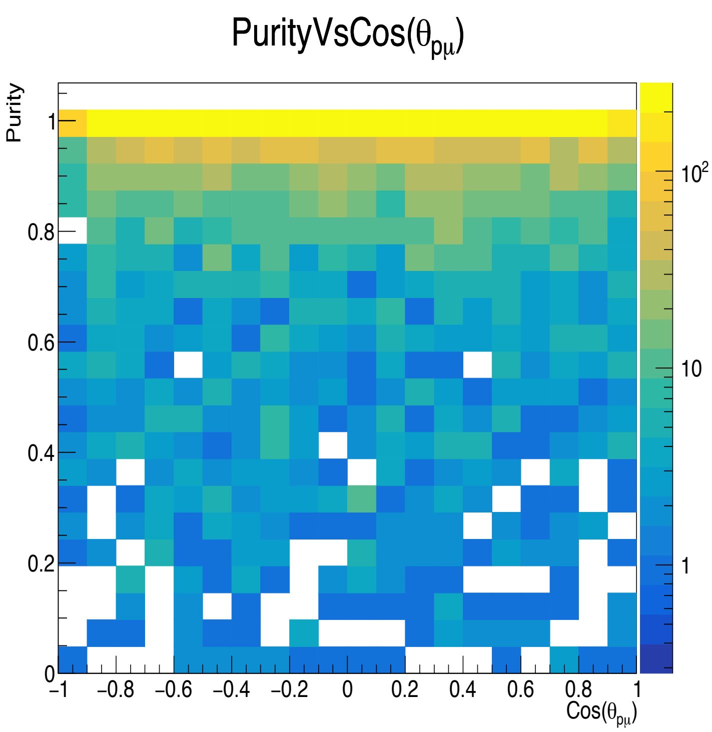


Figure 18

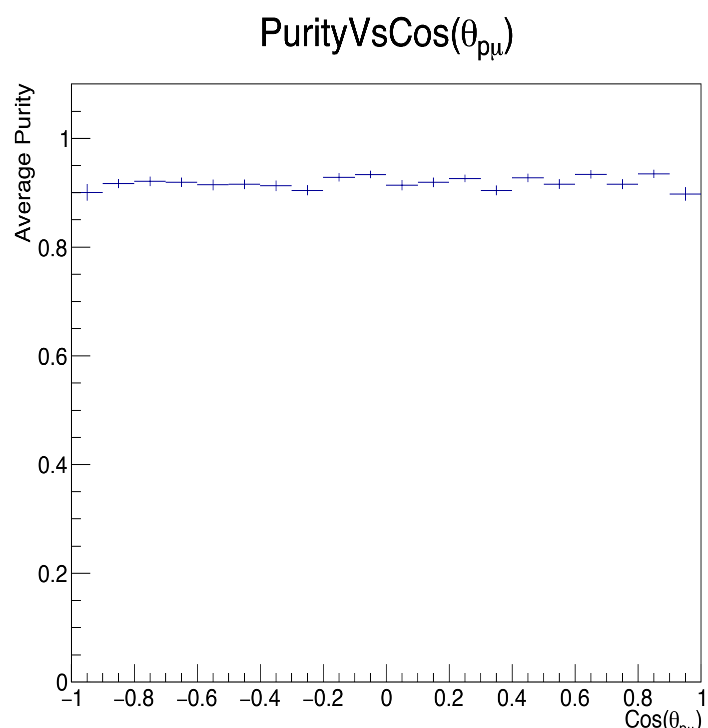


Figure 19

While the average purity seems to be constant for every value of opening angle, why it drops at low energy is not very clear, because even if in this case the proton's track is short and more difficult to be well detected, the bad-tracked protons have been removed from the analysis
The reason why completeness seems to have a peak at very low energies isn't very clear as well, and needs some further investigation. The first guess was that, at low energy, the proton starts to be associated to the muon because of the shortness of his track, resulting in a fake increase in completeness accompanied by a bad purity. This behavior though, should be avoided requiring the "well-tracked protons" condition (first plots, that showed the same trend, didn't include this condition).
The leading hypothesis right now, not yet verified, is that the little bump in completeness, as well as the average purity below 0.5 in the first bin (proton with purity under 0.5 should be labeled as not tracked by the software), could be due to a secondary scatter of the proton This would reproduce this behavior, the only problem being that low energetic scattering proton are less likely to be tracked, and so it's difficult to guess how much they affect the analysis.

Figure 20 and 21 are the same histograms shown for the muon, and essentially represent a synthesis of the previous 2D histograms, showing a pretty much everywhere good level of completeness, and a drop in purity at low energies.

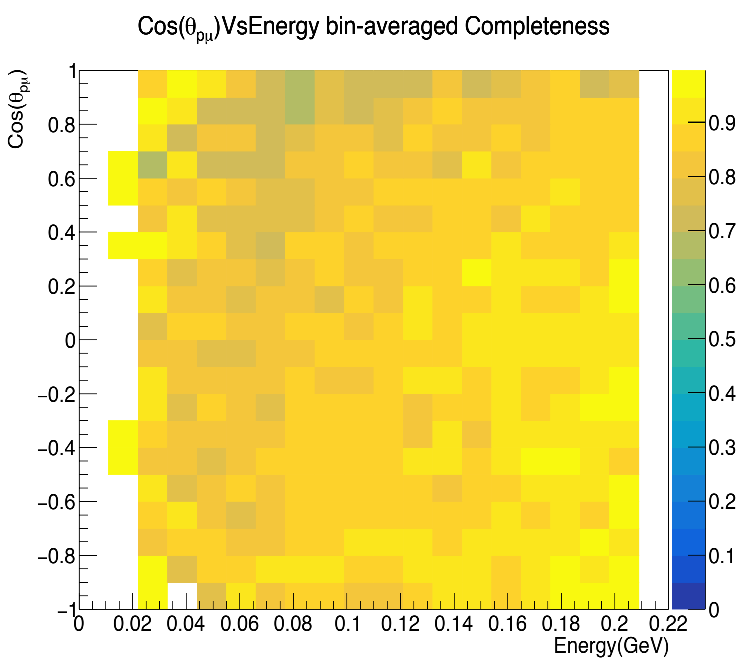


Figure 20

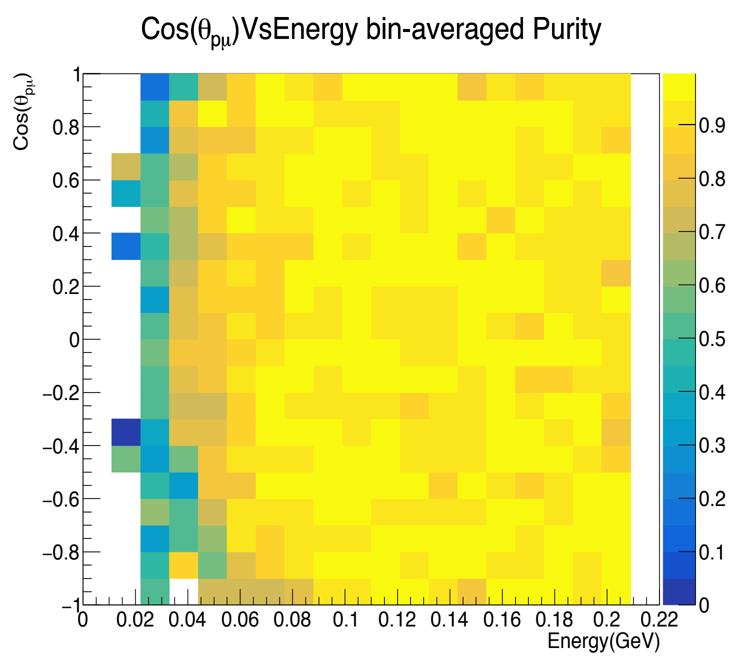


Figure 21

The most interesting results of this analysis though, are the ones on tracking efficiency. Tracking efficiency is by definition the ratio of the number of tracked protons over the total number of events and so just a single number; anyway, it's more significant to look at efficiency as a function of different variables, like proton energy or opening angle.

$$e_t = \frac{N_{tracked}(E_p, \theta_{\mu p}, ..)}{N_{total}(E_p, \theta_{\mu p}, ..)}$$

For the energy dependence, this is obtained by plotting the kinetic energy histogram that includes the bad tracked protons (figure 22), the energy histogram of only the tracked ones (figure 23), and dividing the second by the first (figure 24). The same has been done with opening angle cosine instead of kinetic energy in figure 25, 26 and 27 respectively. As can be seen from figure 27, efficiency drops down for cosine near 1 or -1, reflecting the difficulty for the software to separate the muon and proton tracks when they are almost collinear (angle equal to 0° or 180°).
In figure 24 instead, efficiency exhibits a plateau until a threshold of about 80 MeV, value under which starts to drop, as the track of the proton get shorter. This is most notable, especially considering that in this preliminary result, the MicroBooNE performance is already comparable with other types of detectors, perhaps even a bit better (as in the case of the MINERvA scintillator detector, which can only track protons about roughly 100 MeV) [3].

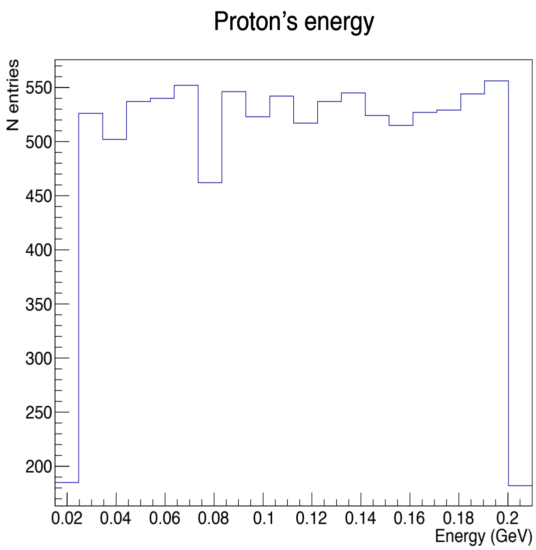

Figure 22

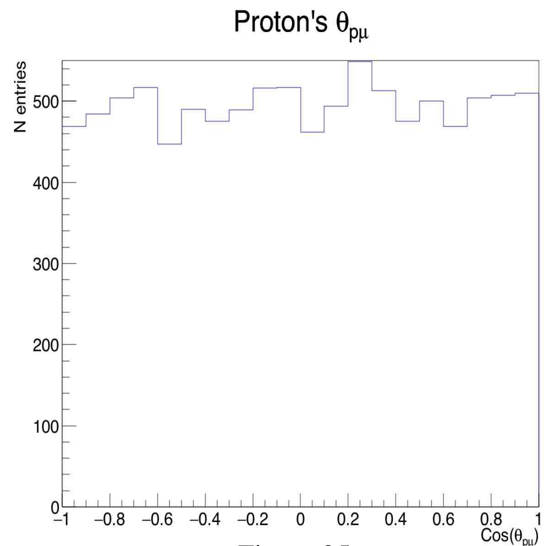

Figure 25

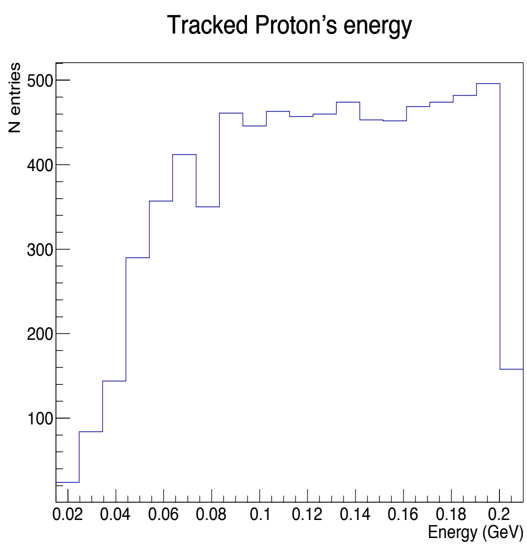

Figure 23

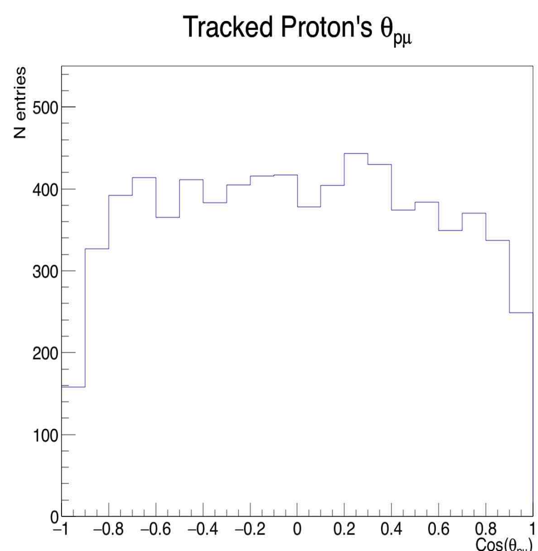

Figure 26

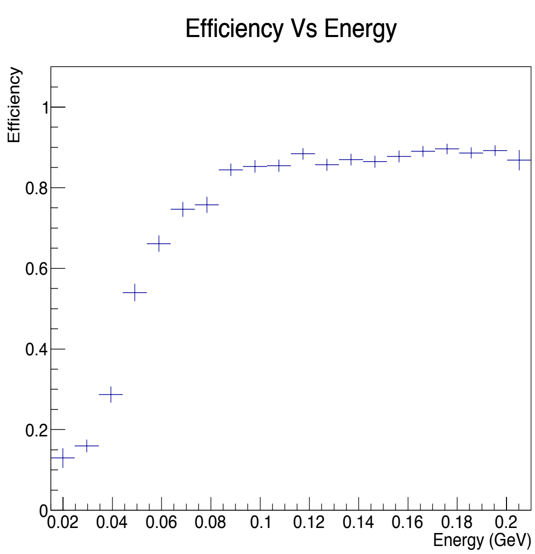

Figure 24

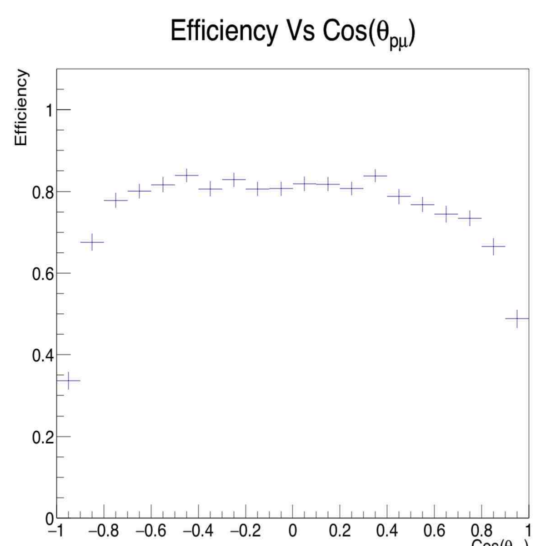

Figure 27

Aware of these results, I modified the script for the event generation to get another sample of just 100 events this time, with the same condition on vertex position and muon kinematic, but with proton having a fixed energy of 50 MeV, and a fixed opening angle between muon and proton of 90°. The purpose of this is to look at the Argo event display a family of events that is not well tracked (and not because of the collinearity of the two particles for sure, given the 90-degree angle) and try to understand why protons are not tracked one time out of two in this "energy-angle" regime, characterizing "by eye" events, by different properties that they could exhibit.
**About Argo:** it's a software that make you see all the three planes of wires together with the 3D reconstruction and gives a great number of information about the event. The most relevant for my "hand-scan" being the MCParticle tracks, raw and calibrated wire data, and reconstructed track
Of these 100 events, 32 of them resulted to have a bad-tracked proton. For these events I've been able to distinguish three categories of "failure":

- **Too few wire crossings:** manually counting the number of wires crossed by each proton, summing over all the planes, I found that 7 of them crossed less than 10 wires cumulatively, that is approximately the threshold the reconstruction software has to being able to identify a particle.
- **Dead/misconfigured regions:** 15 of them had a sufficient number of wires crossed, but it happened that in one or more plane the vertex fell in a dead or misconfigured region, i.e. a part of the detector for which a certain number of wires doesn't work (dead), or is working in a wrong way (misconfigured). The reconstruction software actually has a way to recover information from misconfigured wires, but, as far as I could see, many time it doesn't work very well, and lot of information is lost.
- **Other:** for 10 events I wasn't able to find an explanation for the reason of bad-tracking.

In the above figures I show examples for each case.
While for protons that cross too few wires there isn't too much that can be done, as well as for the cases in which the vertex is in a dead region, improving the recalibration of misconfigured wires would probably be reflected into an efficiency improvement. It also has to be pointed out that for the analysis of the efficiency of an ideal LArTPC, dead and misconfigured region should not at all be considered, i.e. 50% of the bad tracked events at this energy and opening angle would instead be tracked (given the so low statistic, this is a number that needs to be taken carefully).
The events that fall in the "other" category also offer a handle to improve the tracking efficiency, once understood why they are not tracked

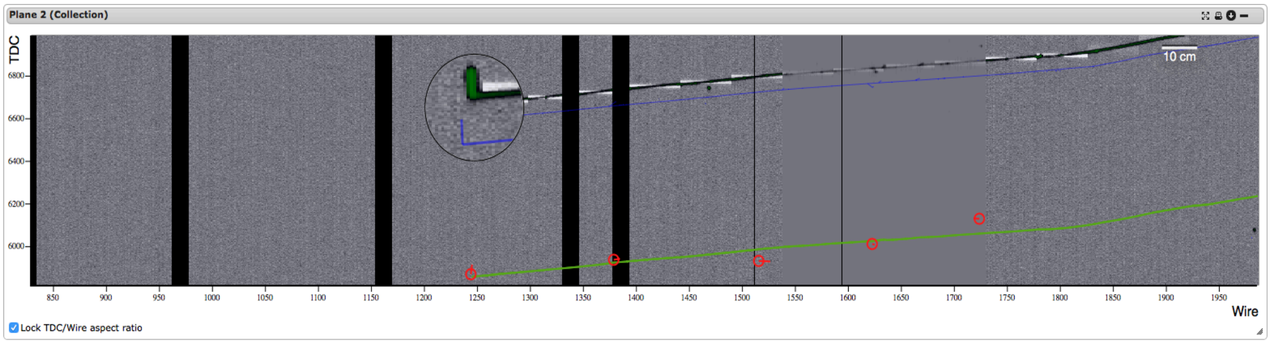


**Figure 27** Entry 61 of the file; example of a proton that cross too few wires in the Y plane(collection) to be tracked. From the bottom to the top we can see: in green the reconstructed track, in blue the MCParticles, in different colors what is obtained from raw wire data. The black vertical bands represent dead regions. Wires crossed from the proton summing over all planes: 6.

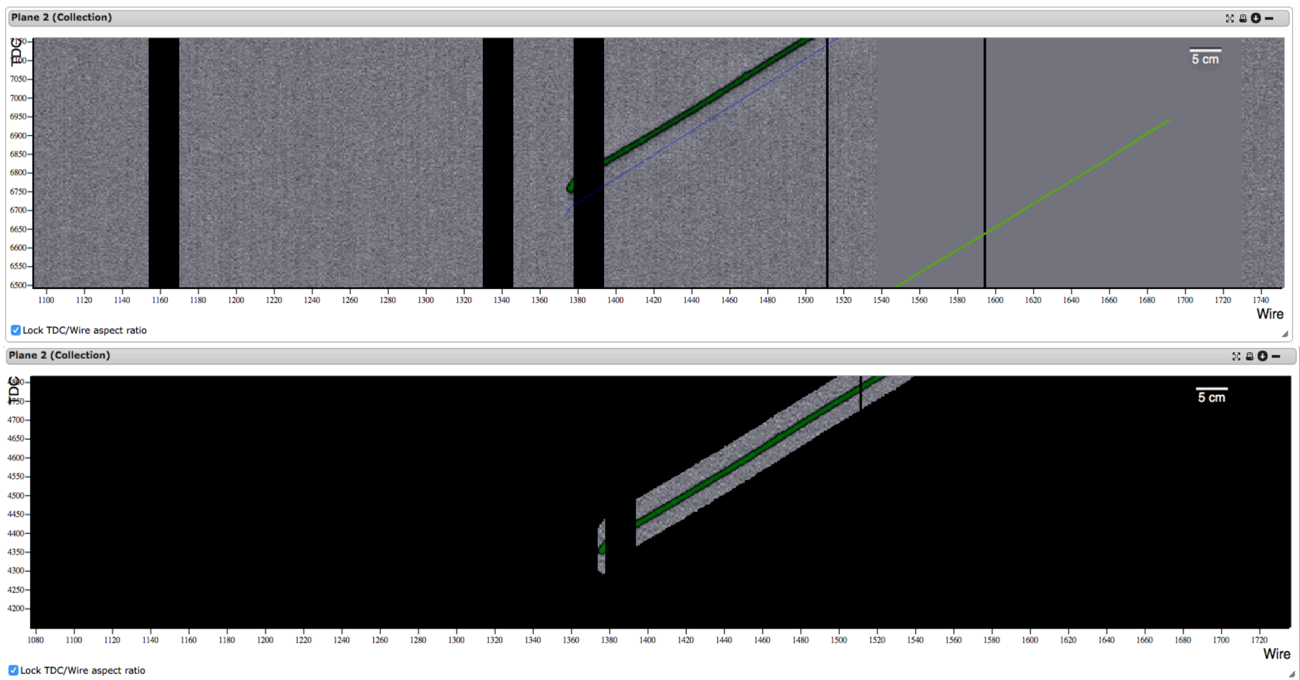


**Figure 28** Entry 46 of the file; example of vertex fallen in a dead region in plane Y (collection). Top: raw wire data; bottom: calibrated wire data mode (i.e. wire data processed to recover information from misconfigured wires). In calibrated data the black doesn't represent dead wires. Wires crossed from the proton summing over all planes: 13. Not counting this one: 9.

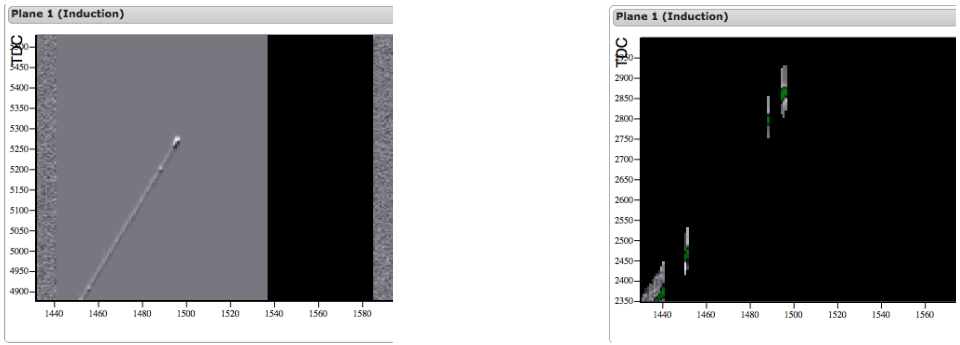


**Figure 29** Entry 31 of the file; example of the calibration operation failure for a vertex fallen in a misconfigured region in plane V (induction). Left: raw wire data; right: same vertex in calibrated wire data mode. Wires crossed from the proton summing over all planes: 13. Not counting this one: 10.

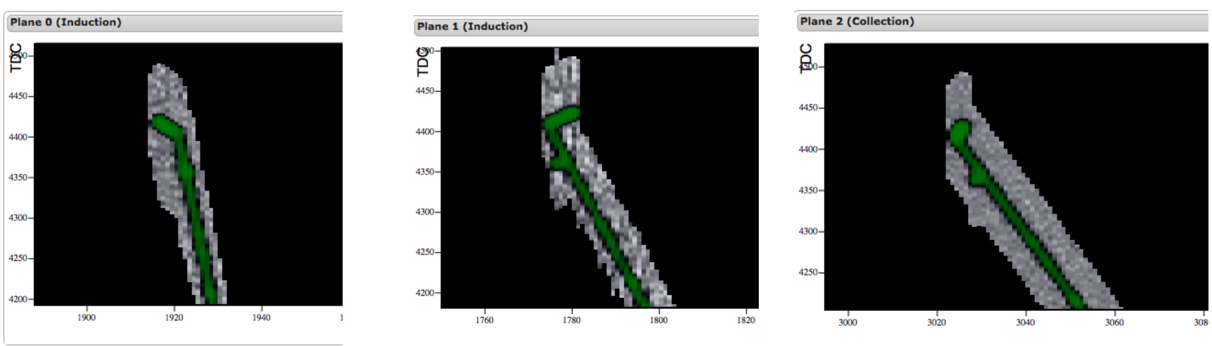


**Figure 30** Entry 16 of the file; example of event not tracked for some unknown reason. Left: plane U calibrated wire data view; center: plane V calibrated wire data view; plane Y calibrated wire data view.

## References

**[1]** MicroBooNE Collaboration, Design and Construction of the MicroBooNE Detector, 2016, **arXiv:1612.05824**

[2] NuSTEC White Paper: Status and Challenges of Neutrino-Nucleus Scattering, 2017, **arXiv:1706.03621**

[3] Measurement of $\nu_\mu$charged-current single $\pi_0$ production on hydrocarbon in the few-GeV region using MINERvA, **2017, arXiv:1708.03723**