Particle Identification by Time of Flight

Muhammad Aiman Bin Zulkipli
Supervisors : Lorenzo Giannessi, Vedantha Srinivas

August 2025

Abstract

The T2K experiment in Japan is an experiment that studies the oscillation of muon neutrinos over a large distance. Accurate identification of the flavor of neutrino interactions—specifically, distinguishing between muons and other particles such as protons produced in the detector—is crucial for the experiment's success. The ND280 near detector is responsible for this particle identification (PID). While the primary PID method in ND280 is based on energy loss per unit length ($\frac{dE}{dx}$), this approach becomes ineffective at momenta around 1.5 GeV, where the energy loss distributions for muons and protons overlap. To address this, time-of-flight (TOF) measurements can be used: by measuring the time a particle takes to traverse the detector, its velocity (β) can be determined. Since particles of different masses have different velocities at the same momentum, combining TOF with dE/dx information allows for improved separation of particle species.

In this project, I analyzed beam Monte Carlo data containing particle momenta and detector hit times. By classifying particles according to their PDG numbers and computing their velocities, I evaluated the separation power between muons, protons, and other species at various momenta. This study quantifies the effectiveness of TOF-based particle identification for ND280, showing that a separation of approximately 2.5σ can be achieved at a momentum range of 1.4 to 1.6 GeV assuming we have sufficiently large statistics, selecting only long reconstructed tracks ($\geq 1.5m$) and a constant time offset which is corrected for in our TOF reconstruction.

Acknowledgment

Contents

1 Introduction

Particle identification (PID) is one the most important things when it comes to high energy physics experiments. Of course, it is important for our detectors to be able to detect for a particle, however, these 'hits' in the detector would not come tagged with the particle name or PDG number, it is our job to develop a method that enables us to separate particles from one another. The first method of particle identification, developed by Charles Thomas Rees Wilson in 1911, is known as the cloud chamber[1]. One would identify particles based on the curvature of their tracks under a magnetic field. In fact, this method of PID was how the first ever antiparticle was discovered[2].

However, we are far removed from 1911. Rephrase and do not use noindent Since then there are far better and more effective methods of PID. The one utilised the most being particle identification by energy loss. This method uses the Bethe-Bloch formula:

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

This formula tells us the energy loss of a given particles is tied to the momentum, charge and the mass of the particle. This method is extremely effective in separating particles from one another. However, at momentum of 1.5GeV certain particles, namely protons, muons, and pions has the same energy deposit, so it is virtually impossible to tell them apart.

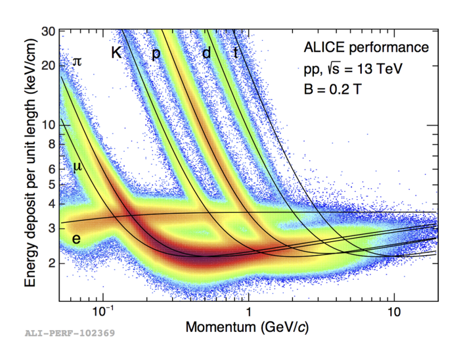


Figure 1: A plot of the $\frac{dE}{dx}$ measured from pp collisions in ALICE

This is the main reason why, as you will see in this report, PID by time of flight is used since it enables us to be able to separate particles from one another based on their velocity.

2 Background

2.1 T2K

The Tokai to Kamiokande experiment, or better known as T2K, is a neutrino experiment that aims to detect neutrino oscillations to better understand the fundamentals of neutrinos. To put it simply, neutrinos (muon neutrinos) are produced from protons that are accelerated to 30 GeV. These neutrinos then travel 295 km to the far detector where data is collected to see signs of neutrino oscillations. However, before hitting the far detector, we need to know for sure that only muon neutrinos are present in the neutrino beam, that is the role of the near detector (ND280).

In ND280, there are 2 million scintillating cubes crossed by optical fibres which generate light when hit by a charged particle (SFGD), these cubes enable us to reconstruct a 3D image of the particle interaction. There are also vertical time projection chambers (vTPC) and high angle time projection chambers(HAT) which measure the momentum of the charged particle using the radius of curvature in the magnetic field. There are two HATs, top HAT(THAT) and bottom HAT(BHAT) which sandwich the SFGD, and all of them are surrounded by these bars made of scintillating material (TOF bars) which would keep track of the position and time of the particle.

The raw data generated from the detector would then be used to reconstruct these events and store them in Flat Trees which has all the relevant information such as the reconstructed momentum, hit position, time, etc. It also has the true information for the particle.

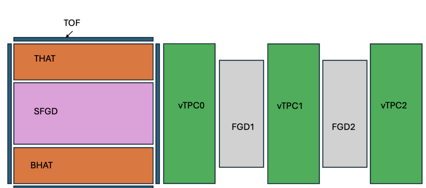


Figure 2: A rough diagram of the internals for ND280

2.2 PID by TOF

Particle identification by time of flight is done by using the velocity of the particle in terms of the speed of light, β, and a given momentum to calculate the mass,

$$\beta = \frac{p}{E} = \frac{p}{\sqrt{p^2 + m^2}} \tag{2}$$

$$m^2 = p^2 \left(\frac{1}{\beta^2} - 1 \right) \tag{3}$$

this means that, in theory we can identify the particle by its mass. Of course theory is simpler than reality, so this all depends on how accurately can we measure the momentum and time of flight of the particle.

We can calculate how well we can separate the particles by finding the 'separation power' using the separation power formula

$$S = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{4}$$

From equation (3) and equation (5), we can find the mass resolution, meaning how accurately can we calculate the mass.

$$\beta = \frac{L}{c * t} \tag{5}$$

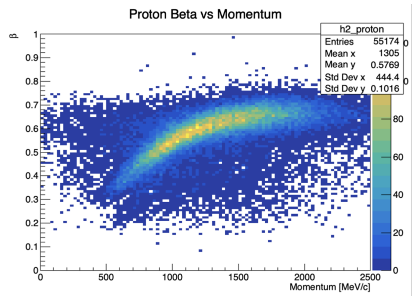
$$\frac{dm}{m} = \frac{dp}{p} + \gamma^2 \frac{d\beta}{\beta} = \frac{dp}{p} + \gamma^2 \left(\frac{dL}{L} + \frac{dt}{t} \right) \tag{6}$$

As we can see, the error on the mass is depends on $\frac{dp}{p}$ which is the momentum resolution and $\frac{dt}{t}$ which is the time resolution. So our ability to separate particles is also tied to how accurately can the detector collect these datasets.

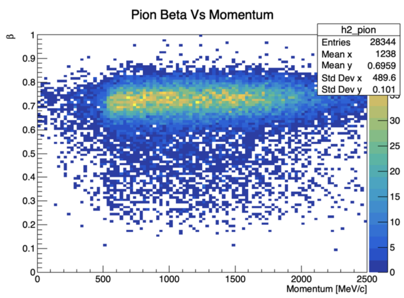
3 Particle Guns

We will simulate charged particles crossing the detector using particle guns. These particle guns only contains information regarding one specific particle (i.e muon particle guns and proton particle guns) This allows us to study how a specific particle would behave in the detector. The data from these particle guns would eventually be used as an 'expected value' when we move on to assigning identities to our particles.

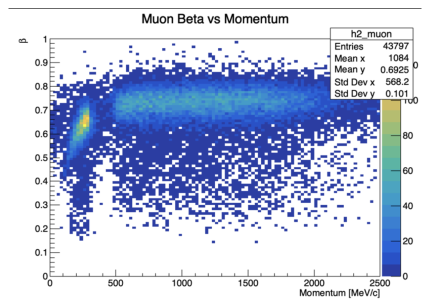
The first step would be to observe the beta of a particle over a momentum range. For our particle guns, it has been created such that the momentum for these particles are defined in the ranges of 500-2000 MeV. Since we are using the reconstructed momentum, meaning that it isn't the true momentum for the particles, there will be entries that are outside our momentum range, but this is not something to be alarmed by. Plotting the beta against the momentum, we get the following



(a) 2D histogram plot for the beta against the momentum for protons (particle guns)



(b) 2D histogram plot for the beta against the momentum for pions (particle guns)



(c) 2D histogram plot for the beta against the momentum for muons (particle guns). The muon particle guns contains low energy muons (300 - 500 MeV/c)

Figure 3: Beta vs momentum plots for particle guns

As an intermediary step, what one could do is create a combined plot for the :

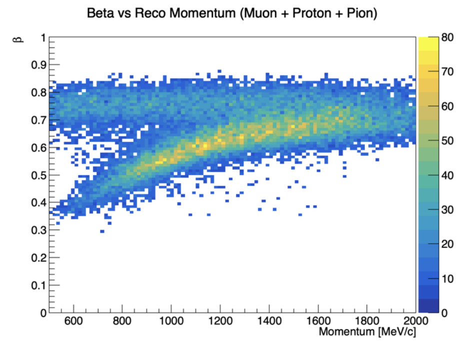
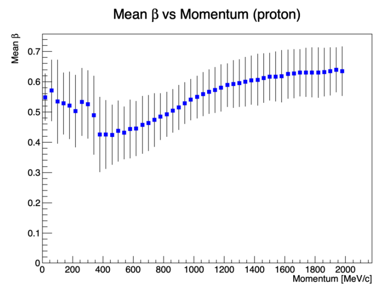


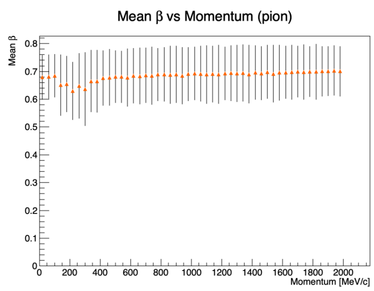
Figure 4: Combined beta vs momentum plots of all particles (protons, muons, and pions)

As expected, we could see that muons and pions have a higher velocity for a given momentum compared to a proton. Something that is worth noting is how low the beta for the particles are. Just by doing rough kinematics we would expect that the beta for pions and muons would be much closer to one, even for low momenta (500 MeV). This is actually a hint that something is off in our calculations, but for now, that is a problem for our future selves.

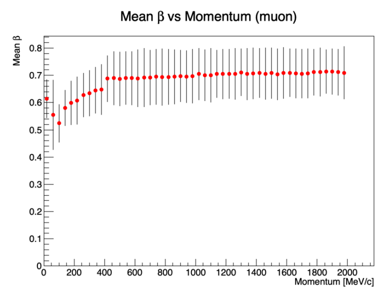
Now that we have the plots of the beta for each particle, we can proceed to the next step; computing the mean and sigma of the beta for each slice. This is done for two reasons, the first being visual clarity. Second, is that the mean and sigma is needed to be able to assign identities for a given particle. Here you can see the plots for protons, pions, and muons.



(a) The mean of the beta vs momentum with σ as the errors for protons



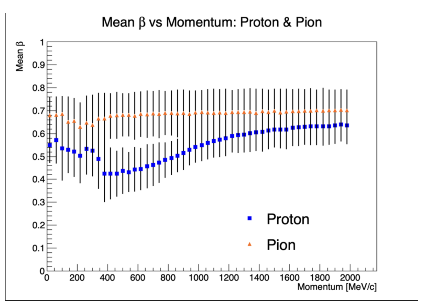
(b) The mean of the beta vs momentum with σ as the errors for pions



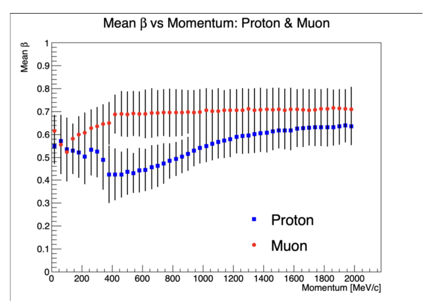
(c) The mean of the beta vs momentum with σ as the errors for muons

Figure 5: Mean beta vs momentum plots with error bars

Now, we can combine these plots to get a better sense of our separation power.



(a) Combined plots of pions and protons for mean of beta vs momentum with error bars

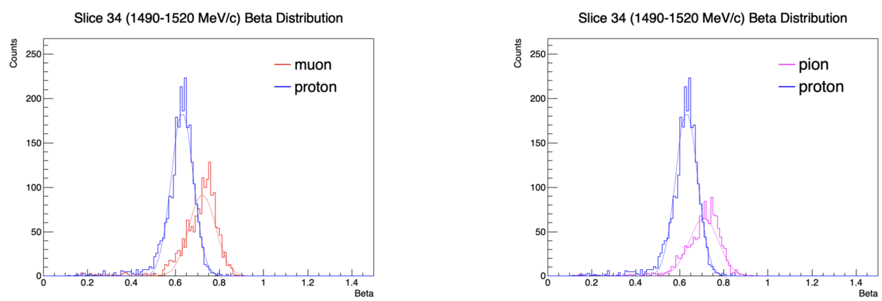


(b) Combined plots of pions and protons for mean of beta vs momentum with error bars

Figure 6: Combined mean of beta vs momentum plots

There is clearly some overlap at higher momenta, due to the error bars, but overall, we can be relatively hopeful in the sense that it is not a complete overlap between the error bars, meaning we would have a reasonably high separation power at this momentum. Since the particle gun are produced with uniform omentum distributed between 500 MeV/c and 2000 MeV/c, the events observed at lower momenta stem from mis-reconstruction of the particle momentum, hence the measured beta cannot be trusted either

Something we could do is plot the mean of the beta for each particle for every slice. This would appear as a gaussian, or at least gaussian-like. We can then combine protons and pions together (do the same for protons and muons). By looking at how much they overlap, this would tell us how well the separation power would be for that slice.



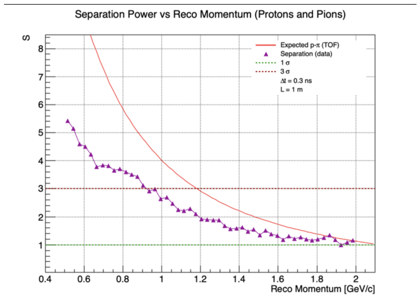
(a) Combined plot for beta distribution slice for muons and protons

(b) Combined plot for beta distribution slice for muons and protons

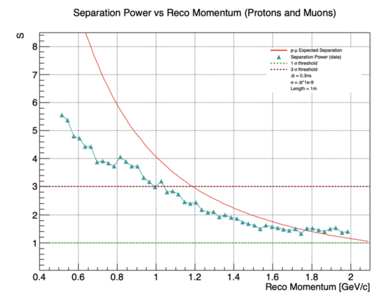
Figure 7: Combined plot for beta distribution in a momentum slice (1490-1520 MeV)

From equation(4), we can calculate the separation between protons and muons and protons and pions. What the separation power tells us is essentially the confidence $(\sigma)^1$ in our ability to distinguish between the two particles. We would ideally like to have around 2σ of separation power in the momentum range of 1400- 1600 MeV.

[1] $3\sigma = 99.7\%$, $2\sigma = 95\%$, $1\sigma = 68\%$



(a) Separation power between protons and pions in a momentum range of 500-2000 MeV

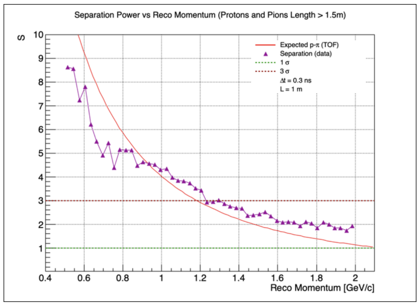


(b) Separation power between protons and muons in a momentum range of 500-2000 MeV

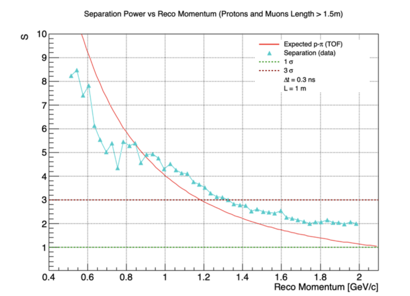
Figure 8: Separation power plots between protons and MIPs

we can see that it drops at higher momentum, as expected, however, at our region of interest, it's around 1.5 σ which is not too bad if we take into account that we also have the energy deposit to help with our particle identification.

However, it is important to note that our length estimation is also not the most accurate considering it is essentially assuming that the particle travels in a straight line from the vertex to the TOF bar. We need to take into account that the particles are subjected to a magnetic field (0.2T), meaning for charged particles, they would have a curved track. This assumption is fine for high momentum particles. However, at lower momentum, the particles actual track length would be significantly different than our estimated one. This causes us to not be able to get the actual speed correctly. Another reason could simply be our momentum resolution. Since our momentum is reconstructed, most likely is that it isn't quite reconstructed accurately. What we can do is only select particle that have long track lengths($\geq$ 1.5m). Longer track length would mean better momentum resolution, which, in theory, would mean we would see better separation power between the particles

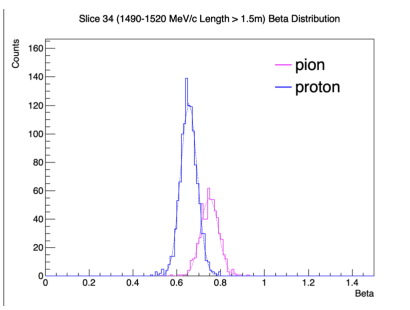


(a) Separation power plots for protons and pions but only selecting entries with track lengths $\geq 1.5m$

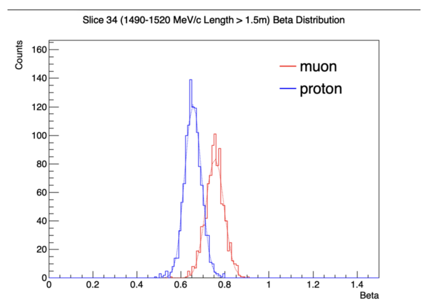


(b) Separation power plots for protons and muons but only selecting entries with track lengths $\geq 1.5m$

Figure 9: Separation power plots with selection criteria



(a) Combined beta distribution plot between protons and pions only using entries with track length $\geq 1.5m$



(b) Combined beta distribution plot between protons and muons only using entries with track length $\geq 1.5m$

Figure 10: Combined beta distribution plots with entry selection

By only selecting entries with track lengths ≥ 1.5m, the separation power greatly increases. We are now looking at a separation power around 2.5σ at momentum of 1400-1600 MeV. This is an excellent result. Meaning, in the best case scenario (not quite as we have not addressed the issue with the underestimated beta), we separate protons from pions/muons 95% of the time, even at 1.5GeV without even making use of dE/dx.

We have now gathered enough data from the particle guns to be able to actually separate particle from a combined dataset. We will not be using the real data just yet, we would be attempting to separate particles from a Monte Carlo.

4 Beam Monte Carlo

A Monte Carlo is a computational techniques that utilize random sampling to simulate the behavior of complex systems, often involving probabilistic or statistical phenomena. This is utilised a lot in high energy physics to simulate how particles would behave in the detector. Unlike the particle guns, it will contain a wide variety of particles, this means the dataset would be a lot more chaotic.

Let's start by plotting the beta vs momentum of all particles.

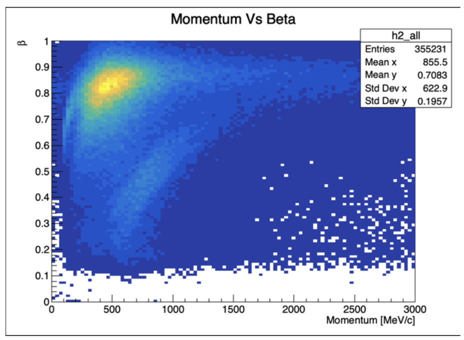
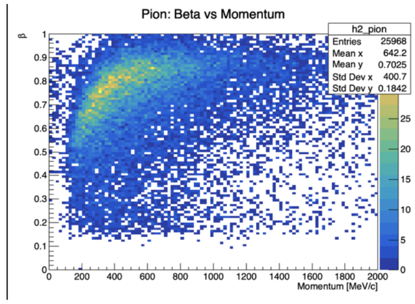
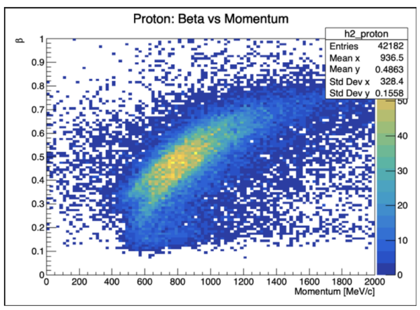


Figure 11: Beta vs momentum plots that include all particles in the beamMC

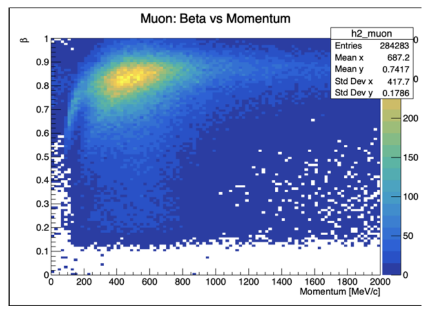
As previously mentioned, figure 9 contains the beta vs momentum plot for all particles. We can of course separate the particles such that each particles has its own beta vs momentum plot. In the dataset, there contains the true PDG number of all the particles in the beamMC file, by only selecting a particle that coresponds to a PDG number of our choice, we can separate the particles from one another. Of course, this can only be done since the true properties of particles are known in the simulation. This technique is often used to evaluate the performances of a detector. In the real data, we would have to separate the particles ourselves. Below are the plots for protons, muons, and pions.



(a) Beta vs momentum plot only for pions in beamMC



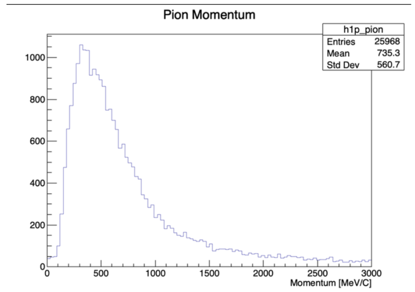
(b) Beta vs momentum plot only for protons in beamMC



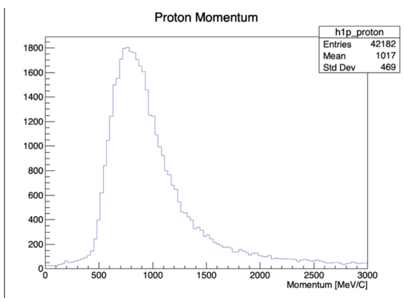
(c) Beta vs momentum plot only for muons in beamMC

Figure 12: Beta vs momentum plots for each particle selected based on PDG number

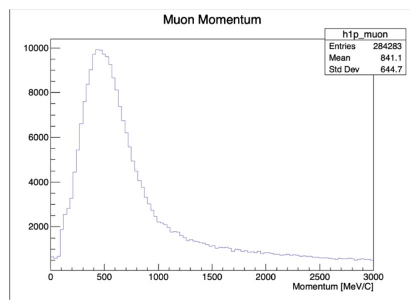
Now that the particles have been successfully separated, we can plot all relevant information of these particles, such as the beta, momentum, and mass.



(a) Momentum distribution plot for pions



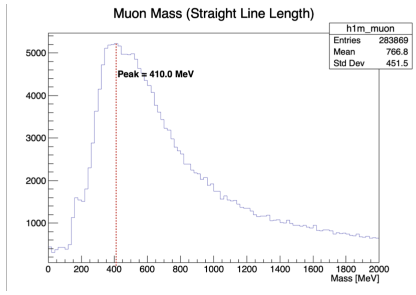
(b) Momentum distribution plot for protons



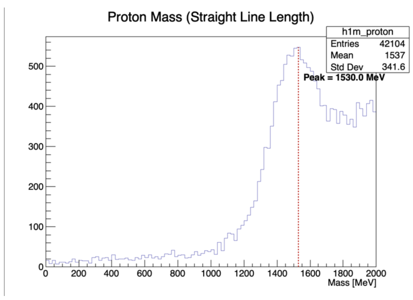
(c) Momentum distribution plot for muons

Figure 13: Momentum distribution plot for each particle selected based on PDG number

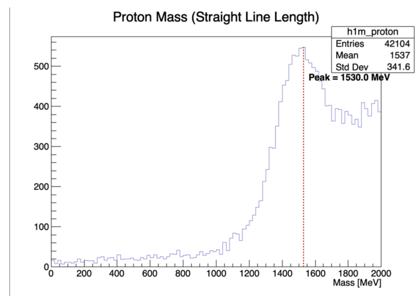
The momentum plots are roughly what we would expect. We see that there are very little low momentum (soft) protons. This is because most of them would decay inside the SFGD and would not be detected. Aside from that there is nothing out of the ordinary.



(a) Muon mass distribution plot, with a red line showing the peak of the mass distribution



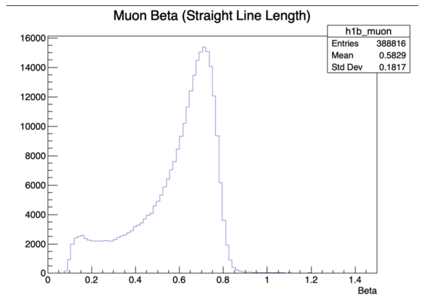
(b) Proton mass distribution plot, with a red line showing the peak of the mass distribution



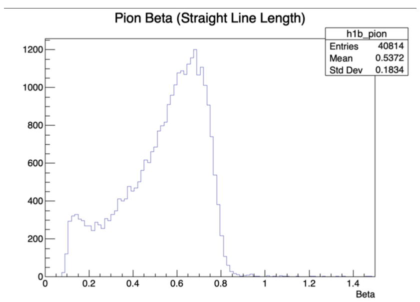
(c) Pion mass distribution plot, with a red line showing the peak of the mass distribution

Figure 14: Mass distribution plots

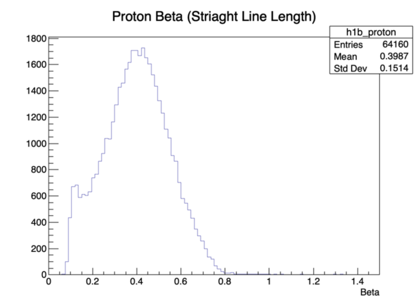
It's obvious from these plots that the masses are drastically overestimated. Of course we don't expect to perfectly reconstruct the masses, however it is nowhere near the actual masses of the particles. From equation (3), we can see that $m \propto \frac{1}{\beta}$. We can then hypothesize that our beta is underestimated. We can confirm this by plotting the beta for each particle:



(a) Muon beta distribution over entire momentum range



(b) Pion beta distribution over entire momentum range



(c) Proton beta distribution over entire momentum range

Figure 15: Beta distribution over entire momentum range

Here we can clearly see that the beta is much lower than what we would expect. There could be two reasons as to why our beta is underestimated, the first is that we are underestimating the length. Second, we could be overestimating the time of flight (TOF) of the particle. Of course, there is also a possibility that there is an issue with both the length calculation and the TOF reconstruction. For simplicity, let's assume that it could only be either a length or a TOF issue. Let's first look at our length calculation; up until now, since the length has been quite simplified and it is estimated as a straight path from the SFGD vertex to the TOF bar hit position, there is a high chance that the underestimated beta is due to our simplilfied length estimation. The assumption that this simplified length is sufficient comes from the fact that at high momentum (what we are more focused on), the particles would not curve that much because of their high velocity. However, it seems that our length is not too simplified for our higher order investigations. Before trying to find the proper calculation for our lengths, we should find the time offset ΔT.

$$\Delta T = T - \frac{length}{c} \tag{7}$$

We can then plot this and get the following:

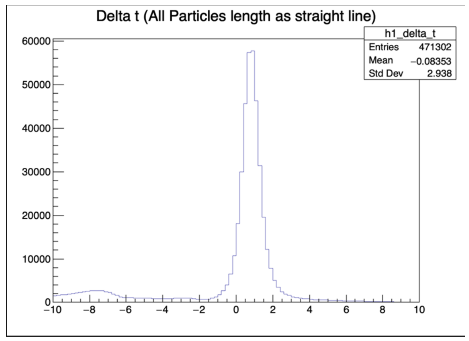


Figure 16: Time offset for all particle when using a length definition that corresponds to a straight line track

If our calculations are accurate, then we should see the plot centered at 0, however, we clearly see that there is an offset of 1 ns. Now, we can proceed with different length estimations. Since we were assuming that particles follow a straight length path, we clearly need to take into account the curved track of the particles. If we recall the structure of the detector, particles would hit multiple nodes in their path. We can sum up the distance of these nodes to get the best track length calculation, however even with this new length definition (Node Length), there is still a significant offset:

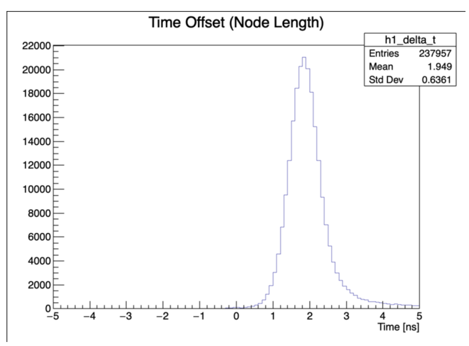


Figure 17: Time offset for all particle when using a node length and reconstructed TOF

Knowing that our new length definition is the most accurate one, this means that there is an issue with our time, or more accurately, our reconstructed time. Since we are not using the true TOF, this means that there is a misalignment in the time in the SFGD and the TOF bar. The beamMC contains all the true properties of the particles, including the true time. Therefore by subtracting the true time with our reconstructed time and plotting it, we get that our time is offset by $\sim 1.8\,\text{ns}$ [2].

[2] We get 1.8ns for our offset is This offset is also applied to the particle guns

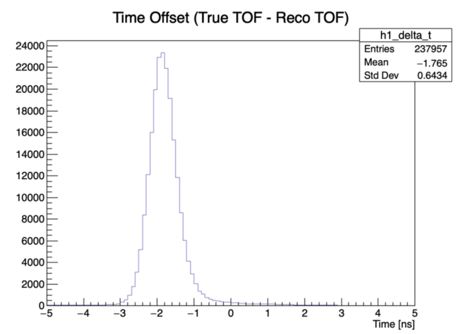
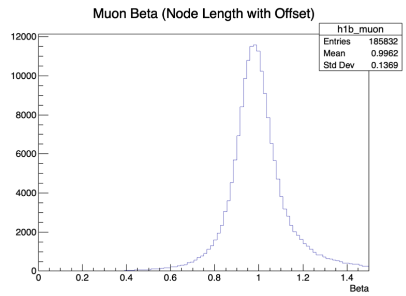


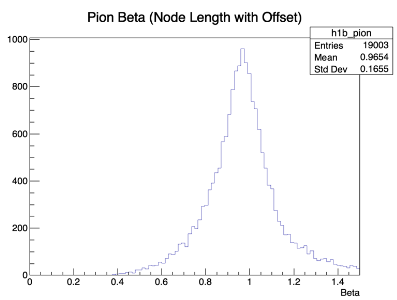
Figure 18: Time offset calculated by True TOF - Reco TOF

By including this offset into our beta calculation, we will now have an accurate beta reconstruction[3]. To check if our beta is accurate enough, we can redo the plots for the beta and masses of the particles. Plotting the beta with the offset, we get the following plots:

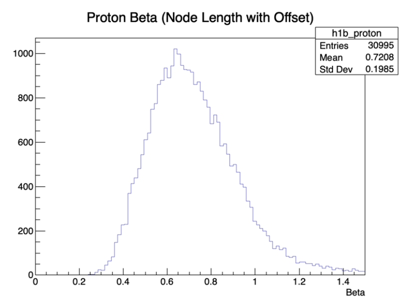
[3] This relies on the fact that it is a fixed time offset, meaning that it is equal for all tracks and particles. This can be understood as a mis-synchronization between detectors. If the time delay between SFGD and TOF depends on specific properties of the event (not having a fixed time offset), the situation would be much more complicated and difficult to correct



(a) Muon beta distribution across all momentum range after using node length and correcting for the offset present in our reconstructed time



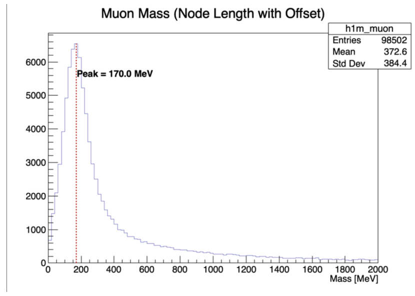
(b) Pion beta distribution across all momentum range after using node length and correcting for the offset present in our reconstructed TOF



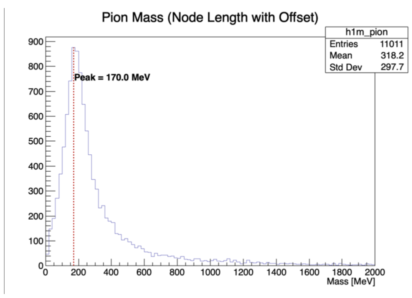
(c) Proton beta distribution across all momentum range after using node length and correcting for the offset present in our reconstructed TOF

Figure 19: Updated particle beta distribution plots

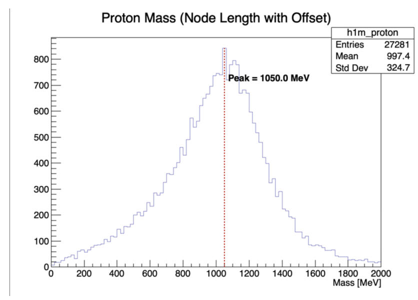
This time, our beta matches what we would expect them to be for each particle. With that it mind, hopefully our masses would also align with their true masses.



(a) Muon mass distribution after using node length and correcting for the offset present in our reconstructed TOF



(b) Pion mass distribution after using node length and correcting for the offset present in our reconstructed TOF



(c) Proton mass distribution after using node length and correcting for the offset present in our reconstructed TOF

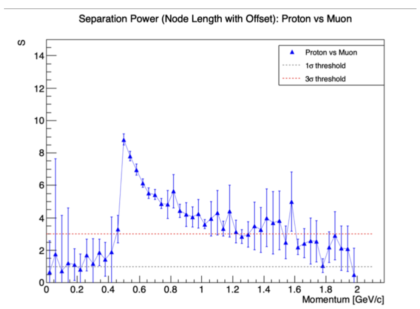
Figure 20: Updated mass distribution after correction to length and TOF calculation

This time our mass plots are significantly better once we included the offset. Of course, it is not 100% accurate, but masses are extremely sensitive quantities to measure. Now, we can proceed with separating the particles based on their TOF.

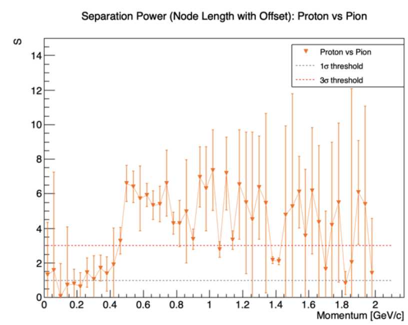
Using the true PDG number of the particles in the beamMC, we can compute the separation power and produce those plots, just like we did for the particle guns. These separation power plots would serve as an indicator to how well we could expect to separate particles. We can also attach error bars for the separation power. The error on the separation power is calculated using the following formula:

$$
\delta S = \left[\left(\frac{1}{\sqrt{\frac{\sigma_1^2+\sigma_2^2}{2}}} \, \delta\mu_1 \right)^2 + \left(\frac{1}{\sqrt{\frac{\sigma_1^2+\sigma_2^2}{2}}} \, \delta\mu_2 \right)^2 + \left(\frac{\mu_1-\mu_2}{2\left(\frac{\sigma_1^2+\sigma_2^2}{2}\right)^{3/2}} \, \sigma_1 \, \delta\sigma_1 \right)^2 + \left(\frac{\mu_1-\mu_2}{2\left(\frac{\sigma_1^2+\sigma_2^2}{2}\right)^{3/2}} \, \sigma_2 \, \delta\sigma_2 \right)^2 \right]^{1/2} \tag{8}
$$

Now we can plot the separation power with the error bars.



(a) Separation power plot between protons and muons in the beamMC with error bars



(b) Separation power plot between protons and pions in the beamMC with error bars

Figure 21: Separation power plot in the beamMC with error bars.

The separation power for protons and pions seem a lot more chaotic compared to protons and muons. This is most likely due to low statistics[4] of pions. This is also the same reason why at higher momentum the error bars for both plots get wider.
What we can do is calculate the 'pull' of a certain particle. One can think of pulls as a way to quantify how close an analyzed particle is to the hypothesized particle. We calculate pulls using the following formula:

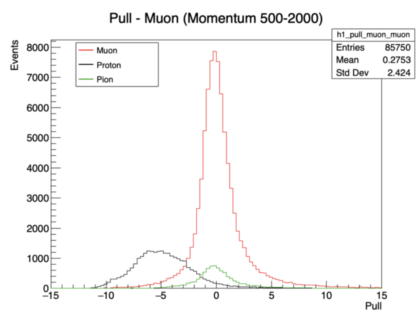
$$
P_i = \frac{\beta_i - \beta_{\exp_i}}{\sigma_{\exp_i}} \tag{9}
$$
[5]

The closer an analyzed particle is to a specific particle, the lower their pull would be. For example, a particle that we analyzed and think is a muon, if we calculated it's muon pull and it is 0, we with relative confidence say it is indeed

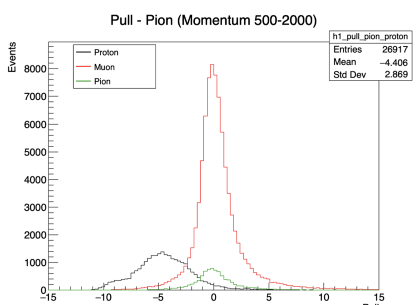
[4]statistics means the number of entries, higher no. of entries means higher statistics

[5]P_i: Particle i pull; β_i: Measured beta; $\beta_{\exp_i}$: expected beta for particle i; $\sigma_{\exp_i}$: Standard deviation of particle$i$.

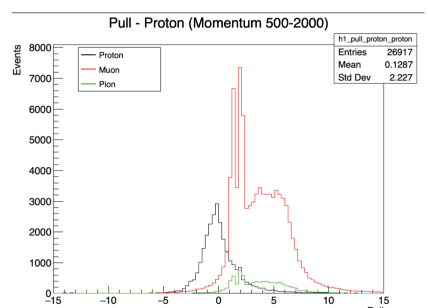
a muon. Based on the puls, one can build a so called discriminator, which would assign an identity to a particle based on a certain criteria. In this project, the discriminator is simply whichever pull is lowest in absolute value. For example, if we have three pulls (proton, muon, and pion), if the muon pull is 1.2, the pion pull is 1.1, and the proton pull is 0.3, we would assign this particle as a proton. From eqn (9) the pull of the particle is determined by the expected β and σ. We get these values from the particle guns for the given particle. It is also important to understand that when using particle guns as the reference, we need to use particle guns that are generated from the same software as the Monte Carlo events. This allows for the datasets to be consistent with one another, this also means that the same problems would occur in both the particle guns and the beamMC. However they would have the same solutions (such as the issue with the offset). From these pulls, we are able to produce the following plots:



(a) Muon pulls for all particles for momentum of 500-2000 MeV



(b) Pion pulls for all particles for momentum of 500-2000 MeV



(c) Proton pulls for all particles for momentum of 500-2000 MeV

Figure 22: The pull for protons at very low momenta (<600 MeV/c) in not defined due to the low statistics. When the pulls are evaluated to assign a particle type, we end up with either muons or pions, which results in the peaks that we see.

The pion and muon pulls are relatively normal, there is a slight bias for the pion and proton pulls (note how the pion pulls for the pions aren't exactly centered at 0). However, the proton pulls is interesting to say the least. There appears to be two peaks where for the pions and muons. We investigate this issue by plotting the proton pulls in a narrow momentum range, let's say every 200 MeV, starting from 500 MeV. Another thing we could do is also look at the beta distribution per slice. It could be that in a certain momentum range, there could be a very large spread in the bets distribution, meaning that the "accepted beta range" for a particle to be identified aa a proton would be quite large for that slice.

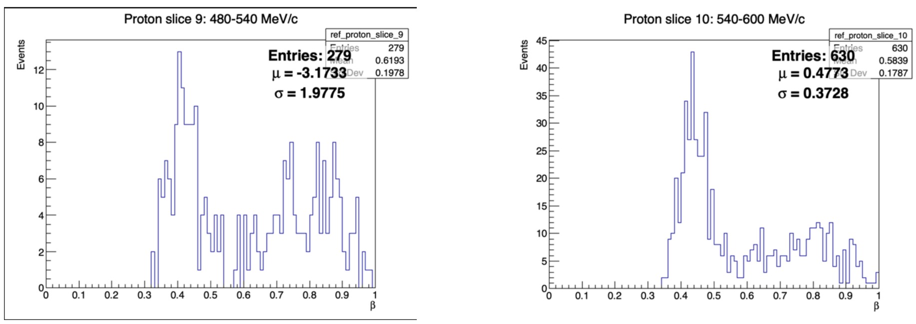


Figure 23: Beta distribution slice for protons in particle guns

We have found the problematic momentum range, we can see for very low momentum($\leq$ 600 MeV), there are not that many entries, but the spread is very large. This causes the mean to shift and of course the large spread leads to pions and muons to be mistaken as protons. By ignoring entries below a momentum of 600 MeV, the weird peak would disappear.

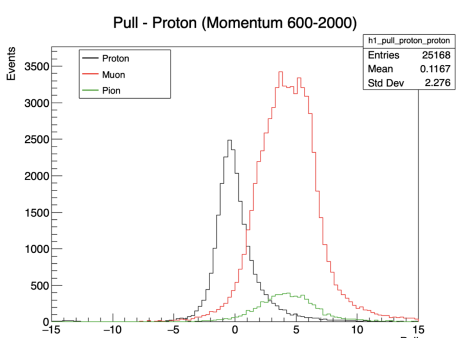
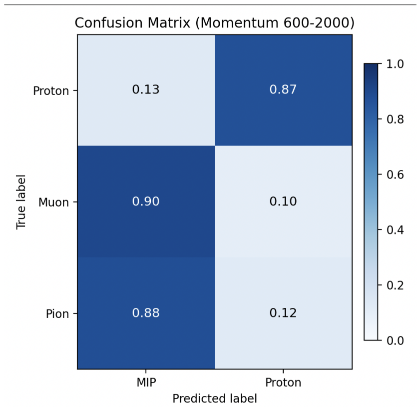


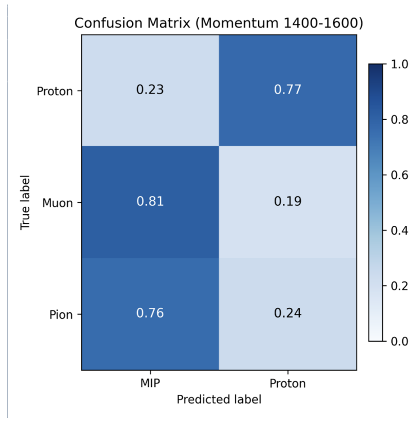
Figure 24: Proton pulls for all particles updated using new momentum range that does not include the problematic region

Indeed, it is no longer present.

Now that we have computed the pulls, we can create a confusion matrix. A confusion matrix can be thought of as a visualization of how well our particle identification is doing. The matrix is made such that it is a 2x3 matrix, where the Y-label represents the true identities of the particles (Protons, Pions, Muons), and the X-label represents the assigned identities of the particles (MIPs[6], Protons), since pions and muons would not be distinguishable from TOF, hence why we combine them as one entry. A perfect confusion matrix would look like a diagonal matrix. The entries in the matrices are also row-normalised to show what percentage of protons, pions, or muons, we are identifying correctly.



(a) Confusion matrix in a momentum range of 600-2000 MeV

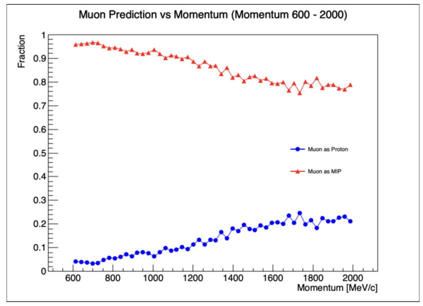


(b) Confusion matrix in a momentum range of 1400-1600 MeV

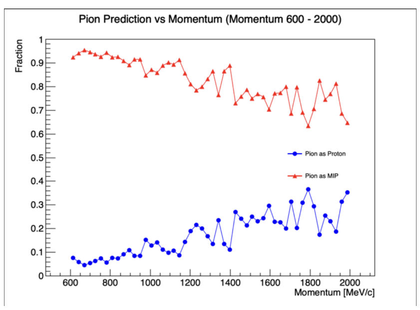
Figure 25: Confusion matrices

We could take it one step further with the confusion matrices. We could for every momentum slice, we could take the predicted results and plot it against the momentum. We will do this for each particle. These plots basically tell us what percentage of particles are we identifying correctly for a given momentum slice.

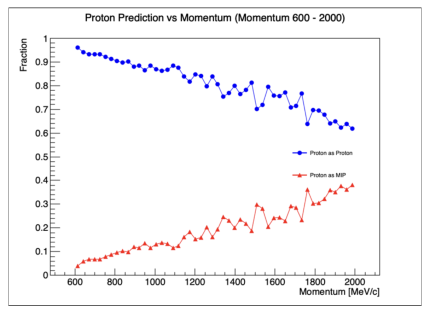
[6] Minimum Ionizing Particles



(a) Plot for prediction of muon identification vs momentum



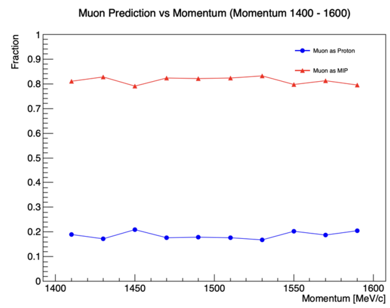
(b) Plot for prediction of pion identification vs momentum



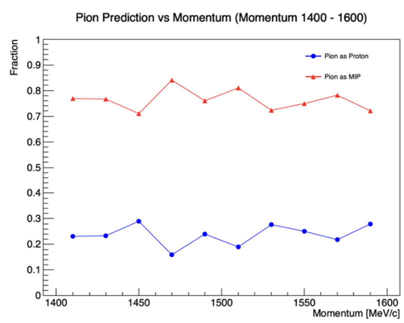
(c) Plot for prediction of proton identification vs momentum

Figure 26: Plot for prediction of particle identification vs momentum

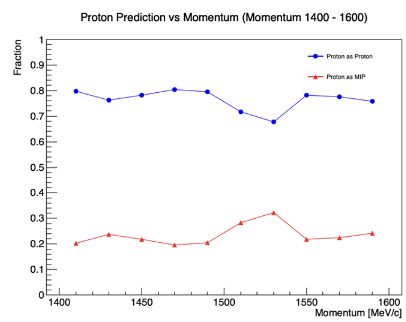
And these are the same plots but restricted to the momentum range that we are interested in:



(a) Plot for prediction of muon identification vs momentum in momentum range of 1400-1600 MeV



(b) Plot for prediction of pion identification vs momentum in momentum range of 1400-1600 MeV

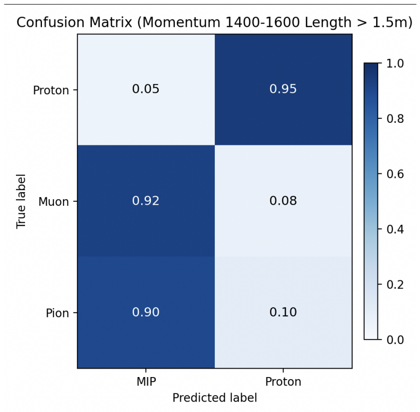


(c) Plot for prediction of proton identification vs momentum in momentum range of 1400-1600 MeV

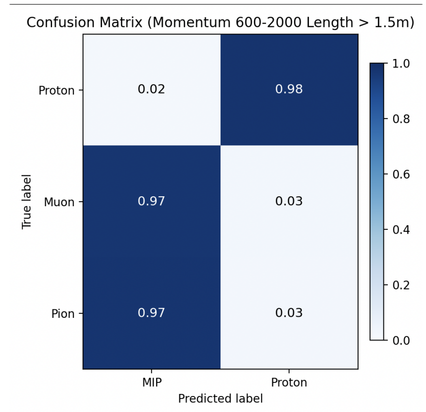
Figure 27: Plot for prediction of particle identification vs momentum in momentum range of 1400-1600 MeV

One might have noticed that the confusion matrix plots are worse than what was indicated by the separation power plots. What is going on? A possible explanation could be is simply that we are using different samples. For the separation power plots, σ_i and μ_i are taken from the beamMC. This is also why at larger momenta, the separation power gets fuzzier and the error bars become larger, due to lower statistics at higher momenta. The confusion matrices are based on the pulls which itself uses the particle guns as the reference. This means that we are using different samples to compare and discriminate particles in the confusion matrix. This is not done in the separation power plots since it gets the σ_i and μ_1 within its own dataset.

Similarly to what we did for the particle guns, we can only select entries with track lengths $\geq$1.5m. This would basically tell us the best performance we can achieve with out PID.

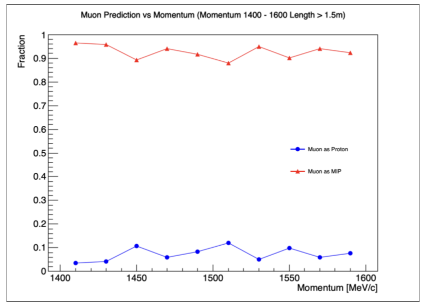


(a) Confusion matrix in momentum range of 1400-1600 MeV only selecting entries with track length $\geq 1.5m$

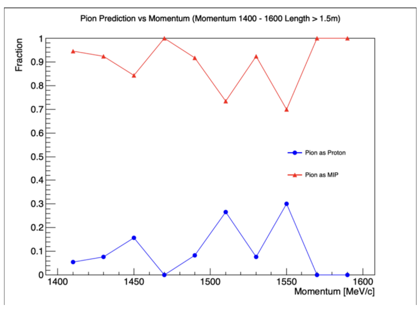


(b) Confusion matrix but only selecting entries with track length $\geq 1.5m$

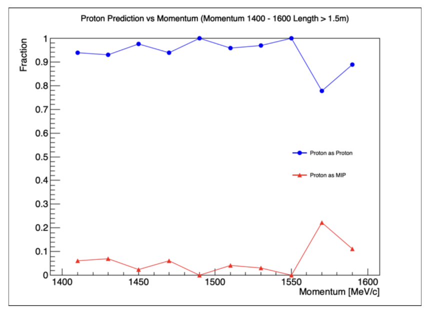
Figure 28: Confusion matrices with selection criteria



(a) Plot for prediction of muon identification vs momentum in momentum range of 1400-1600 MeV only selecting entries with track length $\geq 1.5m$

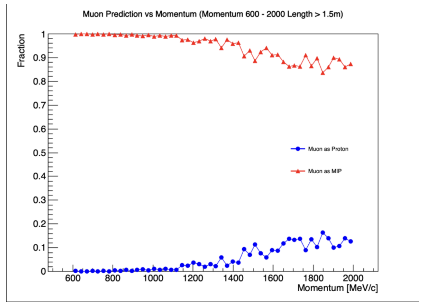


(b) Plot for prediction of pion identification vs momentum in momentum range of 1400-1600 MeV only selecting entries with track length $\geq 1.5m$

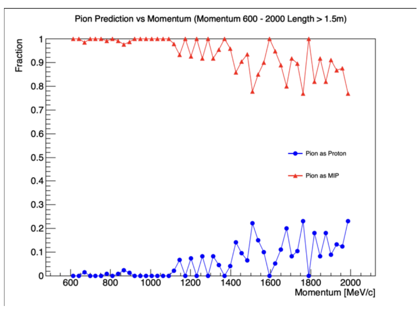


(c) Plot for prediction of proton identification vs momentum in momentum range of 1400-1600 MeV only selecting entries with track length $\geq 1.5m$

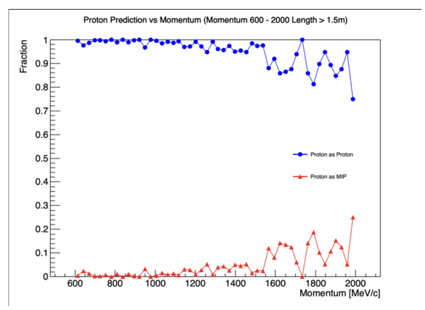
Figure 29: Plot for prediction of particle identification vs momentum in momentum range of 1400-1600 MeV with selection criteria



(a) Plot for prediction of muon identification vs momentum, only selecting entries with track length $\geq 1.5m$

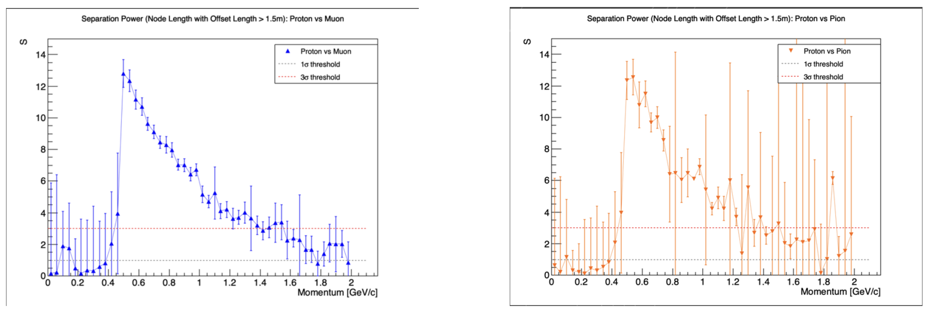


(b) Plot for prediction of pion identification vs momentum, only selecting entries with track length $\geq 1.5m$



(c) Plot for prediction of proton identification vs momentum, only selecting entries with track length $\geq 1.5m$

Figure 30: Plot for prediction of particle identification identification vs momentum, with selection criteria



(a) Separation power of protons and muons in beamMC using node length, reconstructed TOF with corrected offset, and selecting entries with track length $\geq 1.5m$

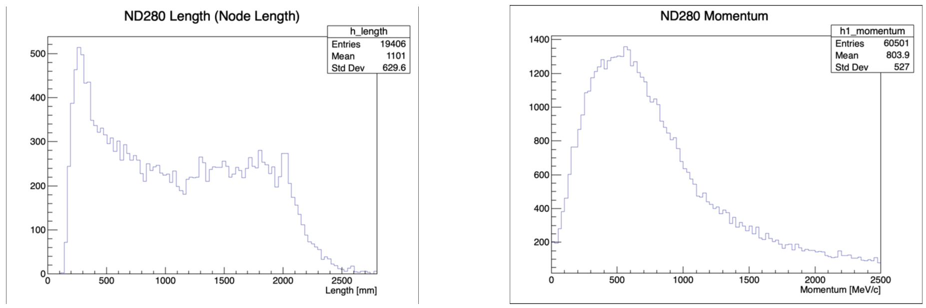
(b) Separation power of protons and pions in beamMC using node length, reconstructed TOF with corrected offset, and selecting entries with track length $\geq 1.5m$

Figure 31: Separation power between protons and MIPs in beamMC using node length, reconstructed TOF with corrected offset, and selecting entries with track length $\geq 1.5m$

Same as with the particle guns, longer track length, means we have better momentum and time resolution. Hence why our ability to identify particles become significantly better. The goal with the beamMC is to develop an understanding of how one would attempt to separate and identify particles from a combined dataset. This is standard procedure in experiments, where we would familiarise ourselves with the Monte Carlo, then use that information to see if the data is consistent with the Monte Carlo. Now that we have essentially performed particle identification using the Monte Carlo datasets, we will proceed to looking at the real data generated by ND280.

5 ND280 Data

Now that we are very acquainted with the data from the Monte Carlo, we can apply what we have done to the the real data. Similarly for the Monte Carlo data, we will plot all relevant characteristics, such as the length, TOF, momentum, beta vs momentum, etc.



(a) Track length distribution in ND280 data

(b) Momentum distribution in ND280 data

Figure 32: Plots for particle properties in ND280 data

These plots for the momentum and length are normal. However, we were unable to plot the beta, TOF, and beta vs momentum. This is because of a time offset.

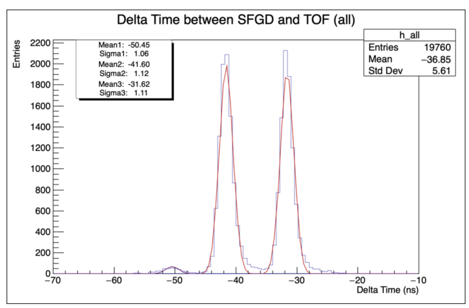


Figure 33: Time offset between SFGD and TOF in ND280

The three peaks for the time offset is interesting to say the least. Normally, if there is a time offset present, we expect it to have one peak, just like in Fig. 17. Moreover, the difference between each peak is almost exactly 10ns. The cause of this is due to a "time slip" effect. To know what is causing this time slip effect, we need to understand how time is being recorded in the detector. To put simply, the TOF and the SFGD would record the time starting on an initial time t_0 which is determined by a signal sent by the Master Clock Board (MCB). However, the TOF and SFGD might receive the signal at different time, for reasons such as having different cable lengths. To fix this, we perform what is called inter-detector calibration to ensure that both SFGD and TOF have the same t_0. This would be fine if it is a consistent offset. But that is not the case with the three peaks. The MCB would send a signal as a square wave every 10 ns. The SFGD and TOF could skip a cycle, if it happens consistently, then we could corrct for that. However, the TOF and SFGD do not process the same signal. SFGD would process it as is, but TOF would downsample it (20 ns ticks). So, if the MCB slips by one tick, SFGD would shift by 10ns, and

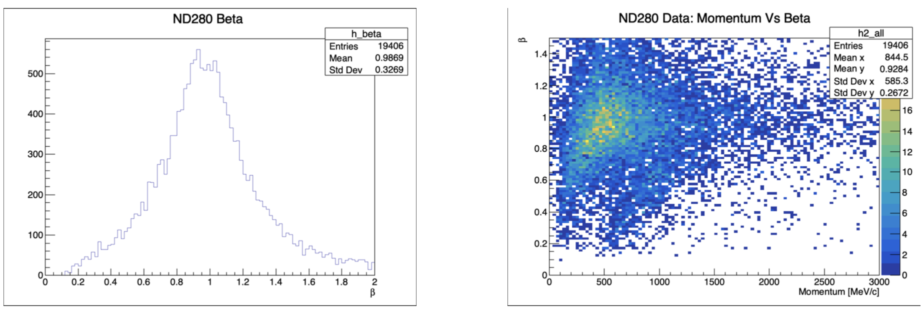
TOF would shift by 20ns, this is why we observe a $\pm 10ns$ between the peaks. The exact reason as to why there is this time slip issue is still as of yet, unknown.

We can still do a simple fix to at least be able to work with the data. This effect can be correct by simply selecting a range of entries in that fall within 4σ of a gaussian. Then, for those entries, we would include an offset[7] (which is just the mean of the gaussian) to the TOF calculation. So our new TOF formula would look something like this:

$$TOF = (T_{TOF} + |\mu_i|) - T_{SFGD} \tag{10}$$

Of course, this does not fix the source of the issue, but it allows us to correct for this offset so we can still analyse the other datasets.

Now that we have corrected for the offset, we can create the plots that require the TOF.



(a) Beta distribution for particle in ND280 data with corrected offset

(b) Beta vs momentum for all particles in ND280 data with corrected

Figure 34: Plots that with offset that has been corrected for

Immediately, we can tell that the data is not as clean as the MC. However, one might be able to make out a slight shoulder on the for the beta. In the beta vs momentum, one can discern two bands, one that corresponds to the muons, and one that corresponds to the protons. We also see that the two bands start to merge around 1.5 GeV, which lines up with what we would expect and also what we see in the beamMC. What we could do to clean up the results a bit is to only select long track in the SFGD, however in this case, we would not be able to do so. The reason why, is simply because of our selection criteria for the data that is being collected. We only select entries that register one hit in the SFGD and one hit in the TOF bar. Protons are normally quite soft, which means they have a short track length. This would lead to us losing the majority if not all the protons, since barely any would have sufficiently long tracks. With increased statistics, this study will be made possible

[7] Left most peak, $\mu = -50.4ns$. Central peak, $\mu = -41.6ns$. Right most peak, $\mu = -31.6ns$

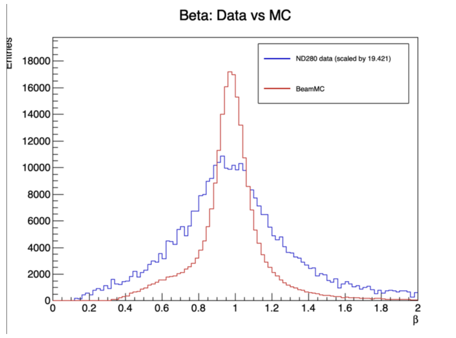
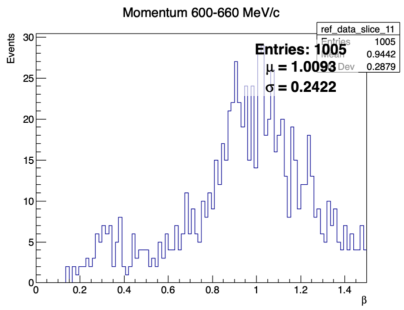


Figure 35: Beta comparison between ND280 data and beamMC

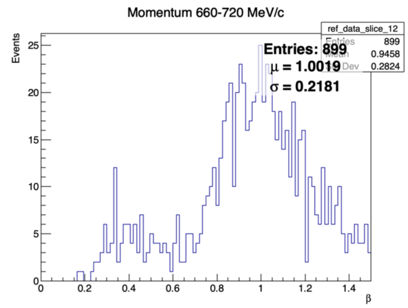
An intermediary step that could be done is that we could overlay (and rescale) the beta distribution plot from ND280 data and beamMC. There are some obvious differences, for example, in the actual data, it's more spread out and has a more subtle shoulder. The data is simply behaving worse than the MC, causing these large spreads[8]. Something interesting to point out is what could potentially be two peaks in the real data. It's hard to say for sure if those are two peaks of if it just flattens out. The reason behind a potential second peak can be chalked up to out offset. recall that we are incorporating 3 different offsets for the events. Since we are these offsets are obviously different, this could lead to potential extra peaks

Similarly to what we did previously, we will select each slice in the 2D histogram and create distribution plots for the beta.

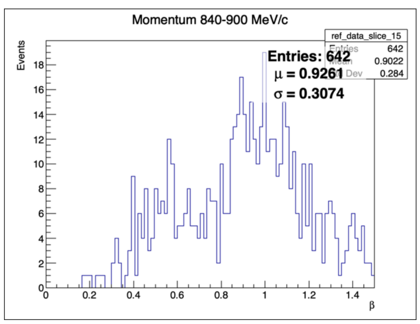
[8]ND280 data in rescaled to an order of 20



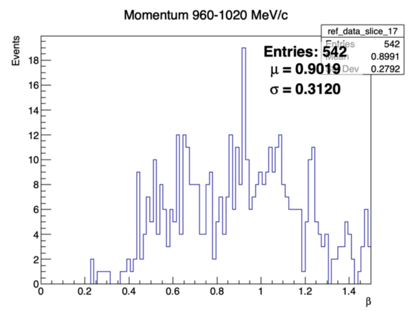
(a) Beta distribution in momentum slice of 600-660 MeV



(b) Beta distribution in momentum slice of 660-720 MeV



(c) Beta distribution in momentum slice of 840-900 MeV



(d) Beta distribution in momentum slice of 960-1020 MeV

Figure 36: Beta distribution in momentum slices of ND280 data

We definitely see the presence of a second peak, which is quite motivating to see. Of course, we do see the two bands in the 2D histogram plot, but due to how messy the data was, it is very encouraging to be able to discern two separate peaks. This means that even at this early stage in ND280, there is an indication that particle identification by time of flight could be done.

6 Further Work

There are still a few smalls tweaks (and some large ones) that could done to the data and the beamMC. The MC should be tuned to better match what we are seeing in the data. Recall the larger spread for the real data compare to the MC. We want the MC to better simulate what we are observing in the data. What one could do is to get the momentum as the particle leaves the SFGD instead of when it first enters. This is because the the particle would lose some of its momentum in SFGD (since it is made of lots of scintillating cubes), by taking the momentum at the exit position/node, we would be able to get more accurate momentum for the particle since it would be traveling in a gaseous TPC[9]. Another thing that could be done is to only select entries with long SFGD tracks, as mentioned before, we would need a larger sample size for the real data. In the real data, what one could hope to do is to take each entry and correct it using its own corresponding offset. This would be quite complicated but it will lead to us having a much more accurate time reconstruction. But the best thing that we could hope to do is to understand the root cause of this issue and fix it, so there would be no correcting that would need to be done. Furthermore, the tree files could be generated again to include the charge information of the particles. This would (in theory) allow us to separate the protons from muons easily, just by looking at their charge. It's not as simple as that, since protons and muons aren't the only particles that are detected.

7 Conclusion

In this report, we have shown that PID by TOF in the ND280 detector is indeed feasible, with an average separation power of around 2.5σ, in the momentum range of 1.4 GeV to 1.6 GeV, between protons and muons, based on the beamMC. These results assume a fixed timing offset of 1.8 ns, requiring long reconstructed tracks, and a large sample size. However, these analyses in the actual data rely on the detector conditions is still riddled with issues, such as inter-detector calibration errors, which cause time slip. Once issues such as the time slip are resolved and larger data samples are available, future work can extend this approach to higher-level analyses, including mass reconstruction and the implementation of pull-based methods for particle discrimination. This report provides a benchmark for TOF-based PID in ND280. In particular, it demonstrates that in momentum ranges where PID by energy loss (dE/dx) cannot separate protons from muons, TOF offers statistically significant discrimination power.

[9]We lose less momenta per material traversed. Energy loss is measured ($\frac{dE}{dx}$ in a TPC)

References

[1] C. T. R. Wilson. On a Method of Making Visible the Paths of Ionising Particles through a Gas. *Proceedings of the Royal Society of London Series A*, 85(578):285–288, June 1911.

[2] Carl D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, Mar 1933.