Maximum Likelihood for Particle Identification in NDLAr

Brooke Schuld

Computational Physics, May 2023

1 Introduction

The Deep Underground Neutrino Experiment (DUNE) is an international collaboration dedicated to measuring neutrino properties. Currently in the developmental phase, the experiment will send a beam of neutrinos produced at the Fermilab accelerator in Batvia, Il to an underground mine in Lead, Sd roughly 800 miles away.

The DUNE near detector is a liquid argon detector located at the Batvia site. This project focuses on track reconstruction in this detector, named NDLAr. There are two main types of reconstructions that occur inside the 2x2 module- shower like and track like. Track like events include those containing particles like protons, muons, and pions. Shower type events are those like electrons and photons. The identification of particles with track-like interactions can be determined through their energy deposits. This code attempts to tag these tracks as proton-like, pion/muon-like and (to a lesser extent) kaon-like.

The characteristic curve for energy deposition inside a material is given by the Bethe-Bloch formula.

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

However, this equation is a function of momentum, which is not a characteristic measured directly by the detector. Integrating the inverse of the Bethe-Bloch equation over the energy gives the distance the particle could travel before it comes to a stop. This quantity is called the residual range.

$$\Delta x = \int_0^{E_0} \frac{-dx}{dE} dE \tag{2}$$

As the particle reaches the end of the track, $\frac{dE}{dx}$ as a function of residual range, creates a unique signature that can be used to identify the particle. The proton and pion tracks differ the most.

2 Monte Carlo Simulations

Since DUNE is not yet collecting data, this code relies on Monte-Carlo generated events produced by GEANT4 and GENIE particle propagation code with the 2x2 module geometry. Pre-generated files are available at

https://portal.nersc.gov/project/dune/data/2x2/simulation/edepsim/

A simple look at the residual range vs energy deposition for protons and pions reveal close agreement between the simulated events and the theory.

3 Code Breakdown

The code itself relies on a few main functions. The first function extracts the simulated dE/dx and residual range for the given particle by looping over the detector segments and looking for places in the detector where the given particle deposited energy. The next two, *bethe_bloch* and *residual* return the dE/dx and residual range for a particle with that momentum. These are used to compare the measured values with the theoretical ones to determine which curve the data best fits. The function *id* calculates the reduced χ^2 value for each of the three curves. The main part of the code loops over each individual events to decide which particles are possible to identify. This includes a few conditions. Firstly, all electrons are immediately skipped. Since they exhibit shower behavior tagging them is beyond the scope of this project. Secondly, the particle must be contained within the 2x2 active volume. The end of the track is the easiest to differentiate, means it must stop within those dimensions. In order to remove particles whose tracks ended because they re-interacted, the reconstructed energy must be at least 90% of the particles initial energy. If these conditions were met, an attempt was made to match the particle to a particular type. If no chi square value was satisfactory, the code returned "no match found". The match was the curve with the best chi-squared result and the code returned the particles expected PDG code.

4 Results

An analysis from one *edep_sim* file particle returned ID's for 1000 particles ($78\%, 8\%, .2\%, 14\%$ protons, pions, kaons, and other respectively). Of those, 79%. 2% were protons were misidentified, as were 43% of pions. Many of these were mistagged as kaons. Additionally, most of the "other" particles were neutrons. These were almost exclusively tagged as pions, drastically affecting the pion success rate. After removing shower particles like electrons, the three track conditions mentioned above removed 99% of the tracks. Of those, a match was made about 52% of the time. When adjusting the track requirement to only include 3 hits, the rejection rate decreased to 96%, but the success rate dropped to only 53%. There are a number of improvements that could be made to this code. The addition of kaons significantly decreased it's effectiveness especially when tagging pions. Other adjustments would need to be made to be applicable to real data, including angle and plane corrections. Additionally, it would be beneficial to standardize a PID to improve certainty in the tagging rather than just relying on chi-squared values.