Cosmic Ray Muon Calibration Update

Rhiannon Jones,
Vitaly Kudryavtsev,
University of Sheffield

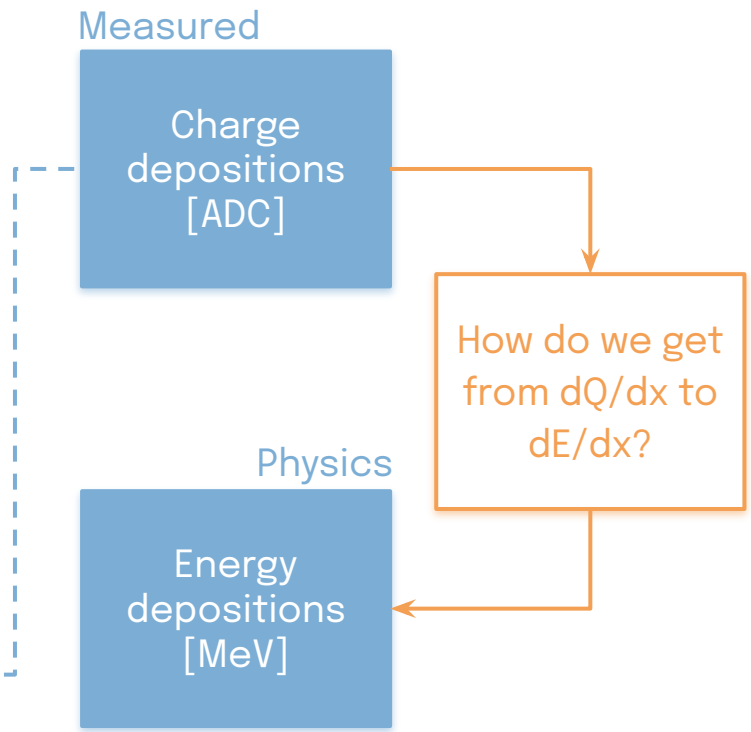
DUNE UK Project Meeting, January 2022

Introduction

- I am aiming to determine the energy scale [MeV/ADC] in a single 17kt SP DUNE FD module
- It is useful to use CR muons for this, since
  - Capable of providing reasonably consistent dE/dx behaviour
  - Available before beam is switched on
  - They are a natural (free) source of calorimetry data
- However, our underground CR muon rate is very low
  - Only ~4,000 per day across the entire module
- Towards this, I have been looking at the dependence of charge depositions on a number of parameters
  - Studied after 2.8 ms lifetime corrections

Viktor has previously done a lot of work towards accounting for the electron lifetime and diffusion in measurements of the dQ/dx of a CR muon.

A recent update, from the May 2021 CM, is **here**.
The most recent calibration WG meeting is **here**.



Samples involved in the calibration studies

Looking at around 400,000 total (true) events which corresponds to around 1 month of data (at ~14000/day total) in one 17 kt FD module.

- See **backup** for a breakdown of the statistics.

CR events were simulated using the Sheffield muon generator, **MUSUN**, and propagated through GEANT4, Detsim and Reconstruction.

Throughout the process of determining a dE/dx calibration procedure, I have compared two methods of reconstructing these events

1. The Pandora v-chain
2. The Pandora CR-chain

I will only be show results with the **CR-chain**.

Truth-level selection and rates

From the GEANT4 tracks, I select:

- Those which enter the TPC geometry
- Primary muons
- Longer than 3m

  See backup for cut motivation
  (defined in reco, same in truth for consistency)

This gives ~110,000 long, primary muons in ~1 month of data.

58% cross both the top and bottom faces & 2% stop.

Reconstructed 'selection' requirements

From the reconstructed tracks, I select:

- Those with a truth-level muon pdg code
  - Can't check whether it is primary yet
- Longer than 3m
- Start Y-position is > 599.5 cm

  See backup for cut motivation

This gives ~110,000 long muons in ~1 month of data.

45% cross both the top and bottom faces & 2% stop.

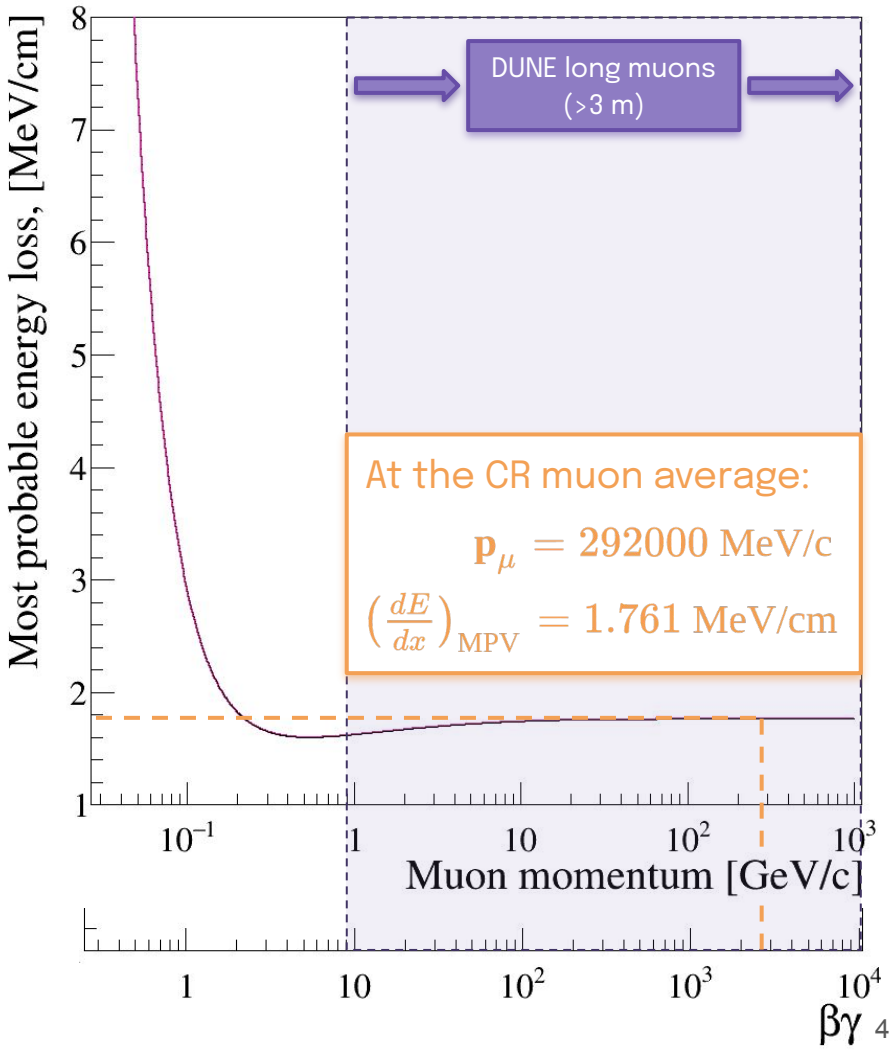
Energy-dependence of the MP muon energy loss

In an ideal world, it would be possible to define the energy-scale calibration in 2 steps:

1. Determine the most-probable reconstructed charge deposition (dQ/dx) for the defined muon sample
2. Convert to dE/dx with a single scaling parameter, calculated using either truth information or theory

In reality, the charge/energy depositions are not independent of energy (see RHS).

I am trying to determine the best way of dealing with this in the calibration.



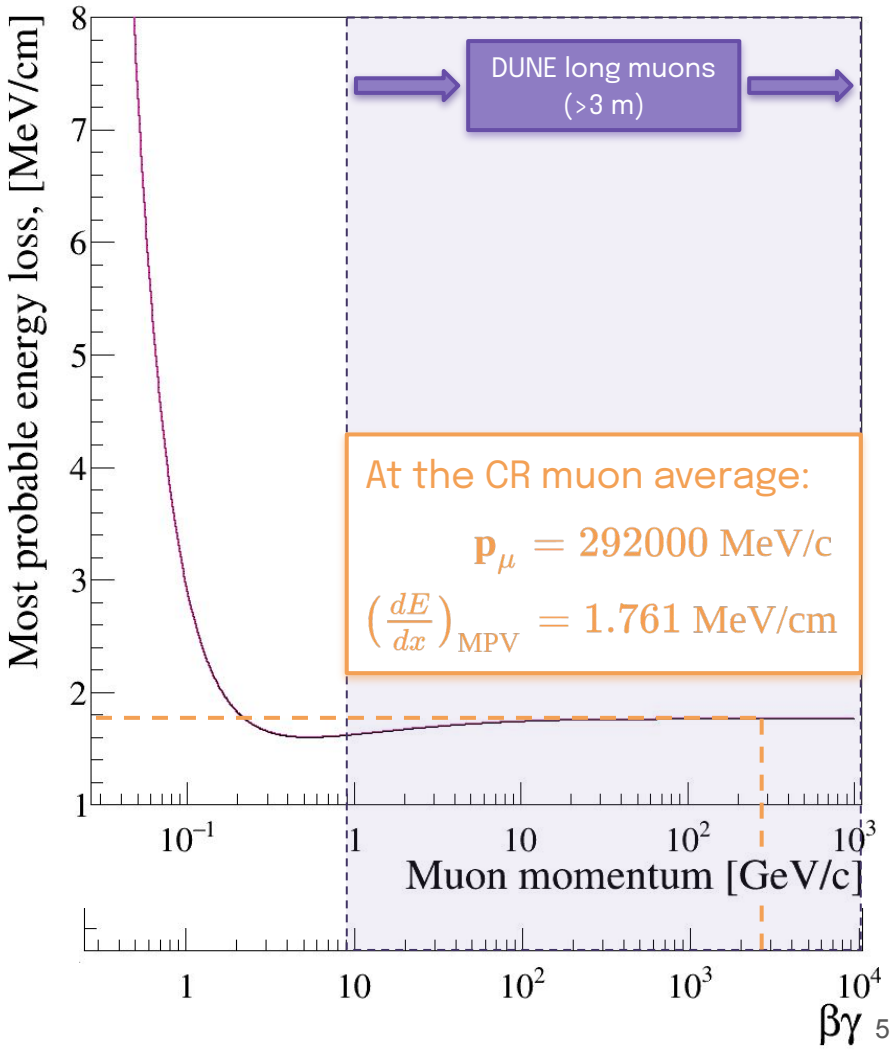
Energy-dependence of the MP muon energy loss

In an ideal world, it would be possible to define the energy-scale calibration in 2 steps:

1. Determine the most-probable reconstructed charge deposition (dQ/dx) for the defined muon sample
2. Convert to dE/dx with a single scaling parameter, calculated using either truth information or theory

In reality, the charge/energy depositions are not independent of energy (see RHS).

I am trying to determine the best way of dealing with this in the calibration.



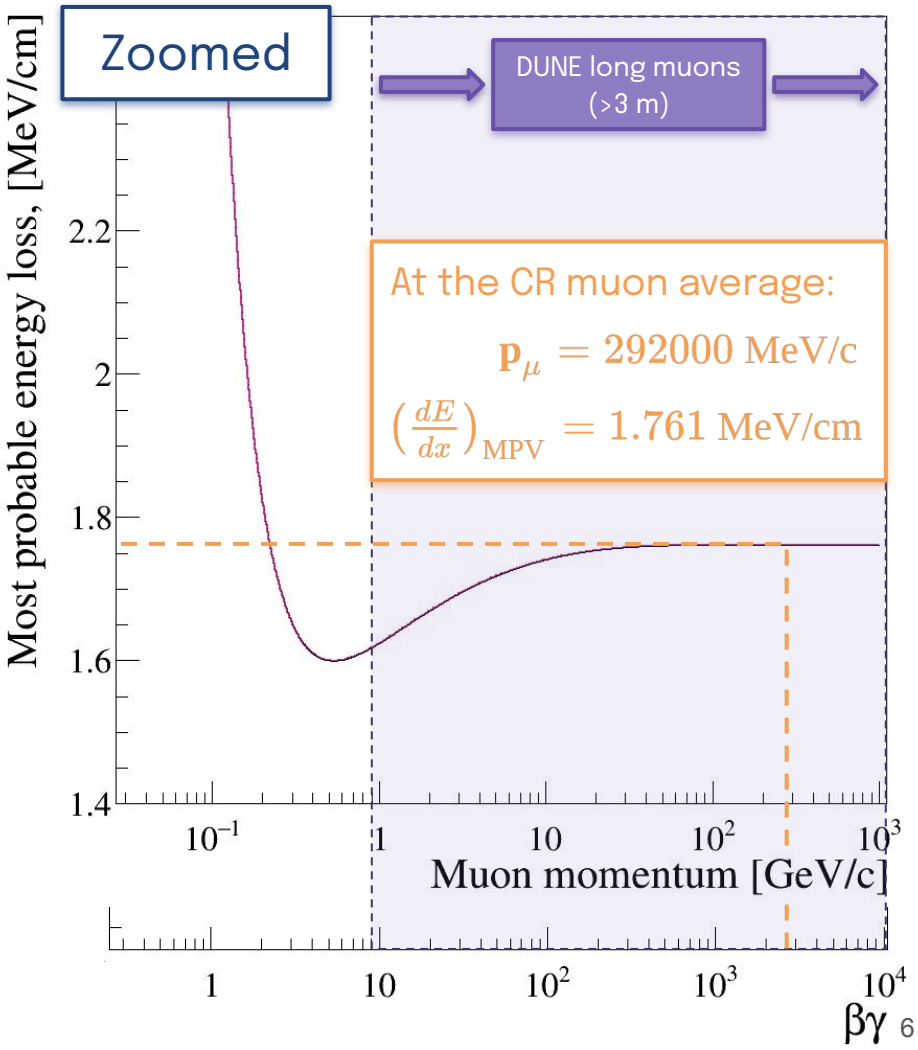
Energy-dependence of the MP muon energy loss

In an ideal world, it would be possible to define the energy-scale calibration in 2 steps:

1. Determine the most-probable reconstructed charge deposition (dQ/dx) for the defined muon sample
2. Convert to dE/dx with a single scaling parameter, calculated using either truth information or theory

In reality, the charge/energy depositions are not independent of energy (see RHS).

I am trying to determine the best way of dealing with this in the calibration.



6

Overview of the studies

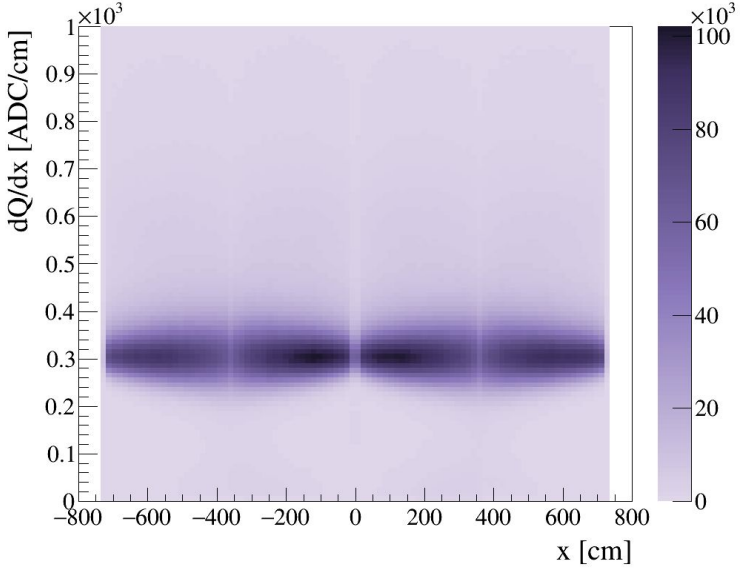
- Utilising the high rate of through-going muons, ~85% of which enter through the top and leave through the bottom of the detector
- Looked at a number of possible parameter dependencies of the reconstructed dQ/dx, in order to determine where any dependencies exist and might need to be addressed
- I began with a simple assessment of the x-position-dependence
  - Given the lifetime correction has been applied and I am not considering any diffusion, dQ/dx should not depend on this parameter
- I have since developed the studies to assess the energy dependence along with the following parameters, chosen based on similar work in SBND and ICARUS by Lan and Gray
  - Residual range
  - Hit width
  - Hit pitch
  - $\cos\theta_{Drift}$

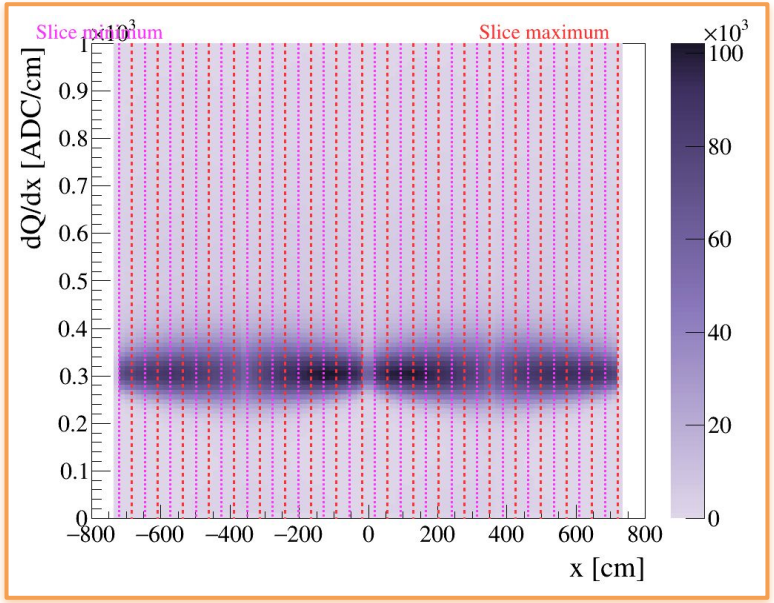
Overview of the method

- I plot the reconstructed, lifetime-corrected dQ/dx against each parameter
- I then slice up the distributions in the parameter space in order to fit a Landau*Gaussian to determine the most probable (MP) dQ/dx in each slice
- I then plot the MP values (MPVs) vs each parameter to determine whether a constant MPV can be extracted, or whether there is some linear or non-linear variation of dQ/dx which might need addressing
- It is then possible to define a scale factor/function by which to translate from the MP charge [ADC/cm] to the 'nominal' dE/dx, 1.761 MeV/cm
  - Value from here for a 292 GeV/c muon with a 0.353 cm hit 'thickness' (pitch+diffusion)
  - The nominal dE/dx value can be interchanged very easily, or could be substituted with a function
  - This step is mainly for demonstration purposes at the moment
- A constant value may not be valid if the depositions depend too strongly on energy
  - This is one of the focal points of my work at the moment

Scaling between dQ/dx and dE/dx, x-dependence

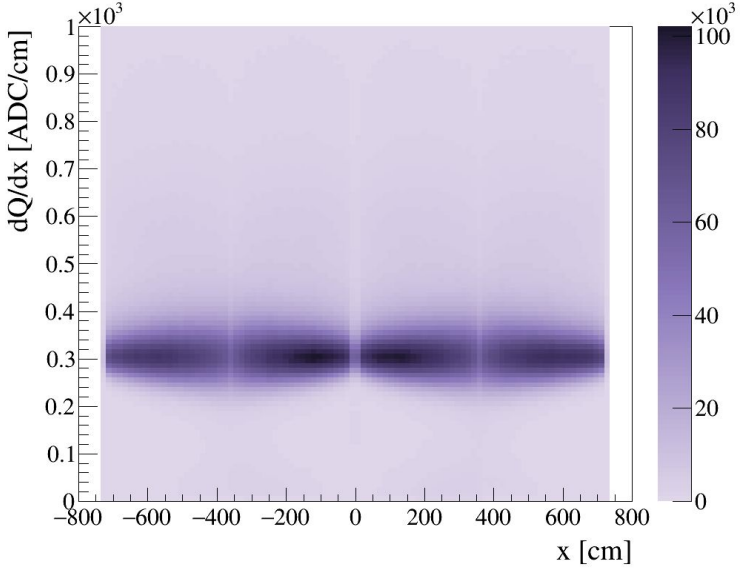
- There are 20 slices between ±720 cm
- Each are 2.5% of the width of the parameter space
- Separated by 2.5% of the width of the parameter space

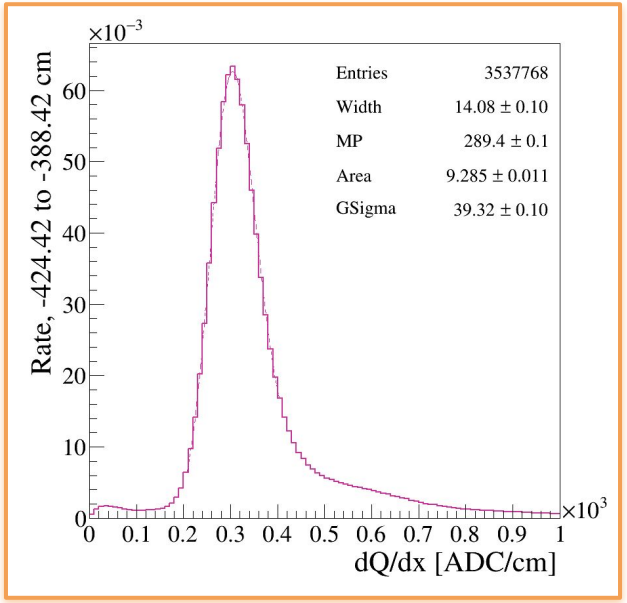




Scaling between dQ/dx and dE/dx, x-dependence

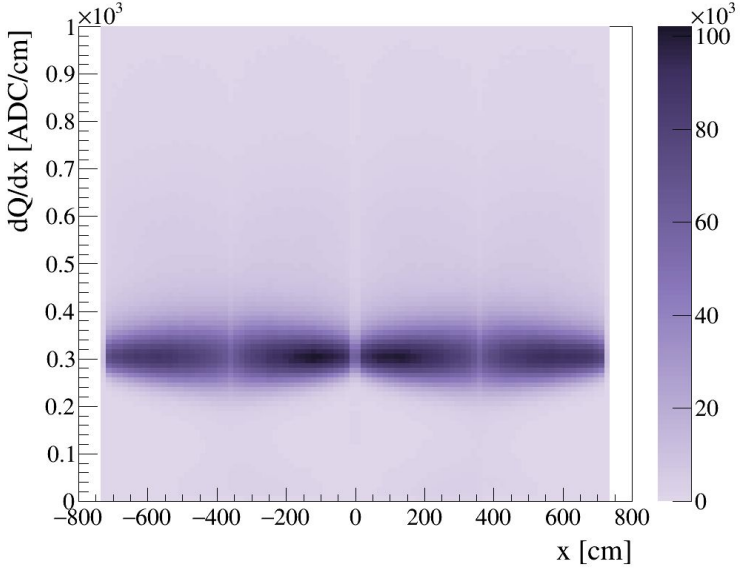
- Fit a Langaus to each slice and determine the MPV in each slice from the MP value given by the Landau portion of the fit and the Gaussian sigma (GSigma)

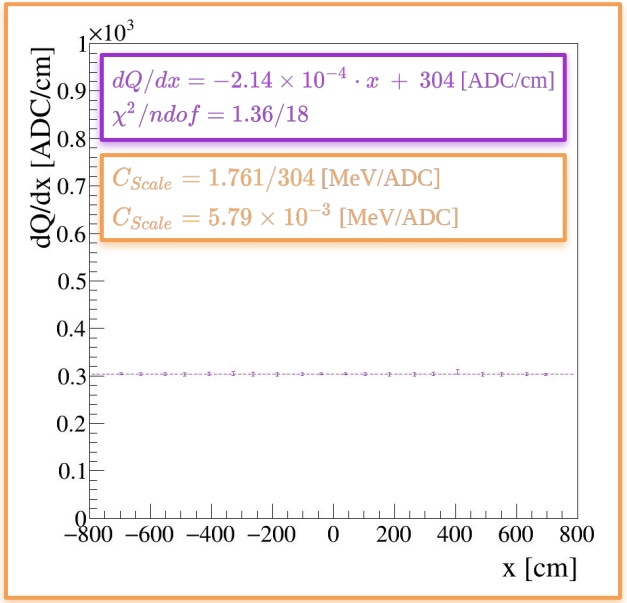




Scaling between dQ/dx and dE/dx, x-dependence

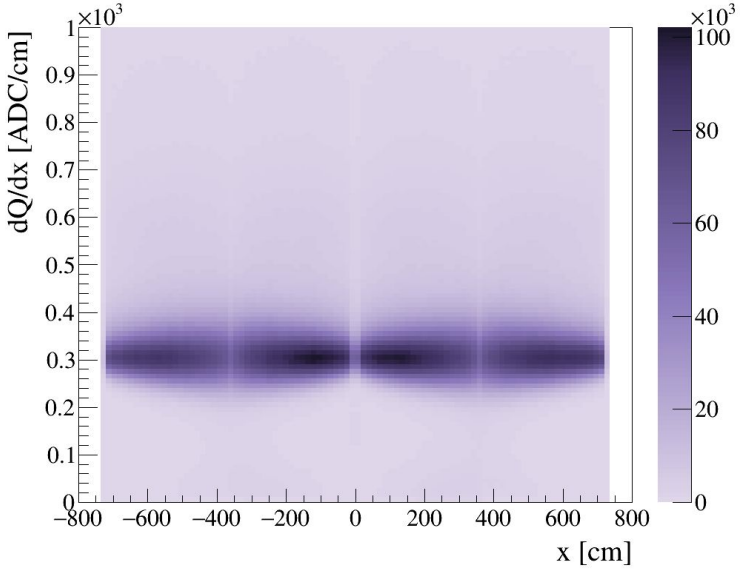
- Plot the MPV's against the x-position and fit a function (in this case just an order-1 polynomial)
- Since the gradient is so small, use the constant of this function to extract the overall MP dQ/dx
- If desired, a scale factor (C_{Scale} [MeV/ADC]) can then be calculated, which would allow for a conversion to dE/dx using the defined true/theoretical MPV for muons (1.761 MeV/cm)

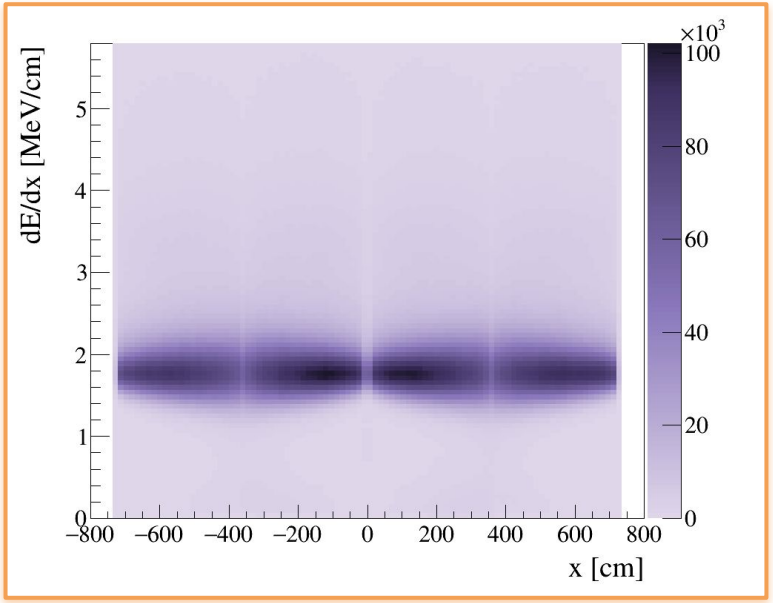




Scaling between dQ/dx and dE/dx, x-dependence

- Convert the dQ/dx distribution to a dE/dx distribution assuming the constant scale factor, $C_{Scale} = 5.79 \times 10^{-3}$ [MeV/ADC]





Looking at additional parameters

True energy

The procedure was repeated for the true energy-dependence of the reco dQ/dx.

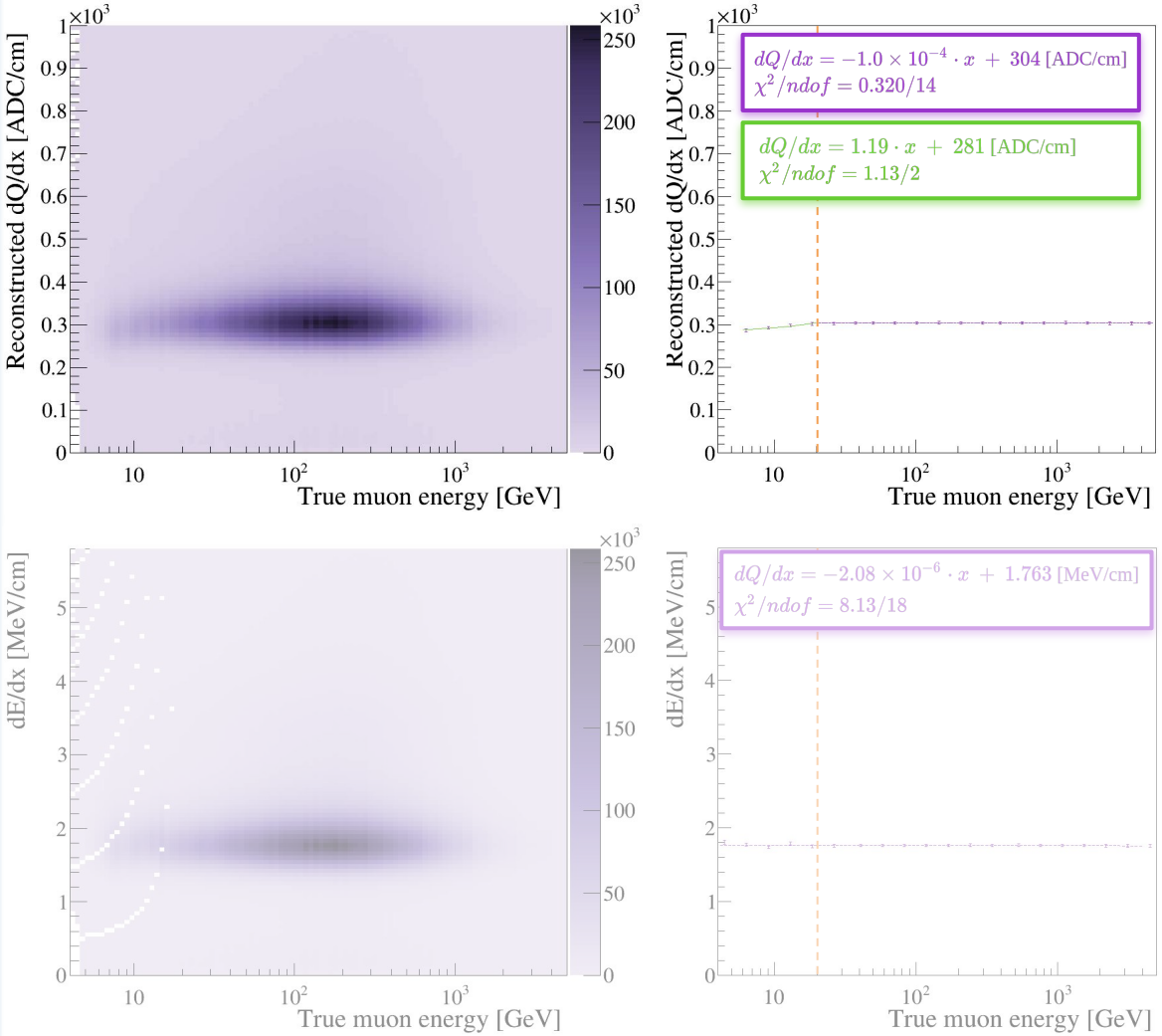
To account for the non-linearity, the parameter space is split at **E = 20 GeV** and two fits are performed.

It is possible to match the constant MPV dQ/dx acquired in x-position space when E > 20 GeV.

Converting to dE/dx with the separate fits returns the input MPV dE/dx (1.76 MeV/cm).

Since this dependence is not flat, and this will not be easily quantifiable in data we need to know if it is possible to,

a) Approximate the dependence as flat and still be comparable to the case in position space
b) Or determine a smaller subset of the parameter space which does have a flat energy dependence



True energy

The procedure was repeated for the true energy-dependence of the reco dQ/dx.

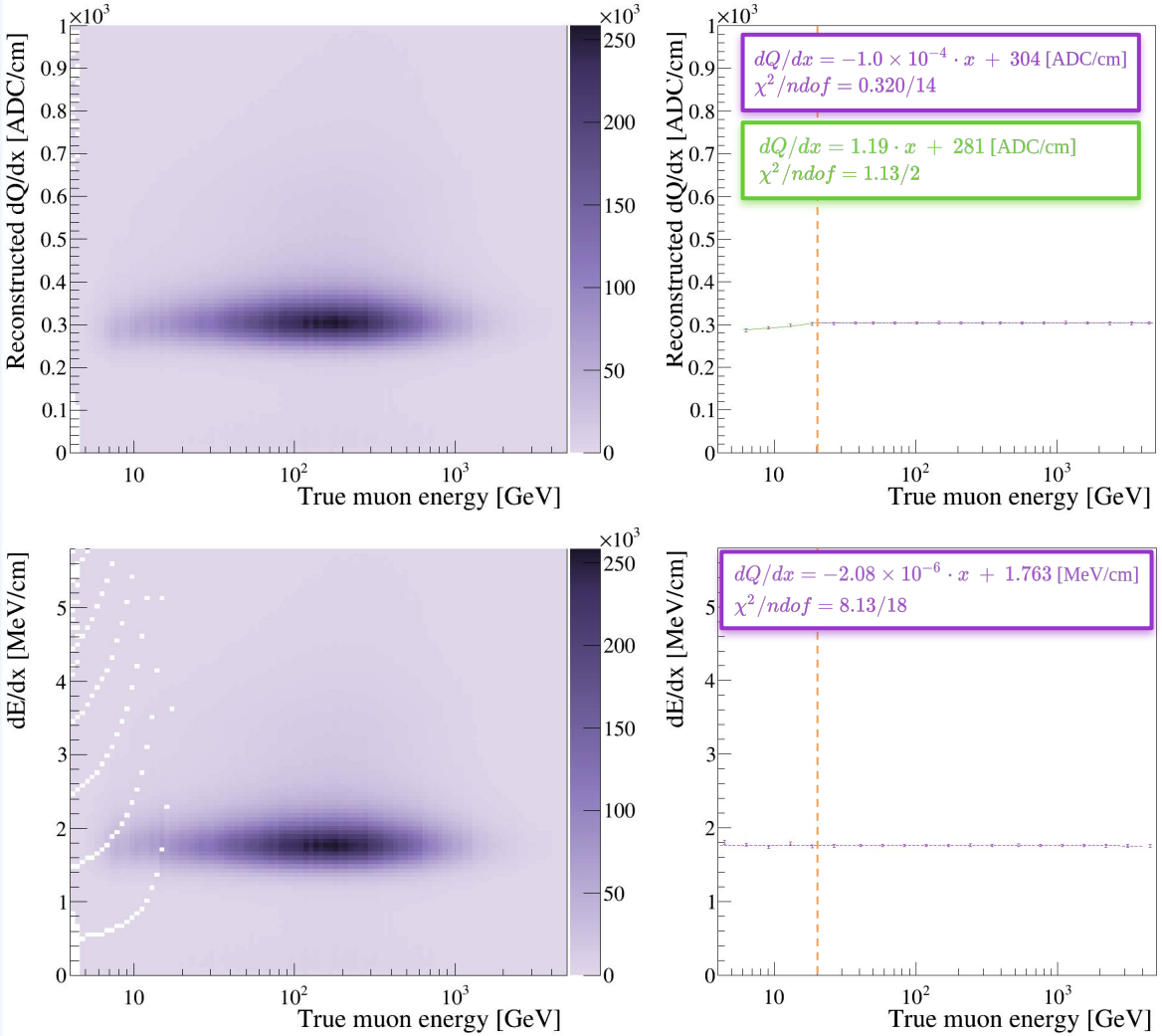
To account for the non-linearity, the parameter space is split at **E = 20 GeV** and two fits are performed.

It is possible to match the constant MPV dQ/dx acquired in x-position space when E > 20 GeV.

Converting to dE/dx with the separate fits returns the input MPV dE/dx (1.76 MeV/cm).

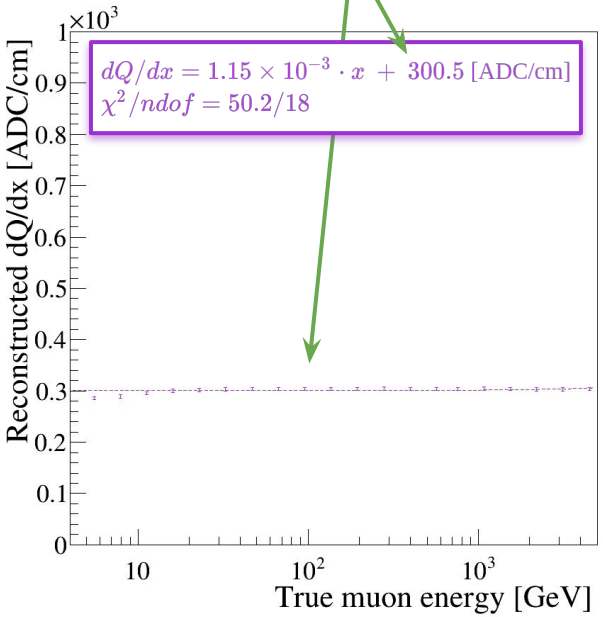
Since this dependence is not flat, and this will not be easily quantifiable in data we need to know if it is possible to,

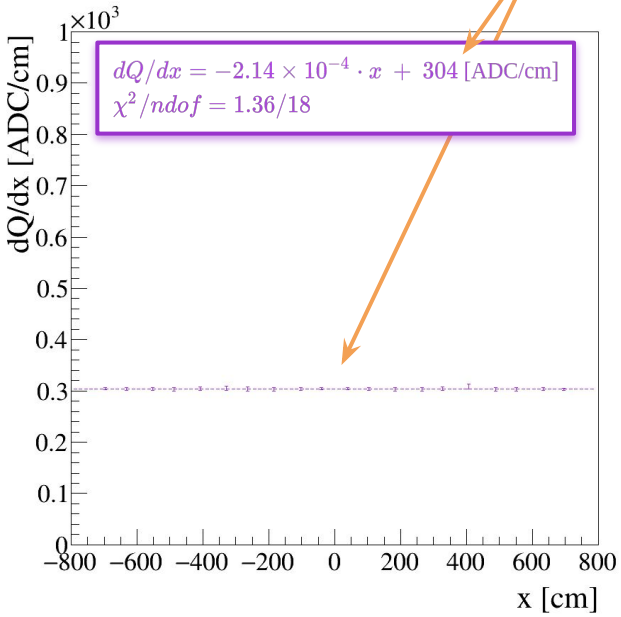
a) Approximate the dependence as flat and still be comparable to the case in position space
b) Or determine a smaller subset of the parameter space which does have a flat energy dependence

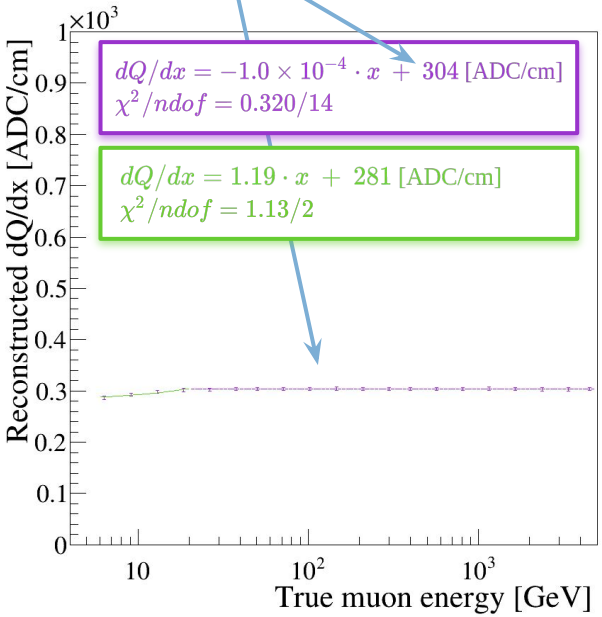


Fitting a single line

Fitting a straight line to the entire E parameter space gives a constant value which is 1.15% lower than the ones obtained in the x-dependent and E > 20 GeV fit







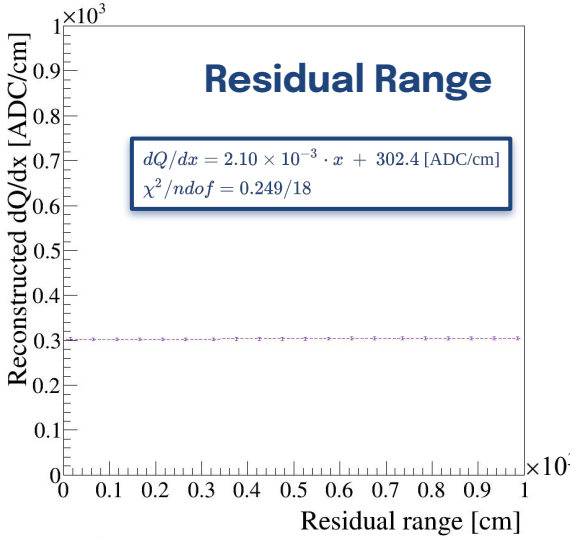
Additional parameter dependencies

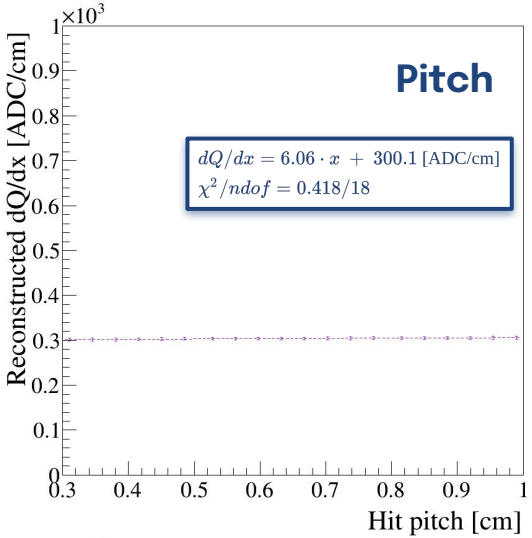
I fit a straight line to the pitch, residual range and hit width distributions, all of which have <5% variation between the maximum and minimum MPVs.

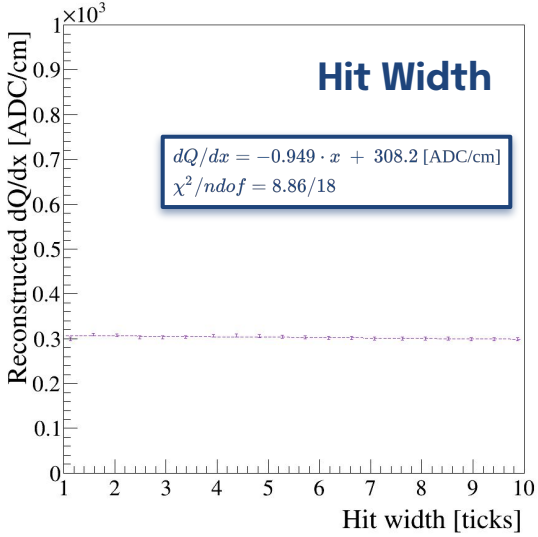
The $\cos\theta_{Drift}$ distribution has a second-order polynomial fit to it, and peaks where the angle is parallel to the APA planes.

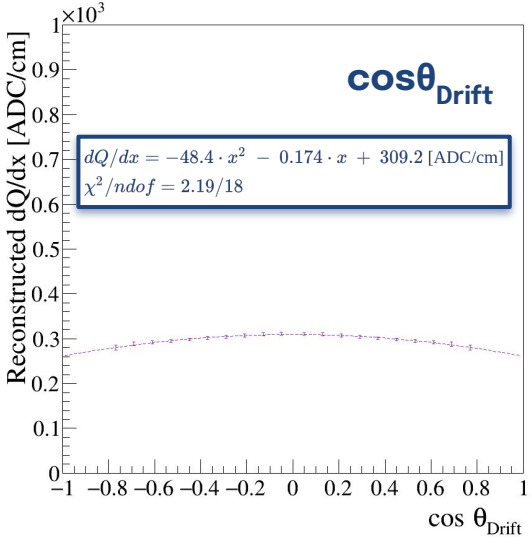
Calculated the region in which the dQ/dx vary by less than 5% from dQ/dx(0) and defined this as a subset of the sample.

$-0.567 < \cos\theta_{Drift} < 0.564$
E > 6.504 GeV









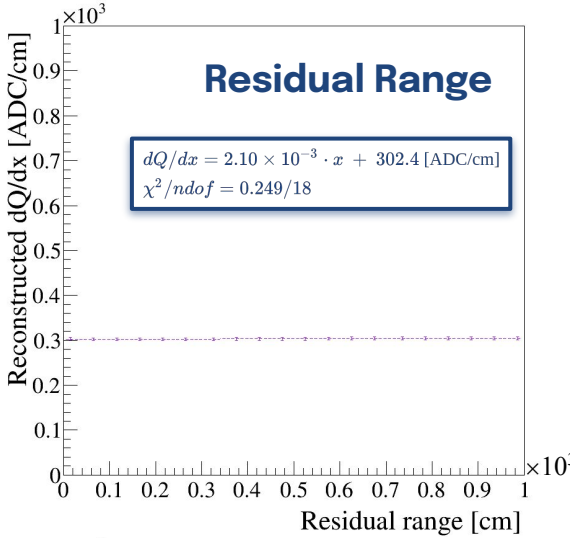
Additional parameter dependencies

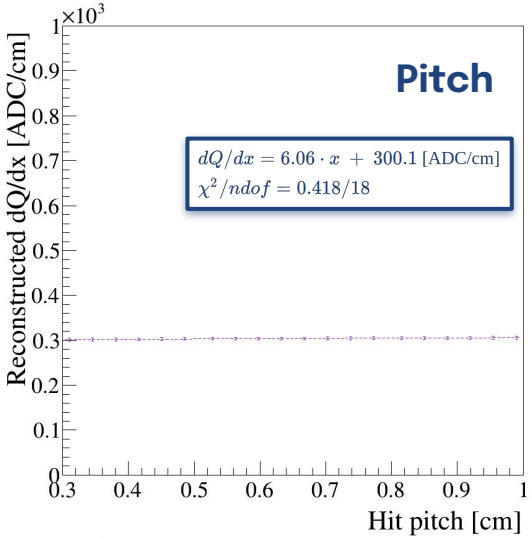
I fit a straight line to the pitch, residual range and hit width distributions, all of which have <5% variation between the maximum and minimum MPVs.

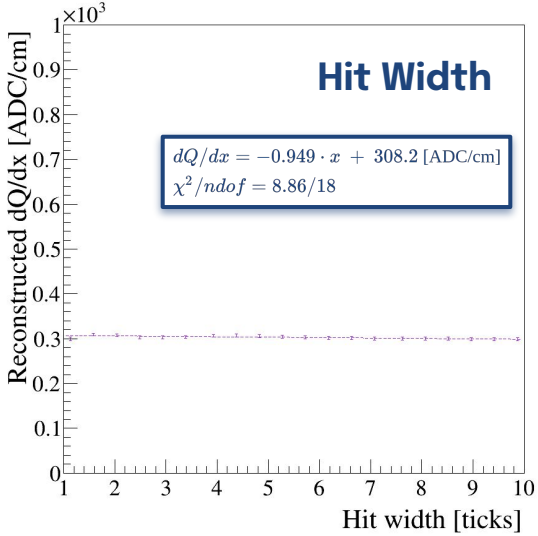
The $\cos\theta_{Drift}$ distribution has a second-order polynomial fit to it, and peaks where the angle is parallel to the APA planes.

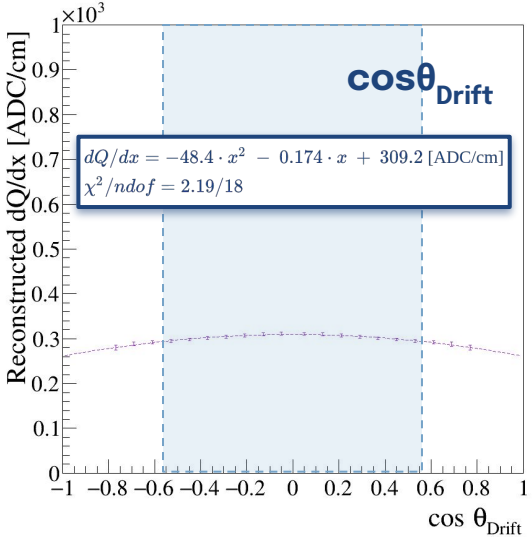
Calculated the region in which the dQ/dx vary by less than 5% from dQ/dx(0) and defined this as a subset of the sample.

$-0.567 < \cos\theta_{Drift} < 0.564$
E > 6.504 GeV



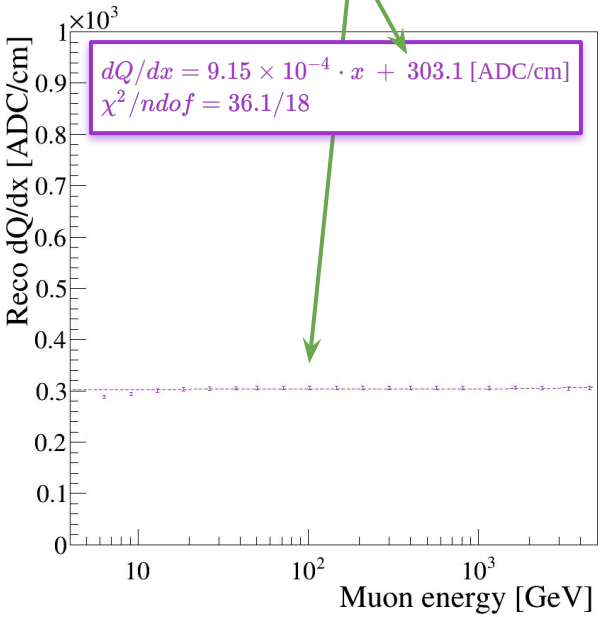


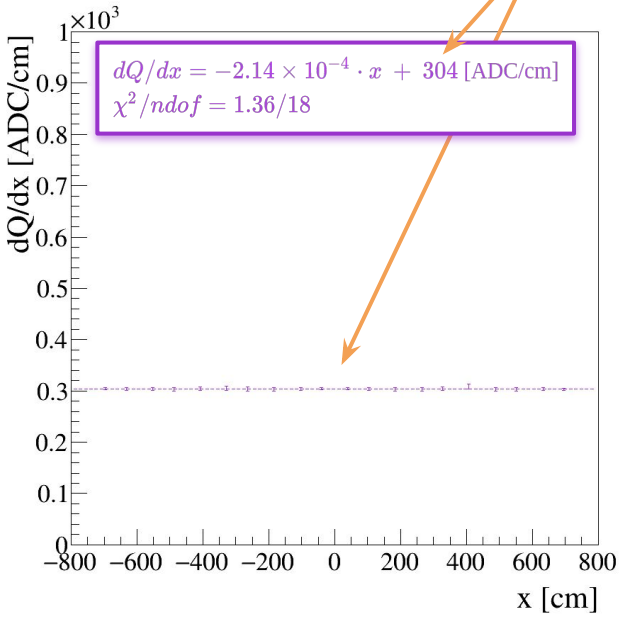


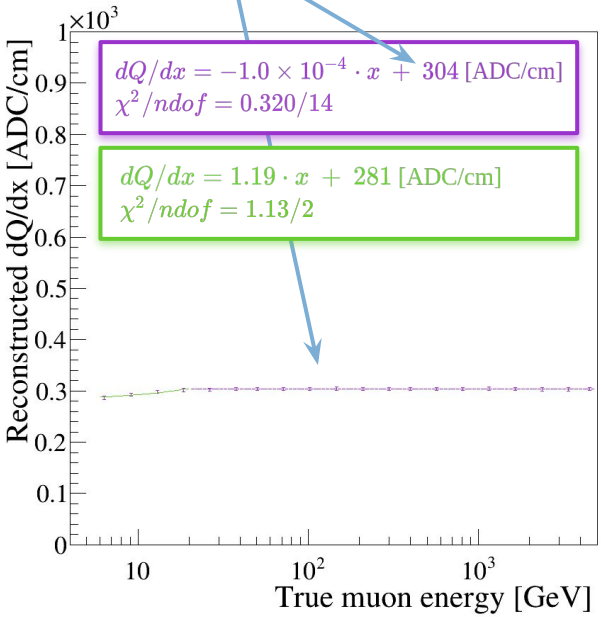


Fitting a subset of the sample

Fitting the sample subset to the entire E parameter space gives a constant value which is 0.296% lower than the ones obtained in the x-dependent and E > 20 GeV fit







Summary and outlook

- Of course the energy cut on the sample is more tricky to implement in data, but I did so to determine whether to move forwards with this approach
- The other parameter dependencies are less worrying since they should be quantifiable in data
- I will assess whether implementing only the angular cut would still result in improvements to the single-line energy fit
  - Or whether the angular-dependence could be used as a handle on the energy dependence in another way
- I am also looking into using truth information instead of theoretical information to determine the value of $(dE/dx)_{MPV}$
- Larger samples and systematics will eventually be incorporated into this method to evaluate its performance

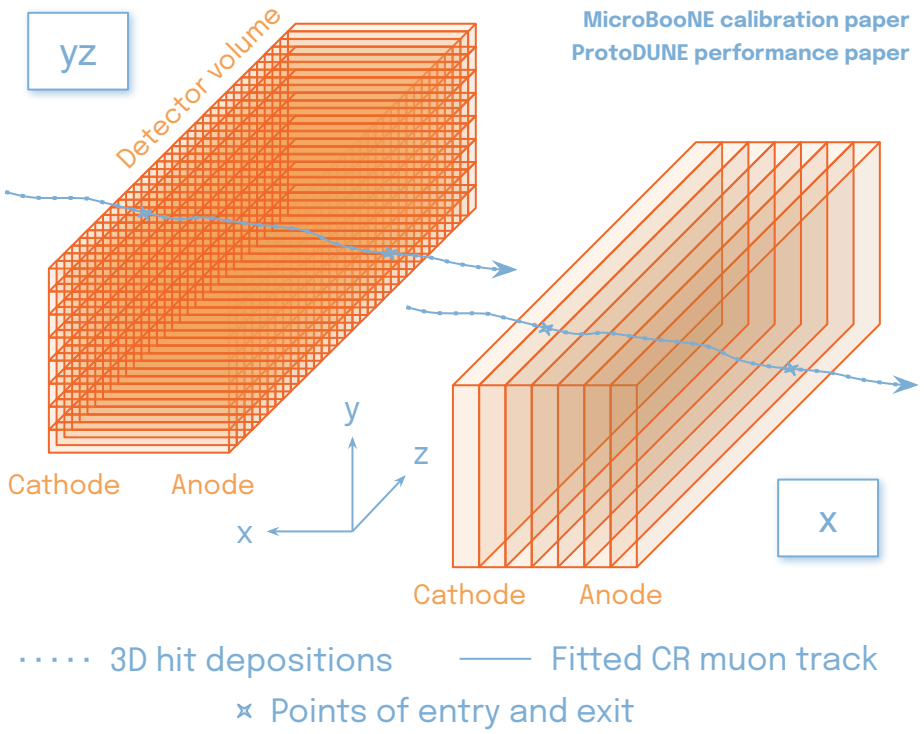
Backup

Discussion of previous approaches in the context of DUNE

- MicroBooNE and ProtoDUNE both used similar methods, but both were able to exploit the huge rate of stopping muons available to a surface detector
- Use Anode-Cathode (AC)-crossing muons to calibrate dQ/dx through normalisations across the entire detector and stopping muons to determine the absolute energy scale
  - AC-crossing muons can aid in the mitigation of drift-dependent effects using the known t_0
- The location of the DUNE FD detectors ~1.5 km underground means the CR muon rate is substantially reduced and the energy profile is altered significantly w.r.t that of MicroBooNE and ProtoDUNE
  - MicroBooNE/ProtoDUNE: A huge flux (5 kHz) of cosmic rays which have a peak energy of around 7 GeV and enter the TPC from a range of angles
  - DUNE: Much lower CR flux (0.05 Hz / module) where the muon energy profile has changed after travelling so far, such the peak is much higher, ≈280 GeV
    - Very few AC-crossing and 'stopping' muons in the TPC volume, muons predominantly enter from the top and leave towards the bottom
    - Muon energy spectrum results in higher muon 'activity'. Many more secondary particles are produced, such as δ-rays, making it potentially harder to reconstruct only the muon hits and calculate the dE/dx

Using AC-crossing and stopping muons for calorimetry calibration

- The detector is first split into 5x5 cm cells in order to calibrate dQ/dx in the yz plane and is then split into 5 cm bins to calibrate dQ/dx in the x dimension
- In these cells & bins, charge depositions are equalised using the ratio of median depositions in the entire plane/dimension (global) to the median in the cell (local)
- Following the calibration of dQ/dx, the absolute energy scale is determined using stopping CR muons
  - 'Modified box model' used
  - Cathode-crossing muons are also required here in order to reconstruct t_0



$$\left(\frac{dE}{dx}\right)_{\text{Calibrated}} = \frac{\exp\left(\frac{\left(\frac{dQ}{dx}\right)''}{C_{\text{Cal}}} \cdot \frac{\beta' W_{\text{ion}}}{\rho\epsilon}\right) - \alpha}{\frac{\beta'}{\rho\epsilon}}$$

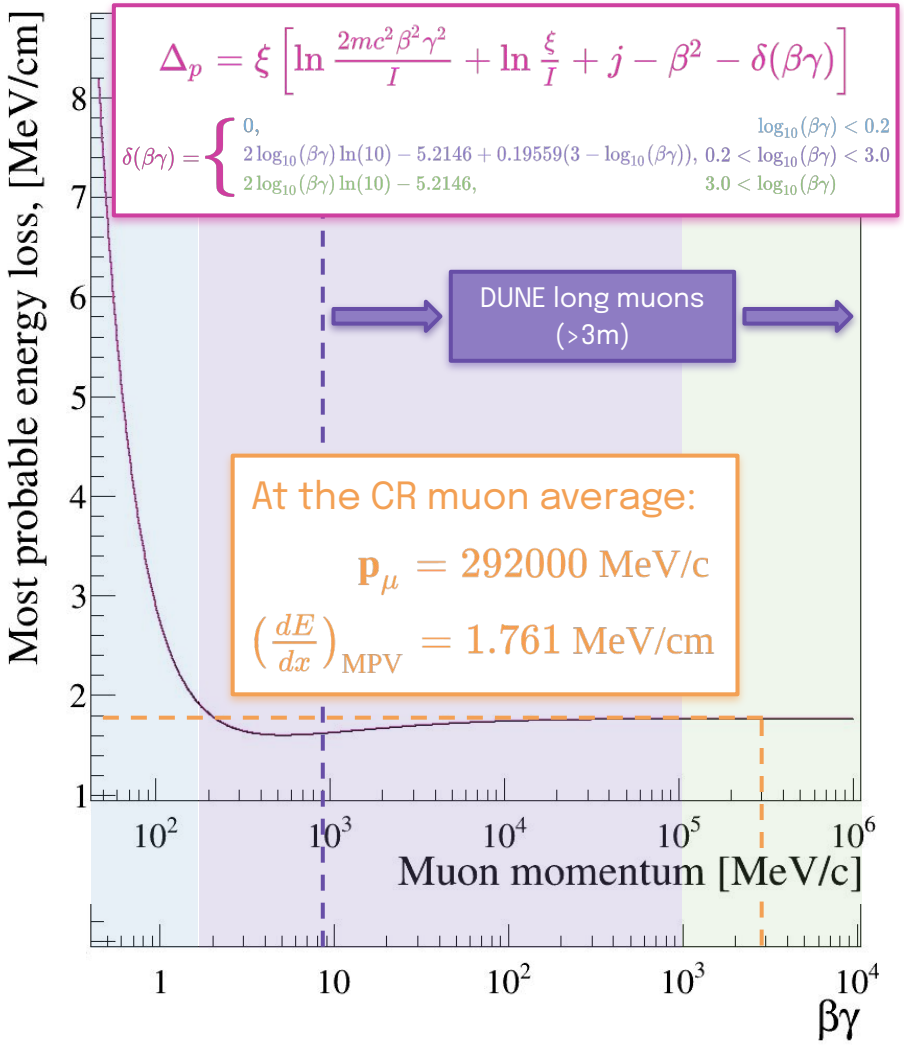
Modified Box Model

Discussion in the context of DUNE

- Unfortunately, it is not possible to simply apply these methods to DUNE, since splitting the detector into yz & x cells will not provide enough hits-per-cell
  - Only around 4000 total events per day in DUNE across the entire module (200 TPCs)
  - The MicroBooNE/ProtoDUNE studies use AC-crossing and stopping muons, which make up 12% and <2% of our primary, long muons respectively
  - Ideally want to be able to calibrate the detector in around 1 day, not 1 month...
- It might instead be possible to use similar techniques across an entire DUNE module (200 TPCs), rather than splitting the detector up into cells
  - I have not looked into this yet
- For now, I am using through-going top-bottom muons since they make up well over half of the sample

Parameterising the energy-dependence of the MP muon energy loss

Param.	Value	Comment
m	0.511 MeV/c^2	Electron mass
I	188x10^{-6} MeV	Mean excitation energy
j	0.2	Value from **here**
δ(βγ)	See RHS of slide, based on **Sternheimer**	Density effect coefficient
β	v/c	Relativistic beta
γ	p/m_μ	Relativistic gamma
ξ	(k/2)*(Z/A)*(x/β^2) MeV	-
k	0.307075	4π N_A r_e m MeV cm / mol
x	(ρ*dp) kg/cm^2	Thickness (density * pitch)
Peak values used to construct plot on RHS		
dp	0.353 cm	Peak pitch from my studies



Samples involved in the calibration studies

Looking at around 450,000 total (true) events which corresponds to around 1 month of data (at 14118/day total from Viktor's studies) in one 17 kt FD module.

CR events were simulated using the Sheffield muon generator, **MUSUN**, and propagated through GEANT4, Detsim and Reconstruction.

Throughout the process of determining a dE/dx calibration procedure, I have compared two methods of reconstructing these events

1. The Pandora v-chain
2. The Pandora CR-chain

For the sake of time, I will only be showing results using the **CR-chain**.

The Pandora v-chain

- Dunetpc version 08_50_00
  - Not the latest sample but is does include fixes to angular issues found in first sample
- MUSUN (larsim) version 08_22_03
- GEANT4 version 04_10_3_p03e
- Detsim (larsim) version 08_22_03
- Pandora version 03_11_01g, v-chain

The Pandora CR-chain

- Dunetpc version 09_10_02
- MUSUN (larsim) version 09_07_02
- GEANT4 version 04_10_6_p01
- Detsim (larsim) version 09_07_02
- Pandora version v03_16_00, CR-chain
  - Includes fixes to track stitching implemented by Isobel Mawby

Muon ‘selection’ criteria

- I currently use truth information to find the muon PDG code, but then use the following method to extract muons which pass as candidates for the calibration studies
- These cuts are defined for all tracks, however sometimes I request only through-going muons such that these cuts have a smaller impact

Truth-level selection requirements

From the GEANT4 tracks, I select:

- Those which enter the TPC geometry
- Primary muons
- Longer than 3m

See backup for cut motivation
(defined in reco, same in truth for consistency)

Reconstructed ‘selection’ requirements

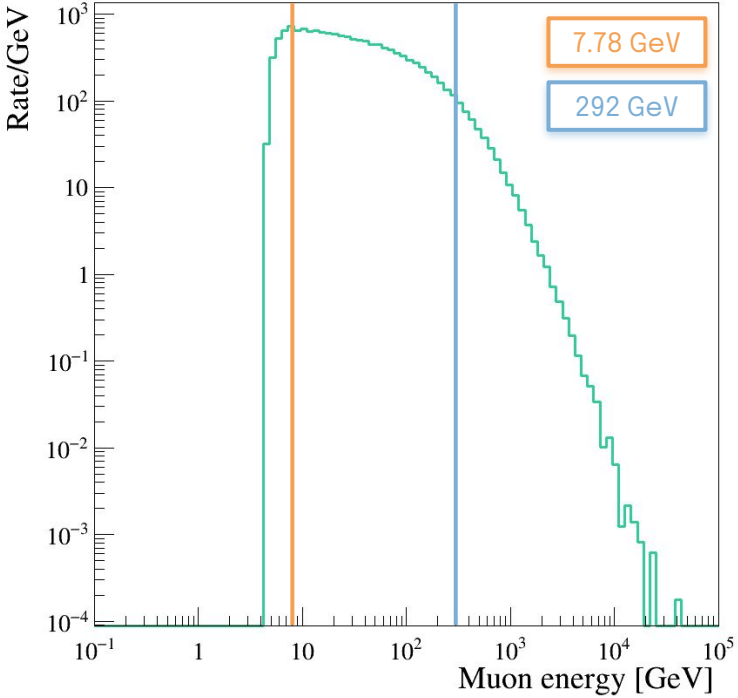
From the reconstructed tracks, I select:

- Those with a truth-level muon pdg code
  - Can’t check whether it is primary yet
- Longer than 3m
- Start Y-position is > 599.5 cm

See backup for cut motivation

Energy distributions passing all muon criteria, CR-chain

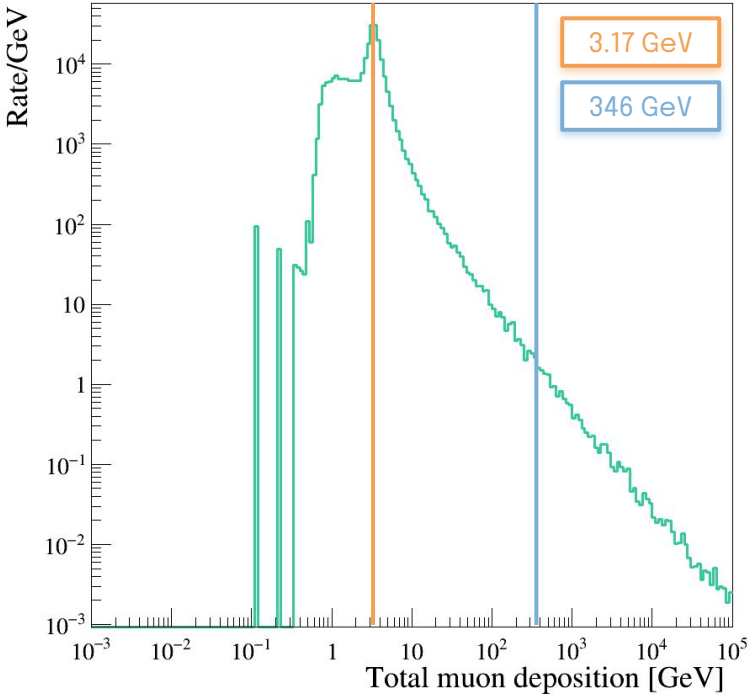
True (T)



Peak

Average
(R-T)/T = 18.5%

Reco (R)



True and reconstructed sample breakdown, CR-chain

Truth-level sample contents

Statistic	Rate / 26.28 Days	% All Events
Events	371,000	-
TPC μ	128,307	34.584 %
Primary TPC μ	125,578	33.849 %
Long, primary TPC μ	112,123	30.222 %
Crosses top or bottom	109,738	29.579 %
Crosses top and bottom	64,988	17.517 %
Crosses $\geq$ 1 APA/CPA	57,548	15.512 %
Crosses $\geq$ 2 APA/CPA	14,066	3.7914 %
Stopping	2253	0.60728 %
Exiting	76,259	20.555 %

I first break down the sample into the various muon contributions.

I then check which planes the muons have crossed,

- 68.0% of all long, primary muons exit the detector
- 85.2% of the exiting muons enter through the top and leave through the bottom of the detector
- Only 11.2% cross both an APA and a CPA

Reconstructed sample contents

Statistic	Rate / 26.28 Days	% All Events
Events	371,000	-
Tracks	1,892,162	510.02 %
Muons	1,793,984	483.55 %
$\mu > 3$ m	112,339	30.28 %
$\mu > 3$ m & $y_i > 599.5$ cm	70,124	18.901 %
Crosses top or bottom	70,016	18.872 %
Crosses top and bottom	49,997	13.476 %
Crosses $\geq$ 1 APA/CPA	47,298	12.749 %
Crosses $\geq$ 2 APA/CPA	12,132	3.2701 %
Stopping	1759	0.47412 %
Exiting	68,297	18.409 %

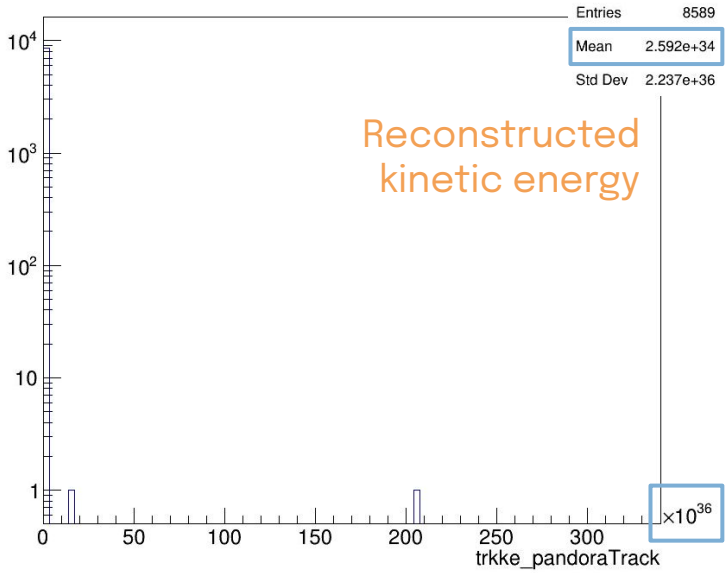
There is a substantial increase in tracks/muons in the reconstructed sample.

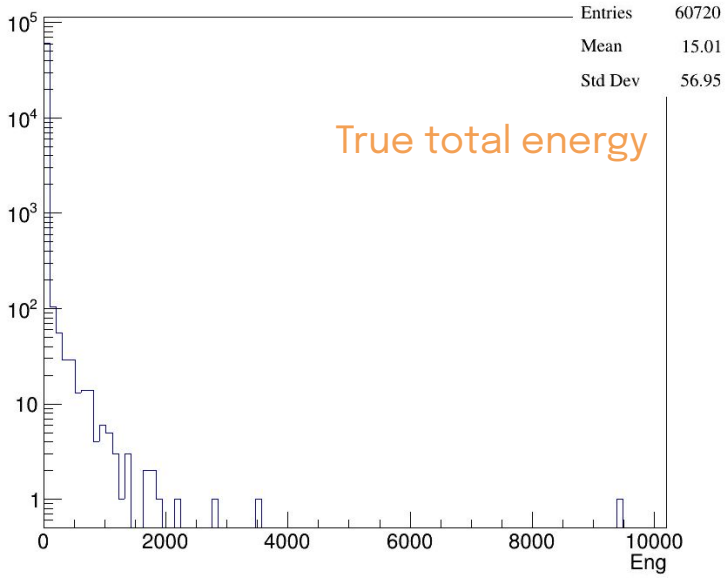
I believe this is caused by reconstructed delta rays which are attributed to the true muon in truth. The significant reduction in rate when applying the muon threshold cuts fits this theory.

60.8% of the long muons exit the detector
73.2% of which enter through the top and leave through the bottom

True and reconstructed energies

- I now request that the reconstructed muon kinetic energy does not exceed the maximum deposited true total energy in the sample (1e4 GeV)
  - There existed some tracks associated with true muons with extremely large energies
  - Plots taken directly from the root file





Reconstructing delta rays and activity dependence

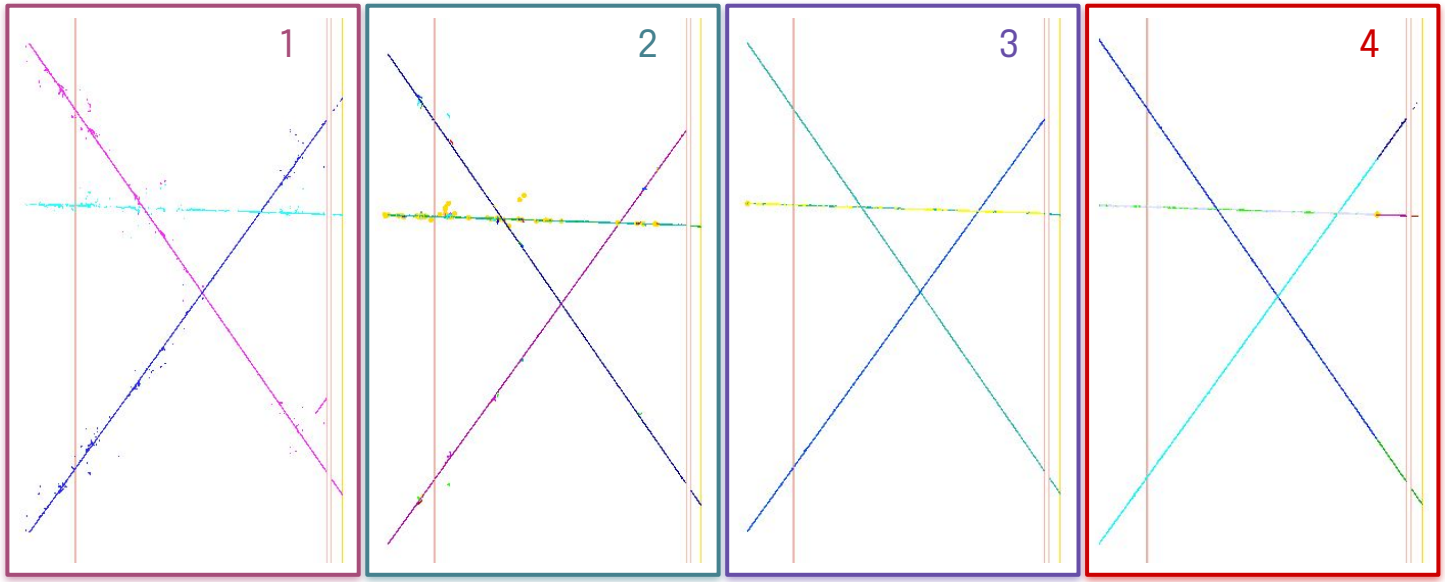
- I was a little confused as to why I am seeing what appears to be very little delta-ray activity w.r.t the muon dQ/dx → dE/dx calibrations studies I've been performing
- At first, I believed this might be because when using the cosmic reconstruction chain, the delta rays may be being correctly separated from the muon track using delta-oriented reconstruction algorithms
  - I then thought that if that was the case, then when I look instead at the reconstructed muons using the neutrino chain, the activity-dependence could begin to appear because the delta ray algorithms are not explicitly called in this chain
  - But that's not what happened
- It turns out there is a set of 'clear track' algorithms which run regardless of whether you are looking at the neutrino or cosmic chain, and the delta-ray algorithms aid in the reconstruction of the delta-rays, rather than contributing to their separation from the track

Looking into what the delta ray algorithms contribute

PDG	Energy	Length	Number of hits			
			Total	U view	V view	W view
-13	443.362 GeV	11.843 m	3,682	1,619	1,723	340

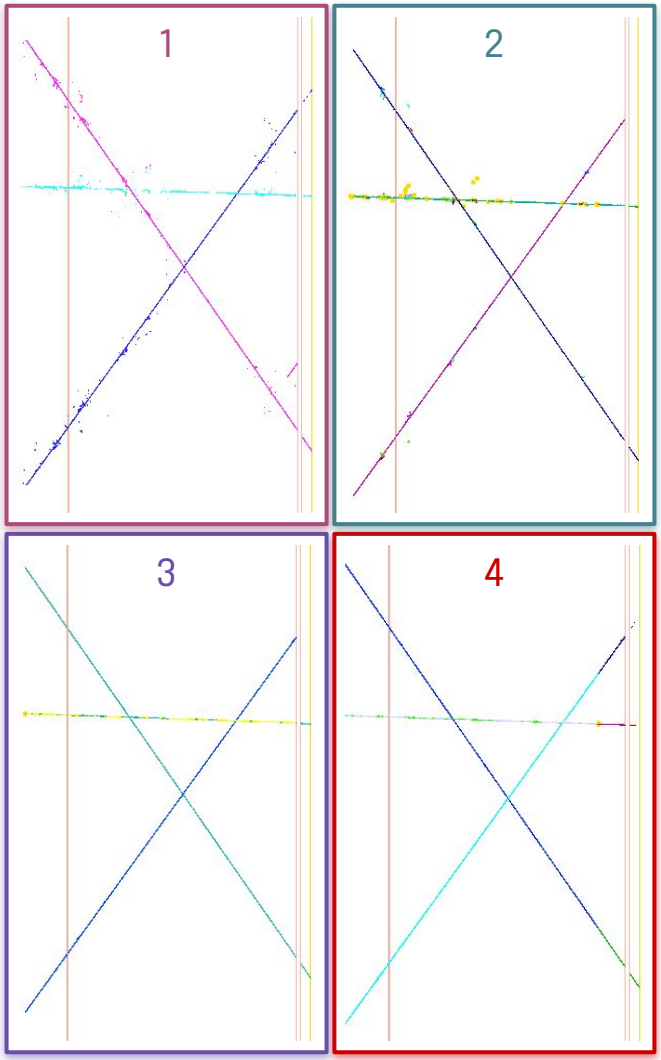
3 views shown as an overlay in all event displays

1. Draw the 'cheated' clustering
2. Run Pandora with the usual CR-chain settings
3. Run Pandora CR with delta-ray algorithms switched off
4. Run pandora with the usual neutrino-chain settings (no delta-ray algorithms by default)



Observations from the event displays

1. There are clearly many delta rays at the truth-level
2. Yellow blobs are candidate vertices, the delta-ray algorithms are finding many of them
3. When the delta-ray algorithms are removed, the 'clear-track' algorithms manage to correctly not reconstruct any of the delta rays
   a. I'm not yet sure what the speckling is all about in the view with the horizontal line
4. In the neutrino reconstruction, because the delta-ray algorithms are also not included by default, the clear-track algorithms appear to be doing a similar thing to the no-delta case
   a. The speckling is there, but the colours have changed which makes it less clear



Draw the 'cheated' clustering

Run Pandora with the usual CR-chain settings

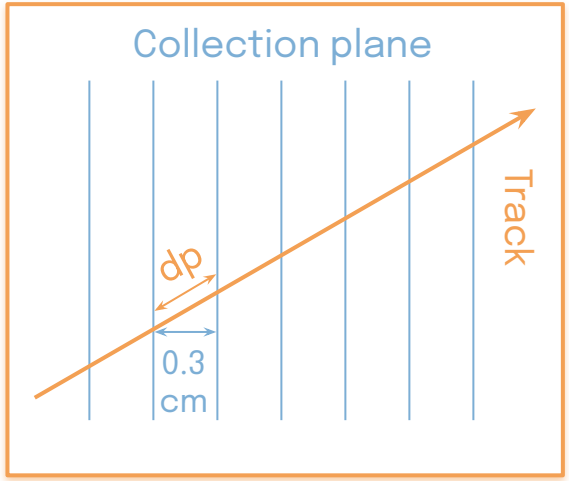
Run Pandora CR with delta-ray algorithms switched off

Run pandora with the usual neutrino-chain settings (no delta-ray algorithms, by default)

Parameter definitions

Reco: ‘trkke’, kinetic energy, $E_K = \sum_{Hits} \frac{dE}{dx} \cdot dp$

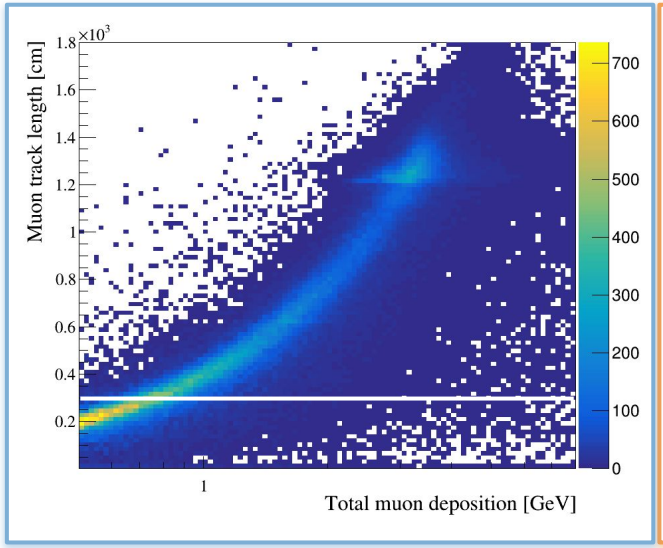
- Calculated from the energy depositions (dE/dx) and the track pitch (dp)
- Track pitch is the length of the track segment between two wires
- Probably better defined as ‘visible energy’ rather than kinetic energy

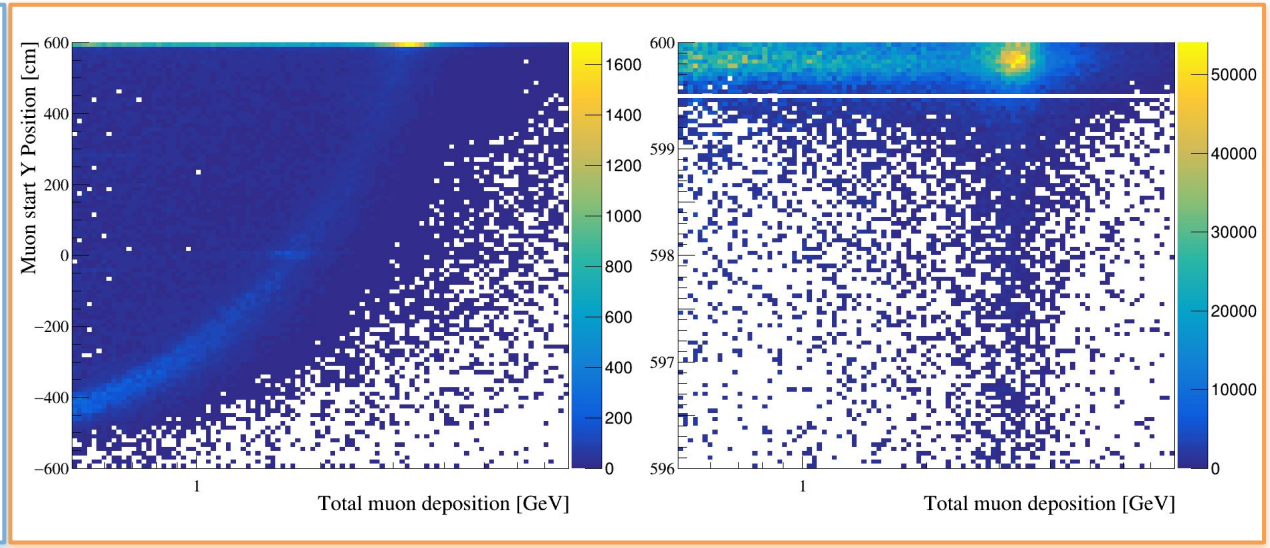


Truth: ‘Eng’

- 1st component of the Lorentz 4 vector from GEANT4 → Total true energy of the track

Length and y start position cut (v08, v-chain)

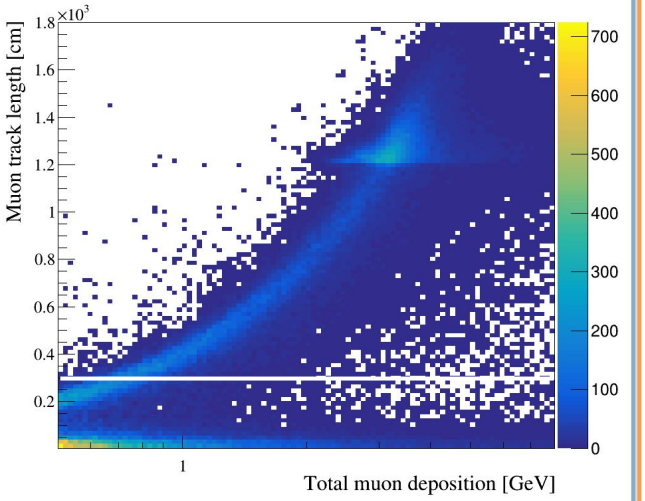


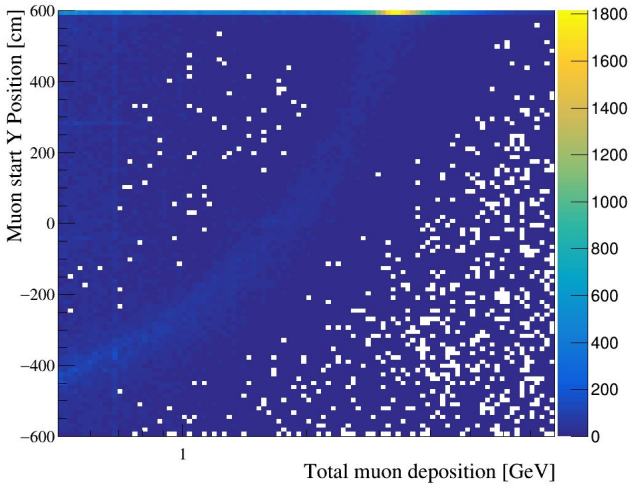


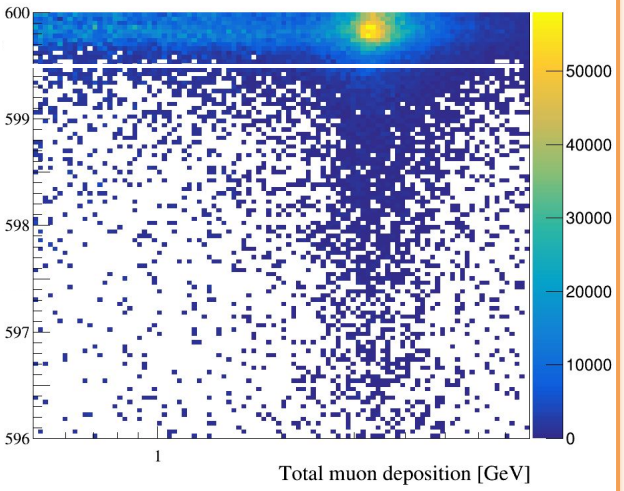
Placing a cut at a track length of 3m instead of 2m, in order to remove a larger portion of the low (reconstructed) total energy deposition distribution.

Also ensuring that the start Y position is right at the top of the detector (> 599.5 cm).

Length and y start position cut (v09, CR-chain)







Placing a cut at a track length of 3m instead of 2m, in order to remove a larger portion of the low (reconstructed) total energy deposition distribution.

Also ensuring that the start Y position is right at the top of the detector (> 599.5 cm).

Converting functional form of dQ/dx to dE/dx

With the aim of removing the parameter (X)-dependence from the energy deposition definition

$$dQ/dx = m \cdot X + c$$

$$C_{Scale} = \frac{1.761 \text{ MeV/cm]}}{dQ/dx} = \frac{1.761 \text{ MeV/cm]}}{m \cdot X + c}$$

$$dE/dx = C_{Scale} \cdot dQ/dx = \frac{1.761 \text{ MeV/cm]}}{m \cdot X + c} \cdot dQ/dx$$

Comparisons with reconstructed and true dE/dx

The substantial differences between the true and converted distributions indicate that the sample I am looking at does not represent the theoretical nominal for through-going muons.

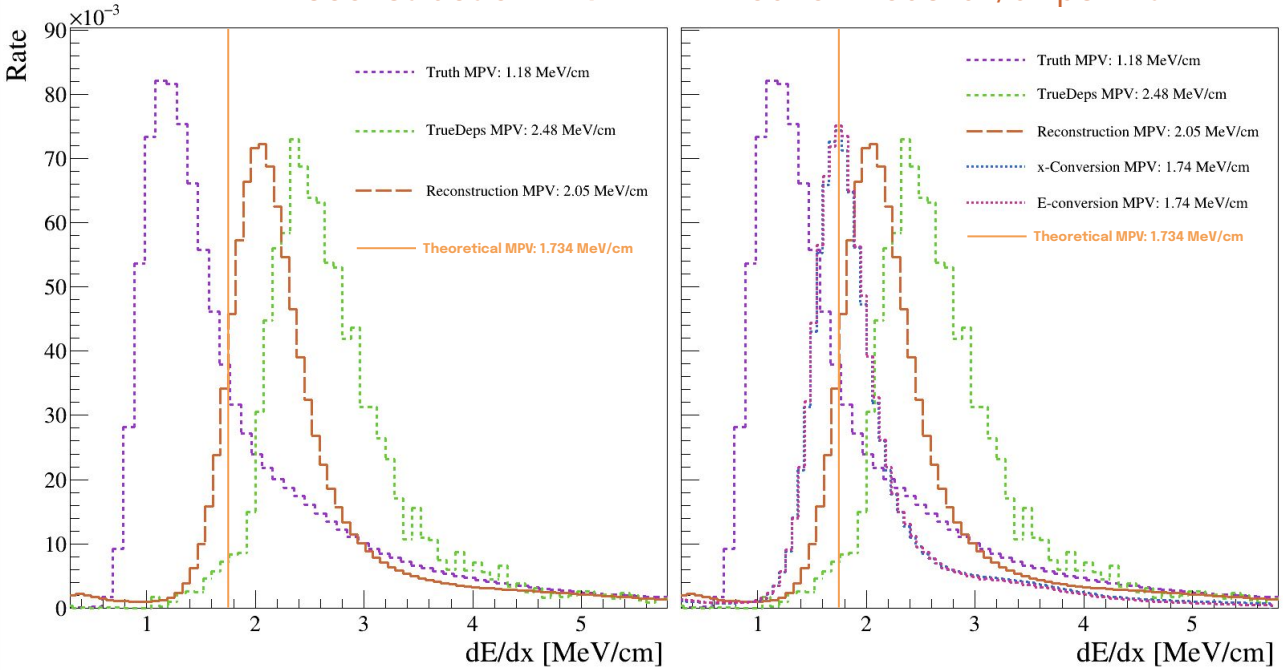
The reconstructed distributions are calculated using the Modified Box model with the configurable parameters given by calculations from ArgoNeuT data.

Once I can verify agreement with truth and theoretical truth, the comparison with reconstruction will be more informative.

Truth MPV = Hit Energy

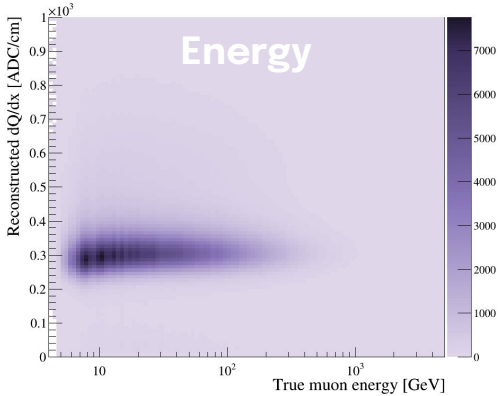
TrueDeps MPV = Total hit depositions / Track L

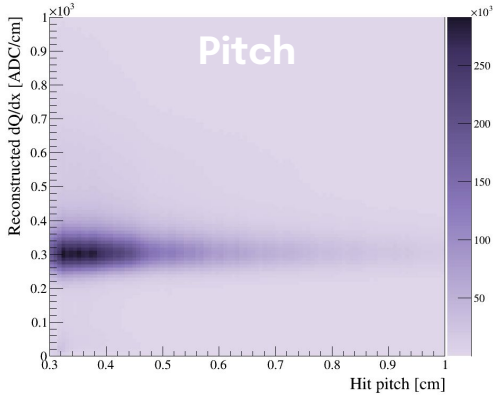
Reconstruction MPV = ModBox model dE/dx per hit

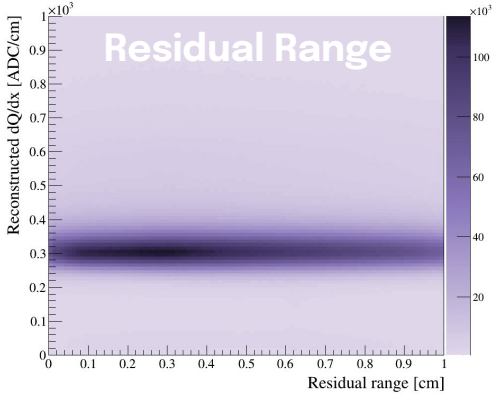


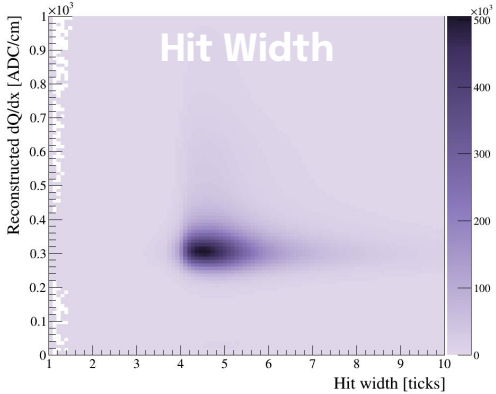
Reconstructed charge deposition parameter dependence

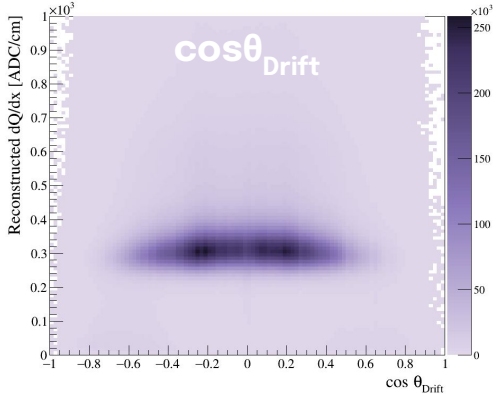
- The dQ/dx distribution is only vaguely dependent on two of these parameters
  - Energy
  - Angle to the APA's
- The next slide will show more clearly the behaviour of these dependencies





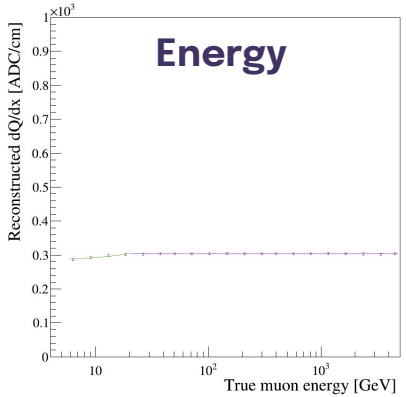


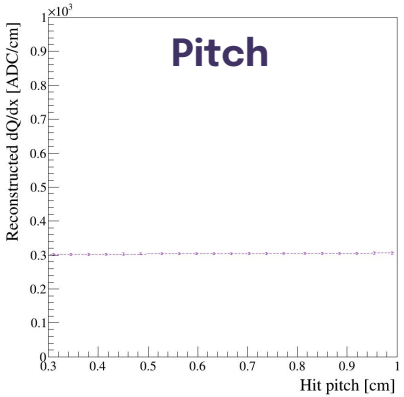


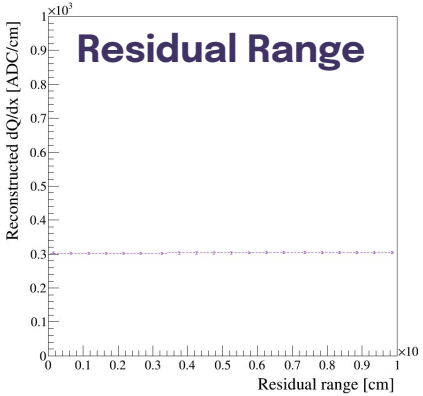


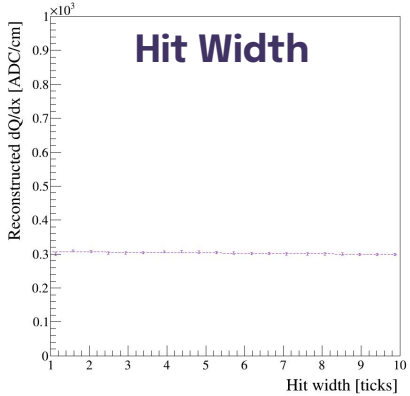
Fit to slices MPVs

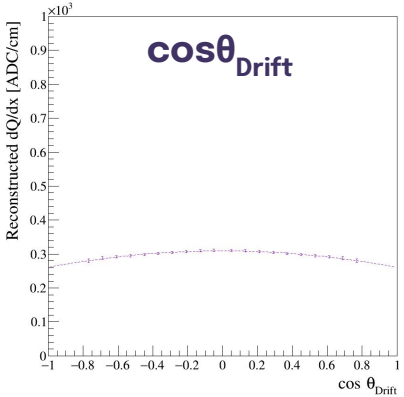
- I fit a straight line to the pitch, residual range and hit width distributions, all of which have negligible gradients
- The energy distribution has two straight lines fit to it, above and below 20 GeV
- The $\cos\theta_{Drift}$ distribution has a second-order polynomial fit to it, and peaks where the angle is perpendicular to the APA planes











Converted energy deposition parameter dependence

- The converted dE/dx distributions are considerably more flat than the input dQ/dx distributions
- It is therefore possible to account for any parameter dependence using this method
- Though I think a better approach will be to define a subset of the sample which removes such dependencies