Pion Cross Section Updates

Jake Calcutt
April 22, 2022

Introduction

Detailing some updates to my pion cross section analysis since my last update

Reconsidering SCE systematic (realized the way I was doing it wasn't very well-motivated)

New approach to beam smearing systematic

Finalizing list of systematics (in my view)

Normalizing MC to data beam momentum

Signal Definition

Note: Considering a threshold of 150 MeV/c on the charged pions due to our inefficiency in identifying these ➔ Signal events can contain charged pions < 150 MeV/c

Absorption:

$$\pi^+ + \mathrm{Ar} \rightarrow \mathrm{N}' + \mathrm{nucleons}$$

Charge Exchange:

$$\pi^+ + \mathrm{Ar} \rightarrow \mathrm{N}' + \mathrm{nucleons} + n\pi^0$$

Other:

$$\pi^+ + \mathrm{Ar} \rightarrow \mathrm{N}' + \mathrm{nucleons} + \mathrm{charged\ pions}$$

Measure exclusive and total (not independent)

Analysis Strategy

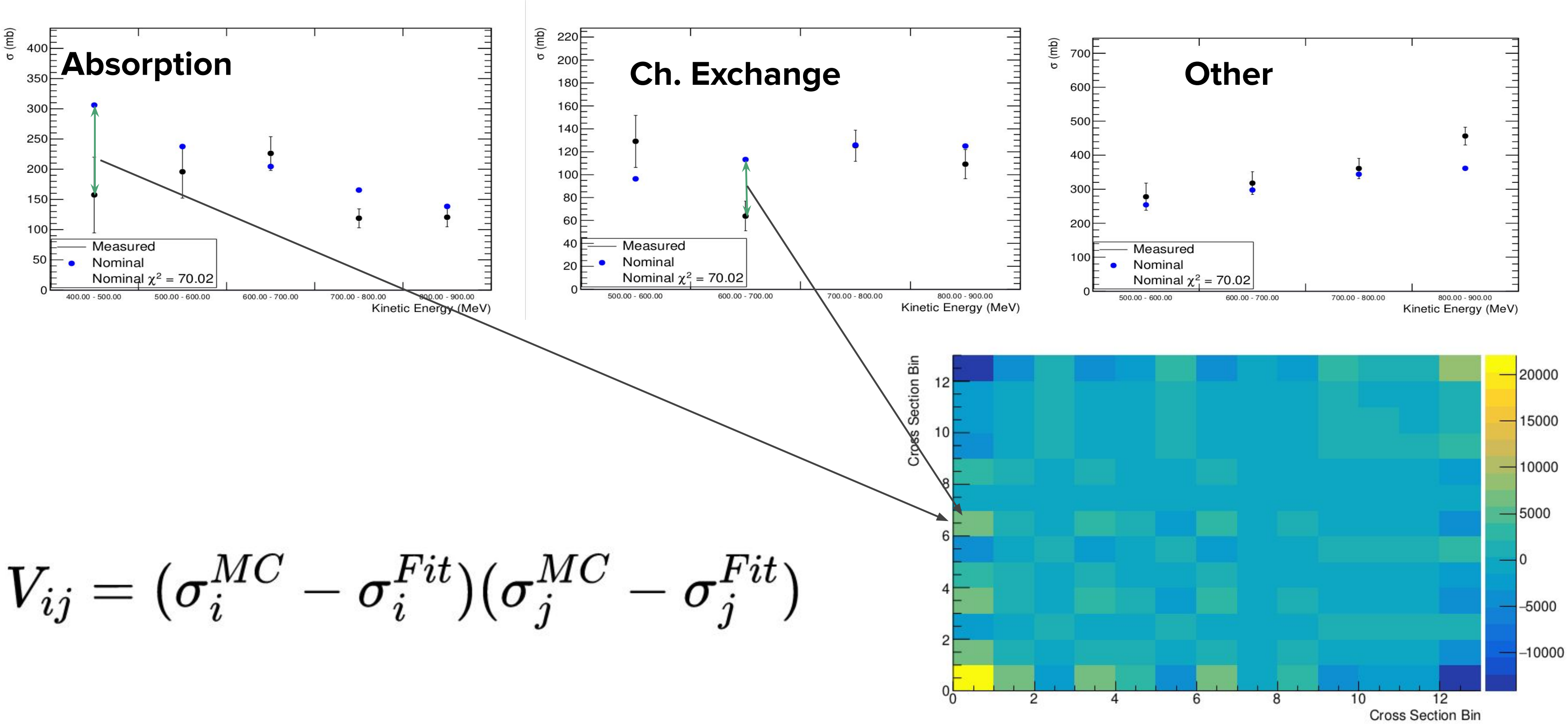
Start with samples of pions/muons from beam and bin events according to event selection (see backup)

- Categories: No beam track, Failed Beam cuts, Past fiducial vol, Interactions, Michels
- Binned in various observables: ending Z position, reconstructed energy at interaction

Parameterize MC according to set of signal, flux, and systematic parameters

- Signal: interactions at specific energy
- Flux: relative number of muons/pions
- Systematics: will discuss later

Previous SCE Syst: Using Diffs. as Uncertainty



$$V_{ij} = (\sigma_i^{MC} - \sigma_i^{Fit})(\sigma_j^{MC} - \sigma_j^{Fit})$$

SCE Syst Discussion

I realized the previous implementation wasn’t very well thought out

- Low stats of the alt-SCE sample
- Is the alt-SCE map even well motivated

Went back to basics to figure out how could SCE affect this analysis

How Could SCE Affect This Analysis?

Directly (in binning)

1. Energy reconstruction
2. End Z position
3. Beam cuts

Indirectly (in reconstruction)

1. Differences in distortions could possibly affect Pandora’s tracking/shower reconstruction and beam-identifying BDT

How Could SCE Affect This Analysis?

Directly (in binning)

1. Energy reconstruction
2. End Z position
3. Beam cuts

Even if the SCE map is wrong, we make dQdX uniform across detector and essentially force dEdX to be right by calibrating charge scale

Indirectly (in reconstruction)

1. Differences in distortions could possibly affect Pandora's tracking/shower reconstruction and beam-identifying BDT

How Could SCE Affect This Analysis?

Directly (in binning)

1. Energy reconstruction
2. End Z position
3. Beam cuts

The overall z-position distortion estimated by the nominal SCE is ~2.5cm

Alt SCE estimates ~2.8cm

➡ Quick check: few (~50/102k) MC events would fall into/out of APA2 selection given a different SCE map

Indirectly (in reconstruction)

1. Differences in distortions could possibly affect Pandora's tracking/shower reconstruction and beam-identifying BDT

How Could SCE Affect This Analysis?

Directly (in binning)

1. Energy reconstruction
2. End Z position
3. Beam cuts

Cuts are done in-situ according to mean SCE corrected positions and direction

Indirectly (in reconstruction)

1. Differences in distortions could possibly affect Pandora's tracking/shower reconstruction and beam-identifying BDT

How Could SCE Affect This Analysis?

Directly (in binning)

1. Energy reconstruction
2. End Z position
3. Beam cuts

Indirectly (in reconstruction)

1. Differences in distortions could possibly affect Pandora's tracking/shower reconstruction and beam-identifying BDT

Tingjun and Yinrui have shown studies where truncated MC and data tracks have different reconstruction & beam ID efficiencies

Could be due to inherent beam differences (position/direction) or (possibly) different SCE distortions

SCE Systematic Effects

If SCE is partly responsible for reconstruction/ID efficiences, treated within systematic parameters that vary these efficiencies

- Two parameters – 1. Rate of matching true beam particle 2. Rate of events having no valid beam track
    - Each as a function of true end Z position (motivated by Tingjun and Yinrui's truncation studies)

I would consider the remaining effects irrelevant

Additional Systematic

Rate of interactions upstream of TPC

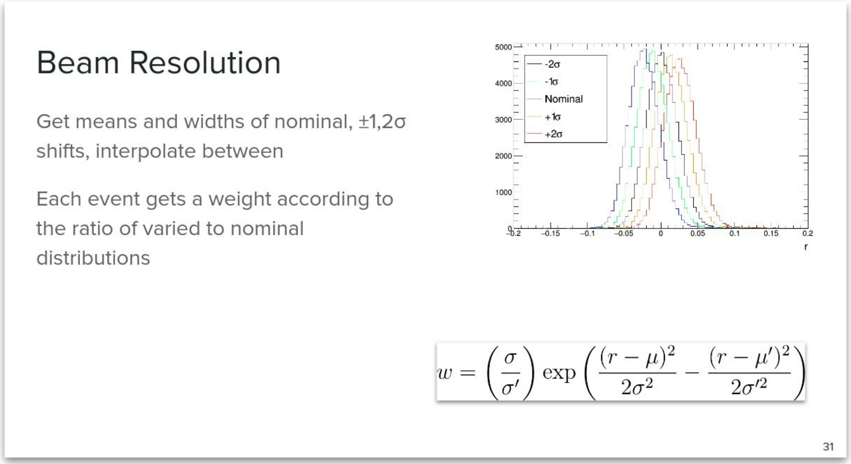
- Partially accounts for data/MC discrepancy in number of events without beam tracks

New Beam Smearing Approach

Previously was reweighting MC events according to shifting beam resolution

New approach: Fit to data with nominal beamline reconstruction as well as data with beam line reconstruction shifted by (~0.7% from effect of profile monitor fiber shift – and possibly 1% magnetic field uncertainty)

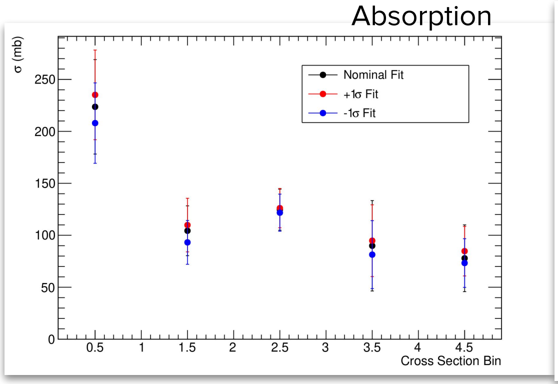
Combine results from three fits if needed (see next slide)

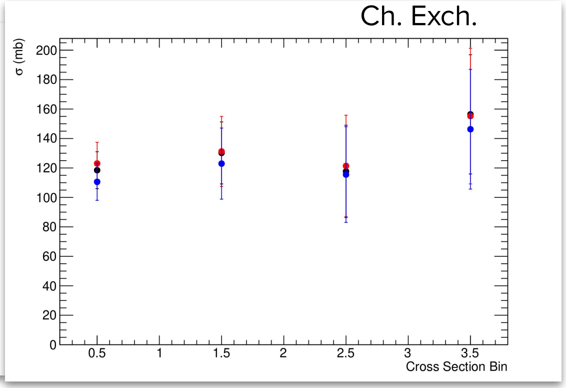


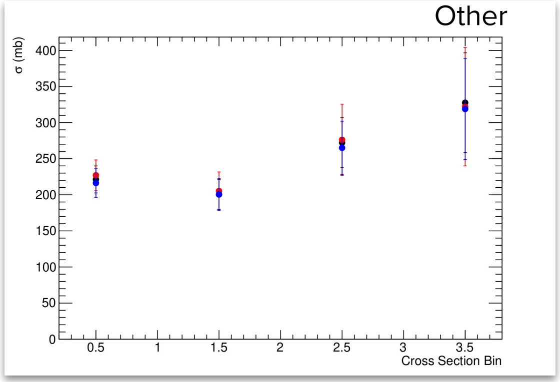
New Beam Smearing Approach

Below: results from fits with reconstructed beam momentum shifted by ±0.7%

Next slide: estimating uncertainty from results





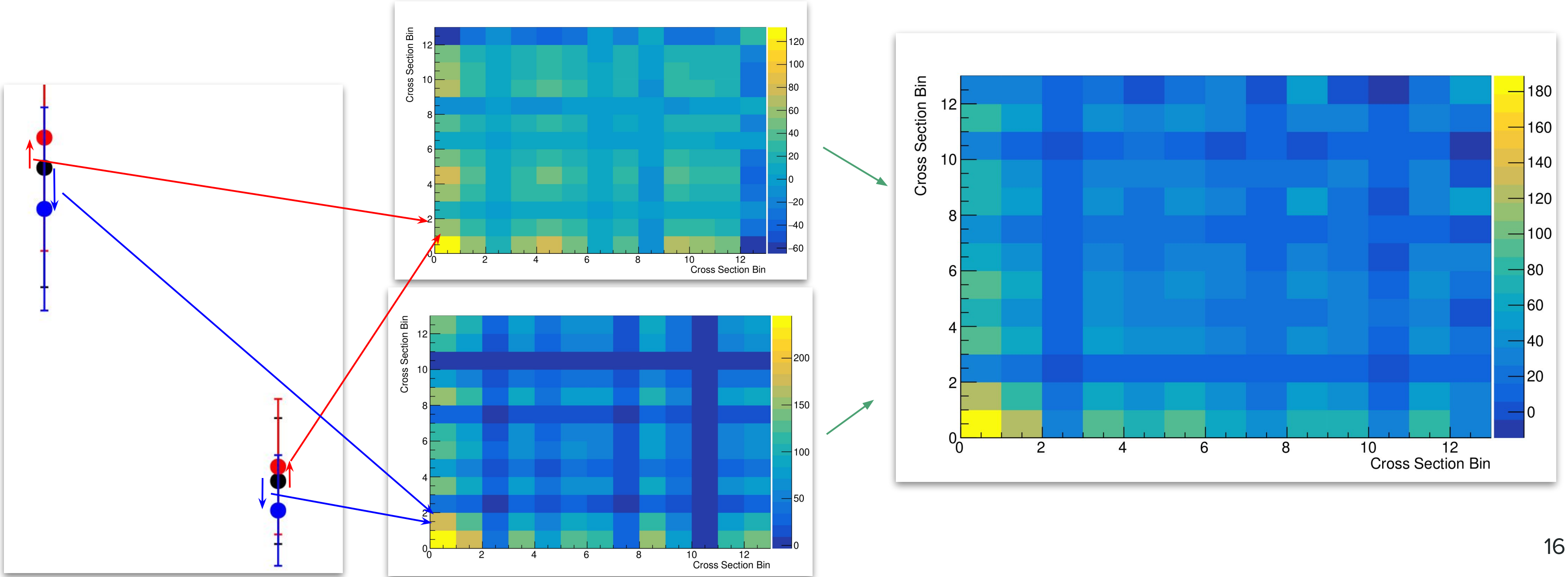


New Beam Smearing Approach

$$V_{ij} = (\sigma_i - \hat{\sigma}_i)(\sigma_j - \hat{\sigma}_j)$$

Bin-by-bin, compute covariance (V_{ij}) separately for the ±1σ results

Average the covariances. Add this to the nominal post-fit covariance



Parameterization

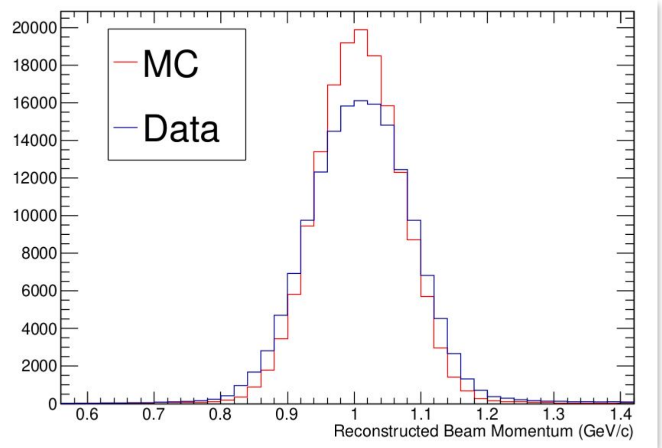
1. Absorption < 400 MeV
2. Abs. 400 - 500 MeV
3. Abs. 500 - 600 MeV
4. Abs. 600 - 700 MeV
5. Abs. 700 - 800 MeV
6. Abs. 800 - 900 MeV
7. Abs. > 900 MeV
8. Charge Exchange < 500 MeV
9. Ch. Exch. 500 - 600 MeV
10. Ch. Exch. 600 - 700 MeV
11. Ch. Exch. 700 - 800 MeV
12. Ch. Exch. 800 - 900 MeV
13. Ch. Exch. > 900 MeV
14. Other Inelastic < 500 MeV
15. Other Inel. 500 - 600 MeV
16. Other Inel. 600 - 700 MeV
17. Other Inel. 700 - 800 MeV
18. Other Inel. 800 - 900 MeV
19. Other Inel. > 900 MeV
20. Muon Fraction
21. Beam Matching Rate
22. dE/dX Calibration
23. Electron Diverter Effect
24. Proton G4RW
25. No Track Rate
26. Upstream Interactions

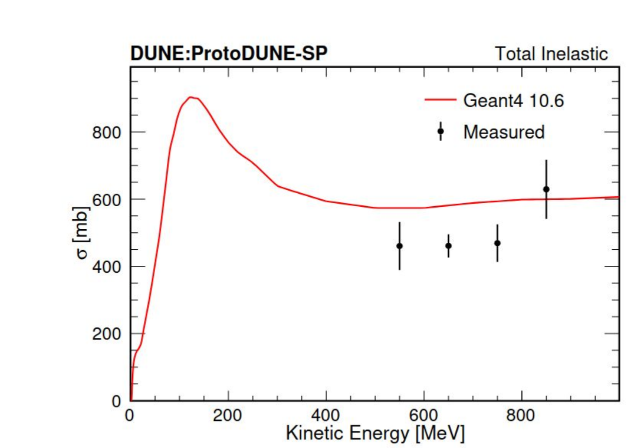
SCE effects

Normalizing to Data – Reco Beam Momentum

Noticed that the high-energy cross section was being pulled high (most prevalent when looking at total inelastic cross section)

Realized this could be due to a difference initial reconstructed momentum profile

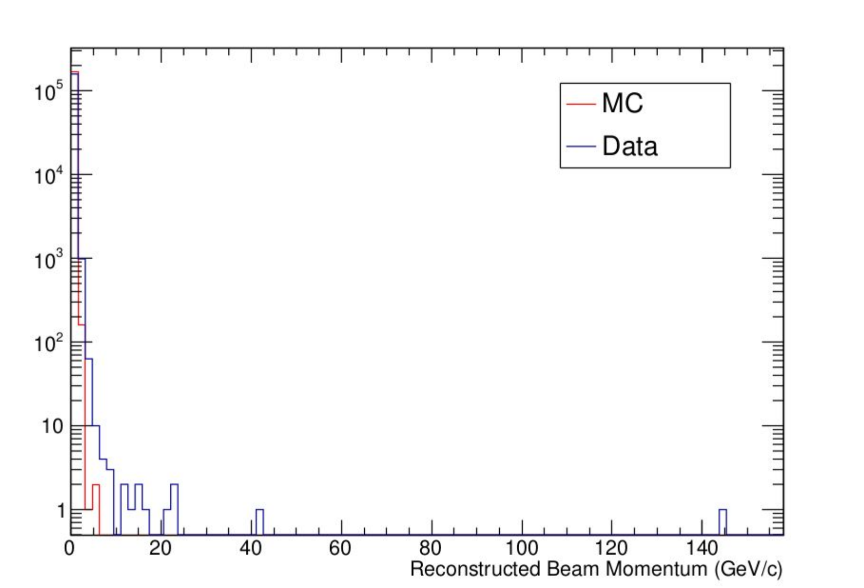




Restricting Reco Beam Momentum

Also, events in data extend to ridiculously large reconstructed beam momentum values

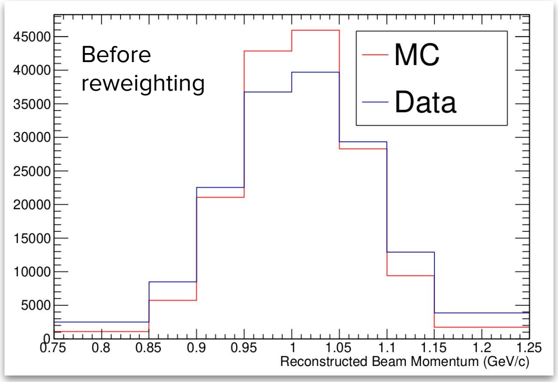
These are obviously erroneous, so I've decided to restrict reco beam momentum in both MC and data

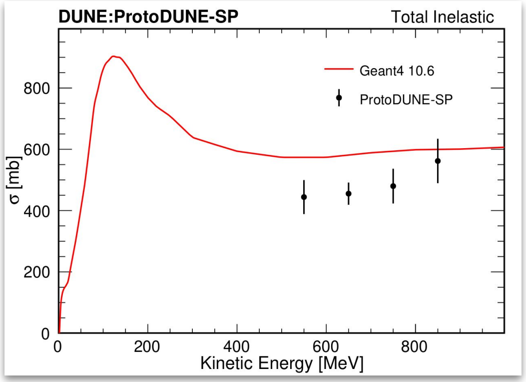


New Reco Beam Momentum Procedure

Prior to fit: restrict available events in data and MC to (750, 1250) MeV/c

Prior to & each step of fit: normalize events to data beam momentum profile





Note: Heng-Ye has previously talked about doing something very similar in his proton analysis

Preliminary Results

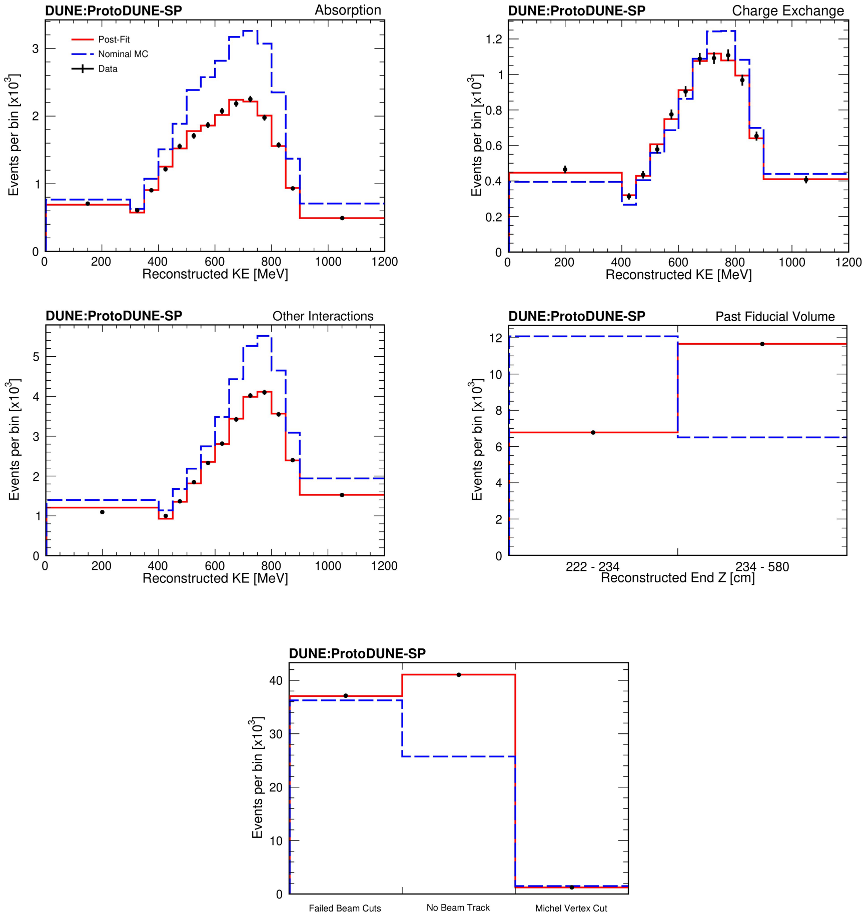
χ^2/ndof = 36.41/22 = 1.66

p-value = .027

Shows a possible residual systematic effect (maybe negligible)

Possibly need to include MC stats in fit (a la Barlow–Beeston)

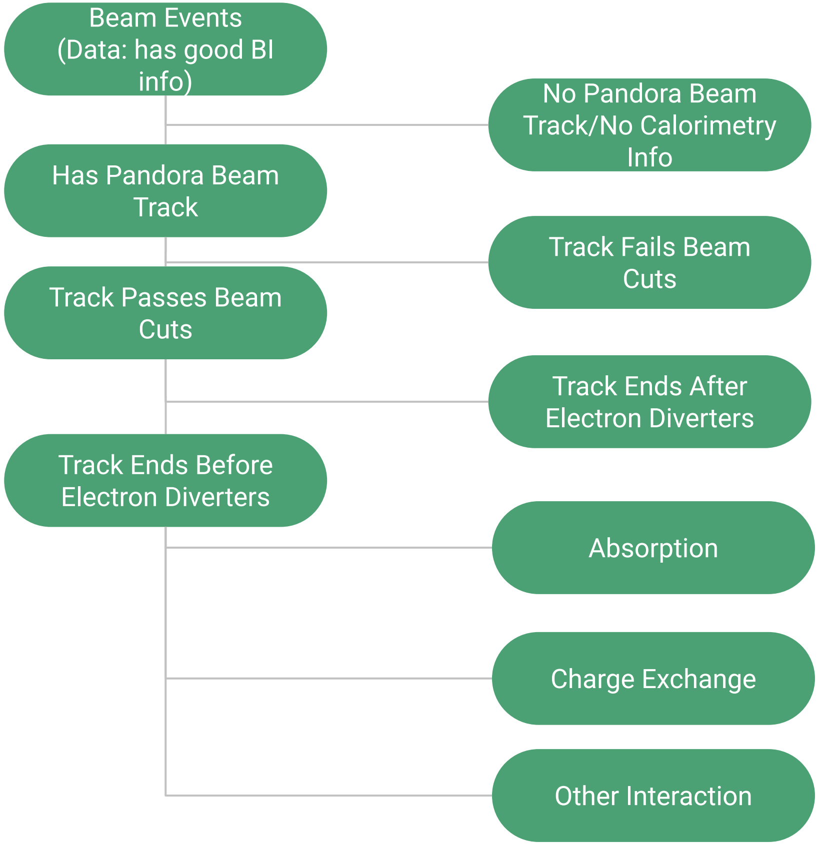
Kolmogorov–Smirnov test gives ~1 (though it's intended for unbinned data)



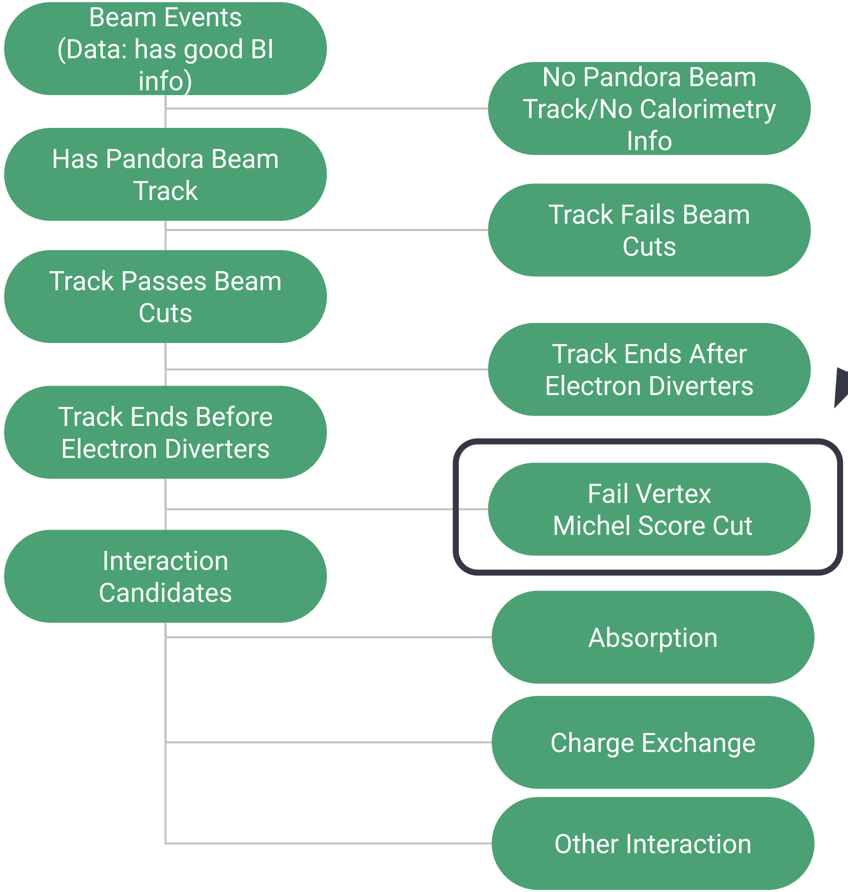
Thank you for listening

Backup Slides

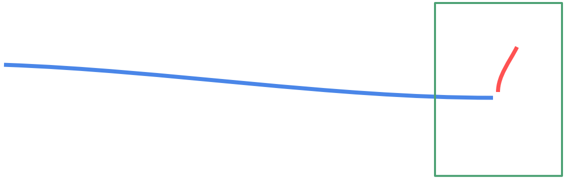
Event Selection - Updated



Event Selection - Updated



Adding another category to cut out muons & stopping pions from the interaction candidates



Look for any hits within a window near the end of the primary track, average their Michel-like CNN score

Event Selection - Updated

Beam Events (Data: has good BI info)

No Pandora Beam Track/No Calorimetry Info

Has Pandora Beam Track

Track Fails Beam Cuts

Track Passes Beam Cuts

Track Ends After Electron Diverters

Track Ends Before Electron Diverters

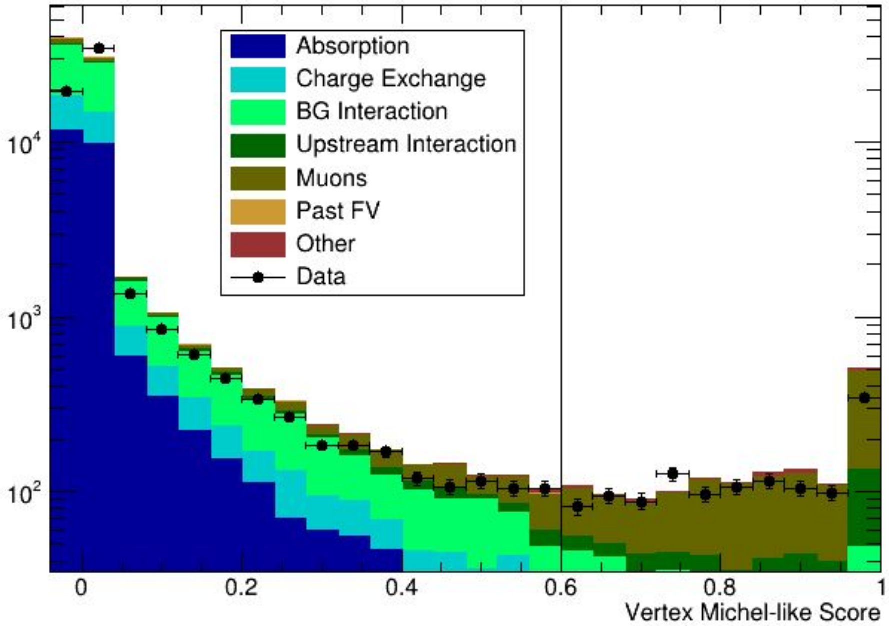
Fail Vertex Michel Score Cut

Interaction Candidates

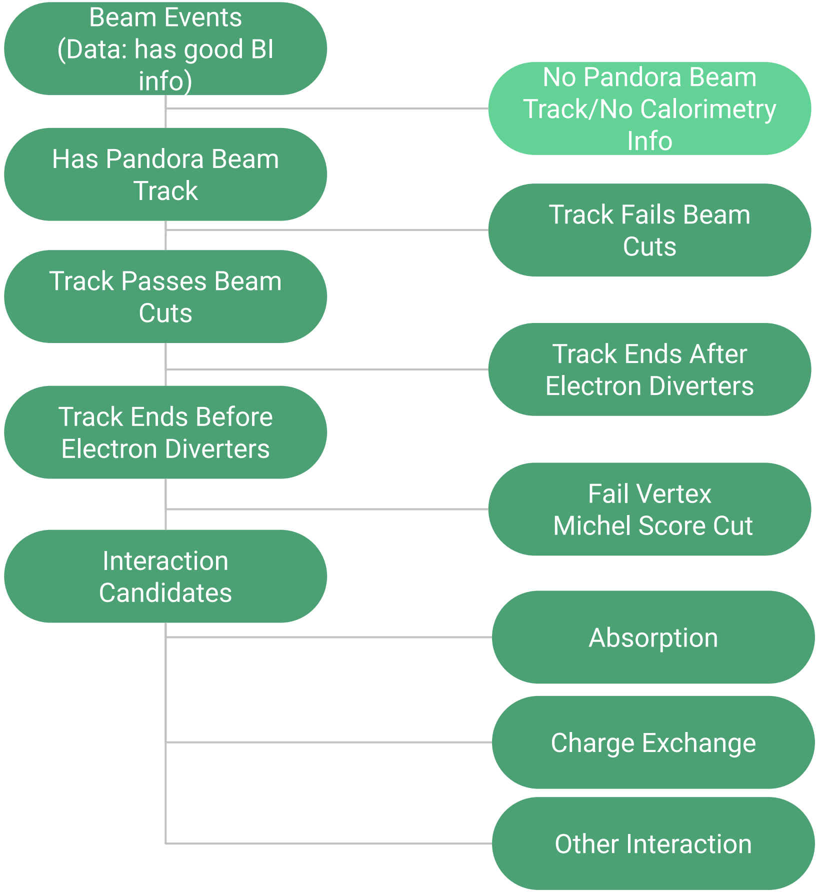
Absorption

Charge Exchange

Other Interaction

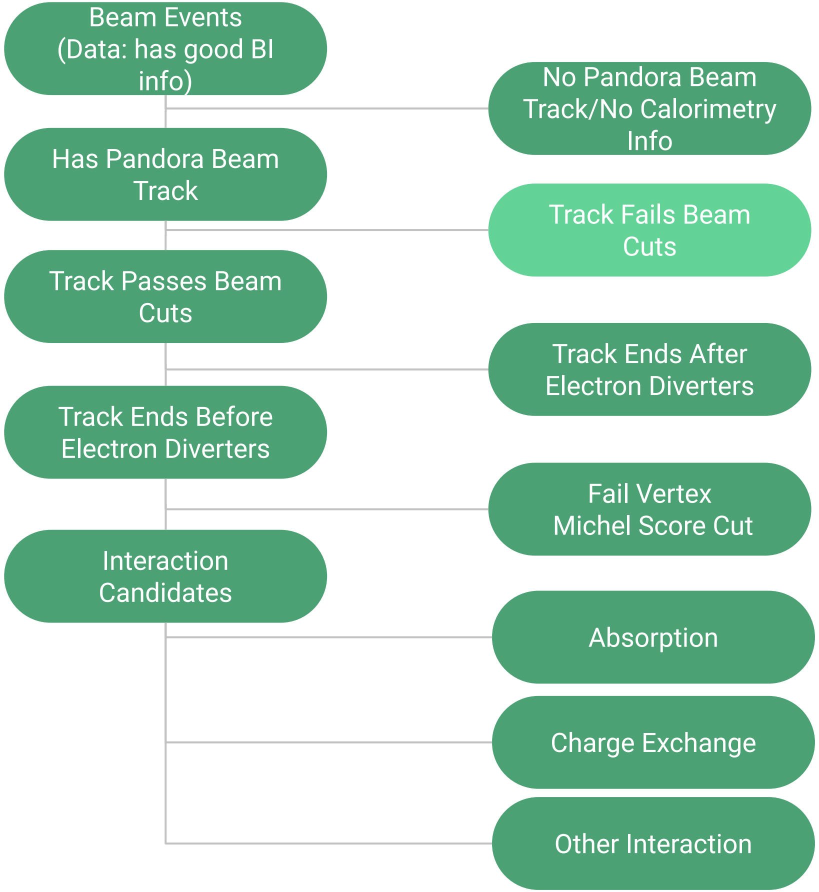


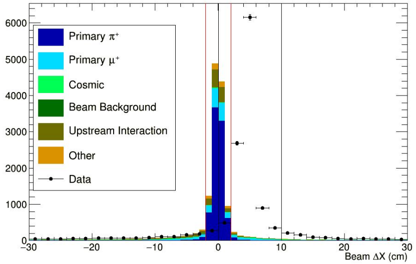
Event Selection

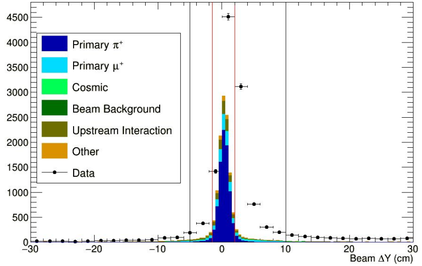


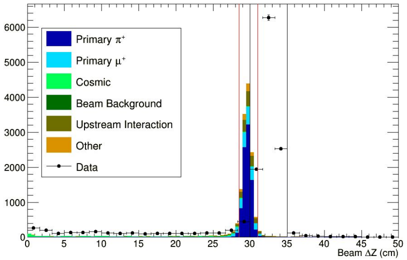
Events where no track was reconstructed in the beam slice by Pandora

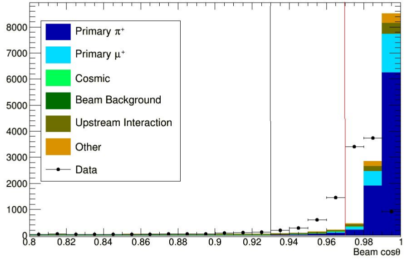
Event Selection



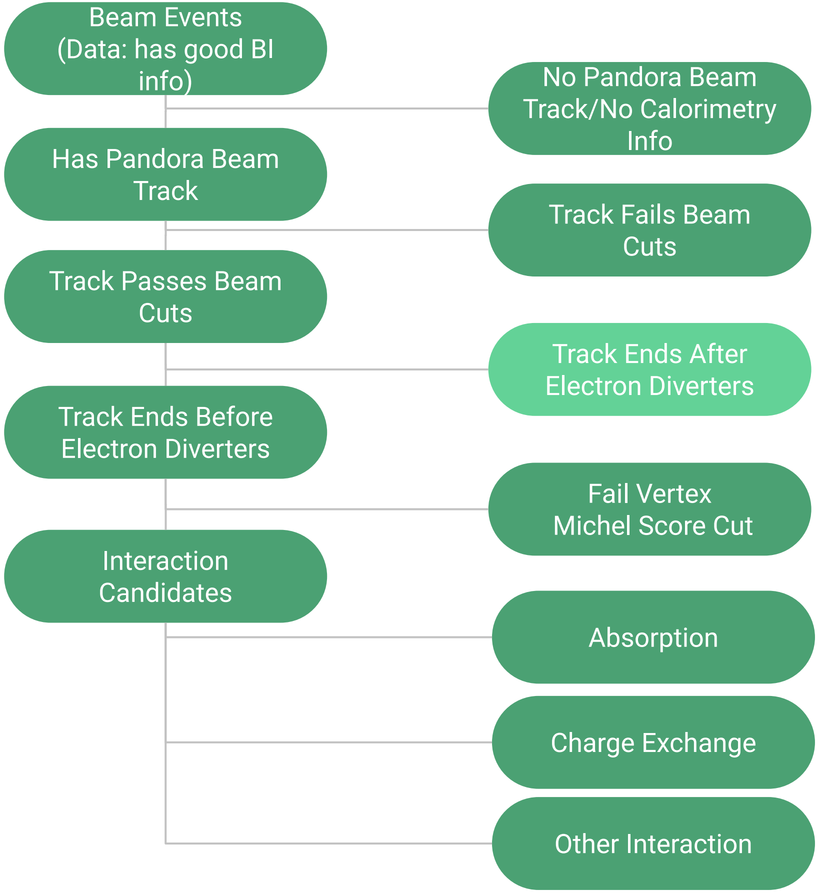


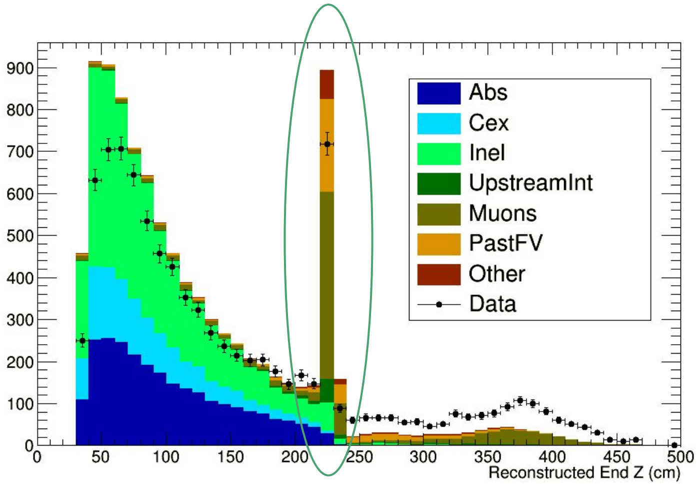


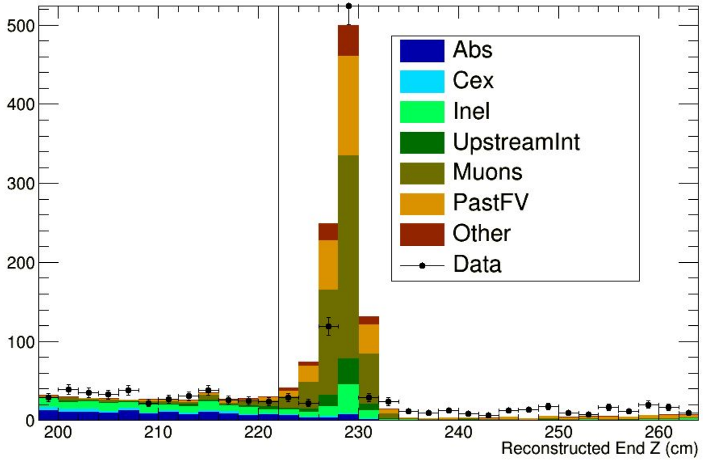




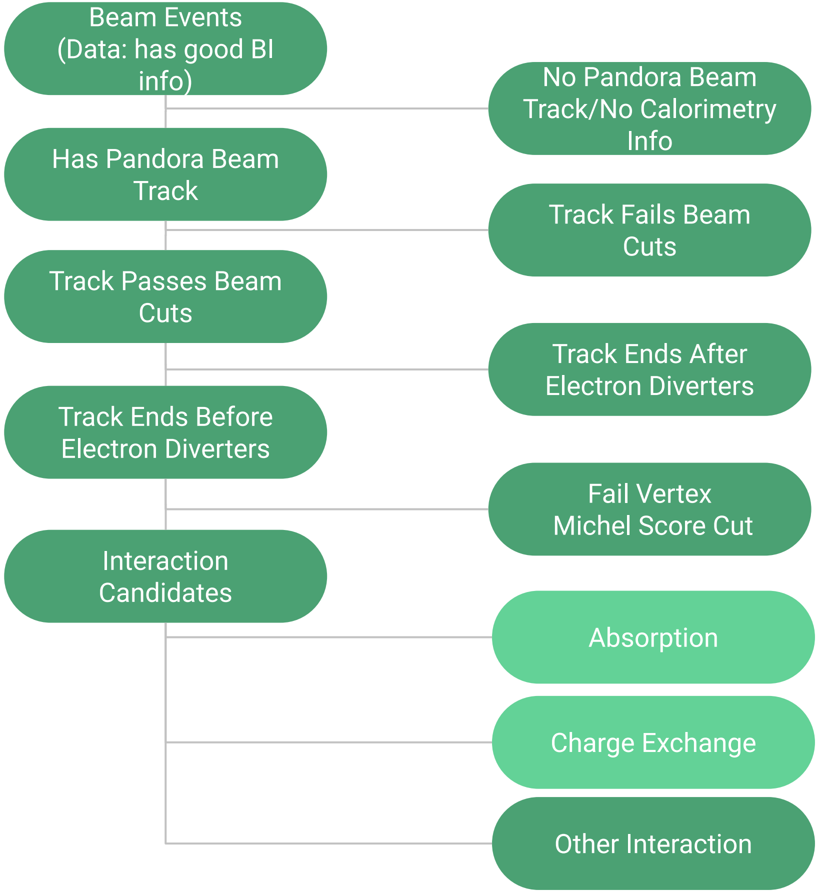
Event Selection





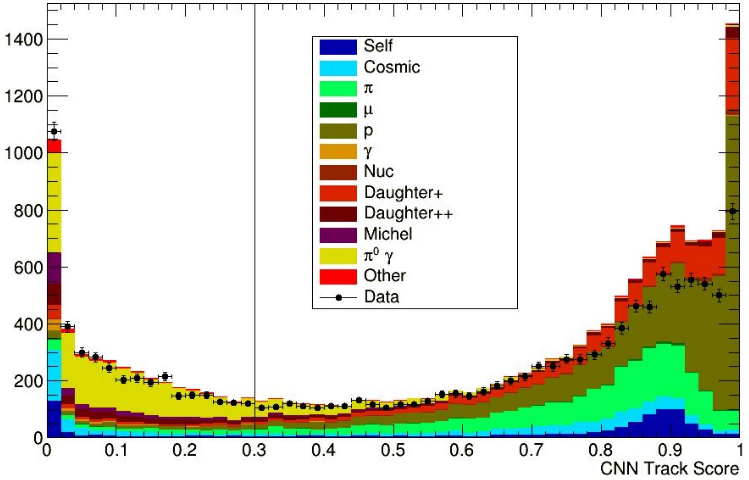


Event Selection

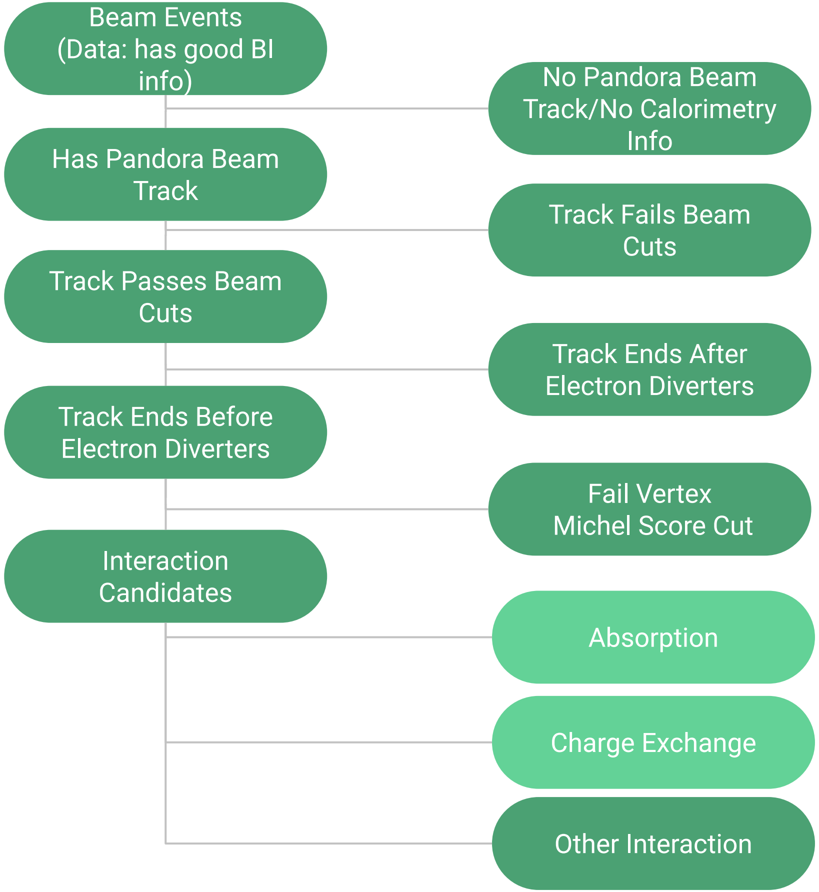


Identify track-like daughters using aggregate CNN scores of particles

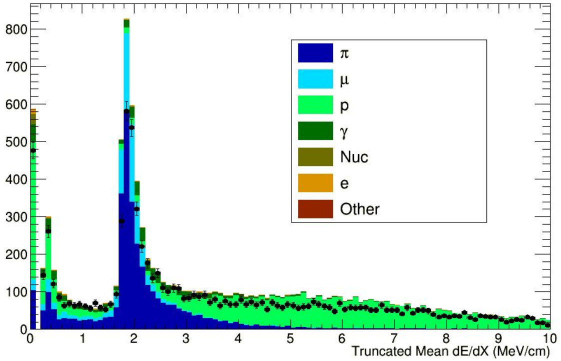
>.3 ➡ Track-like

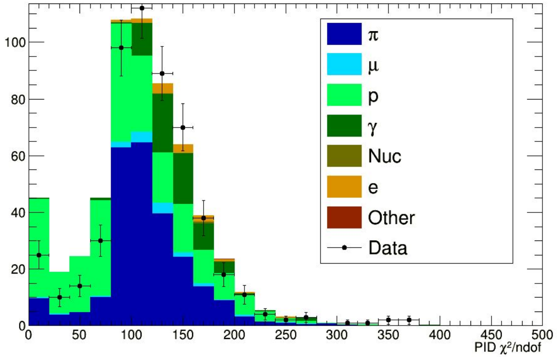


Event Selection

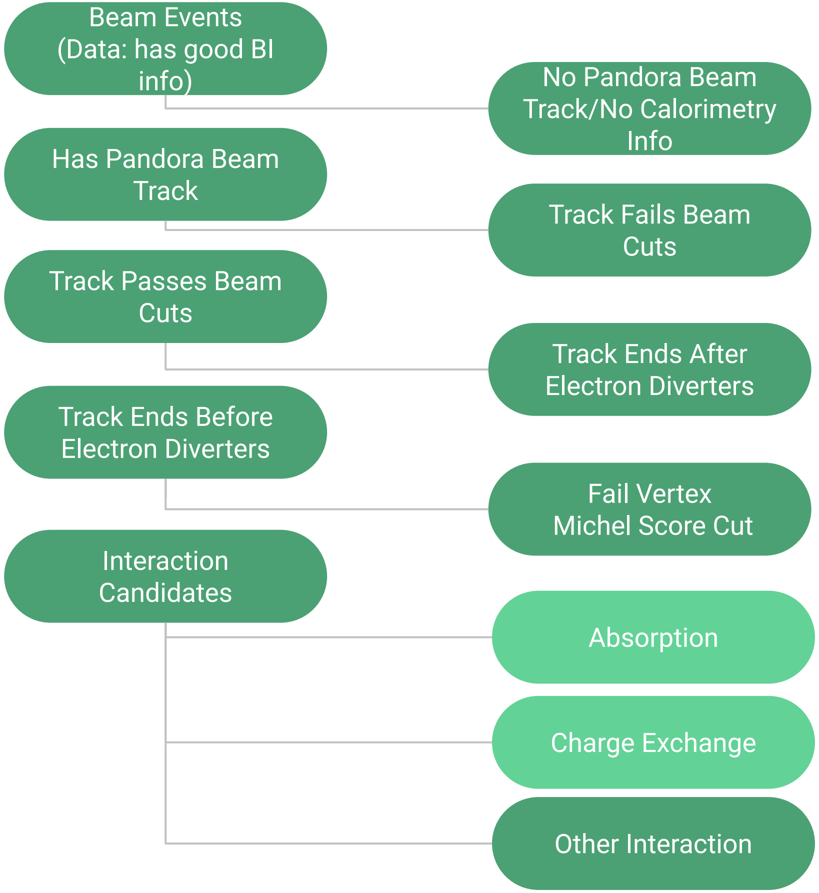


Use calorimetry information to identify charged pions within tracks

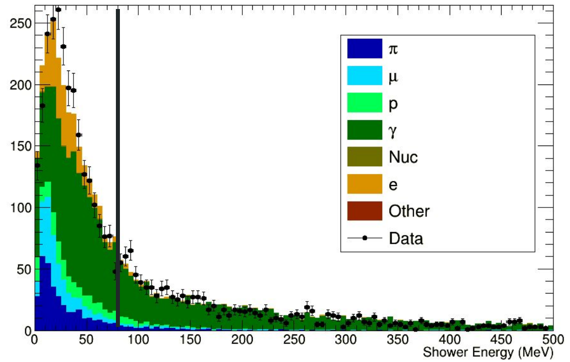


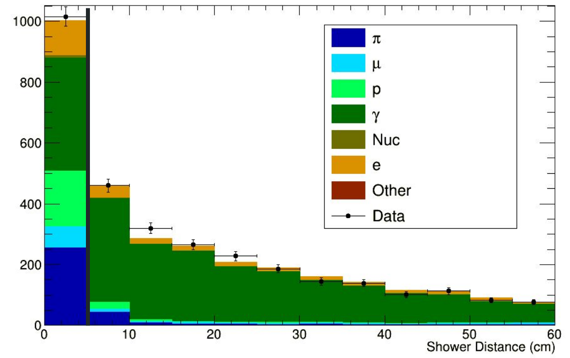


Event Selection

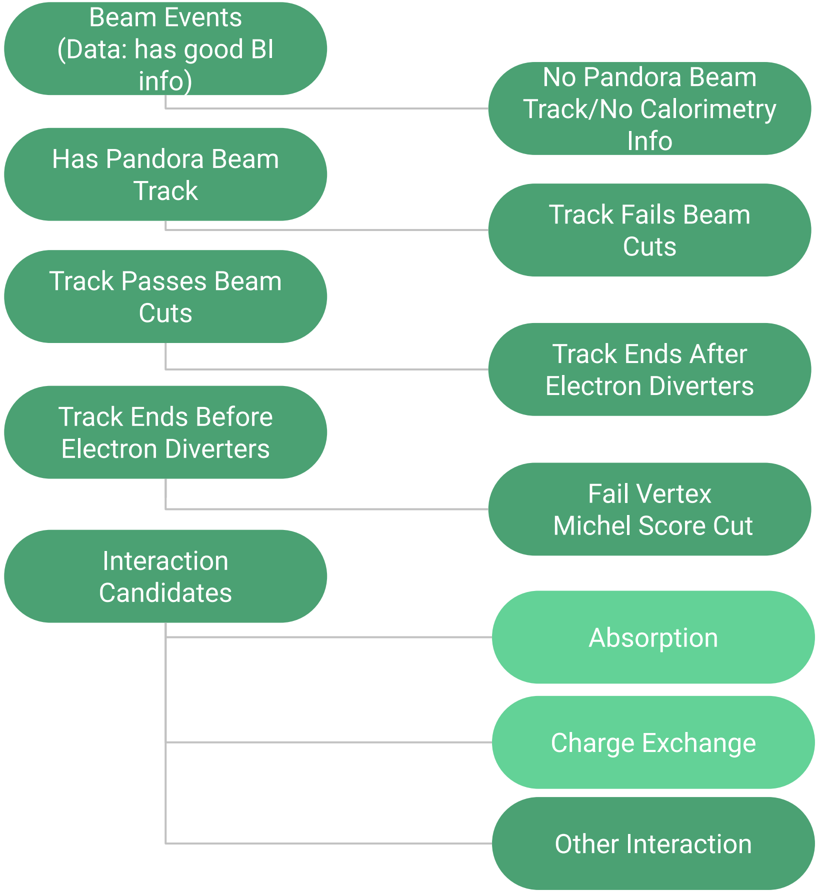


Was previously using two 1D cuts on energy and distance-to-vertex of shower-like reco daughter particles

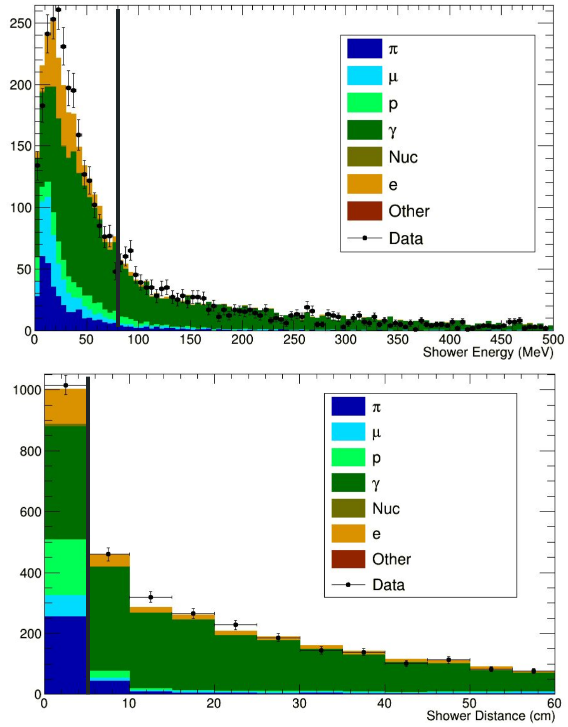




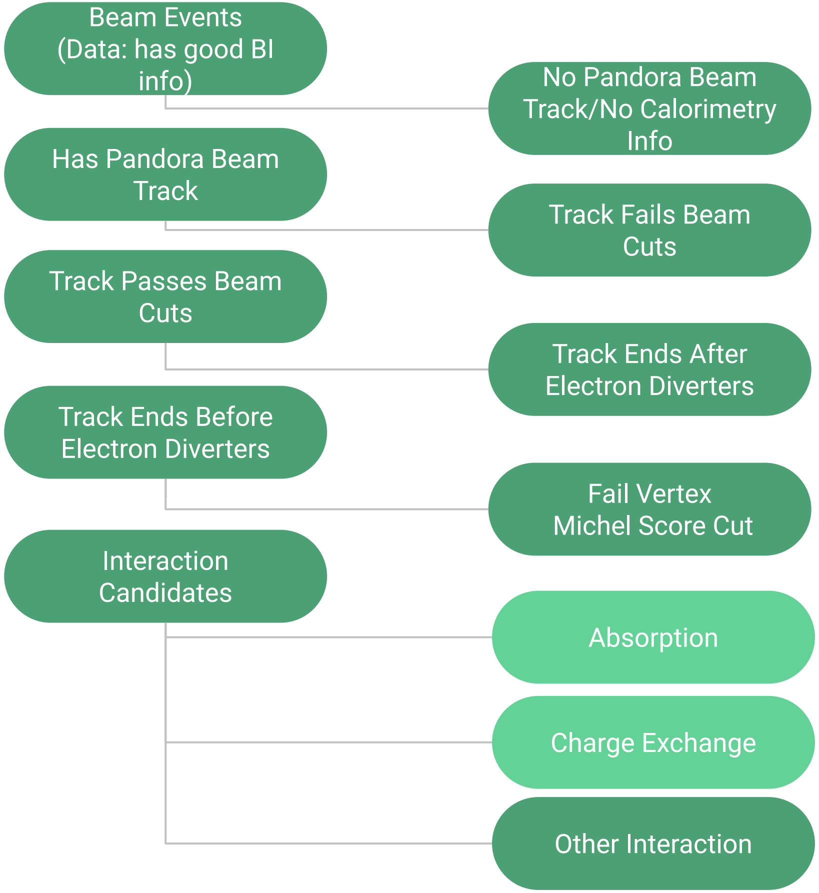
Event Selection



Realized a set of 2D cut would be better



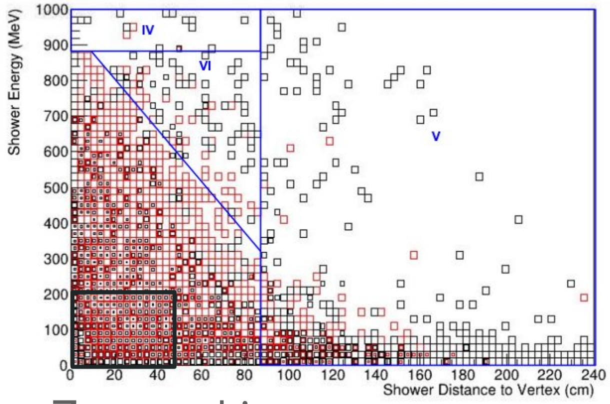
Event Selection



In this:

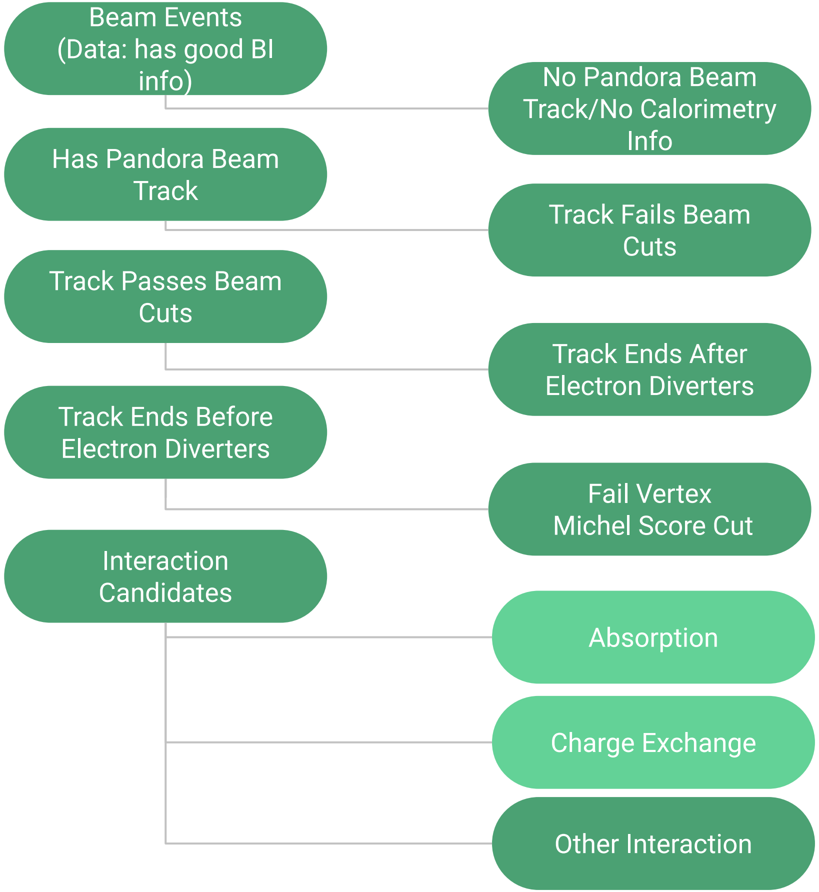
- size of square = fraction of particles in bin
- Red: True π^0 (signal)
- Black: Other (background)

Cut out areas where (generally) black > red



Next slide → Zoomed in

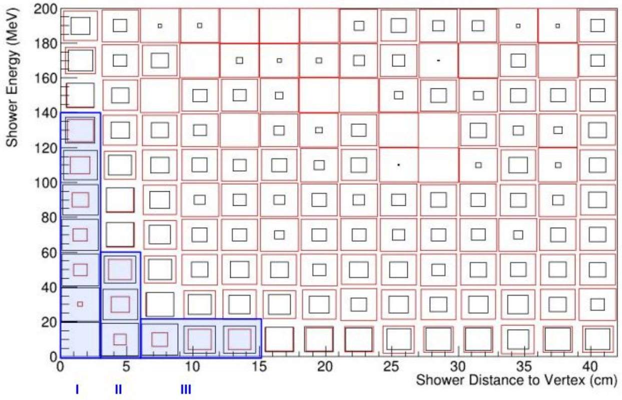
Event Selection



In this:

- size of square = fraction of particles in bin
- Red: True π^0 (signal)
- Black: Other (background)

Cut out areas where (generally) black > red



Beam Resolution

Implement as affecting the smearing from true to reco (*r*)

Magnetic field: direct 1% uncertainty on p_{Reco}

$$r = \frac{p_{\text{Reco}} - p_{\text{True}}}{p_{\text{True}}}$$

Shift: determine from nominal beam MC

➡ 0.7% uncertainty on p_{Reco}

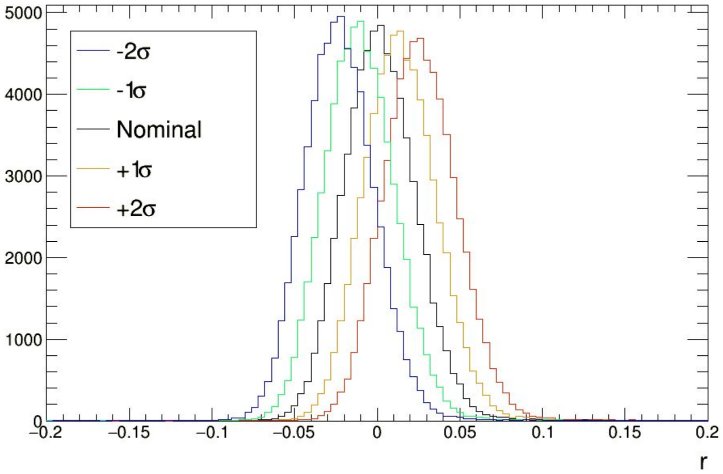
Add in quadrature

➡ overall 1.2% uncertainty on p_{Reco}

Beam Resolution

Get means and widths of nominal, ±1,2σ shifts, interpolate between

Each event gets a weight according to the ratio of varied to nominal distributions

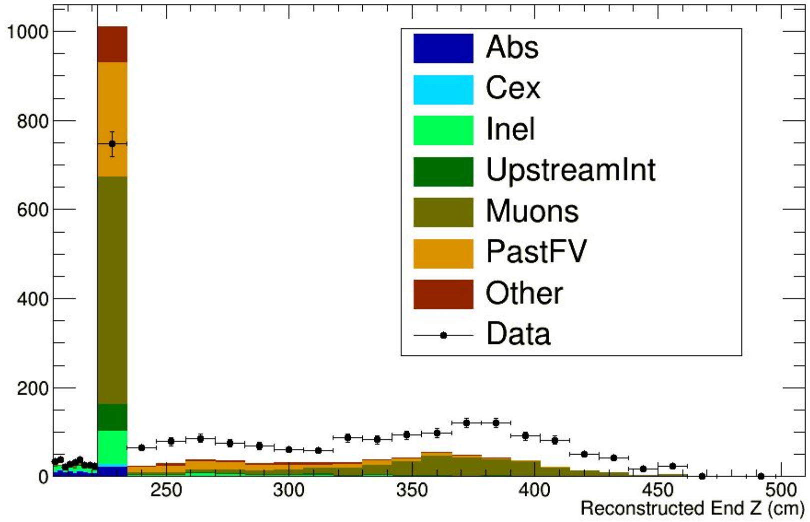


$$w = \left(\frac{\sigma}{\sigma'}\right) \exp\left(\frac{(r-\mu)^2}{2\sigma^2} - \frac{(r-\mu')^2}{2\sigma'^2}\right)$$

Electron Diverter Effect

Prod4a includes a simulation of the electron diverters (thanks to Tom Junk)

But the overall effect seems overestimated ➜ need to account for the uncertainty in rate of track breakage



Electron Diverter Systematic Implementation

$$f_{\text{Break}} = \frac{N_1}{N_1 + N_2}$$

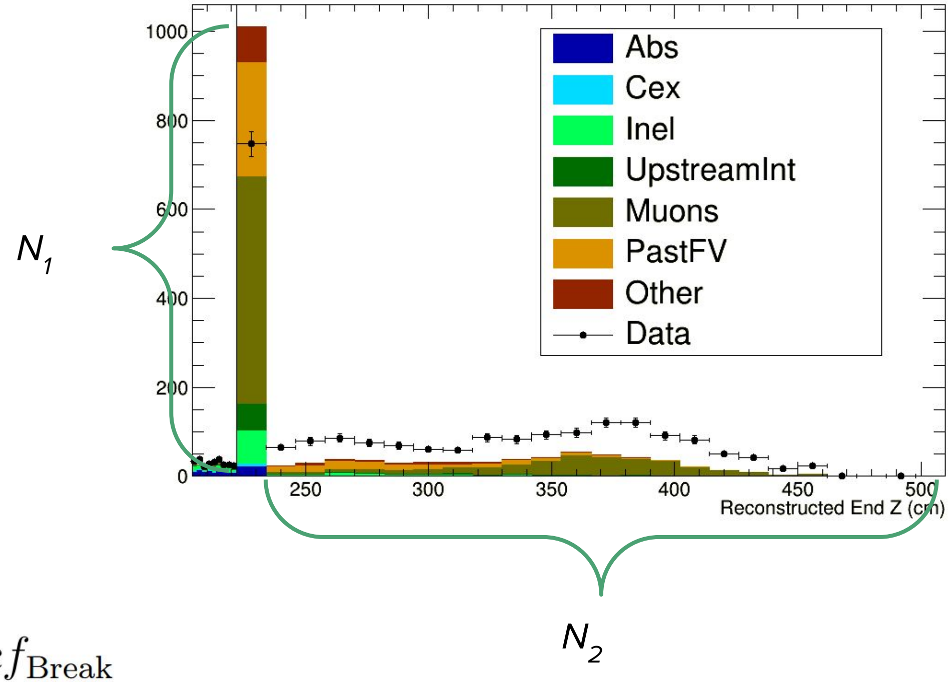
$$f_{\text{Break}} \to f'_{\text{Break}} = c f_{\text{Break}}$$

If track ends in break region (220-234 cm)

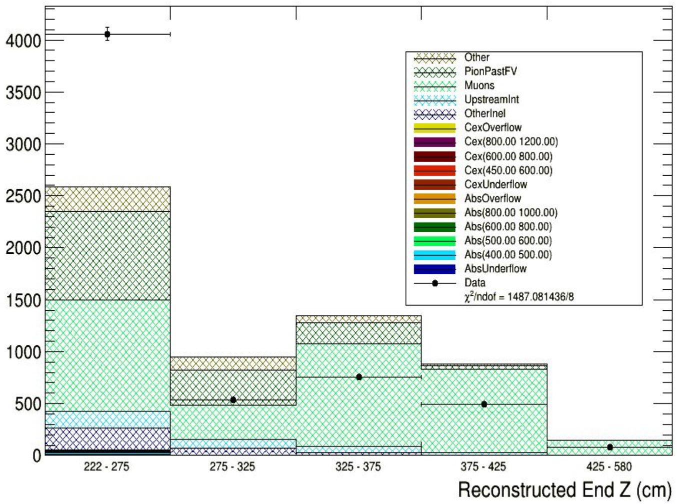
$$W_1 = \frac{f'_{\text{Break}}}{f_{\text{Break}}} = c$$

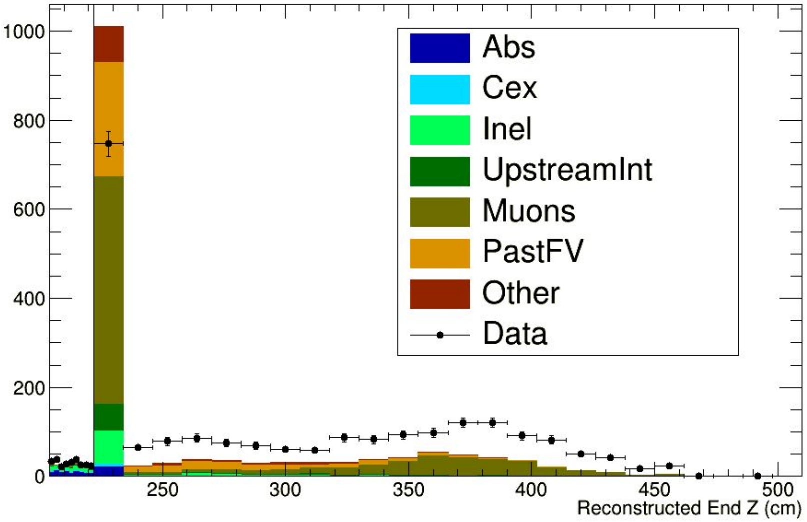
If track ends past break region

$$W_2 = \frac{1 - f'_{\text{Break}}}{1 - f_{\text{Break}}} = \frac{1 - c f_{\text{Break}}}{1 - f_{\text{Break}}}$$



Electron Diverter Prior Uncertainty





Data: nominal MC
Stacks: f_{Break} reduced by 50%

Electron Diverter Prior Uncertainty

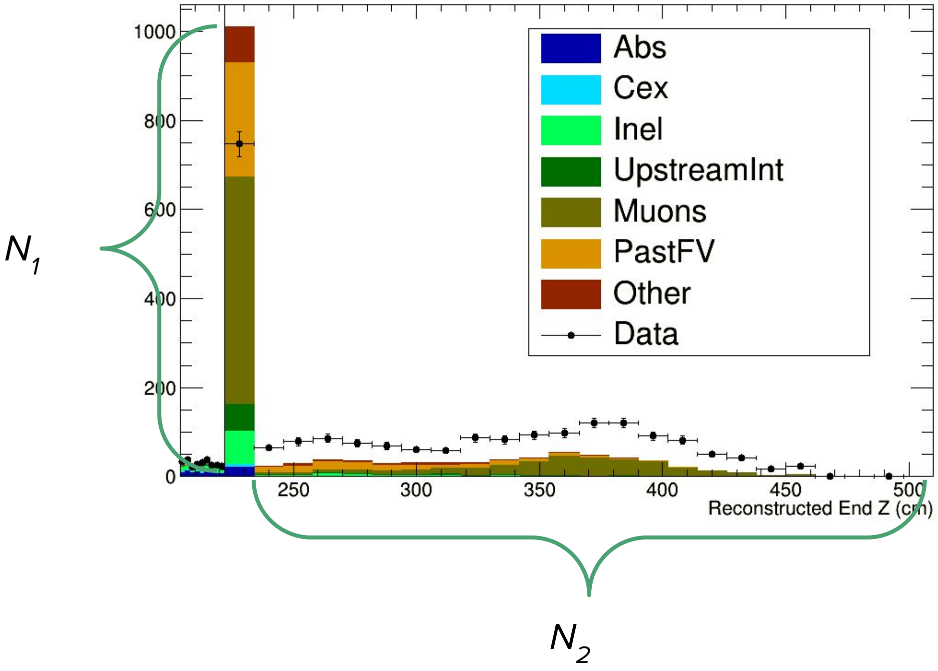
Data:

- $N_1 = 696$
- $N_2 = 1414$
- $f_{Break} = 0.330$

MC*:

- $N_1 = 3627$
- $N_2 = 2287$
- $f_{Break} = 0.613$

$\sigma_c = 50\%$



* Note: Stated MC rates unnormalized

Pandora & Beam Cut Efficiencies

Data-MC differences:

1. Fraction of events with a beam track reconstructed by Pandora
2. Fraction of events passing beam quality cuts

Allow for freedom in fit to vary these

	Total	Pandora	Calo size	Beam quality
Data	18289	14003	13639	9485
Total MC	18289	15549	15255	11035

From Tingjun's talk

Pandora & Beam Cut Efficiencies -- Implementation

Event categories:

1. No beam track
2. Failed beam cuts
3. “Good” events

Consider variation to these fractions:

$$f_1 \rightarrow f_1' = c_1 f_1$$

$$f_2 \rightarrow f_2' = c_2 f_2$$

$$f_3 \rightarrow f_3' = 1 - c_2 f_2 - c_1 f_1$$

Weight each event according to what category it is:

$$W_1 = \frac{f_1'}{f_1} = c_1$$

$$W_2 = \frac{f_2'}{f_2} = c_2$$

$$W_3 = \frac{f_3'}{f_3} = \frac{1 - c_2 f_2 - c_1 f_1}{1 - f_2 - f_1}$$

Note on Beam Resolution Systematic

The beam resolution systematic was causing instability in the fit during validation

- Fake data created by throwing systematics to prior uncertainties would sometimes create giant weights for large variations of the beam resolution parameter

Fixed parameter before fit, then added prior uncertainty in quadrature to post-fit covariance

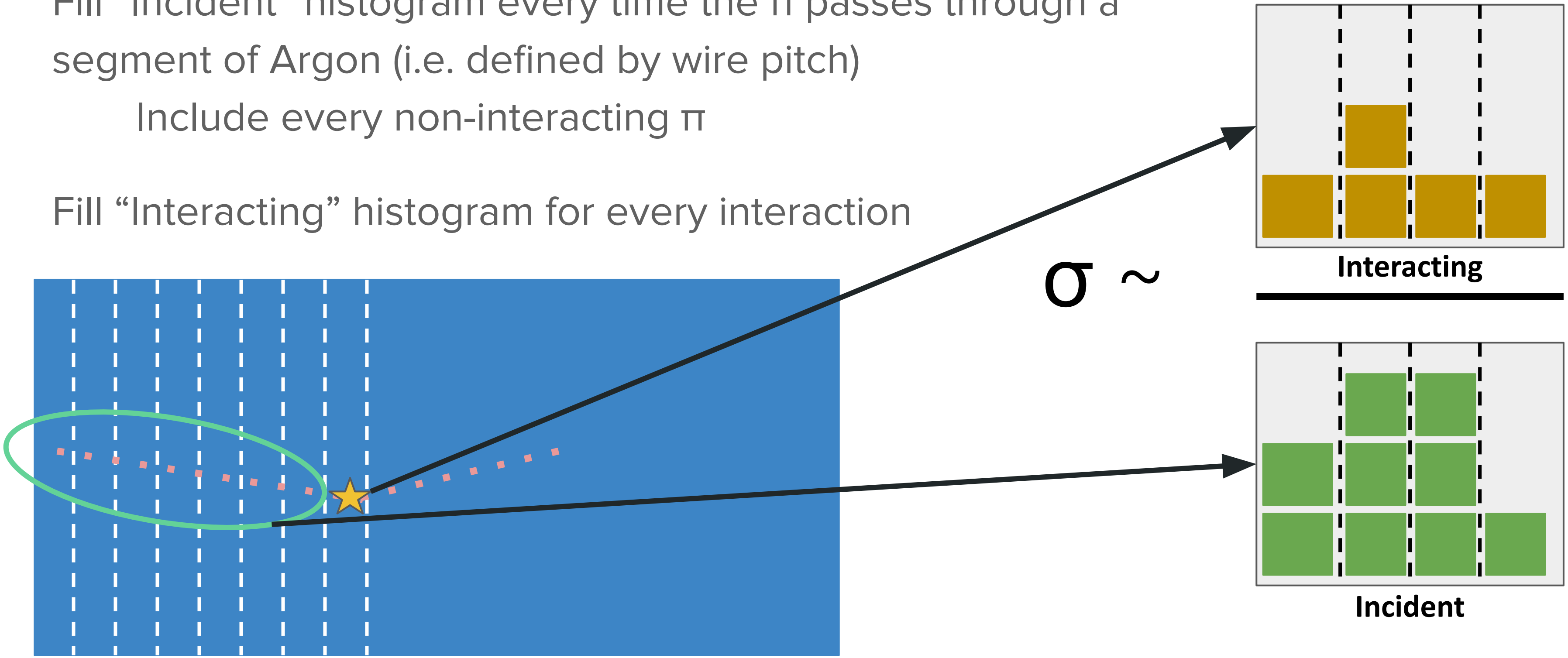
- Prior uncertainty still used within error propagation procedure (will describe later)

Thin Slice Method

Fill “Incident” histogram every time the π passes through a segment of Argon (i.e. defined by wire pitch)

Include every non-interacting π

Fill “Interacting” histogram for every interaction



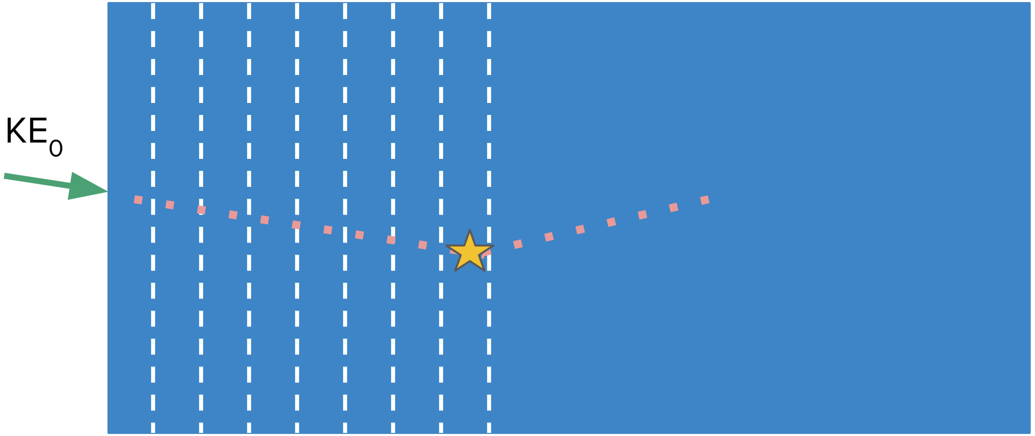
Thin Slice -- True Slices

To calculate the cross section, ‘slice’ up the path of the simulated pion to create a sequence of thin target scattering experiments.

Using the true energy at the start of LAr, and the energy of the MC trajectory points: calculate the energy incident in each of the slices

Use these to create the incident histogram

Reminder: Essentially the same as a flux in a ‘classic’ thin target experiment.



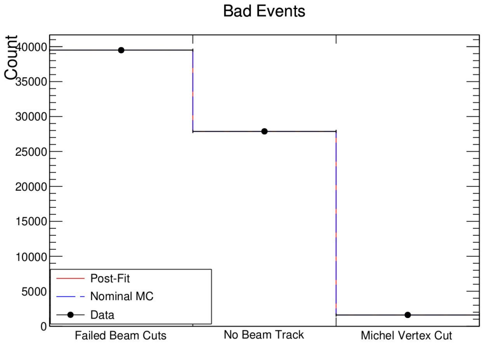
Measurement Strategy

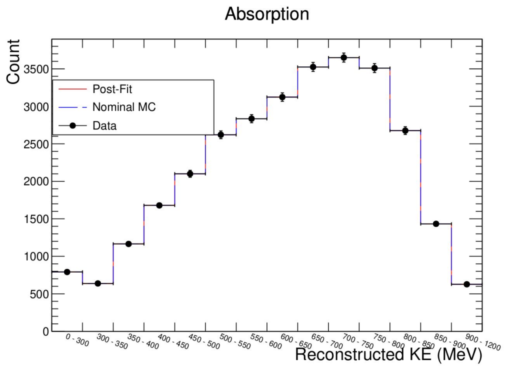
1. **Fit** to the number of selected interactions in reconstruction
   a. The fit varies the number of true signal interactions (binned in true energy)
   b. Has a resulting change on the reconstructed distributions
   c. An alternative technique to unfolding
   d. Best-fit results will be a set of varied MC events

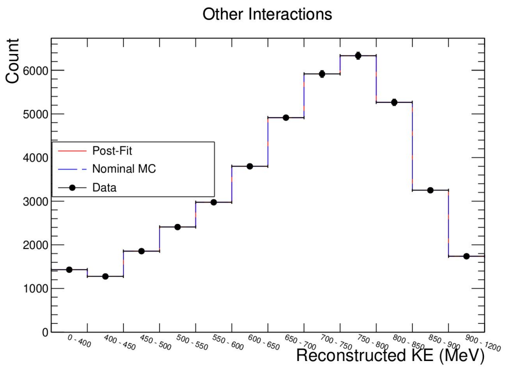
2. **Extract the cross section** from the varied MC
   a. Using the ‘thin slice method’ on varied truth information

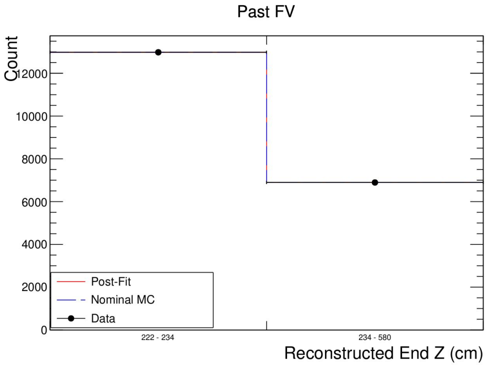
Fit Validation

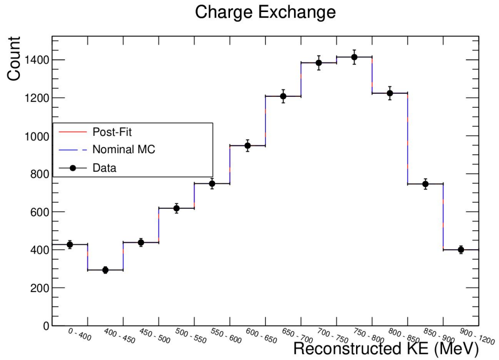
Asimov Results





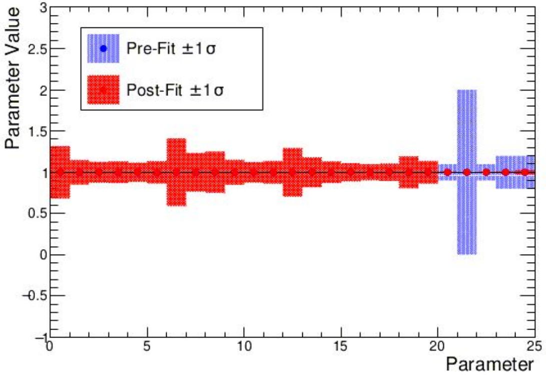


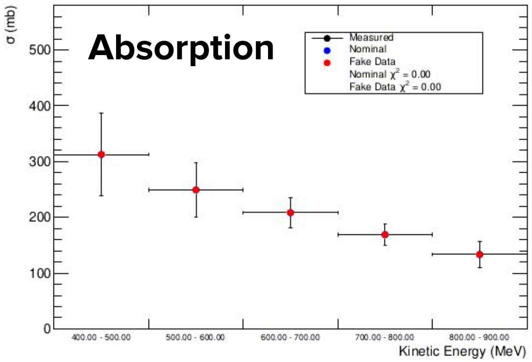


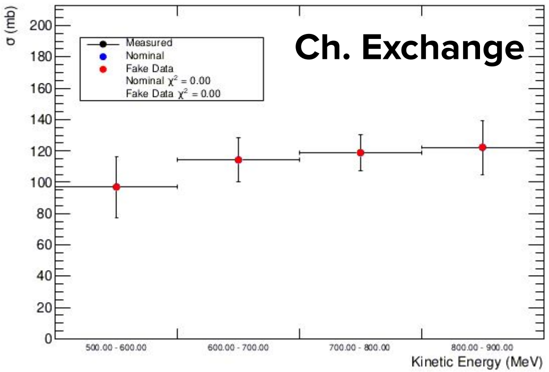


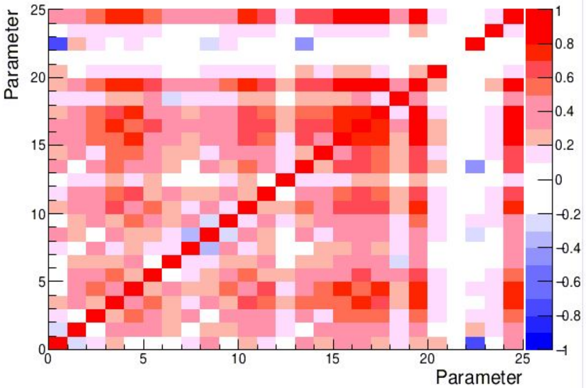
See prev. talk for error propagation procedure

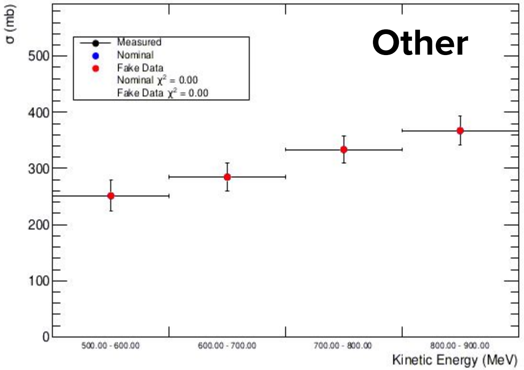
Asimov Results

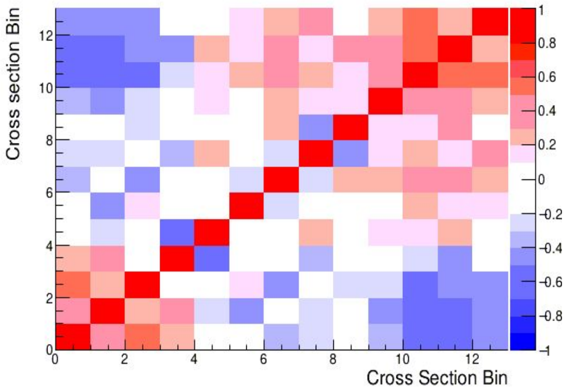








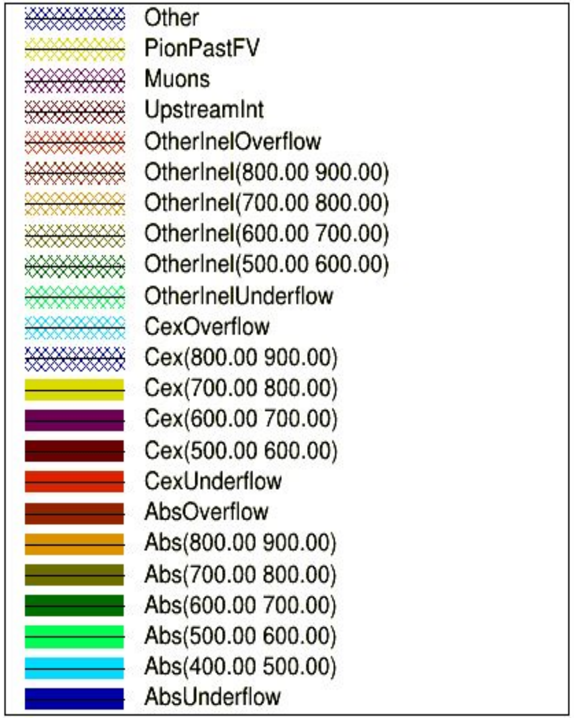


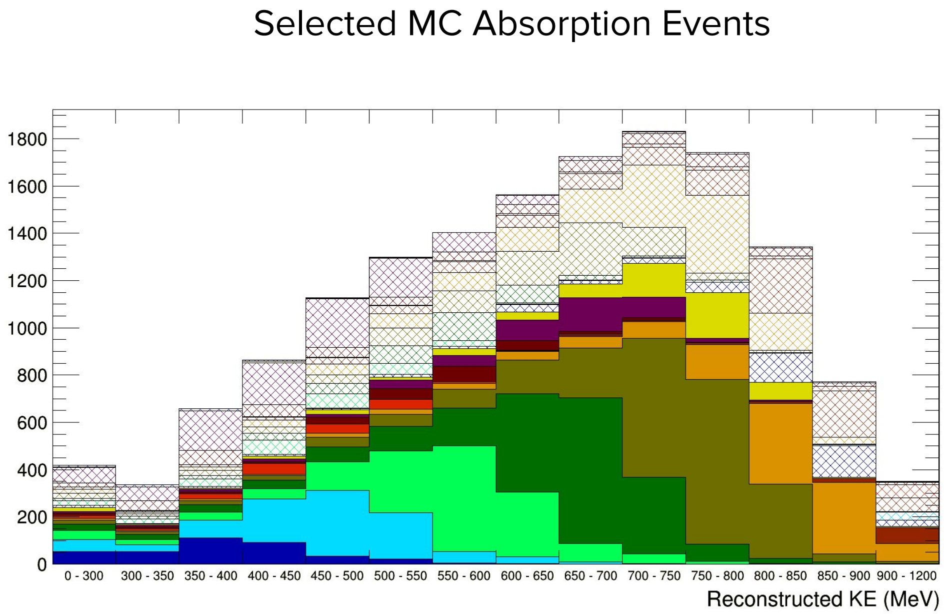


Truth Categories

Background events

Signal events in true bins





Fit Statistic

λ ➔ Likelihood ratio

Statistical term -- Multinomial statistics

Systematic term

$$-2\ln\lambda = 2\sum_j n_j \ln\frac{n_j}{y_j} + \sum_{i,j}(p_i - q_i)(V_{\mathrm{Cov}}^{-1})_{ij}(p_j - q_j)$$

Data

MC

Nominal systematic parameter values

Input systematic covariance matrix

Loop over all covariance bins

Varied systematic parameter values

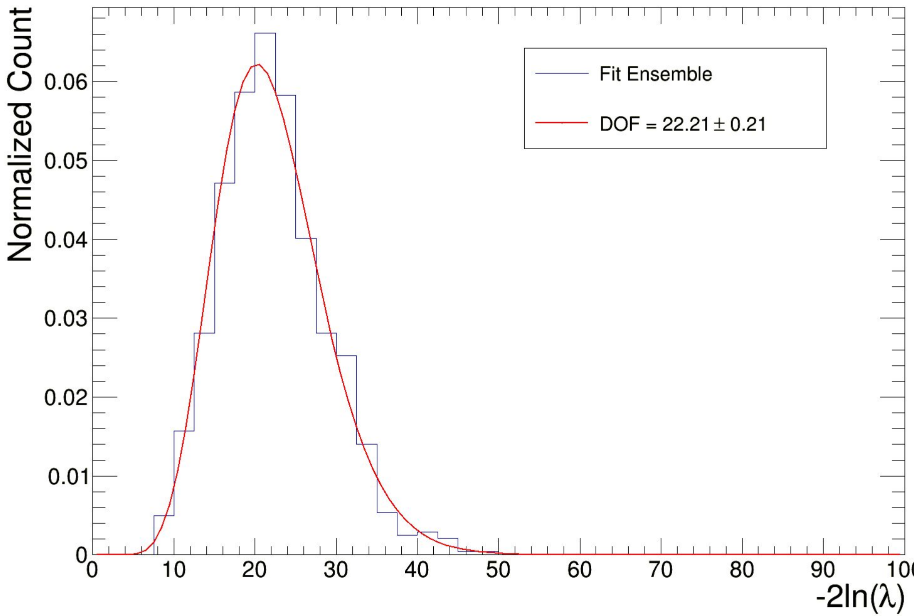
Systematic Uncertainties

Systematic Uncertainties

- dE/dX Calibration
  - Affects energy reconstruction
- Beam Resolution
  - Varies smearing between true and reconstructed beam line momentum
- Electron Diverter Effect
  - Varies how likely tracks are to break due to electron diverters
- Pandora Beam Track Efficiency
  - Varies how (un)likely Pandora is to identify a beam track
- Beam Cuts
  - Varies the fraction of events failing the beam cuts

Metrics -- Fit performance

Ensemble of toy fits appears χ^2-distributed so can use a simple χ^2/dof test to measure fit performance

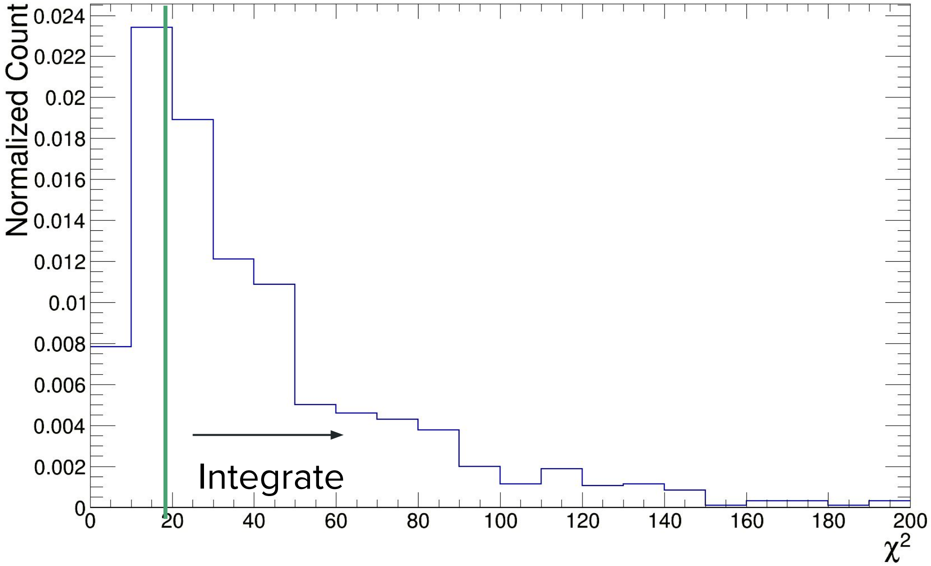


Metrics -- Cross Sections

Compare extracted cross section to nominal/fake data using post-fit covariance

Not exactly χ^2-distributed (some assumptions regarding the extracted errors are failing)

Calculate p-value rather than simple check against degrees of freedom



$$\chi^2_\sigma = \sum_{i,j} (\sigma_i - \bar{\sigma}_i)(V^\sigma)^{-1}_{i,j}(\sigma_j - \bar{\sigma}_j)$$

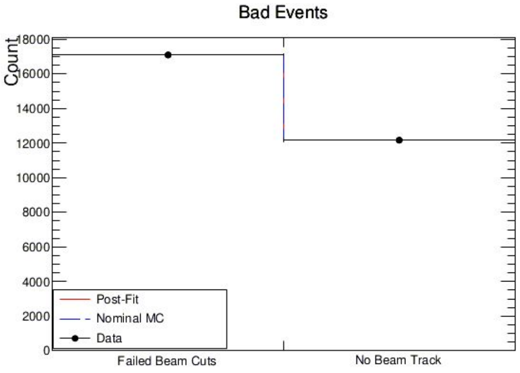
Extracted

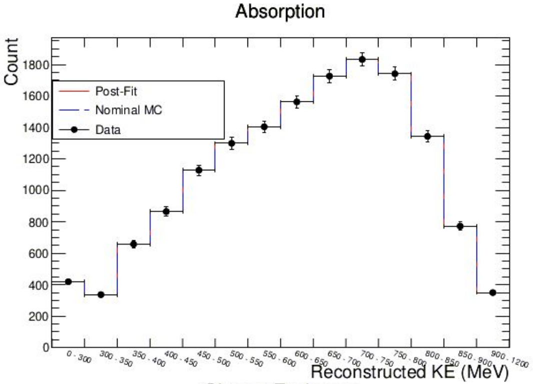
Post-fit covariance

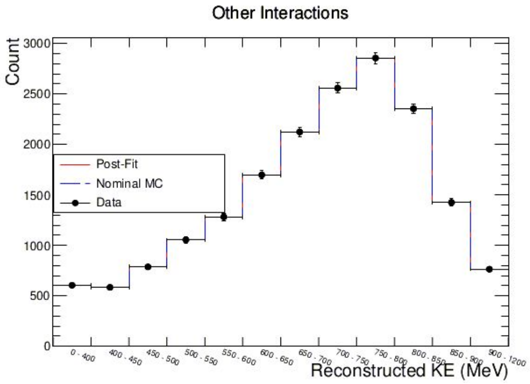
Nominal/Fake data

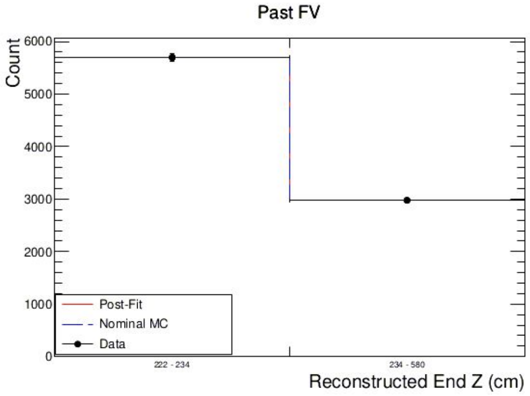
Asimov Fit

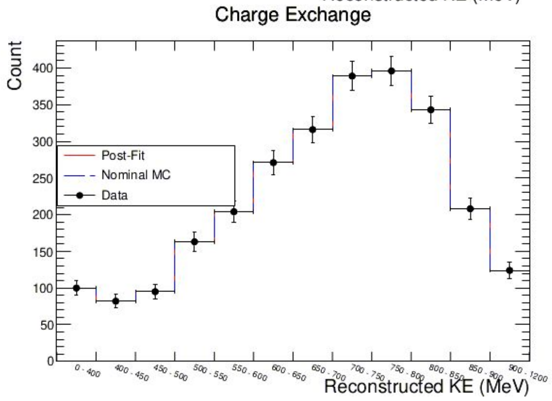
Asimov Results











Geant4Reweight Fake Data

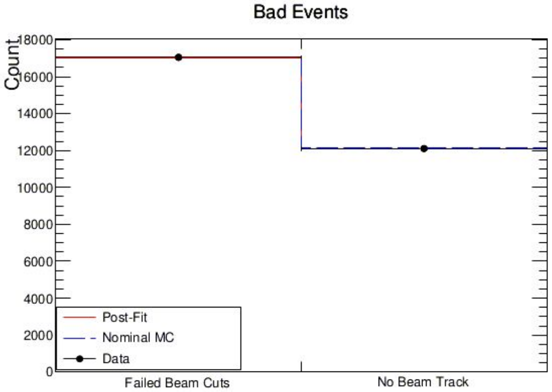
Geant4Reweight Fake Data

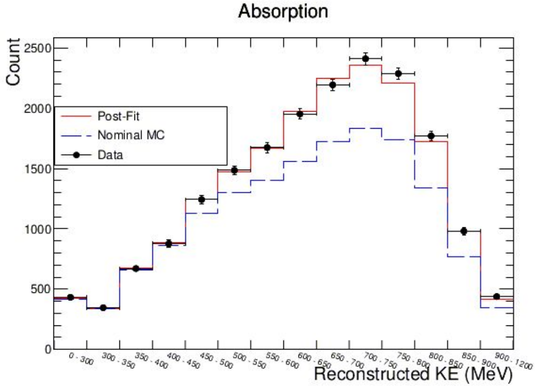
Create fake data by using Geant4Reweight to vary cross sections

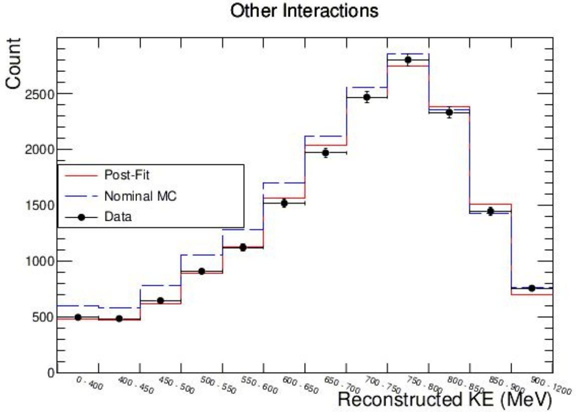
2 sets

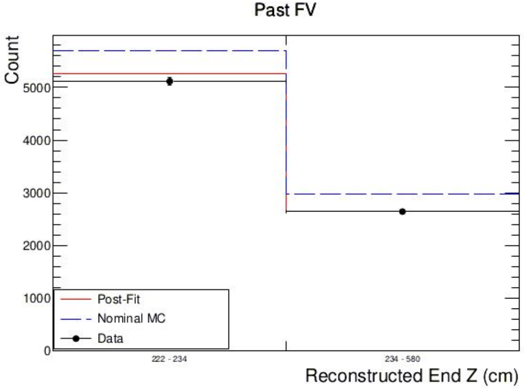
1. Increase absorption by 70%, reduce charge exchange by 60%
2. Vary total cross section: increase by 80% below 800 MeV/c, reduce by 60% above

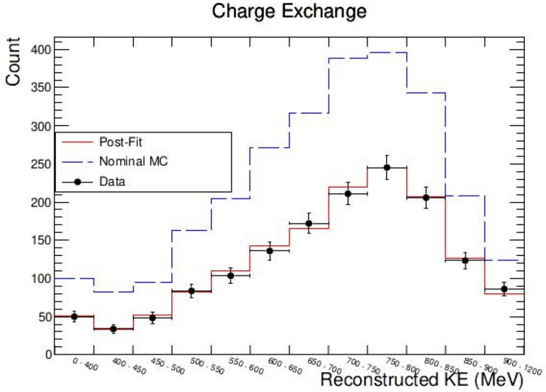
Geant4Reweight Fake Data 1







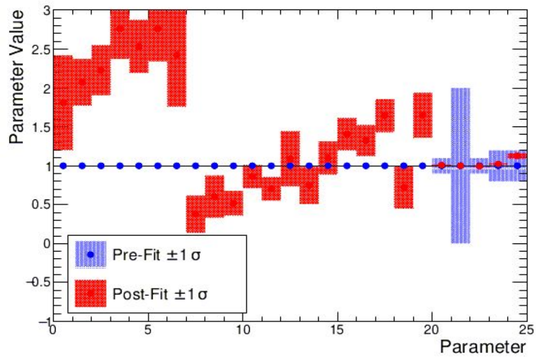


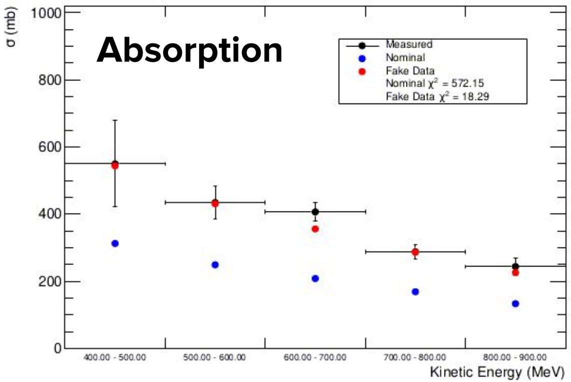


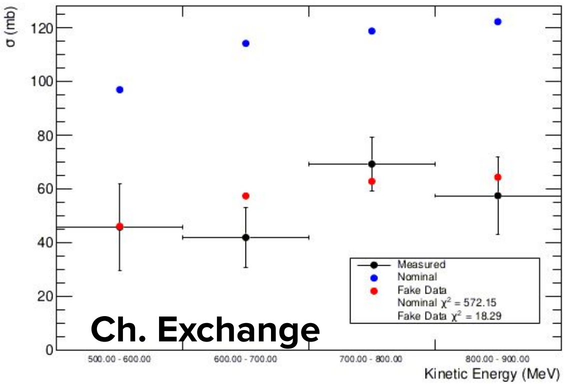
Pre-fit $-2\ln\lambda_{Stat}$	1617.04
Post-fit $-2\ln\lambda_{Stat}$	9.00
Post-fit $-2\ln\lambda_{Syst}$	0.43

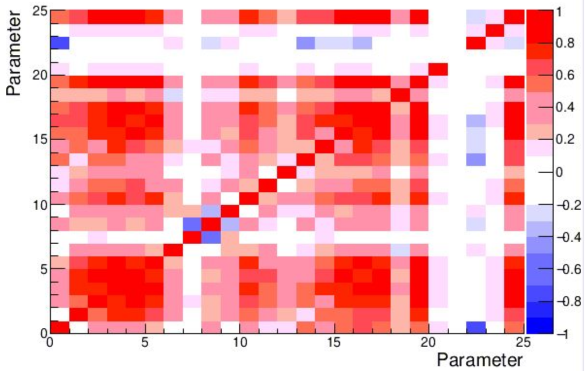
Geant4Reweight Fake Data 1

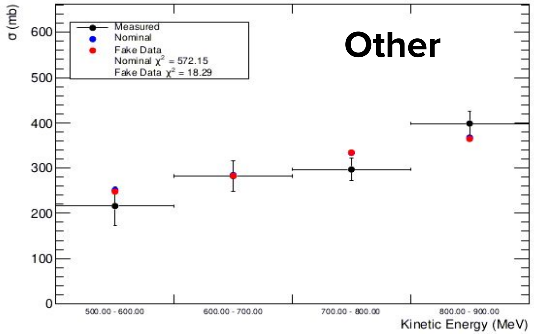
Fake Data p-value	0.72
Nominal p-value	0.00

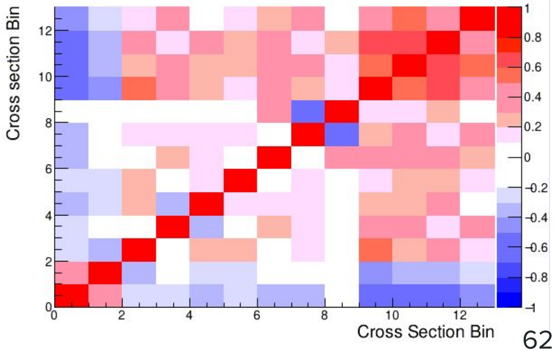




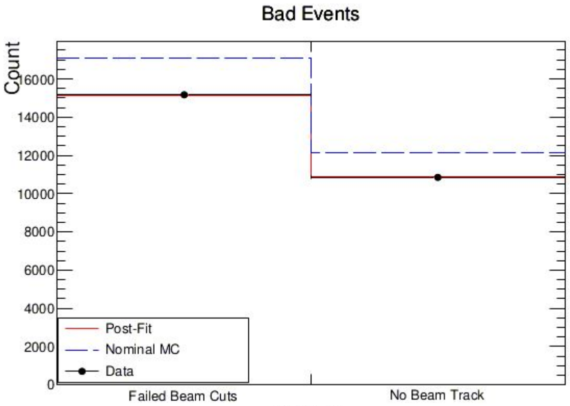


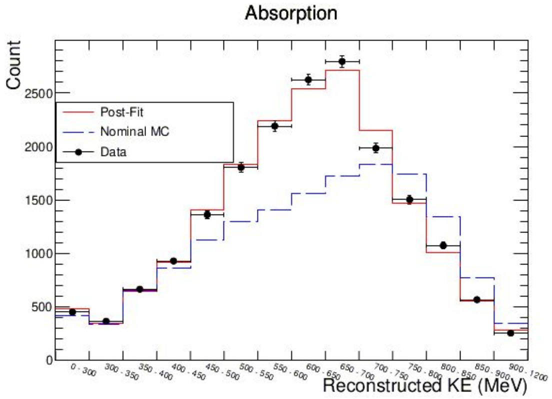


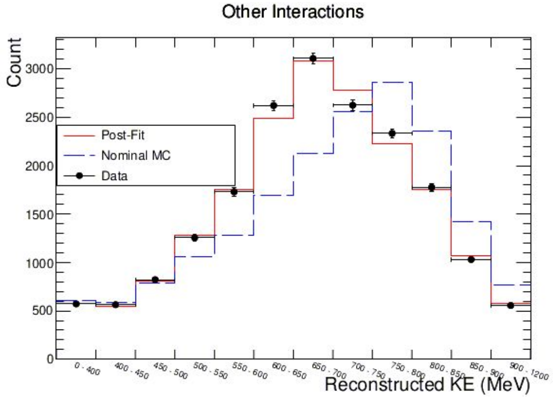


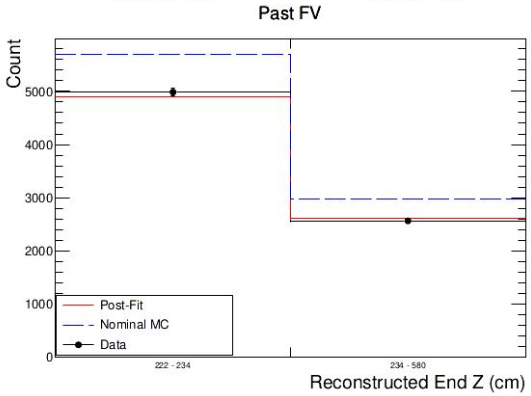


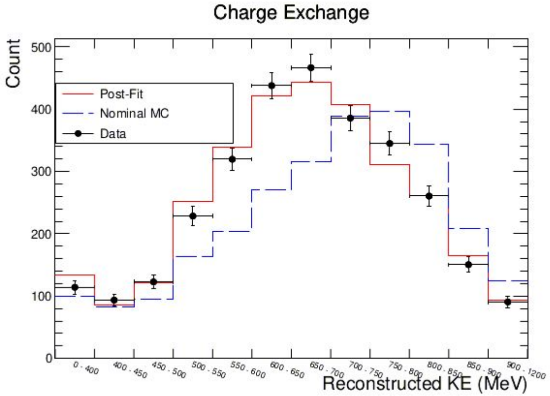
Geant4Reweight Fake Data 2







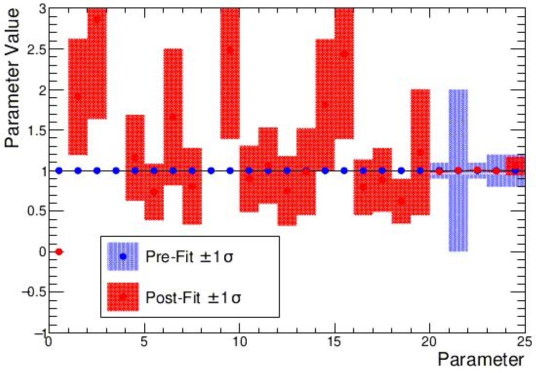


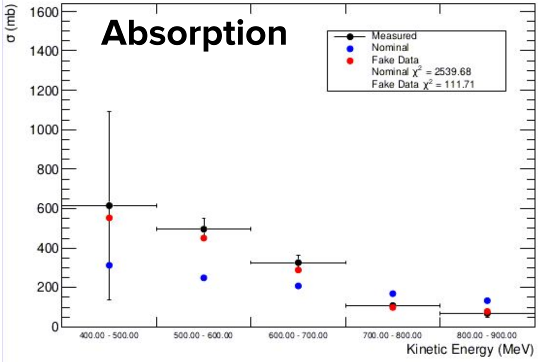


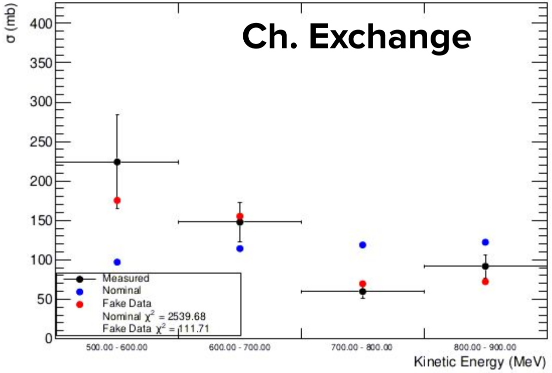
Pre-fit $-2\ln\lambda_{Stat}$	4299.83
Post-fit $-2\ln\lambda_{Stat}$	85.21
Post-fit $-2\ln\lambda_{Syst}$	0.10

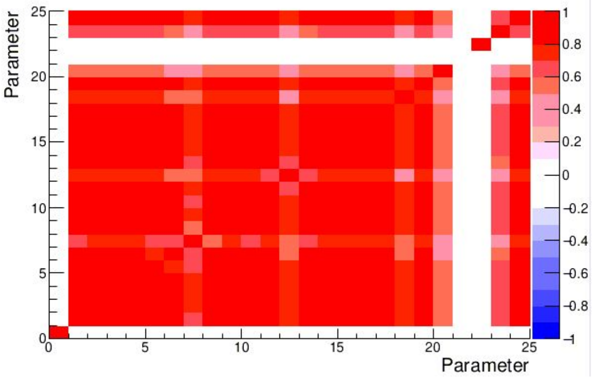
Geant4Reweight Fake Data 2

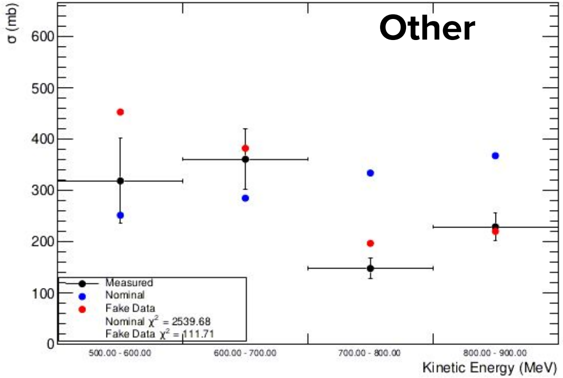
Fake Data p-value	0.07
Nominal p-value	0.00

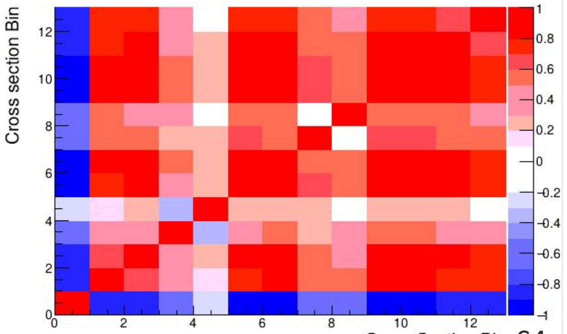












Geant4Reweight Fake Data 2 Discussion

Parameterization can not fit the variation applied

- Results in a bad fit p-value

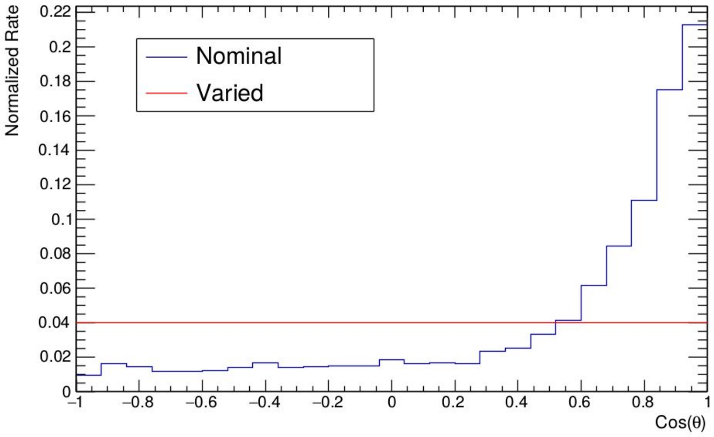
Example of how a bad data fit can be identified

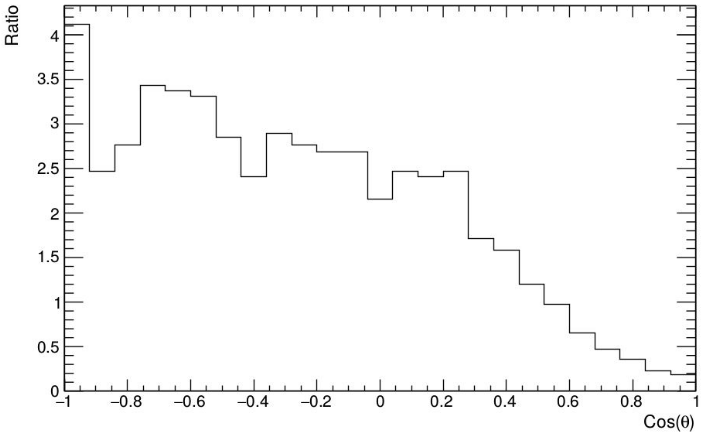
Angular Variation Fake Data

Angular Variation Fake Data

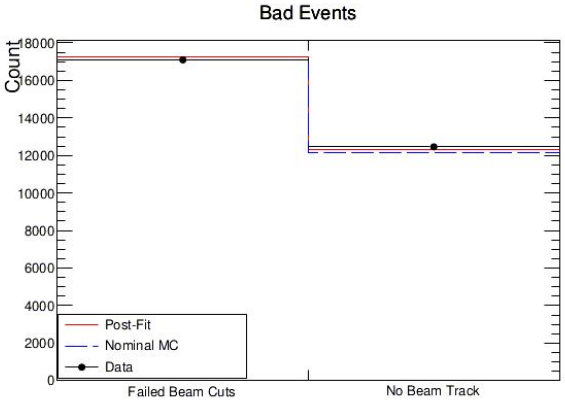
Create fake data by varying the outgoing angle of leading-momentum pions resulting from primary pion interactions

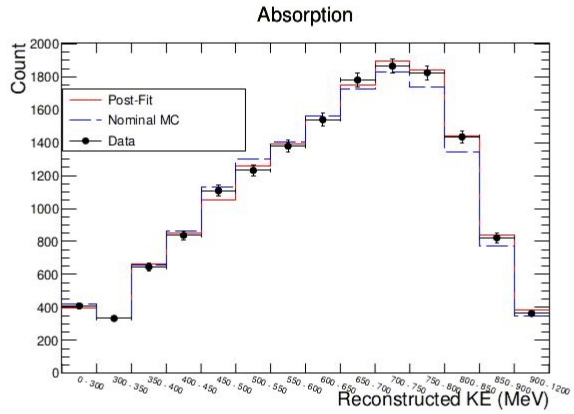
Create distribution by hand (e.g. flattened distribution), use ratio as event weights

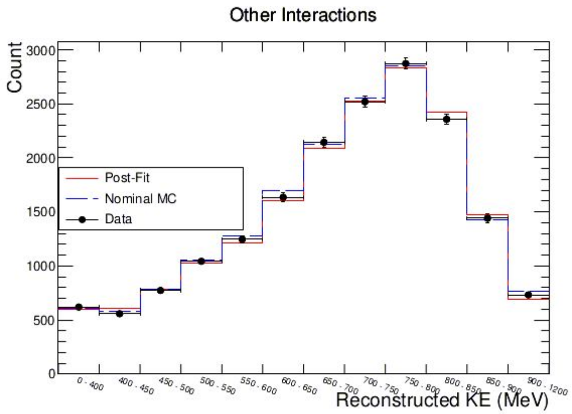


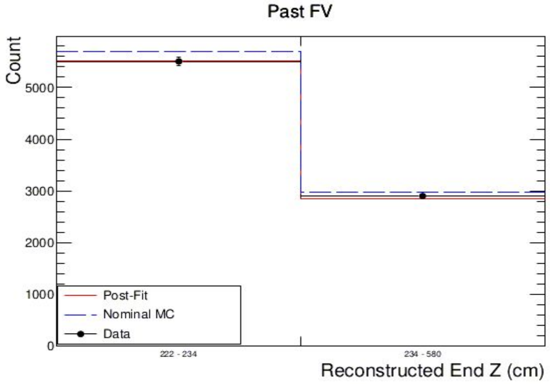


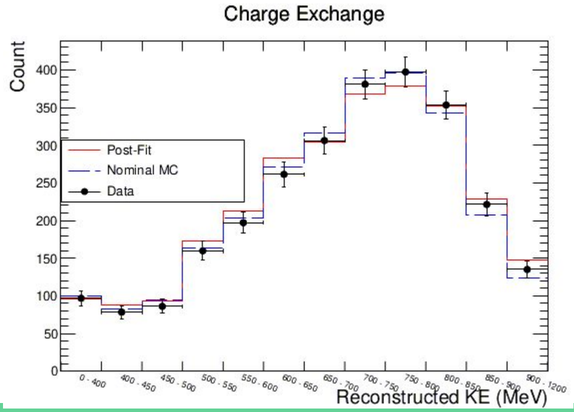
Angular Variation Fake Data







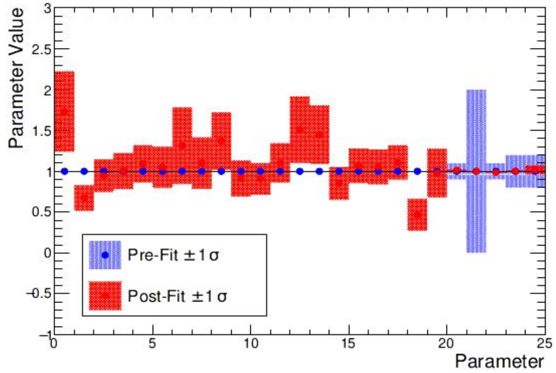


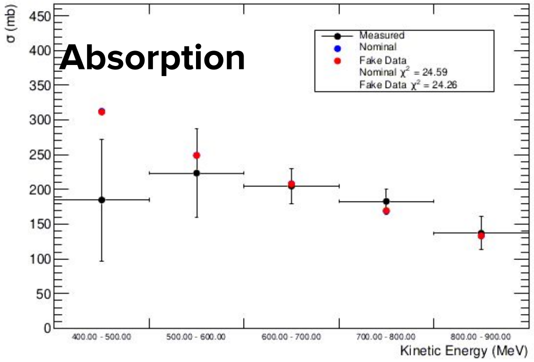


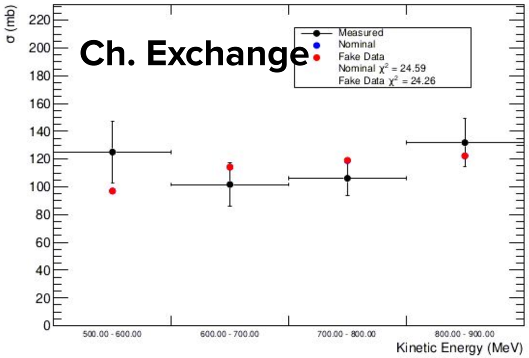
Pre-fit $-2\ln\lambda_{Stat}$	108.00
Post-fit $-2\ln\lambda_{Stat}$	27.31
Post-fit $-2\ln\lambda_{Syst}$	0.04

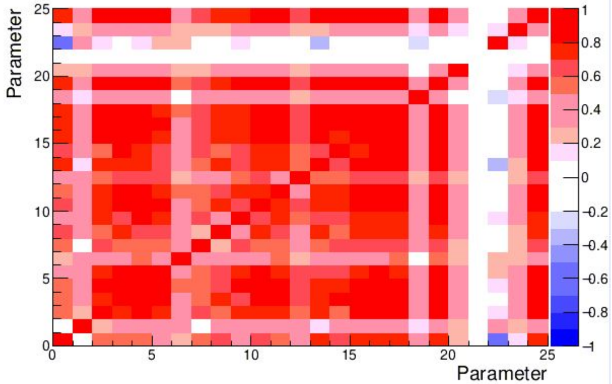
Angular Variation Fake Data

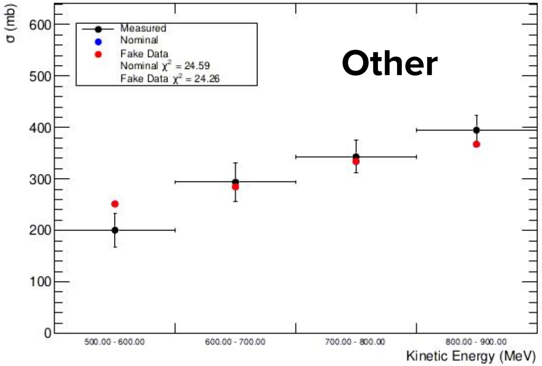
Fake Data p-value	0.60
Nominal p-value	0.60

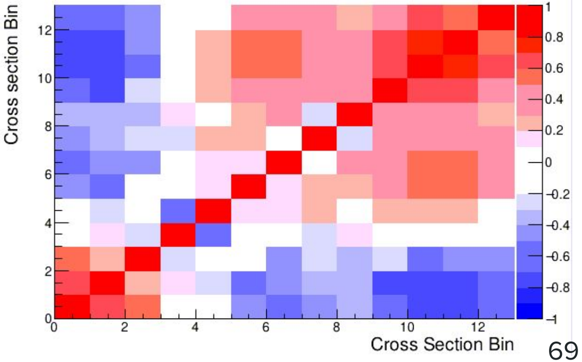










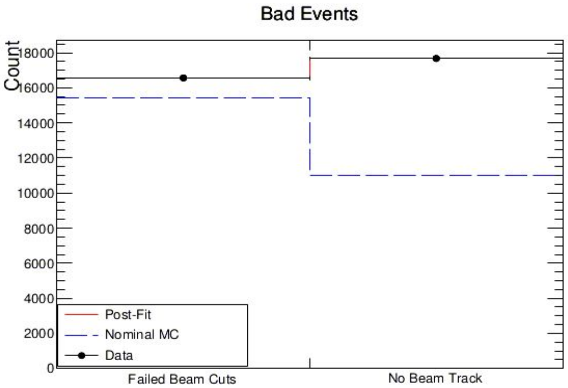


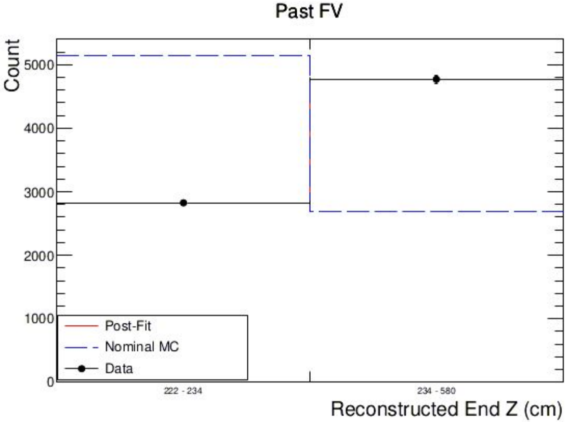
Angular Variation Fake Data Discussion

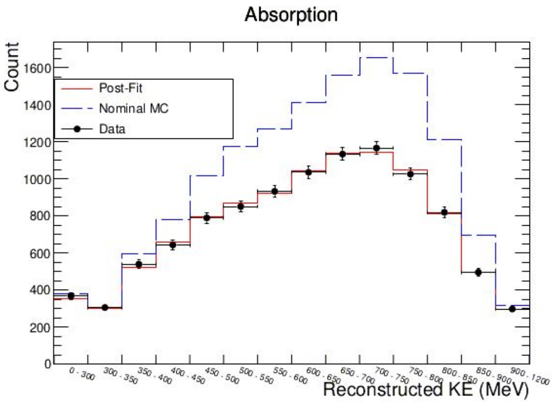
Successful fit shows robustness against mismodeling of outgoing pion kinematics

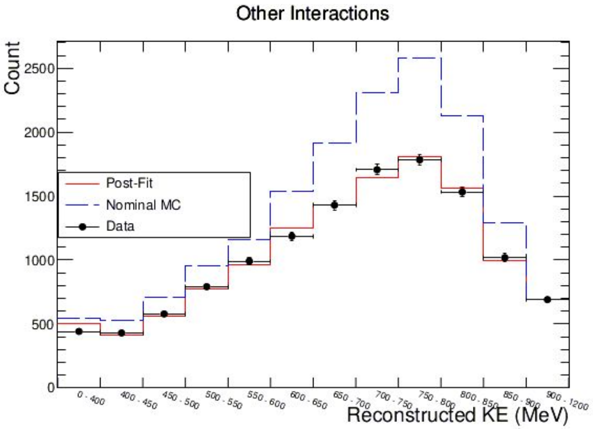
Results on 1 GeV Data

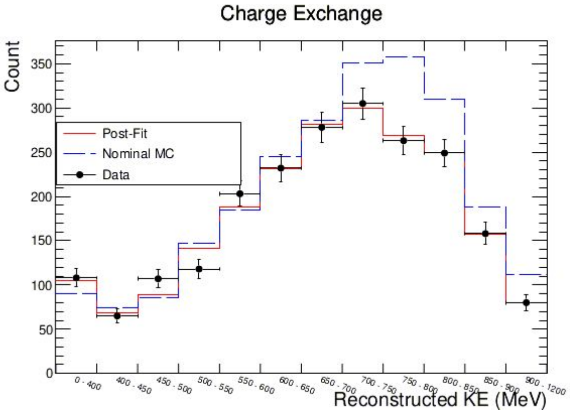
Fit to Data







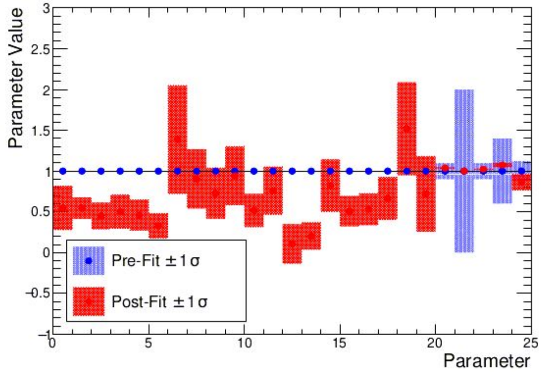


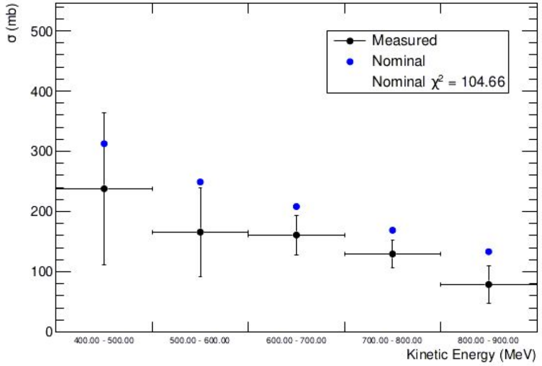


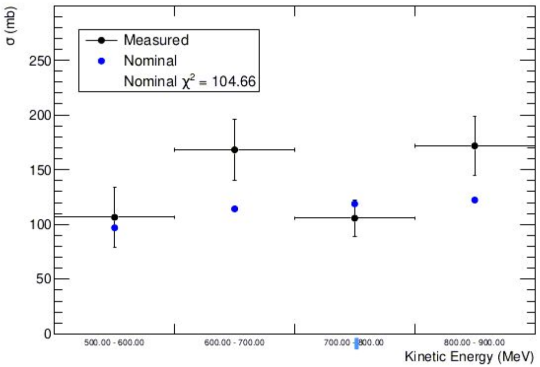
Pre-fit $-2\ln\lambda_{Stat}$	8293.26
Post-fit $-2\ln\lambda_{Stat}$	29.27
Post-fit $-2\ln\lambda_{Syst}$	1.46

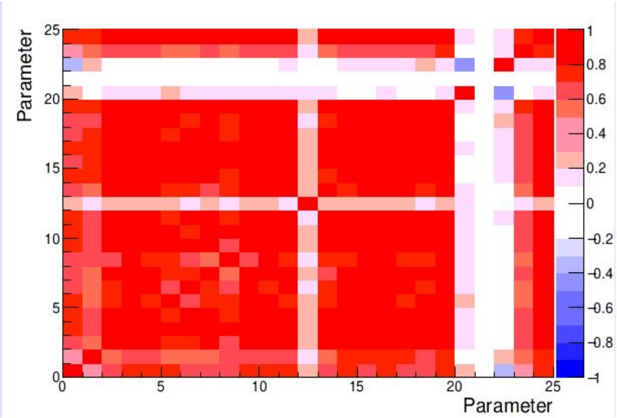
Nominal p-value	0.08

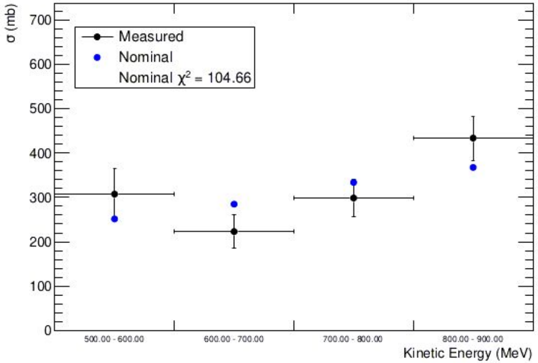
Fit to Data

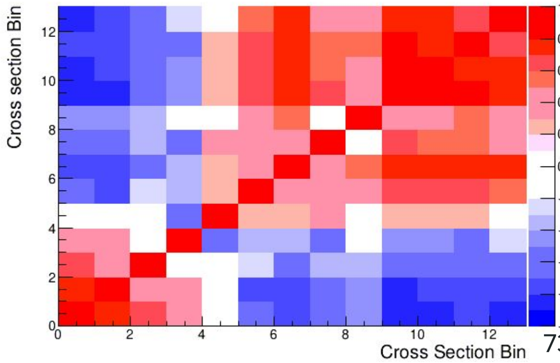






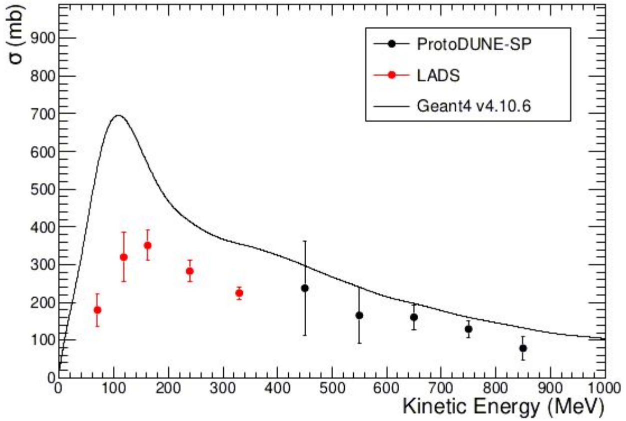




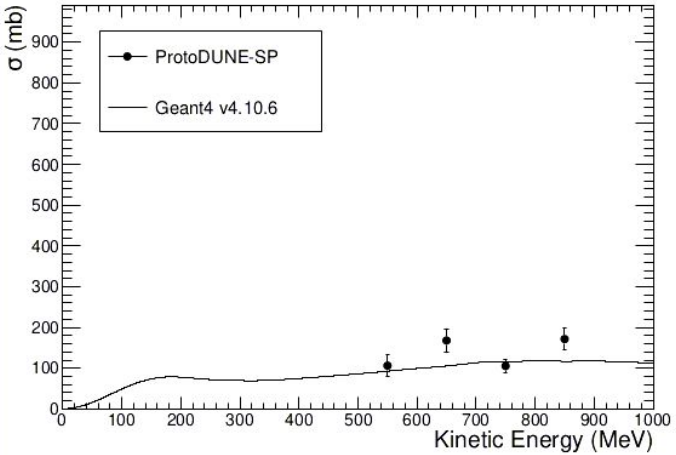


Fit to Data

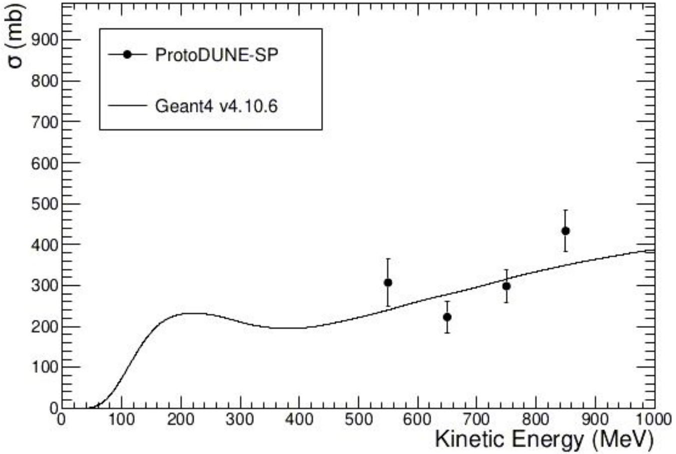
Absorption



Ch. Exchange



Other



Summary

Presented end-to-end pion cross section analysis

Showed current, preliminary results fitting to 1 GeV/c data

Future work

- Implementing SCE systematic
- Understanding underlying issues behind Pandora's beam inefficiency (see backup)