Pion Cross Section Updates

Jake Calcutt
February 10, 2022

Introduction

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

- Slightly changed the parameterization to increase fit stability
- New selection category using Michel-like CNN score of hits near vertex
- Slightly updating the cuts used to identify π^0-like photon showers
- SCE systematic uncertainty from fit to alternate SCE map sample
- Testing whether a proton cross section systematic parameter is needed
  - Fit to fake data with proton cross section variation
  - Ran toy tests with proton cross section parameter
- Discuss removing the topological-based signal definitions

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)

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. dE/dX Calibration
22. Beam Resolution
23. Electron Diverter Effect
24. Pandora Tracking Efficiency
25. Beam Cut Efficiency

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

Had over/underflow bins to cover any discrepancies outside of signal regions

Too much freedom in fit, these would sometimes be pushed to near 0. within toy studies

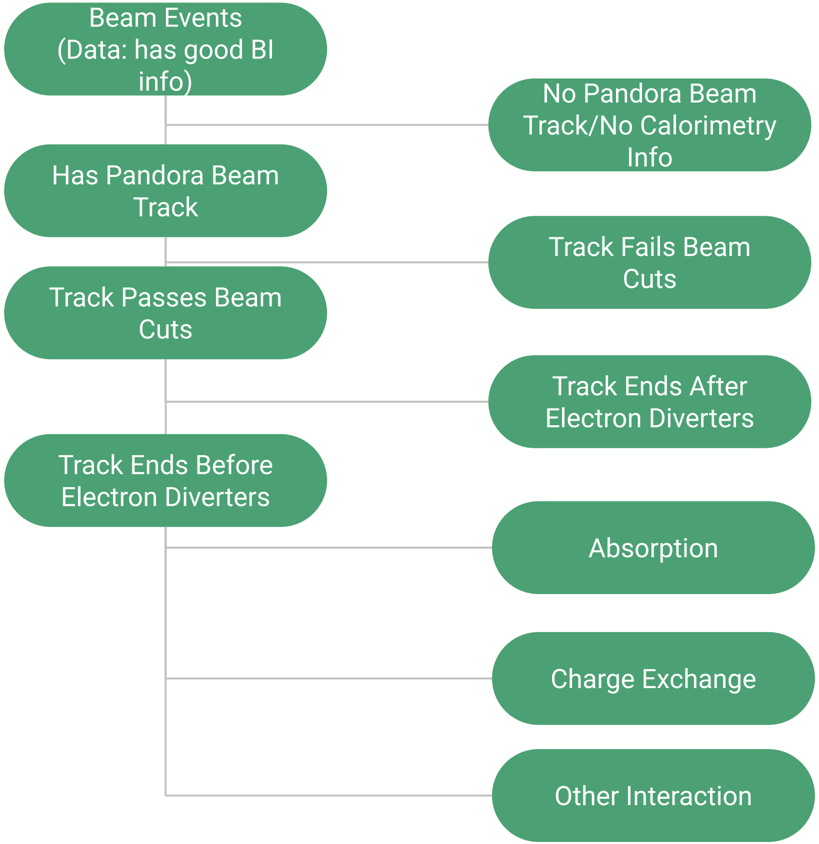
Added ability to 'tie' the over and underflow components of the signal categories

‘Tied’ Parameterization

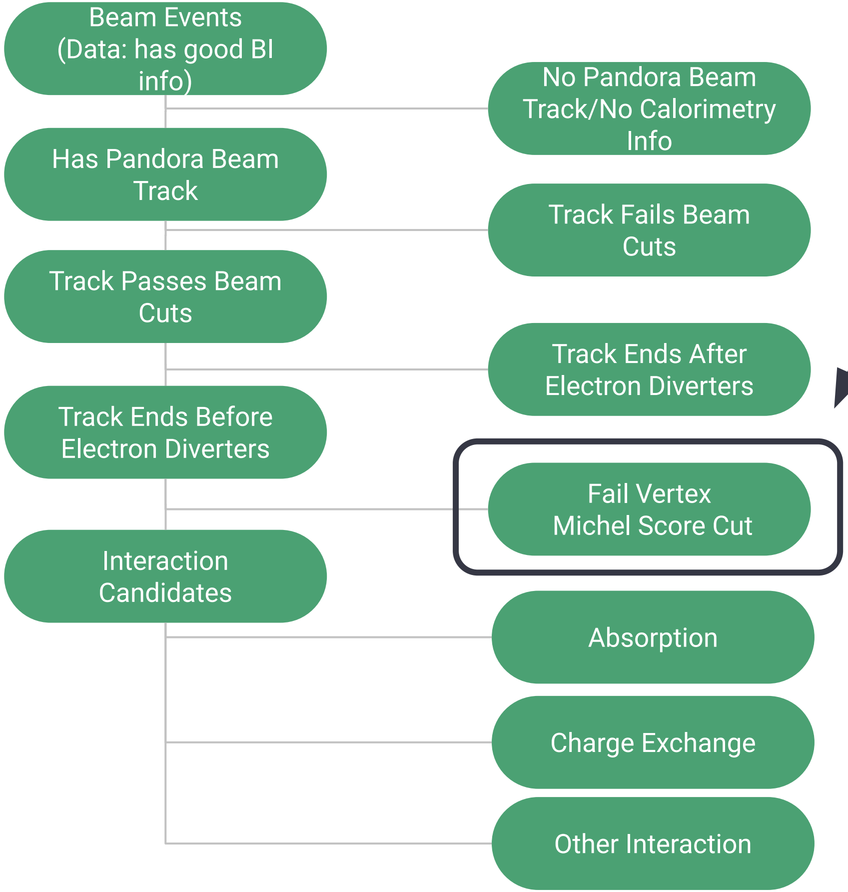
1. Absorption $\notin$ [400, 900] MeV
2. Abs. 400 - 500 MeV
3. Abs. 500 - 600 MeV
4. Abs. 600 - 700 MeV
5. Abs. 700 - 800 MeV
6. Abs. 800 - 900 MeV
7. Charge Exchange $\notin$ [500, 900] MeV
8. Ch. Exch. 500 - 600 MeV
9. Ch. Exch. 600 - 700 MeV
10. Ch. Exch. 700 - 800 MeV
11. Ch. Exch. 800 - 900 MeV

14. Other Inelastic $\notin$ [500, 900] 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
20. Muon Fraction
21. dE/dX Calibration
22. Beam Resolution
23. Electron Diverter Effect
24. Pandora Tracking Efficiency
25. Beam Cut Efficiency

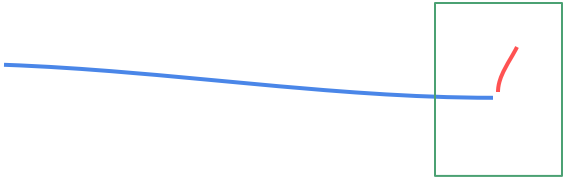
Event Selection - Updated



Event Selection - Updated

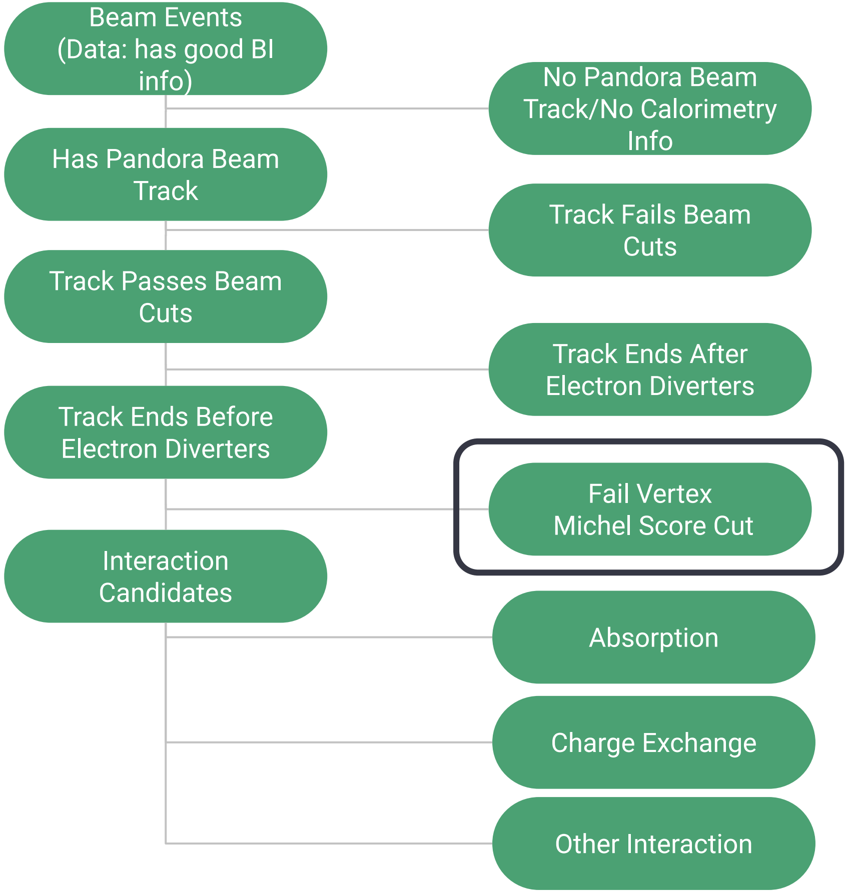


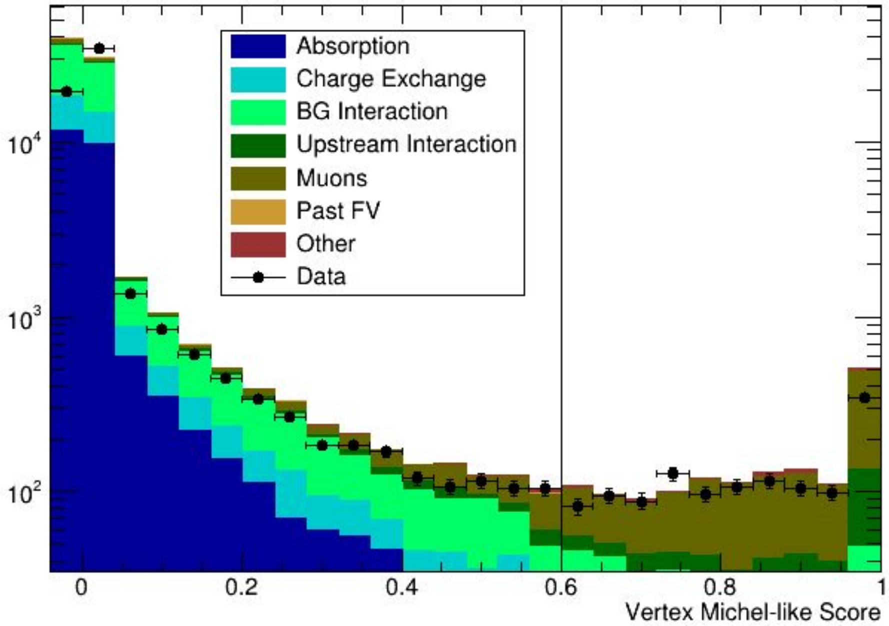
Adding another category to cut out muons & stopping pions from the interaction candidates



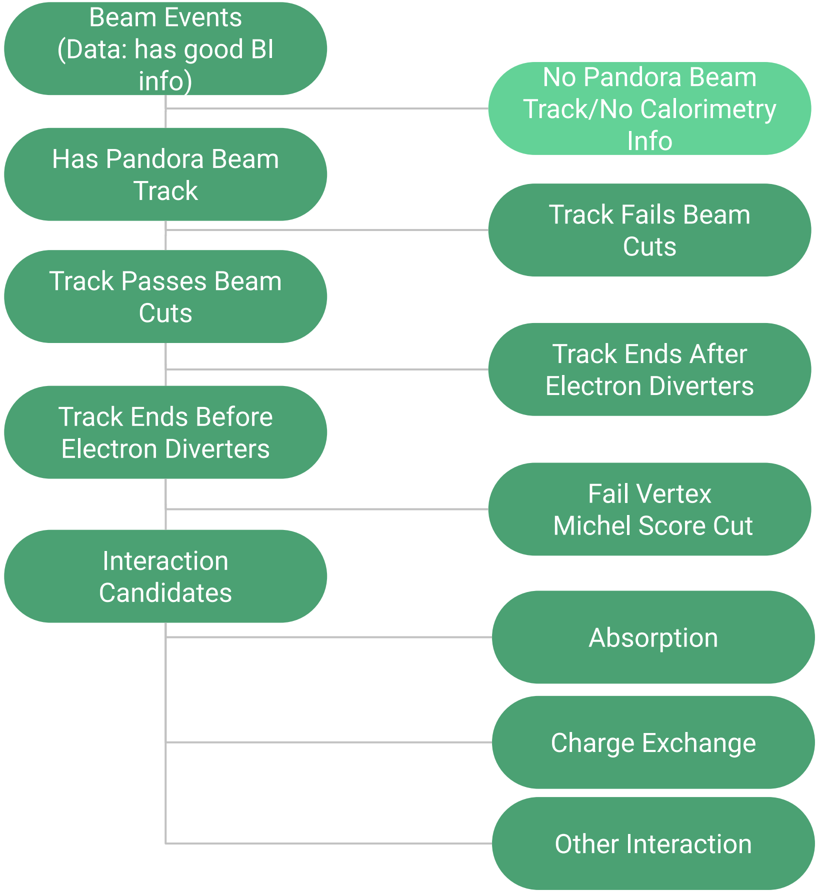
Look for any hits within a window near the end of the primary track, average their Michel-like CNN score

Event Selection - Updated



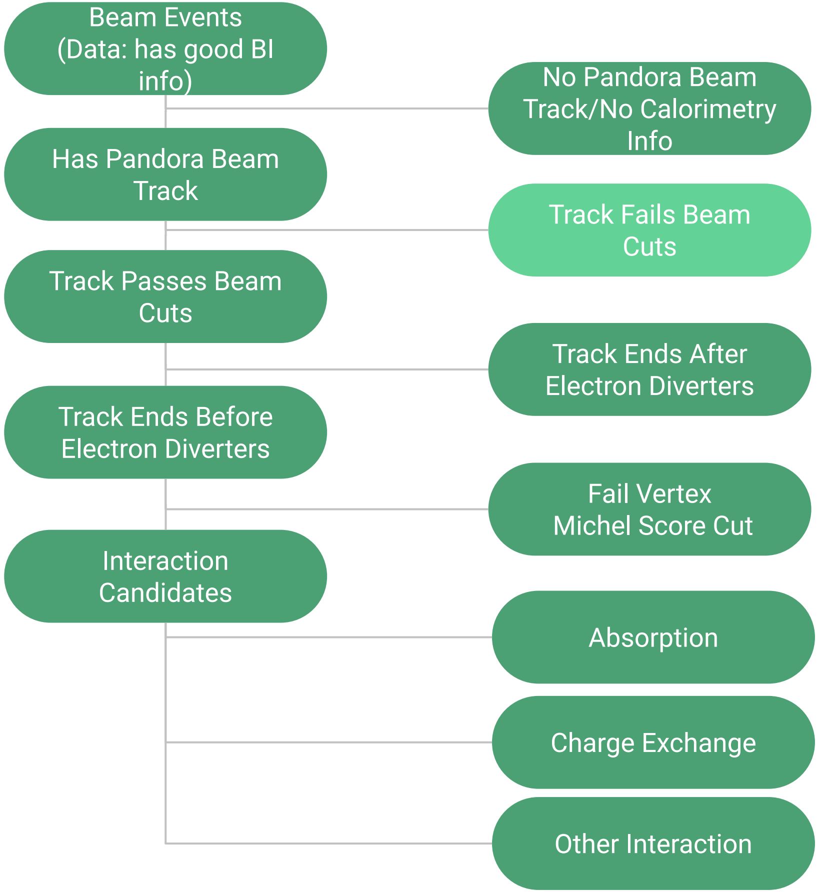


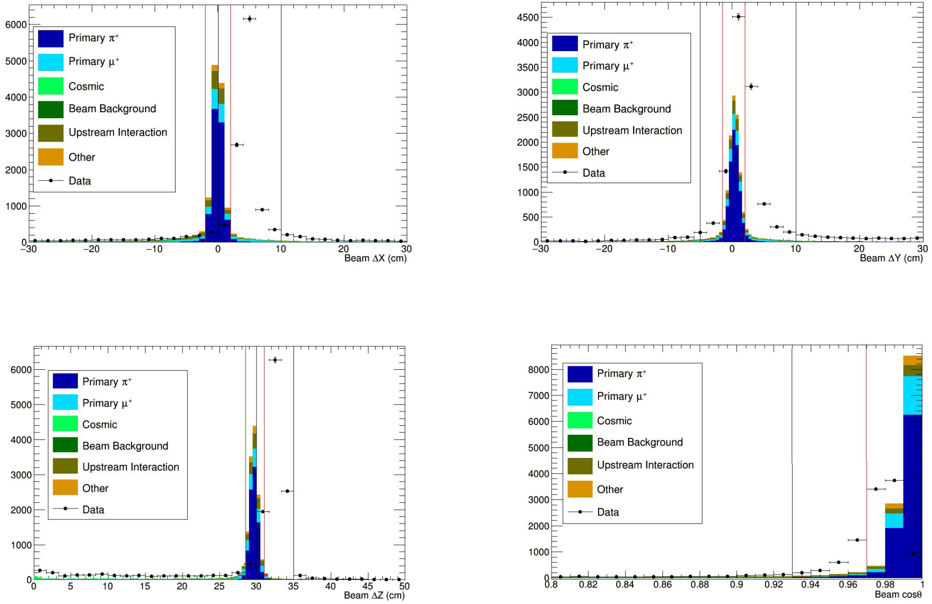
Event Selection



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

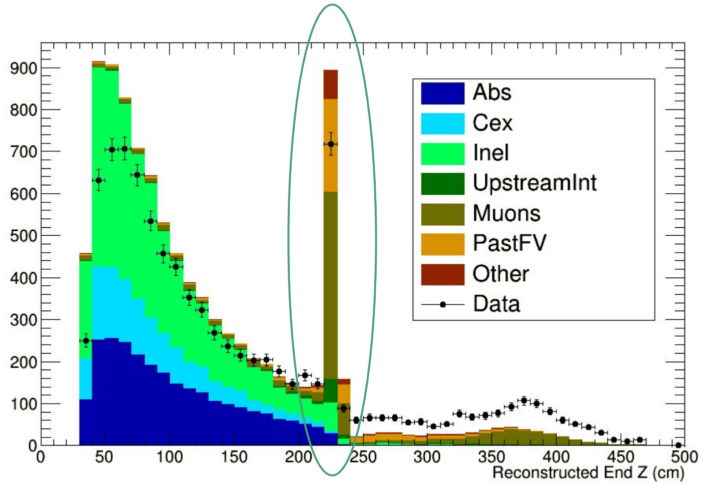
Event Selection

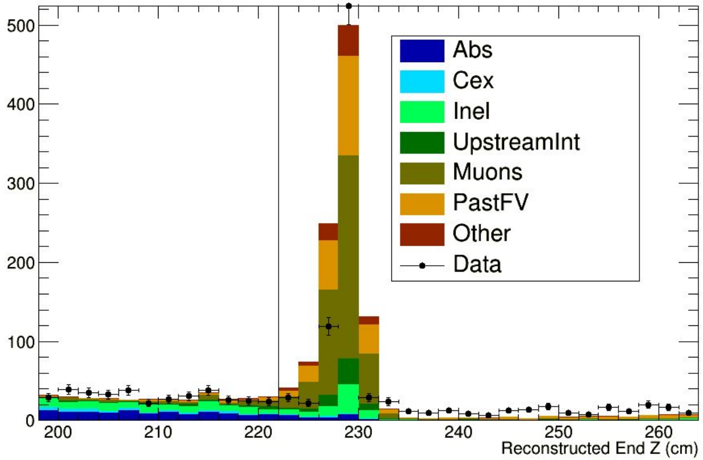




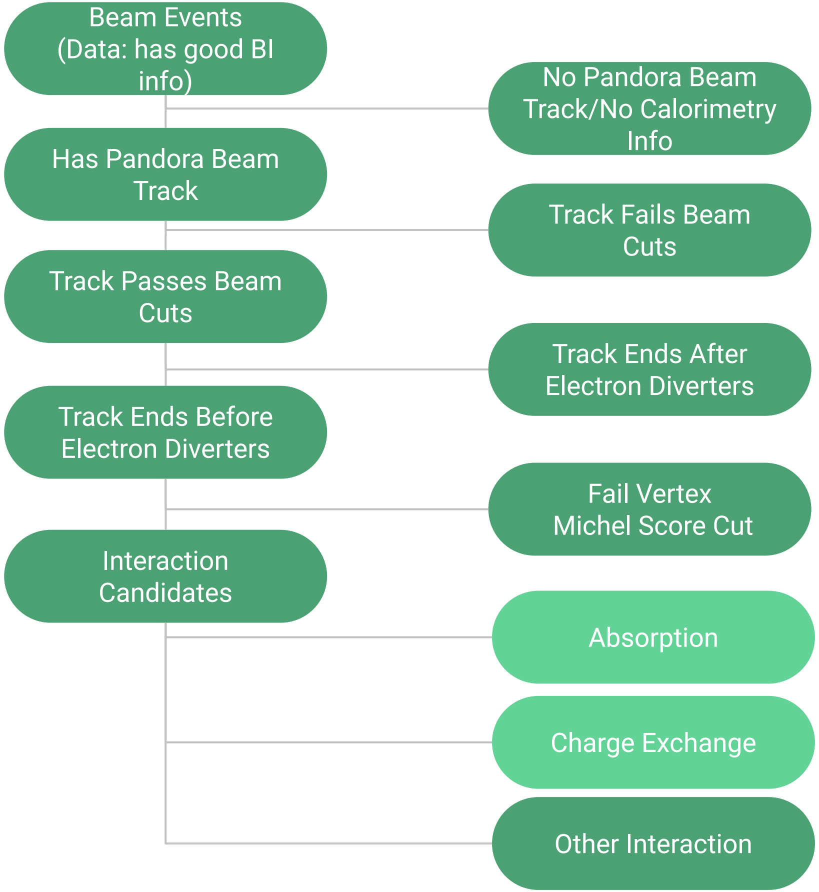
Event Selection

- 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



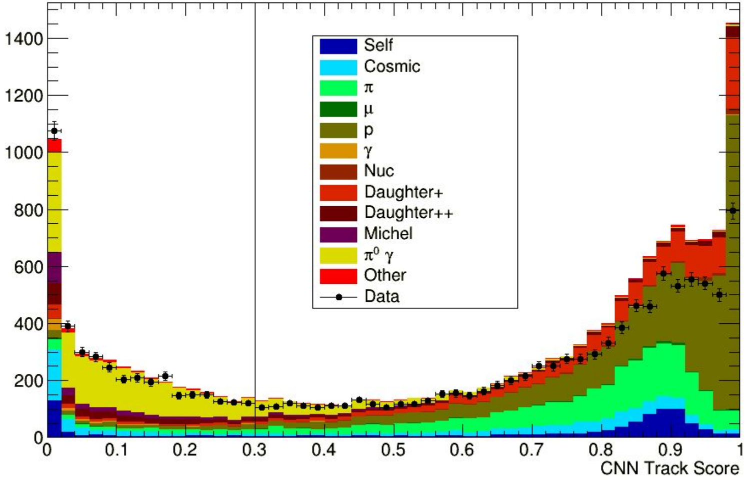


Event Selection



Identify track-like daughters using aggregate CNN scores of particles

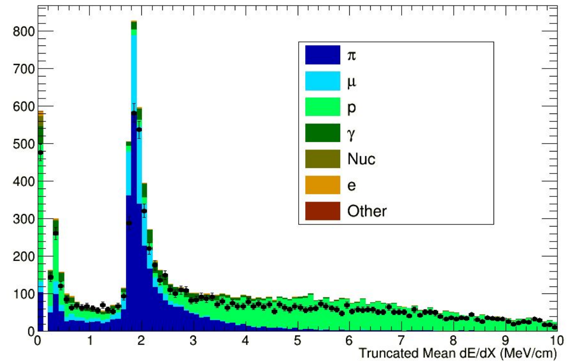
>.3 ➔ Track-like

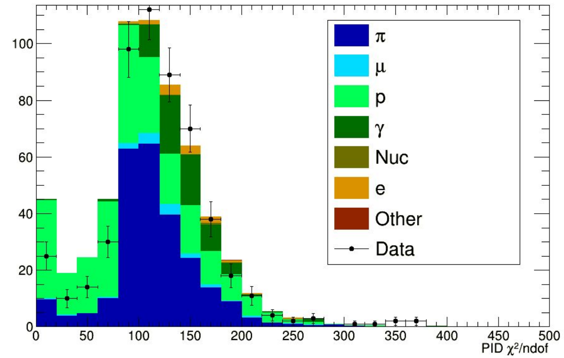


Event Selection

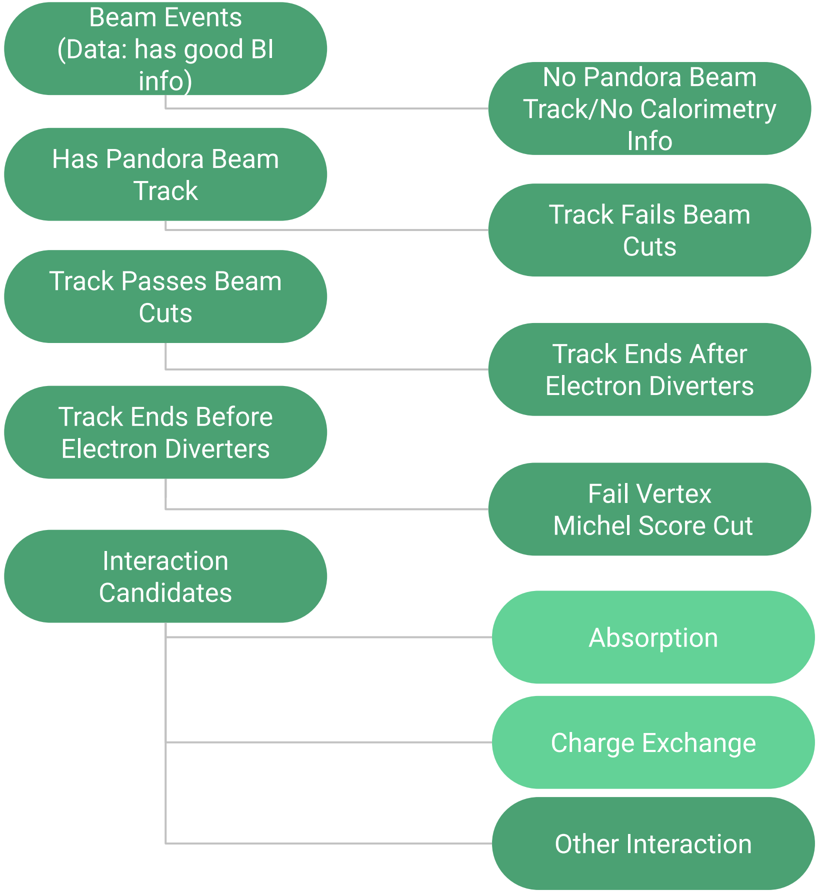
Use calorimetry information to identify charged pions within tracks

- 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

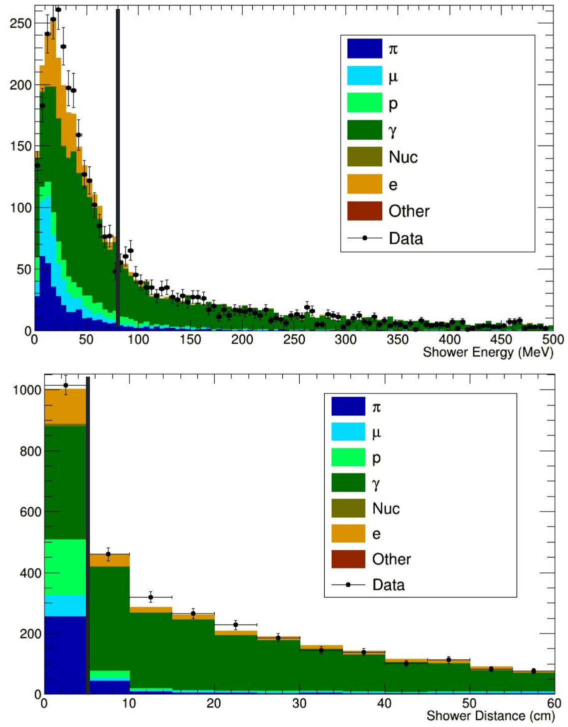




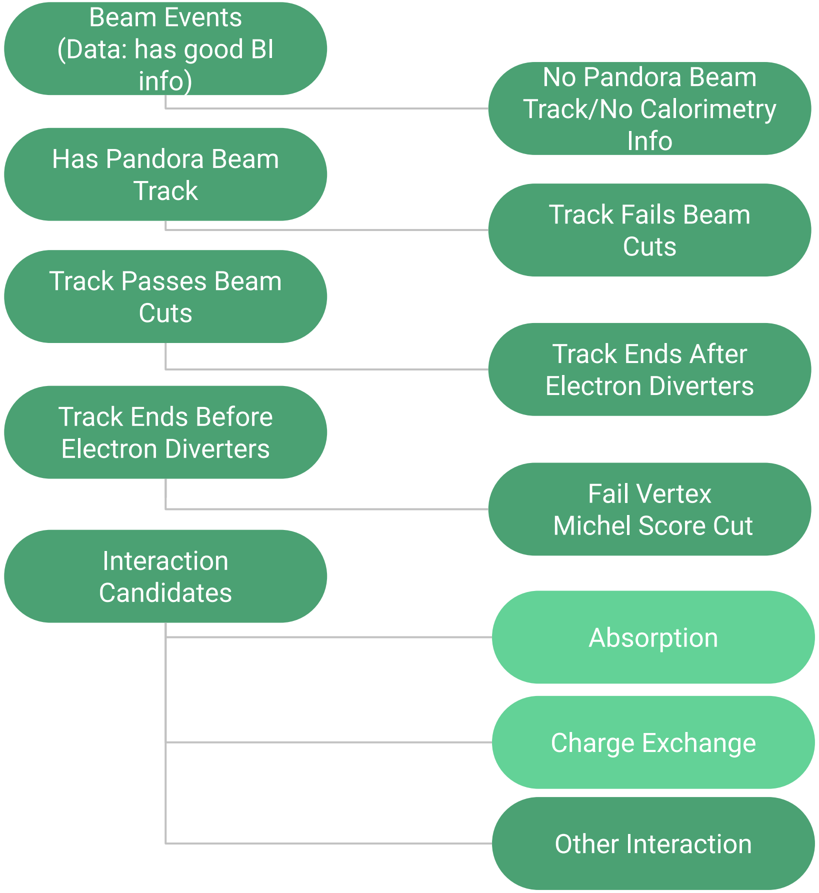
Event Selection



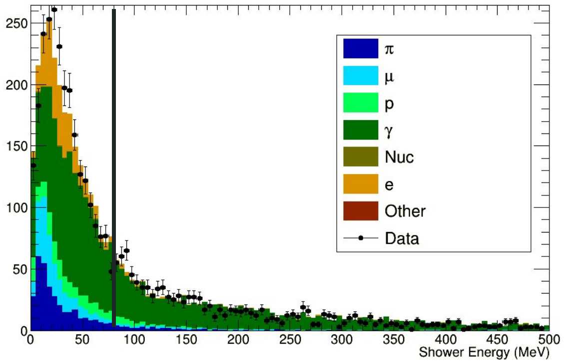
Was previously using two 1D cuts on energy and distance-to-vertex of shower-like reco daughter particles

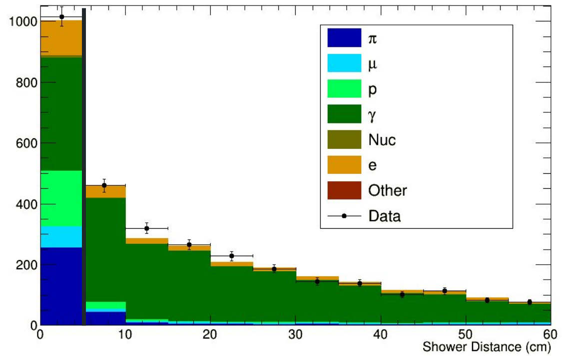


Event Selection

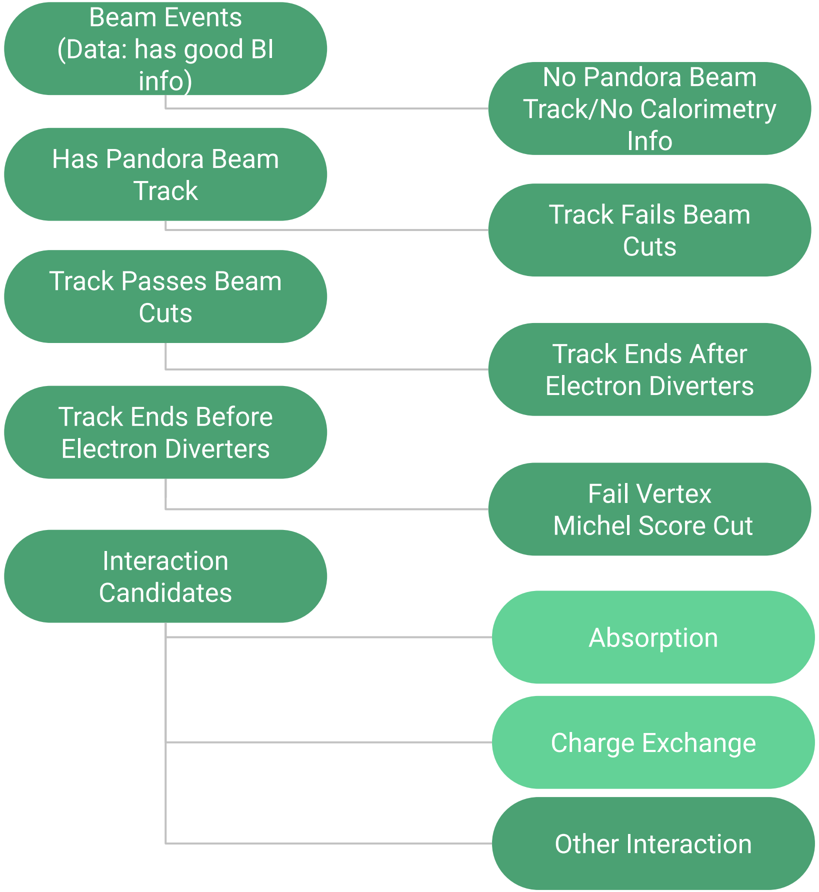


Realized a set of 2D cut would be better





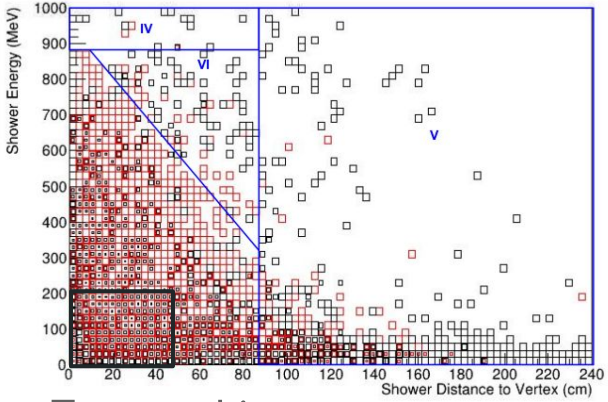
Event Selection



In this:

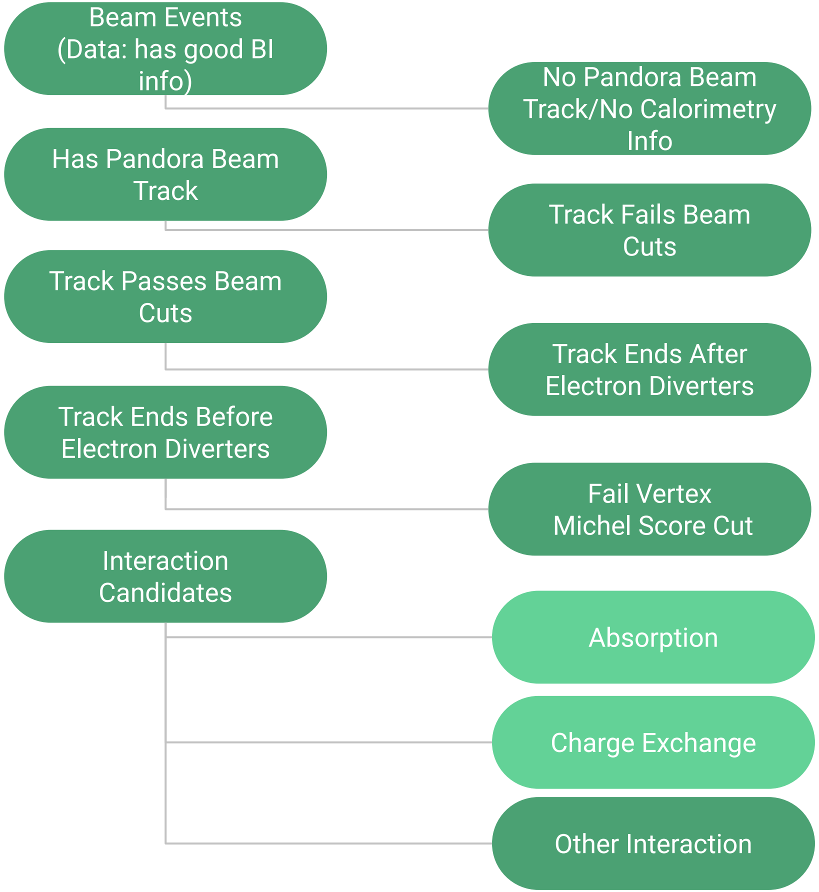
- size of square = fraction of particles in bin
- Red: True π^0 (signal)
- Black: Other (background)

Cut out areas where (generally) black > red



Next slide → Zoomed in

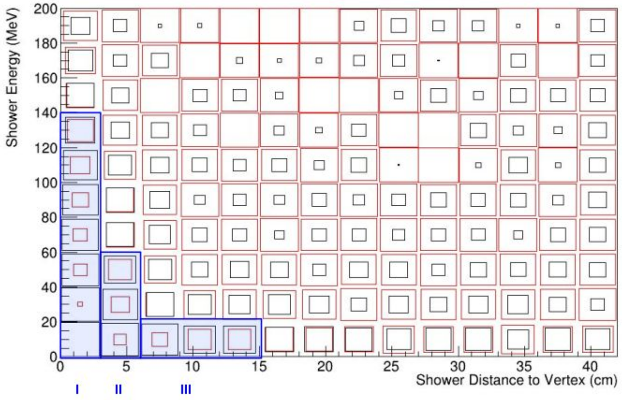
Event Selection



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



Testing Proton Cross Section Variation

Originally planned for proton cross section systematic parameter

Wanted to test whether it's necessary to have

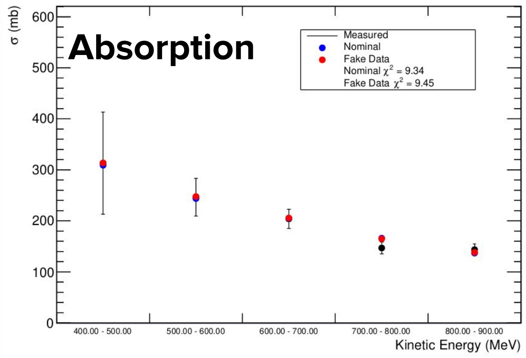
Next slides: presenting two fake data studies

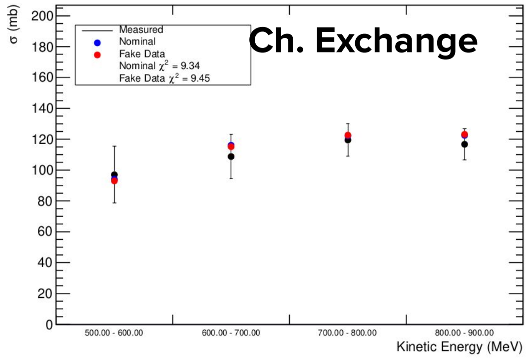
1. Increased proton cross section by 60%
2. Decreased proton cross section by 40%

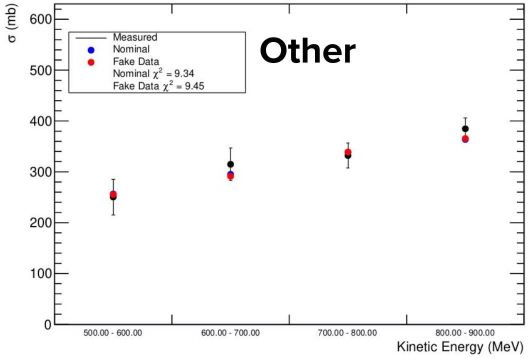
Fit to both with nominal MC

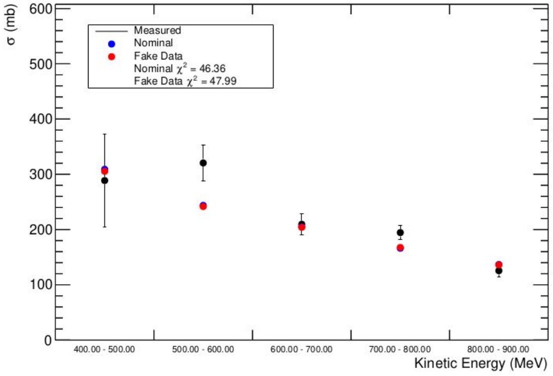
Top: proton xsec raised
Bottom: lowered

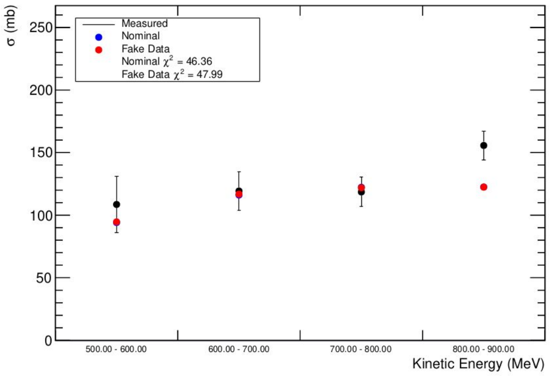
Proton Variation Results

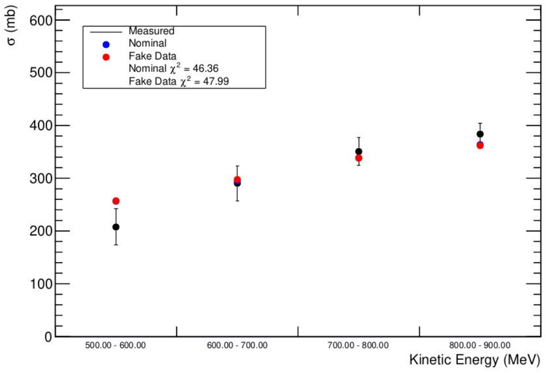












Proton Cross Section Variation – Conclusions

Changing the total proton cross section can possibly affect the results

Have reformatted an old parameter to better perform the weighting using geant4reweight variations calculated during ntuple production

- Because you have to choose specific geant4reweight variations at ntuple production runtime, the variation surface for each event is not smooth
  - Previously: at fit-time create splines for broad categories of events across variations, use these for weighting events
  - Now: the g4rw variations for each event are fit during ntuple production to produce a polynomial describing the weighting ➔ true event-by-event reweighting

Proton Cross Section Systematic

Preliminary results for toy systematic tests are promising

Cross section parameter seems stable

- Previous implementation would often fail to approach toy value – sometimes reaching parameter limits

Will include the new proton systematic in fit to data

SCE Systematic Uncertainty

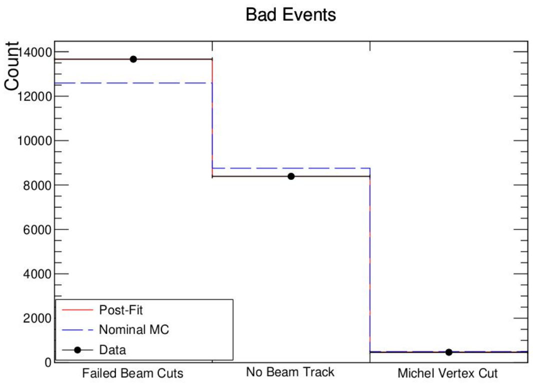
Production team created MC samples with alternate space charge maps used for forward distortions but nominal maps for corrections

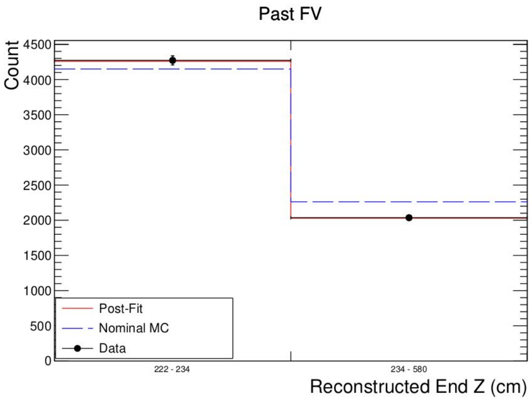
Approximates what's occurring in data (MC is idealized – same maps for distortion and corrections)

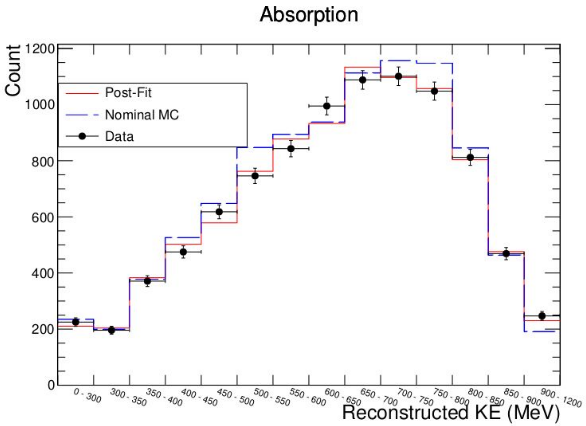
Use alt SCE sample as fake data, fit with nominal MC

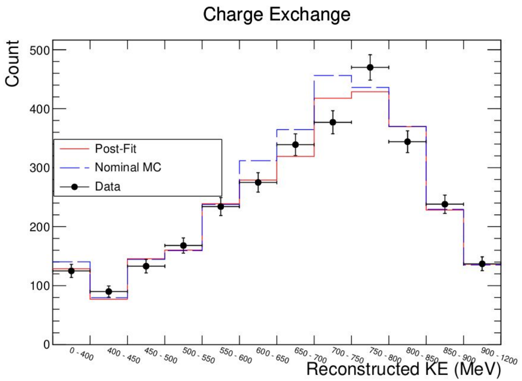
- Squared difference to nominal cross section taken as uncertainty

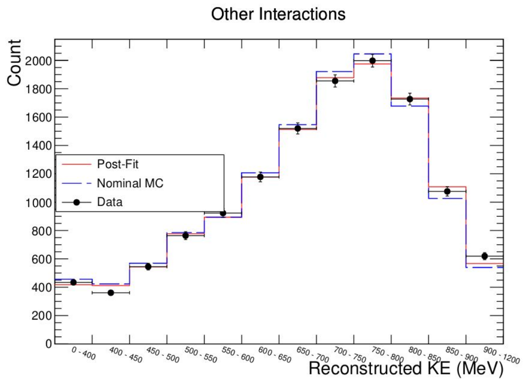
Alt SCE Fit





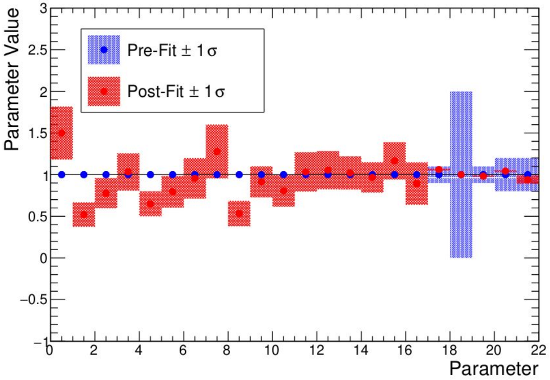




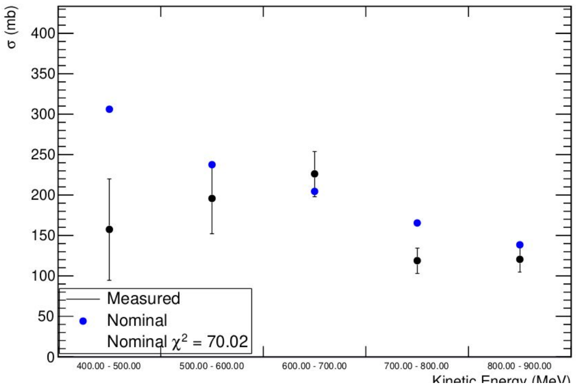


Pre-fit $-2\ln\lambda_{Stat}$	251.34
Post-fit $-2\ln\lambda_{Stat}$	45.19
Post-fit $-2\ln\lambda_{Syst}$	0.53

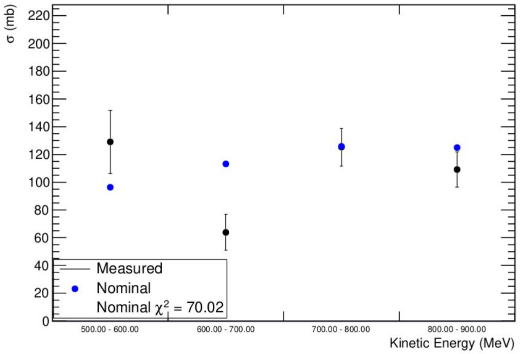
Alt SCE Fit



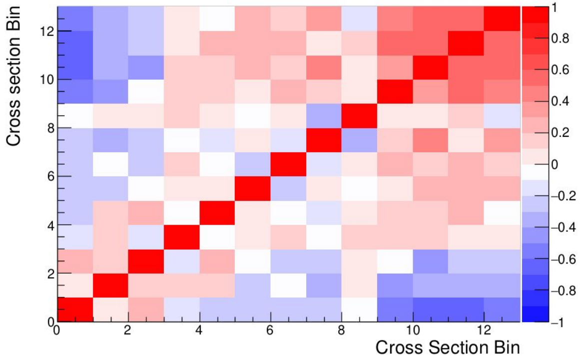
Parameter Value
Pre-Fit ± 1 σ
Post-Fit ± 1 σ
Parameter



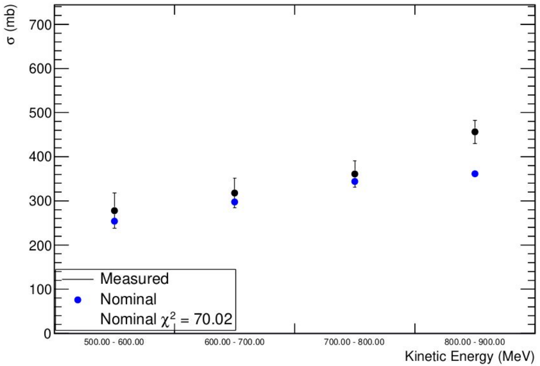
σ (mb)
Measured
Nominal
Nominal χ^2 = 70.02
400.00 - 500.00
500.00 - 600.00
600.00 - 700.00
700.00 - 800.00
800.00 - 900.00



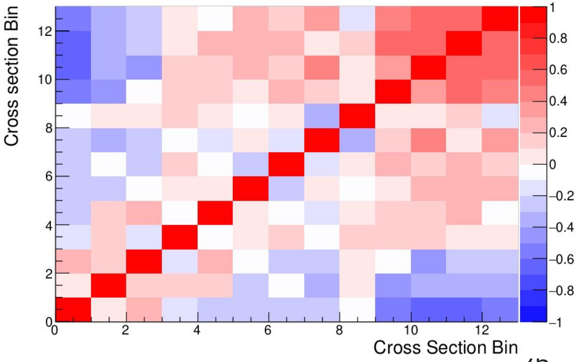
σ (mb)
Measured
Nominal
Nominal χ^2 = 70.02
500.00 - 600.00
600.00 - 700.00
700.00 - 800.00
800.00 - 900.00
Kinetic Energy (MeV)



Cross section Bin
Cross Section Bin

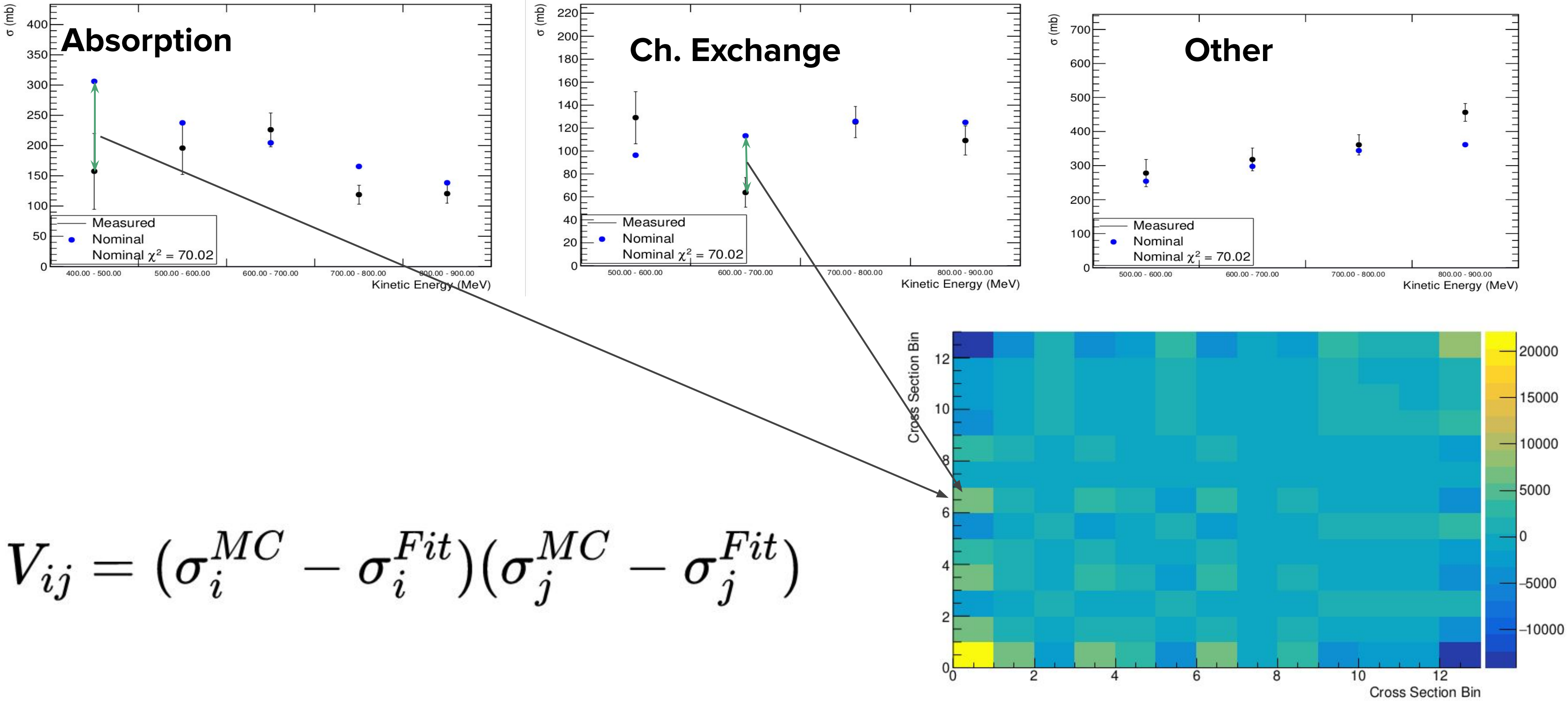


σ (mb)
Measured
Nominal
Nominal χ^2 = 70.02
500.00 - 600.00
600.00 - 700.00
700.00 - 800.00
800.00 - 900.00
Kinetic Energy (MeV)



Cross section Bin
Cross Section Bin

Using Differences as Uncertainty



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

Thank you for listening

Backup Slides

Beam Resolution

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

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

Shift: determine from nominal beam MC

➡ 0.7% uncertainty on p_{Reco}

Add in quadrature

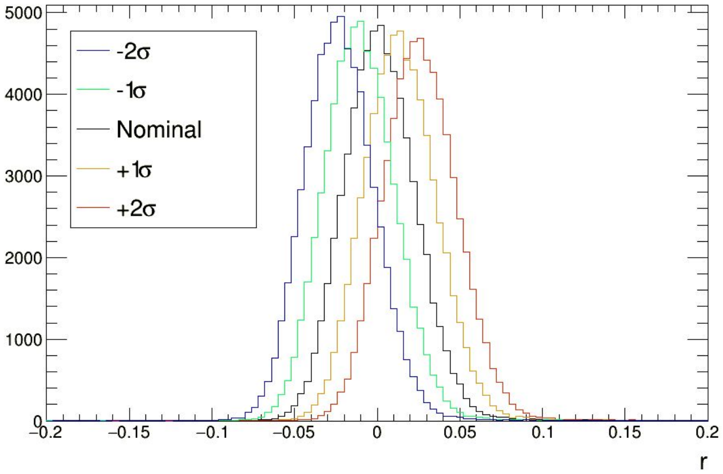
➡ overall 1.2% uncertainty on p_{Reco}

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

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

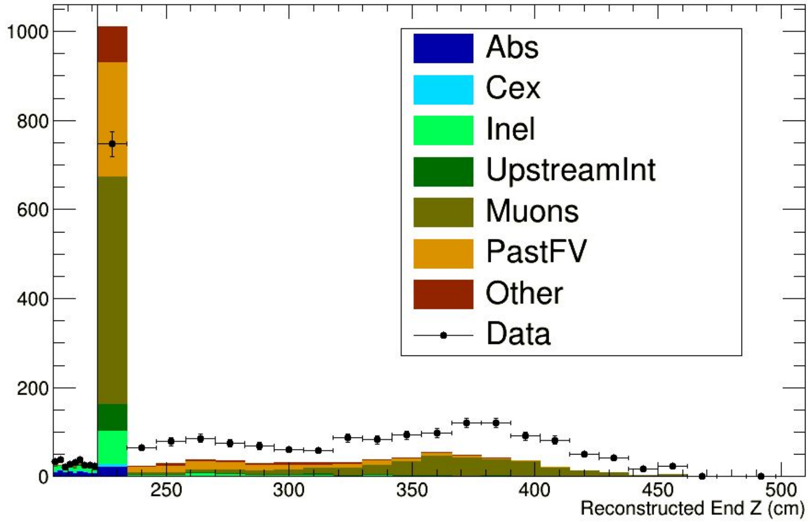


$$w = \left(\frac{\sigma}{\sigma'}\right) \exp\left(\frac{(r-\mu)^2}{2\sigma^2} - \frac{(r-\mu')^2}{2\sigma'^2}\right)$$

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

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

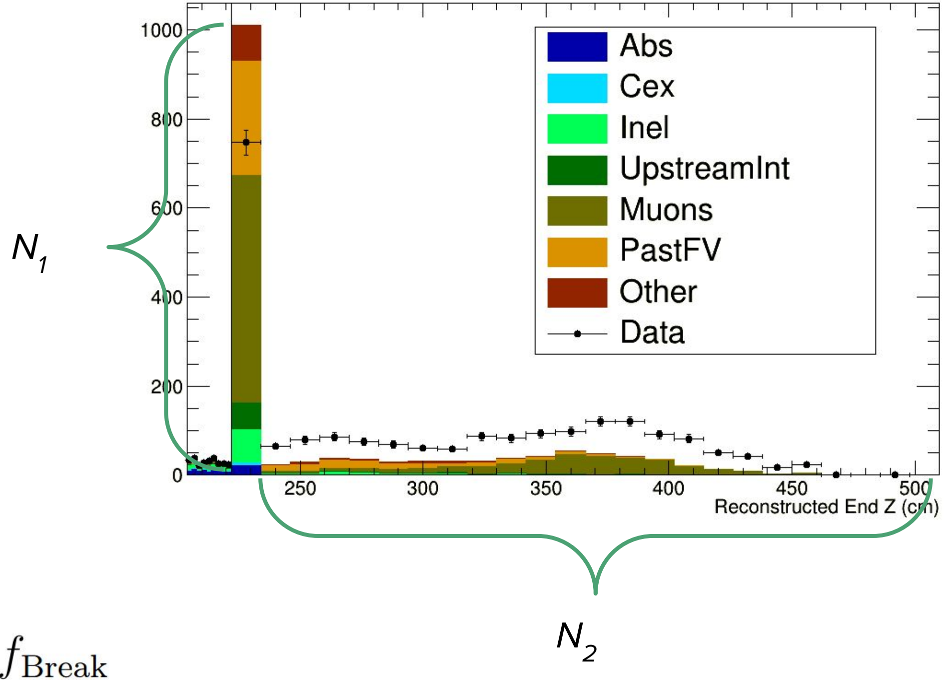
$$f_{\text{Break}} \rightarrow f'_{\text{Break}} = c f_{\text{Break}}$$

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

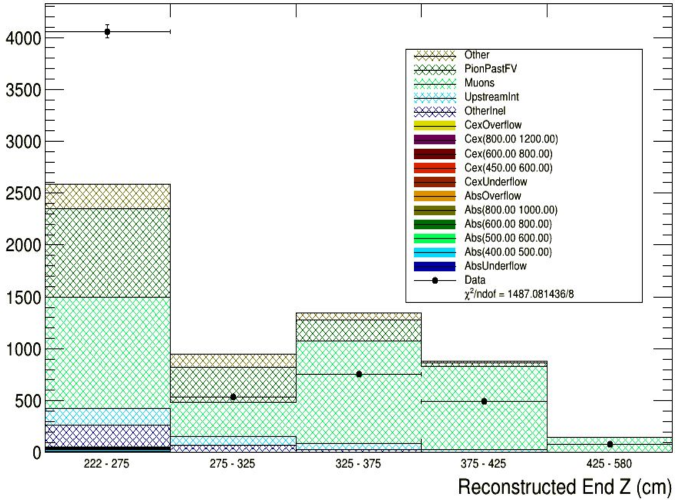
$$W_1 = \frac{f'_{\text{Break}}}{f_{\text{Break}}} = c$$

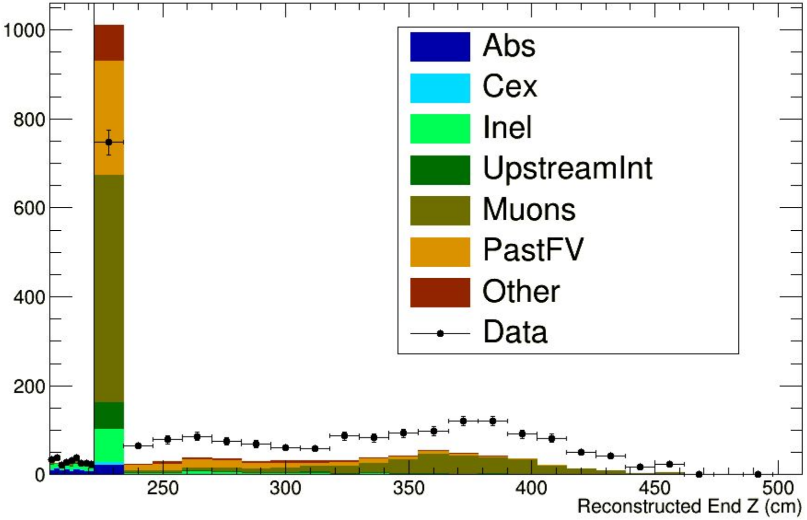
If track ends past break region

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



Electron Diverter Prior Uncertainty





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

Electron Diverter Prior Uncertainty

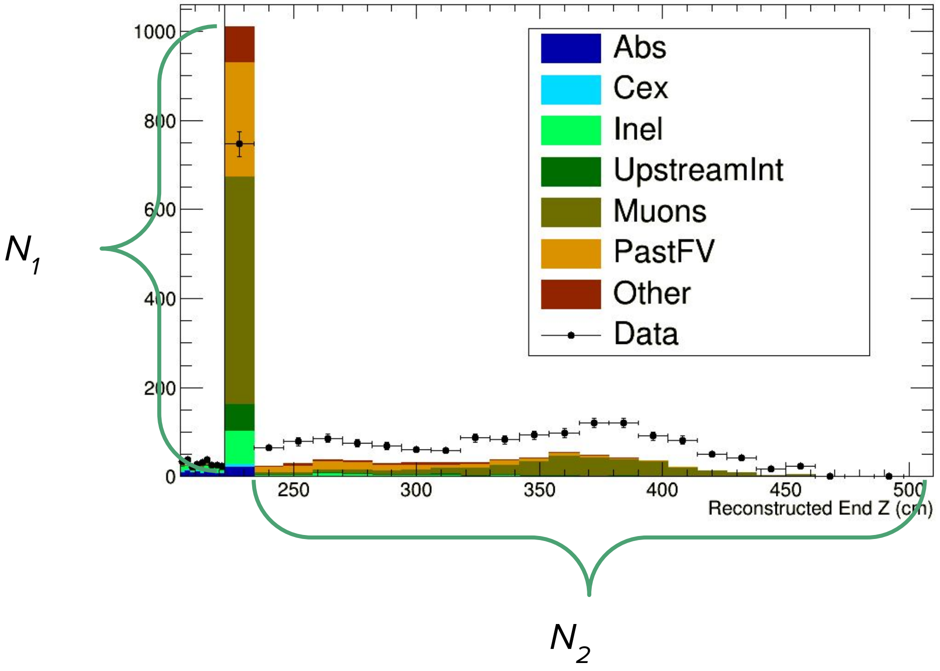
Data:

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

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

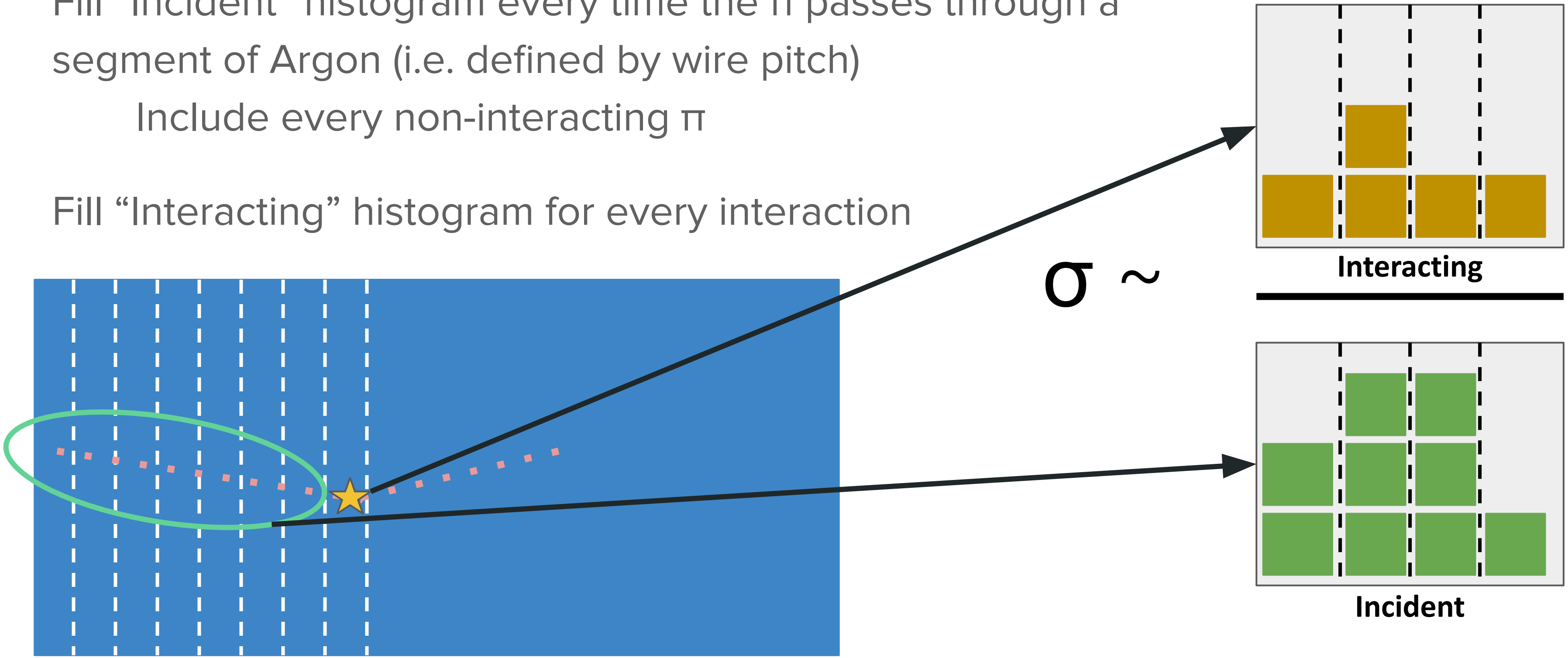
- Prior uncertainty still used within error propagation procedure (will describe later)

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



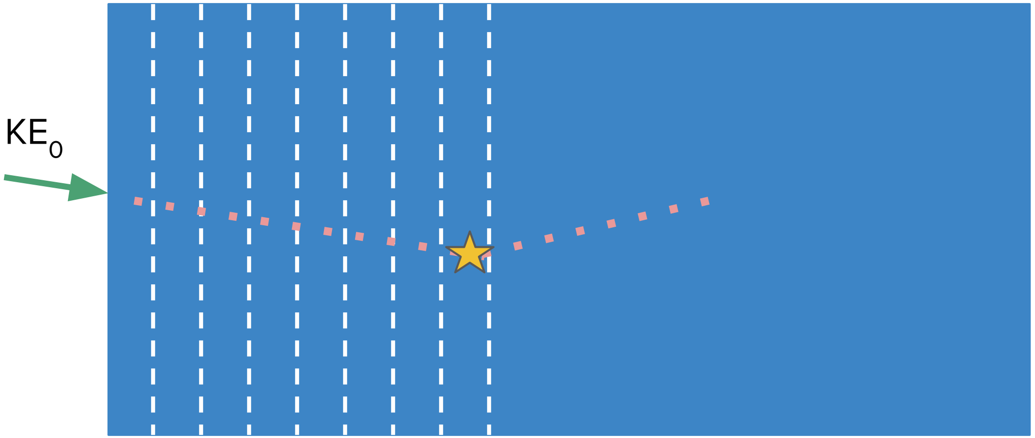
Thin Slice -- True Slices

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.



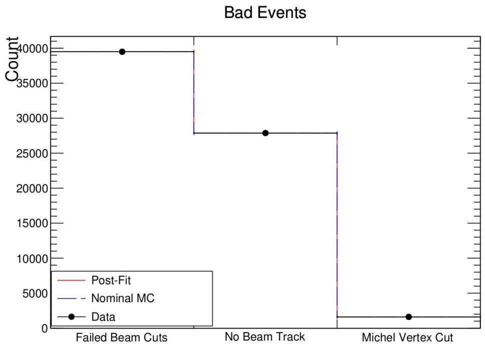
Measurement Strategy

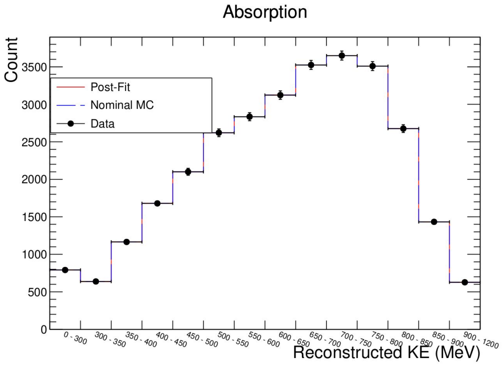
1. **Fit** to the number of selected interactions in reconstruction
   a. The fit varies the number of true signal interactions (binned in true energy)
   b. Has a resulting change on the reconstructed distributions
   c. An alternative technique to unfolding
   d. Best-fit results will be a set of varied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

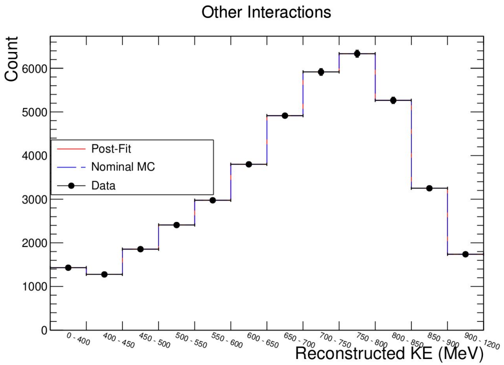
2. **Extract the cross section** from the varied MC
   a. Using the ‘thin slice method’ on varied truth information

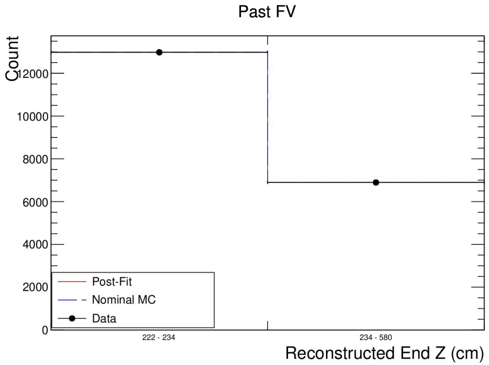
Fit Validation

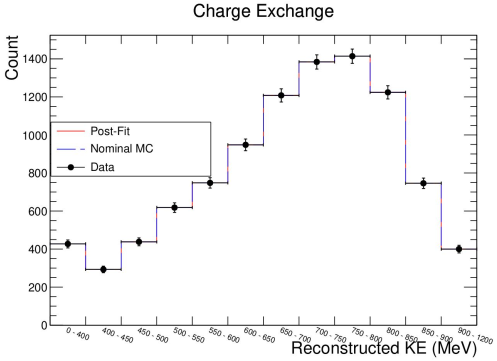
Asimov Results





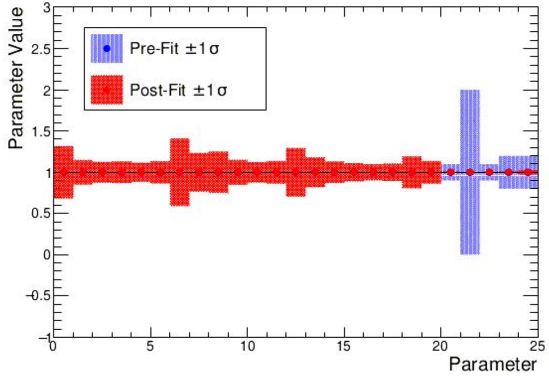


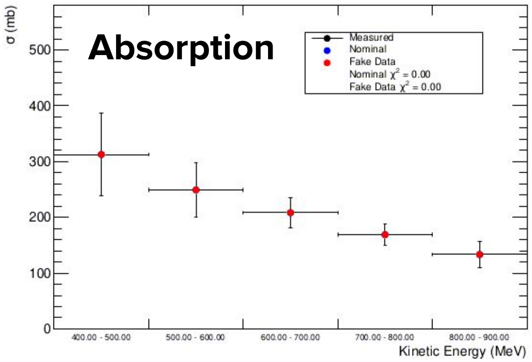


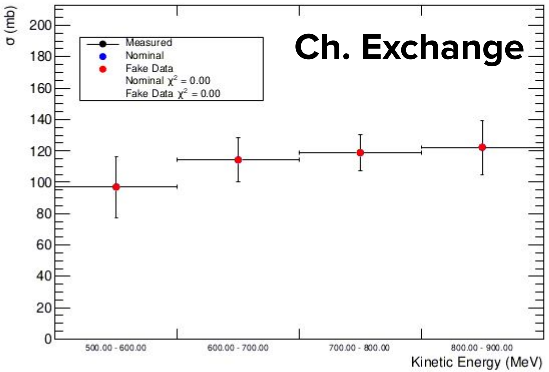


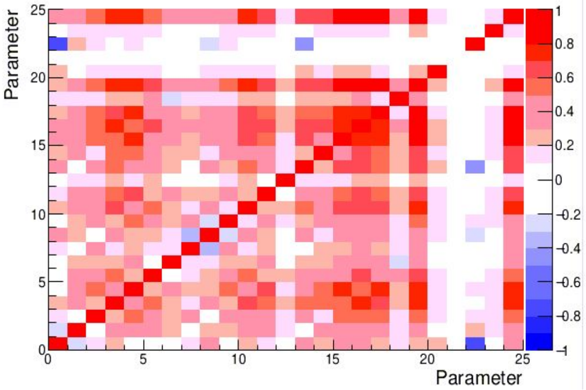
See prev. talk for error propagation procedure

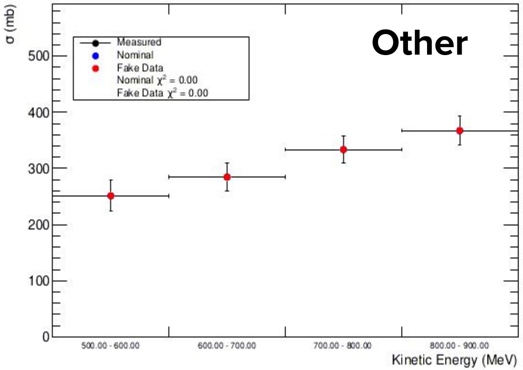
Asimov Results

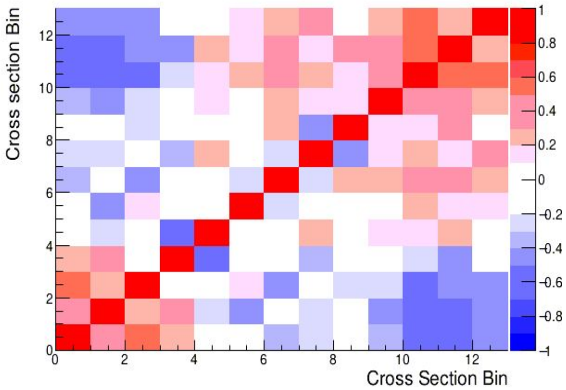








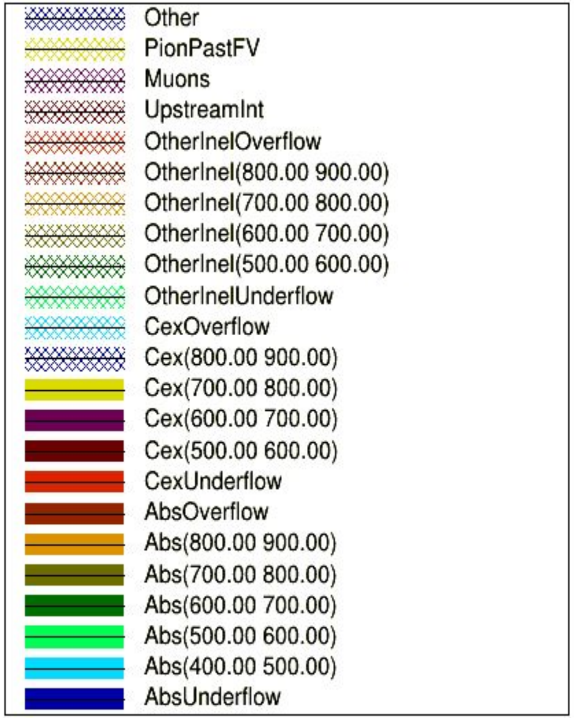


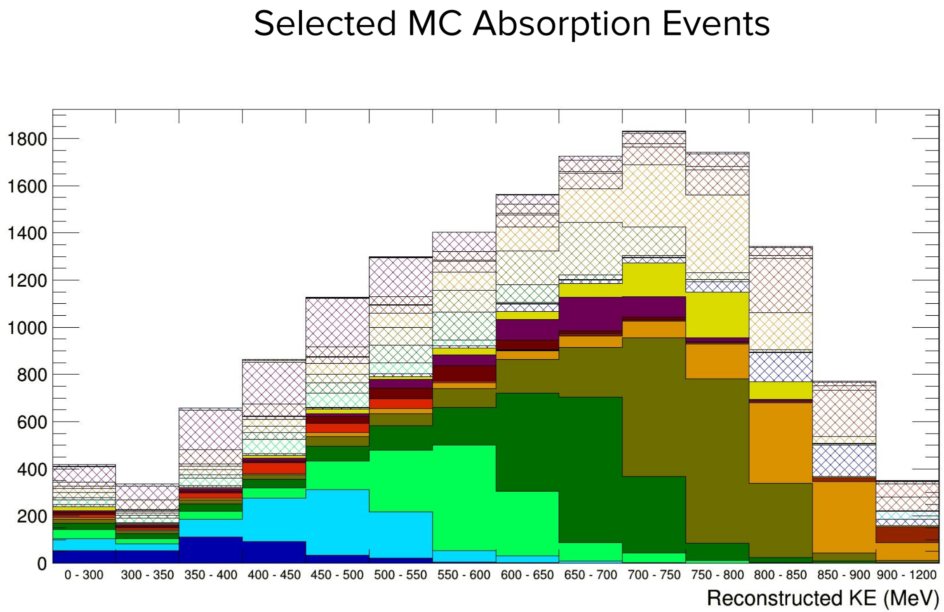


Truth Categories

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

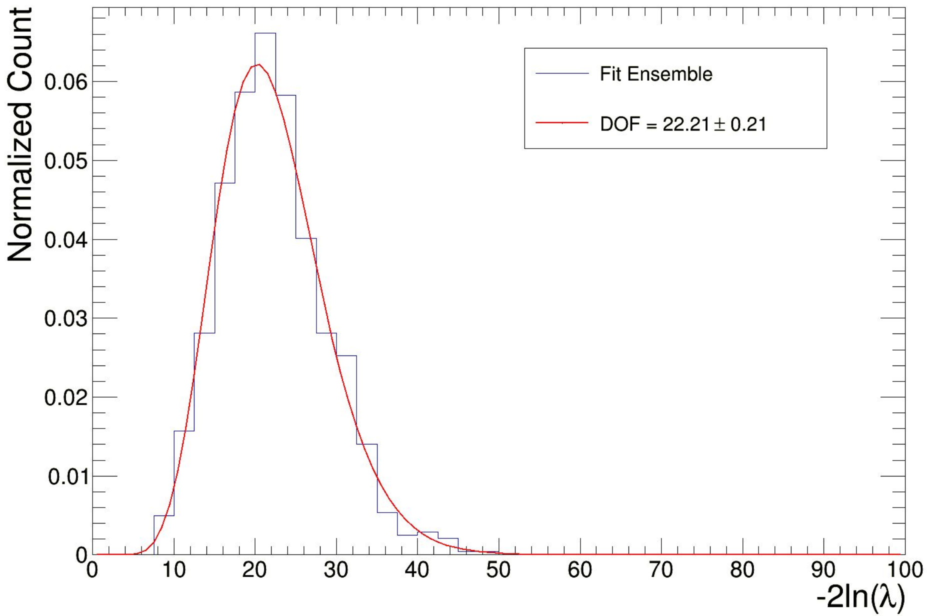
Systematic Uncertainties

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

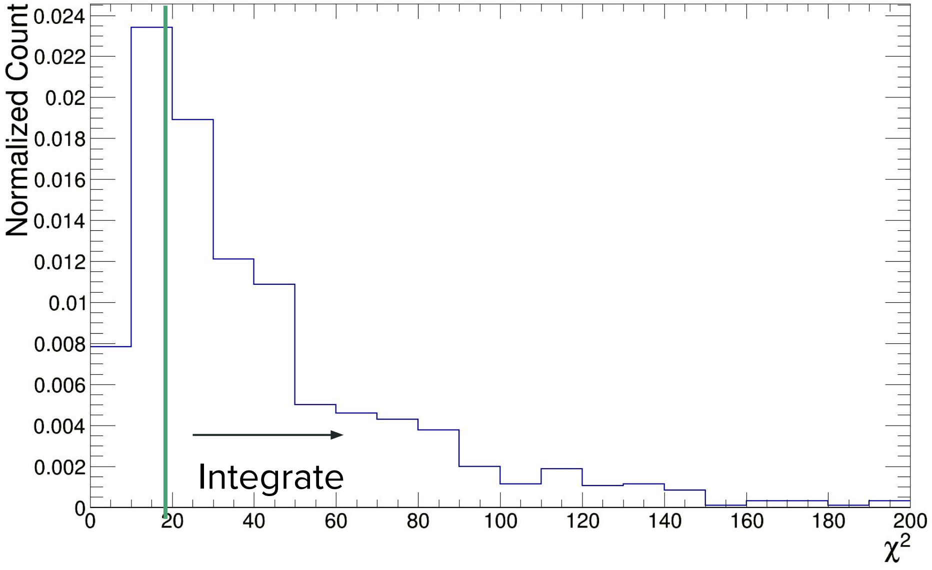


Metrics -- Cross Sections

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



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

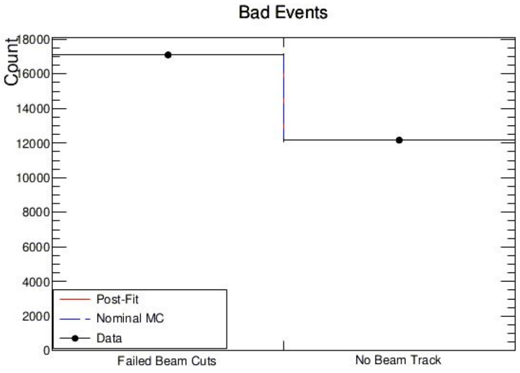
Extracted

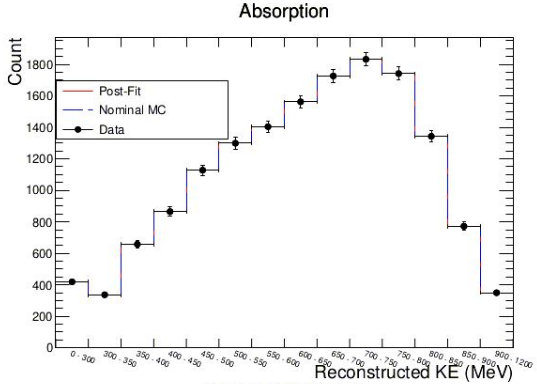
Post-fit covariance

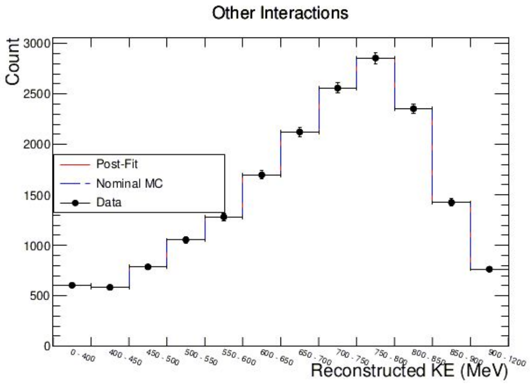
Nominal/Fake data

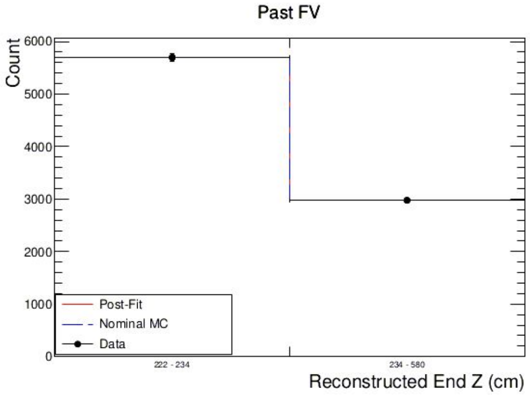
Asimov Fit

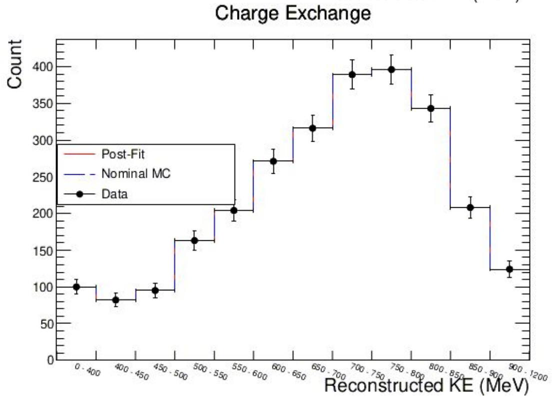
Asimov Results











Geant4Reweight Fake Data

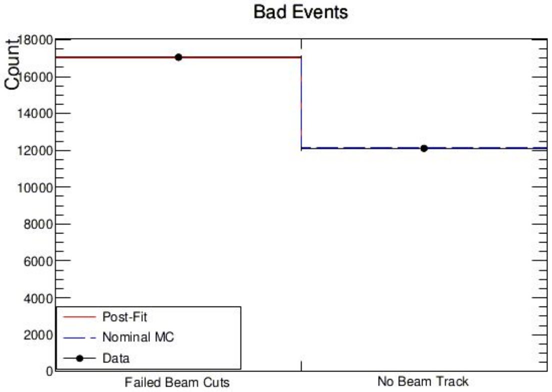
Geant4Reweight Fake Data

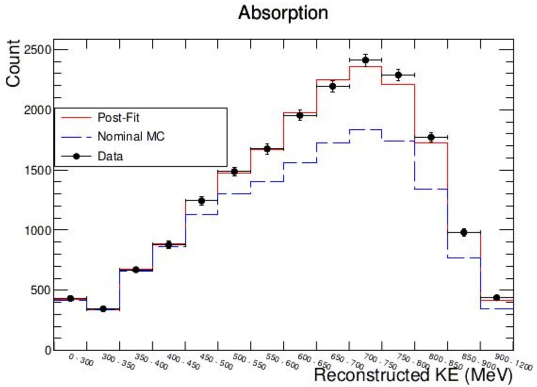
Create fake data by using Geant4Reweight to vary cross sections

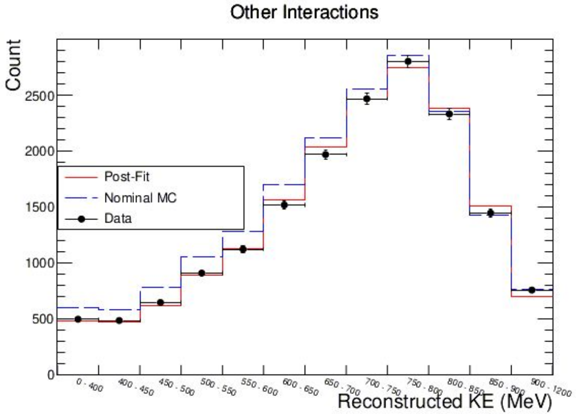
2 sets

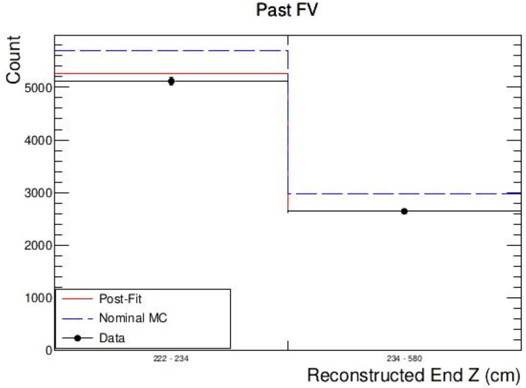
1. Increase absorption by 70%, reduce charge exchange by 60%
2. Vary total cross section: increase by 80% below 800 MeV/c, reduce by 60% above

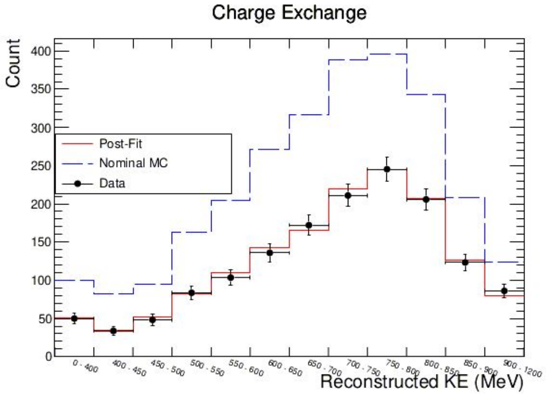
Geant4Reweight Fake Data 1







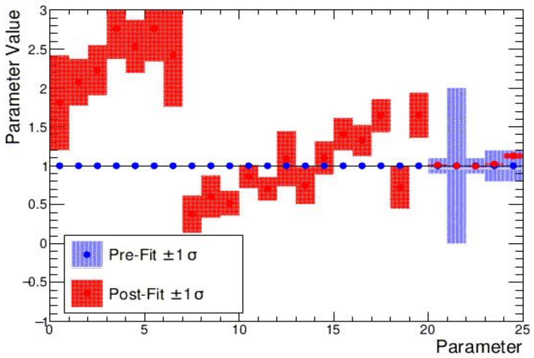


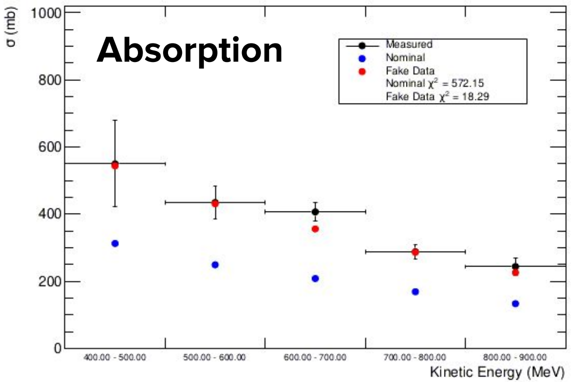


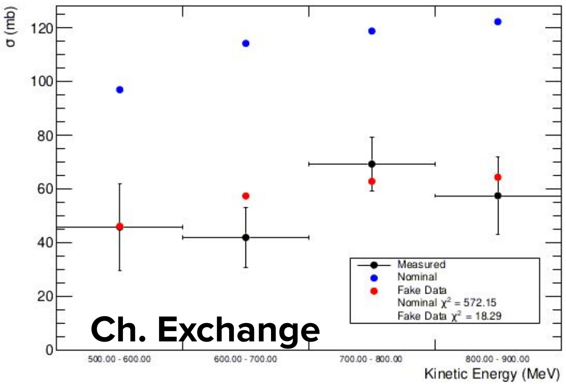
Pre-fit $-2\ln\lambda_{Stat}$	1617.04
Post-fit $-2\ln\lambda_{Stat}$	9.00
Post-fit $-2\ln\lambda_{Syst}$	0.43

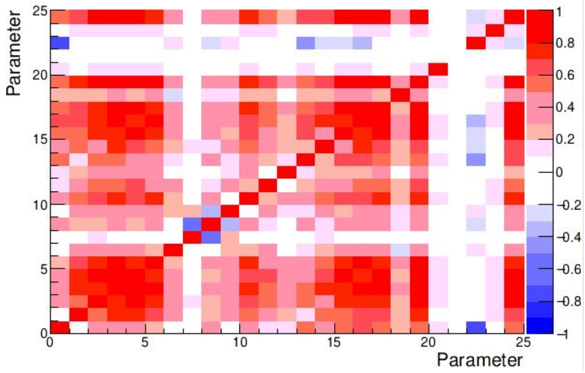
Geant4Reweight Fake Data 1

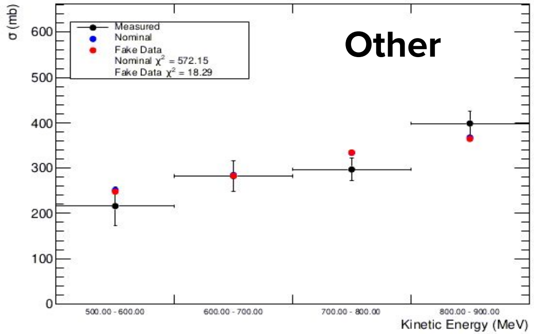
Fake Data p-value	0.72
Nominal p-value	0.00

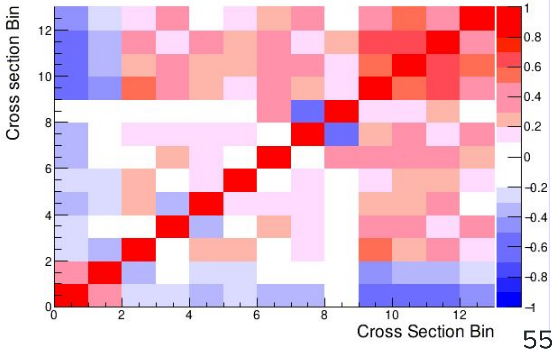




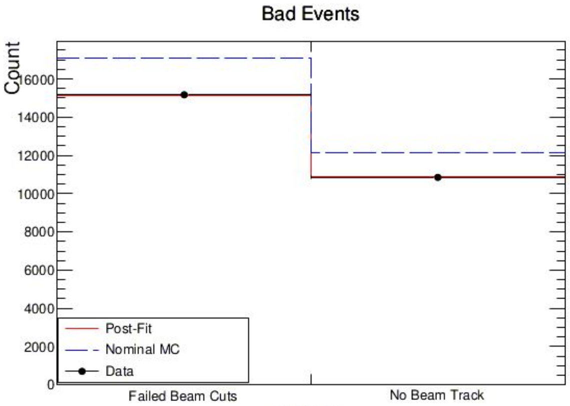


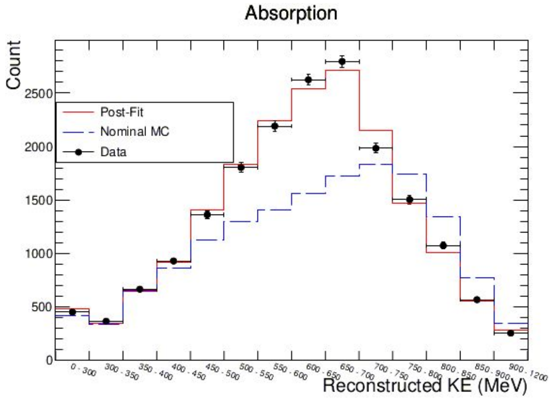


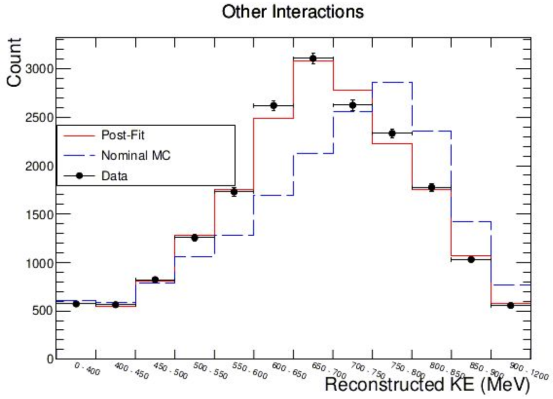


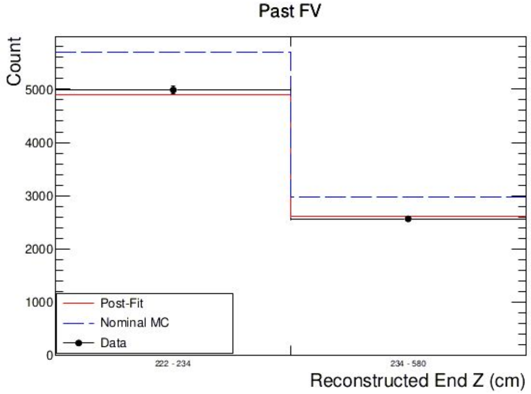


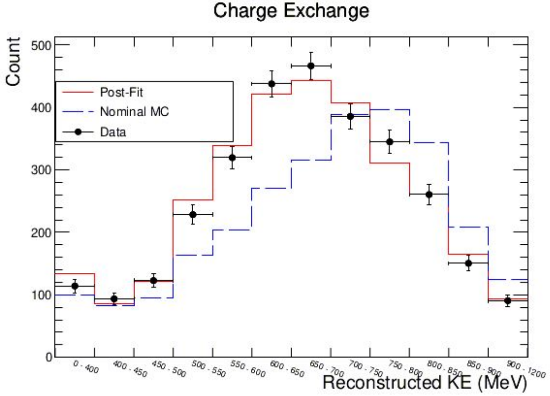
Geant4Reweight Fake Data 2







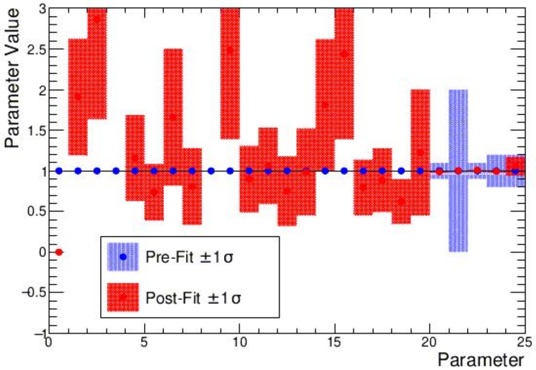


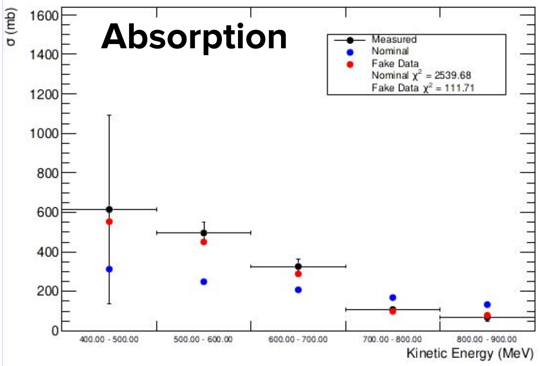


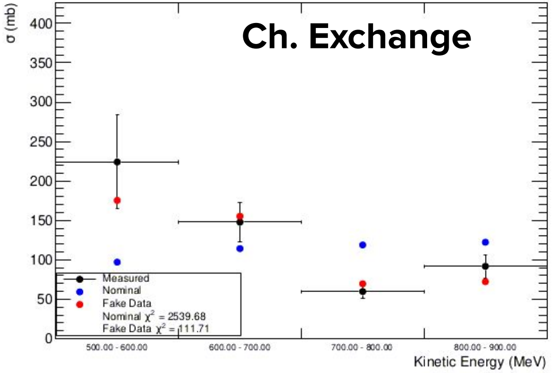
Pre-fit $-2\ln\lambda_{Stat}$	4299.83
Post-fit $-2\ln\lambda_{Stat}$	85.21
Post-fit $-2\ln\lambda_{Syst}$	0.10

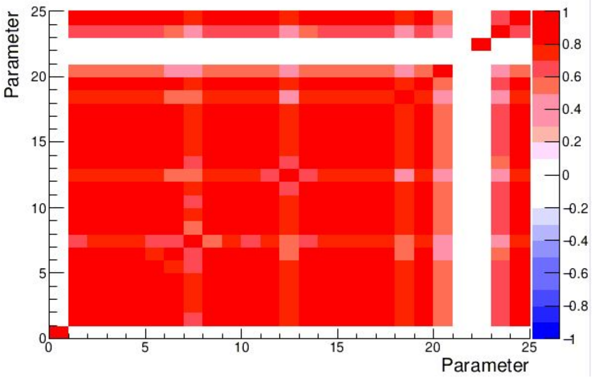
Geant4Reweight Fake Data 2

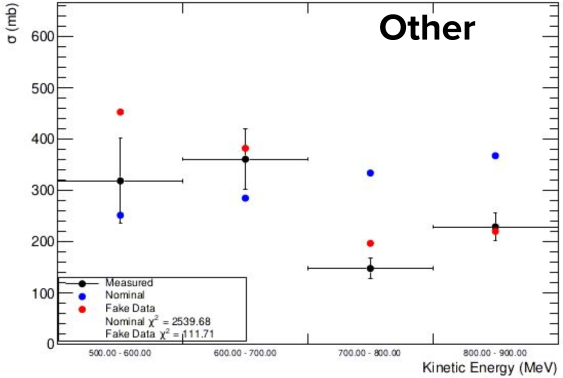
Fake Data p-value	0.07
Nominal p-value	0.00

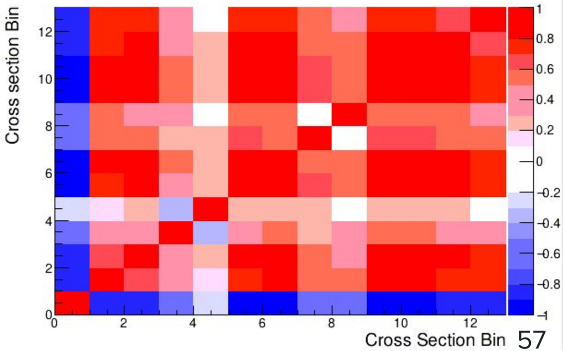












Geant4Reweight Fake Data 2 Discussion

Parameterization can not fit the variation applied

- Results in a bad fit p-value

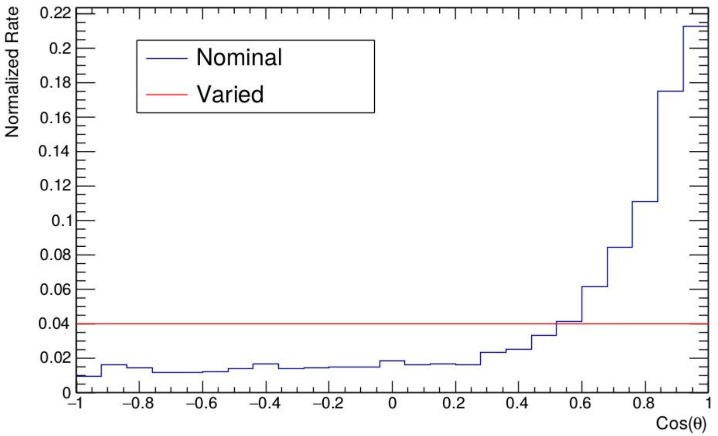
Example of how a bad data fit can be identified

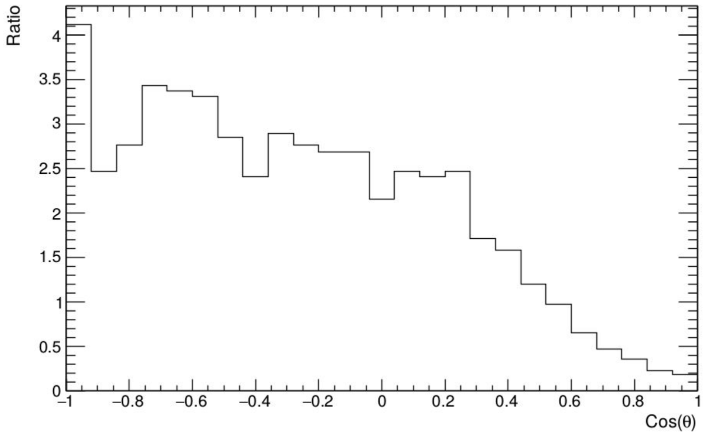
Angular Variation Fake Data

Angular Variation Fake Data

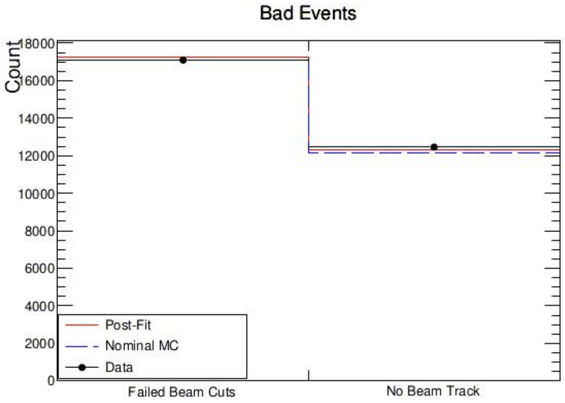
Create fake data by varying the outgoing angle of leading-momentum pions resulting from primary pion interactions

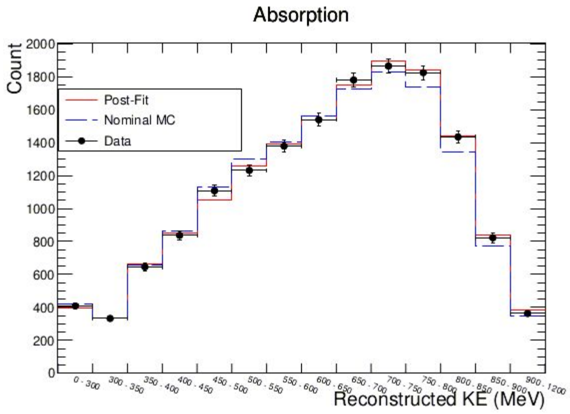
Create distribution by hand (e.g. flattened distribution), use ratio as event weights

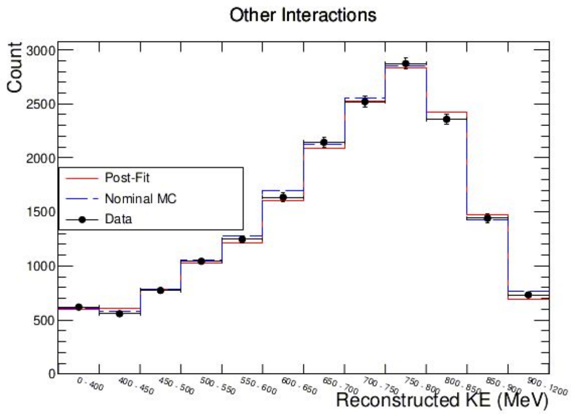


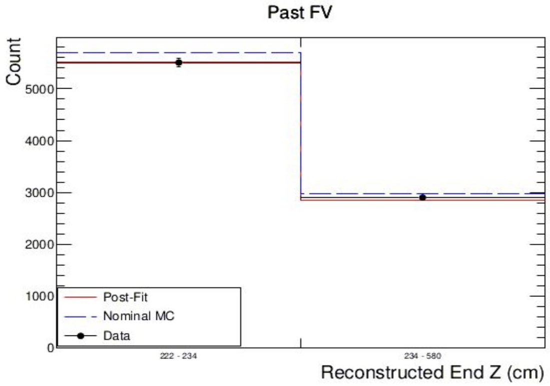


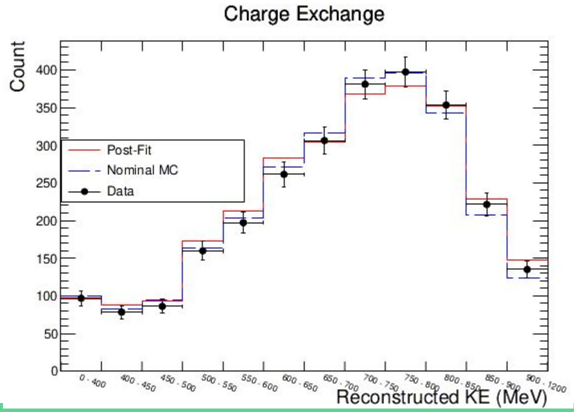
Angular Variation Fake Data







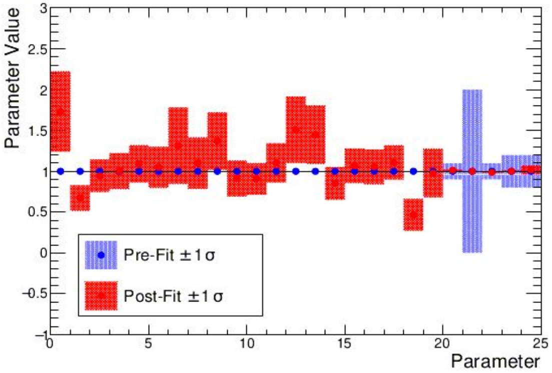


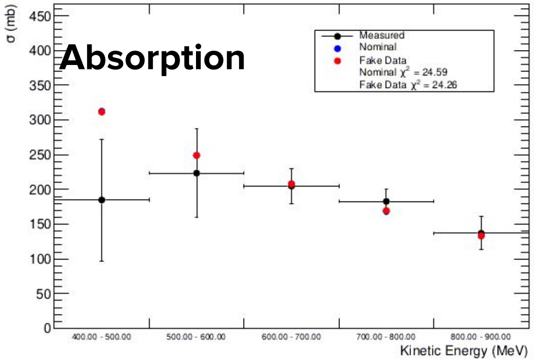


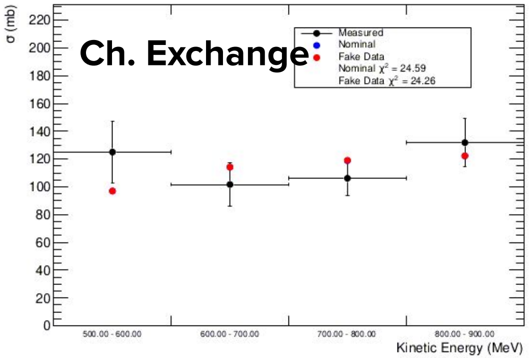
Pre-fit $-2\ln\lambda_{Stat}$	108.00
Post-fit $-2\ln\lambda_{Stat}$	27.31
Post-fit $-2\ln\lambda_{Syst}$	0.04

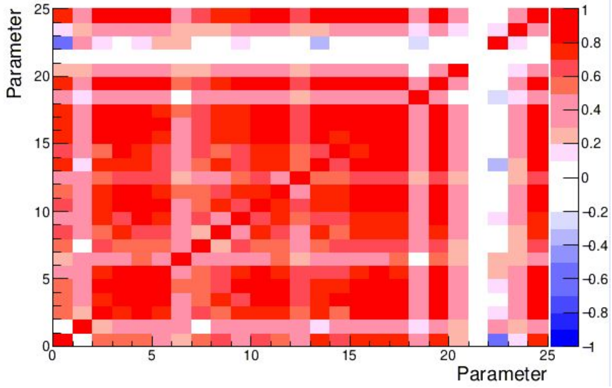
Angular Variation Fake Data

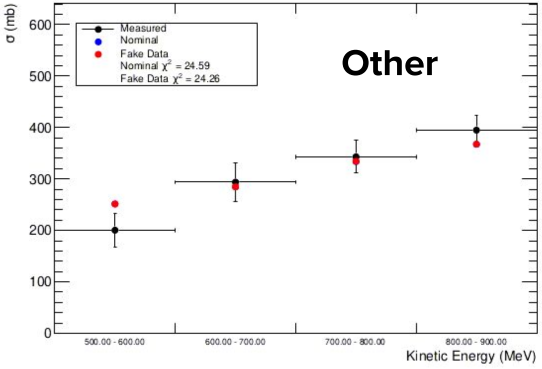
Fake Data p-value	0.60
Nominal p-value	0.60

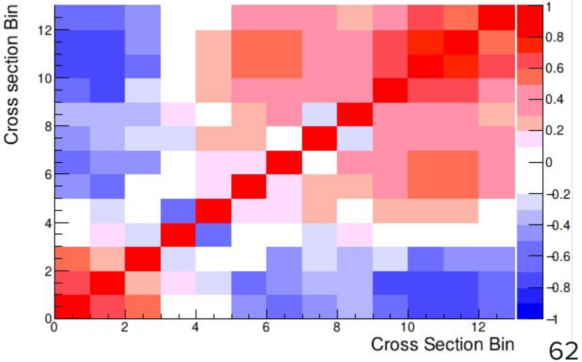










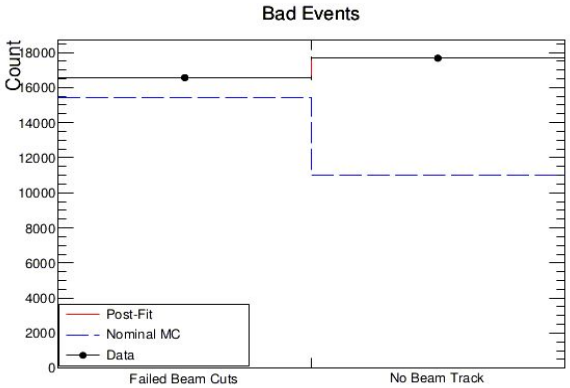


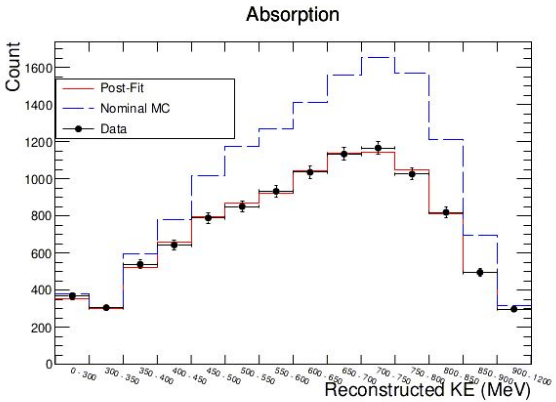
Angular Variation Fake Data Discussion

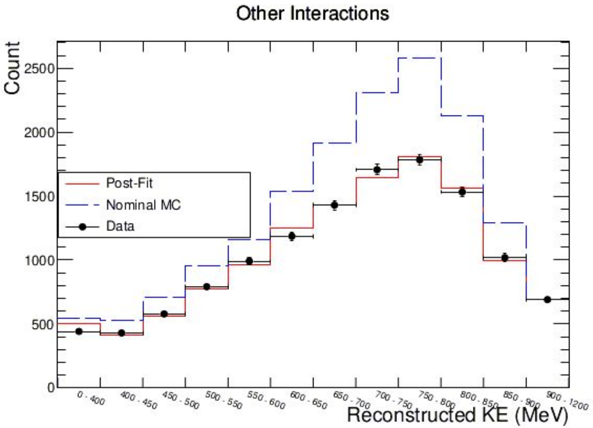
Successful fit shows robustness against mismodeling of outgoing pion kinematics

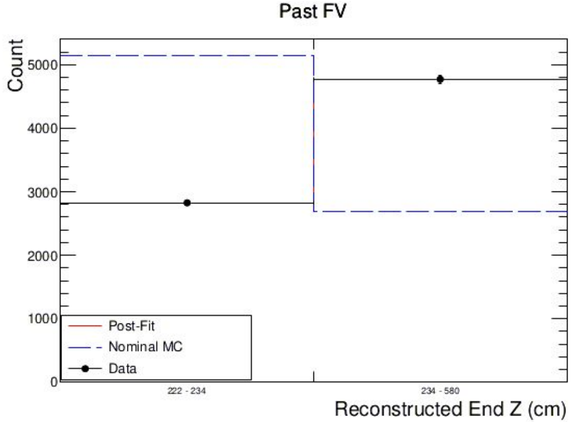
Results on 1 GeV Data

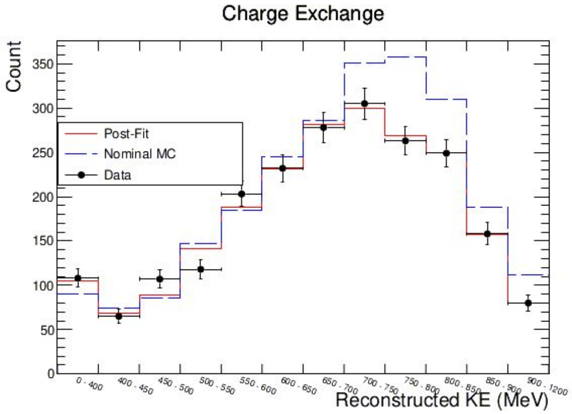
Fit to Data







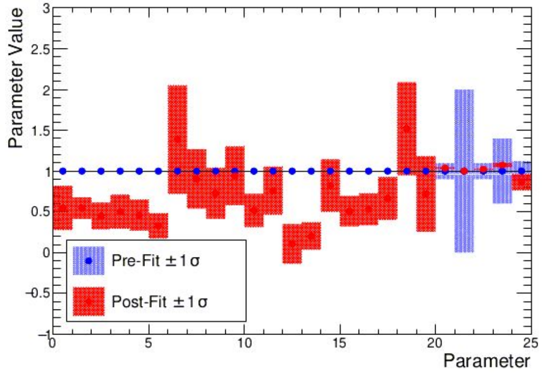


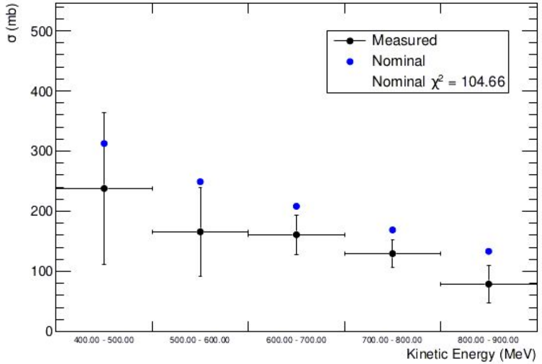


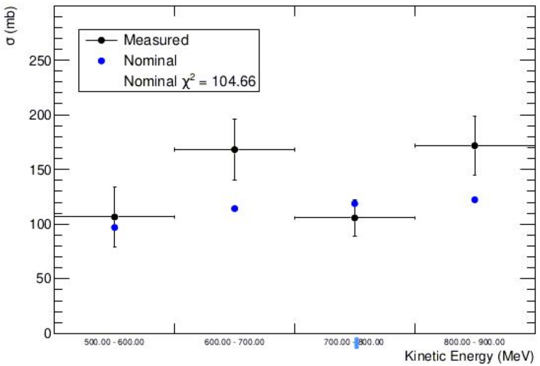
Pre-fit $-2\ln\lambda_{Stat}$	8293.26
Post-fit $-2\ln\lambda_{Stat}$	29.27
Post-fit $-2\ln\lambda_{Syst}$	1.46

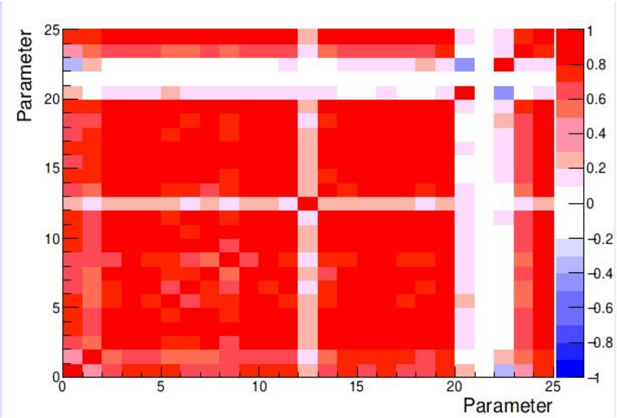
Nominal p-value	0.08

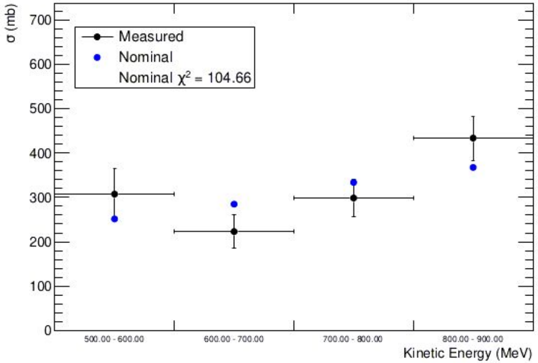
Fit to Data

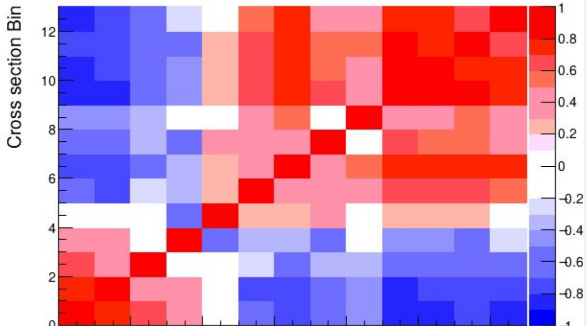






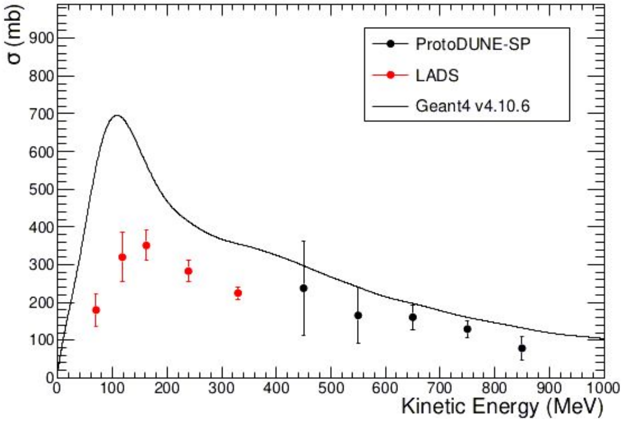




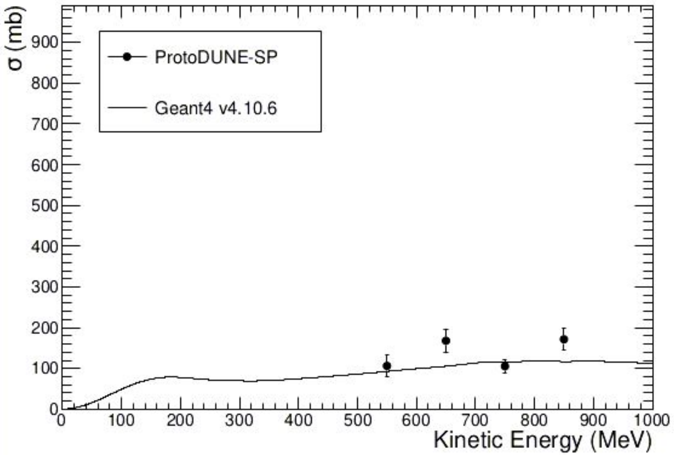


Fit to Data

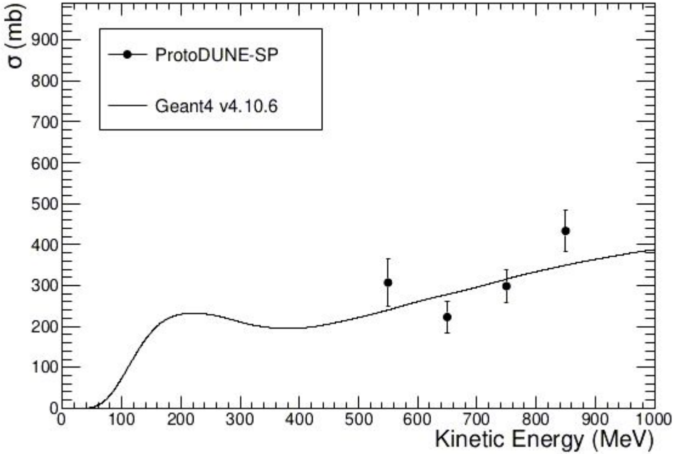
Absorption



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)