Updates on π^+-Ar inclusive cross-section measurement

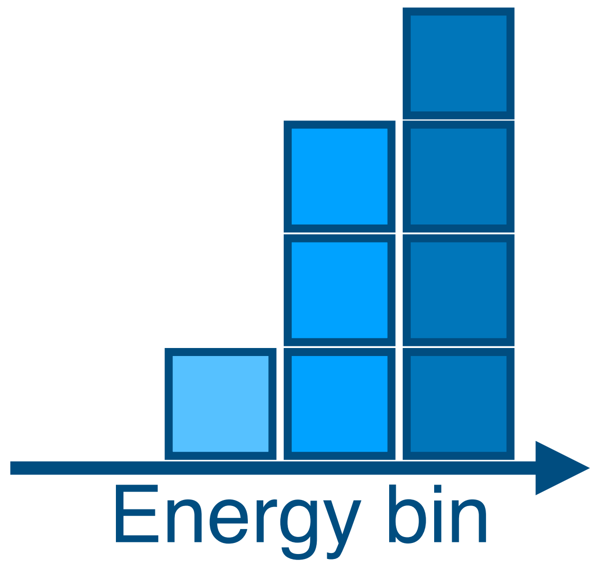
Yinrui Liu[1]
Sept 8, 2022 @ HadAna meeting

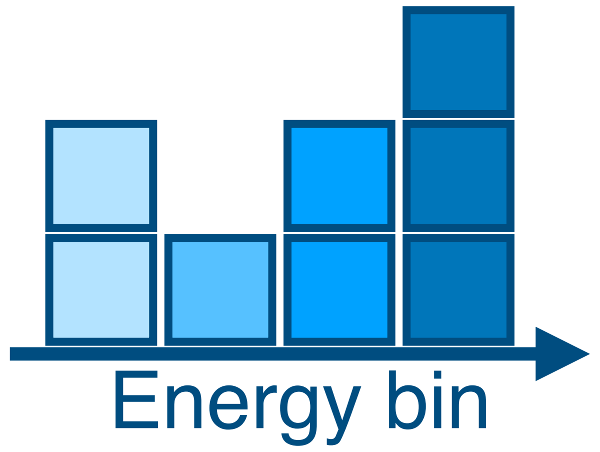
[1]University of Chicago

Energy-slicing method

Each selected beam track has an **initial slice** and an **end slice**.

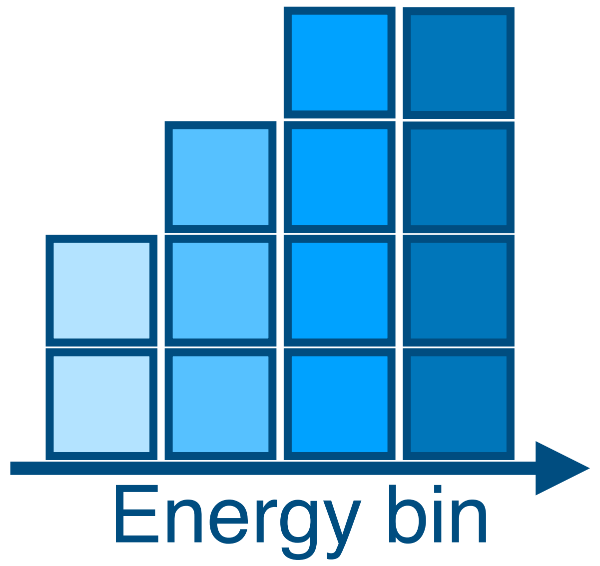
N_{initial} $\qquad$ N_{end}





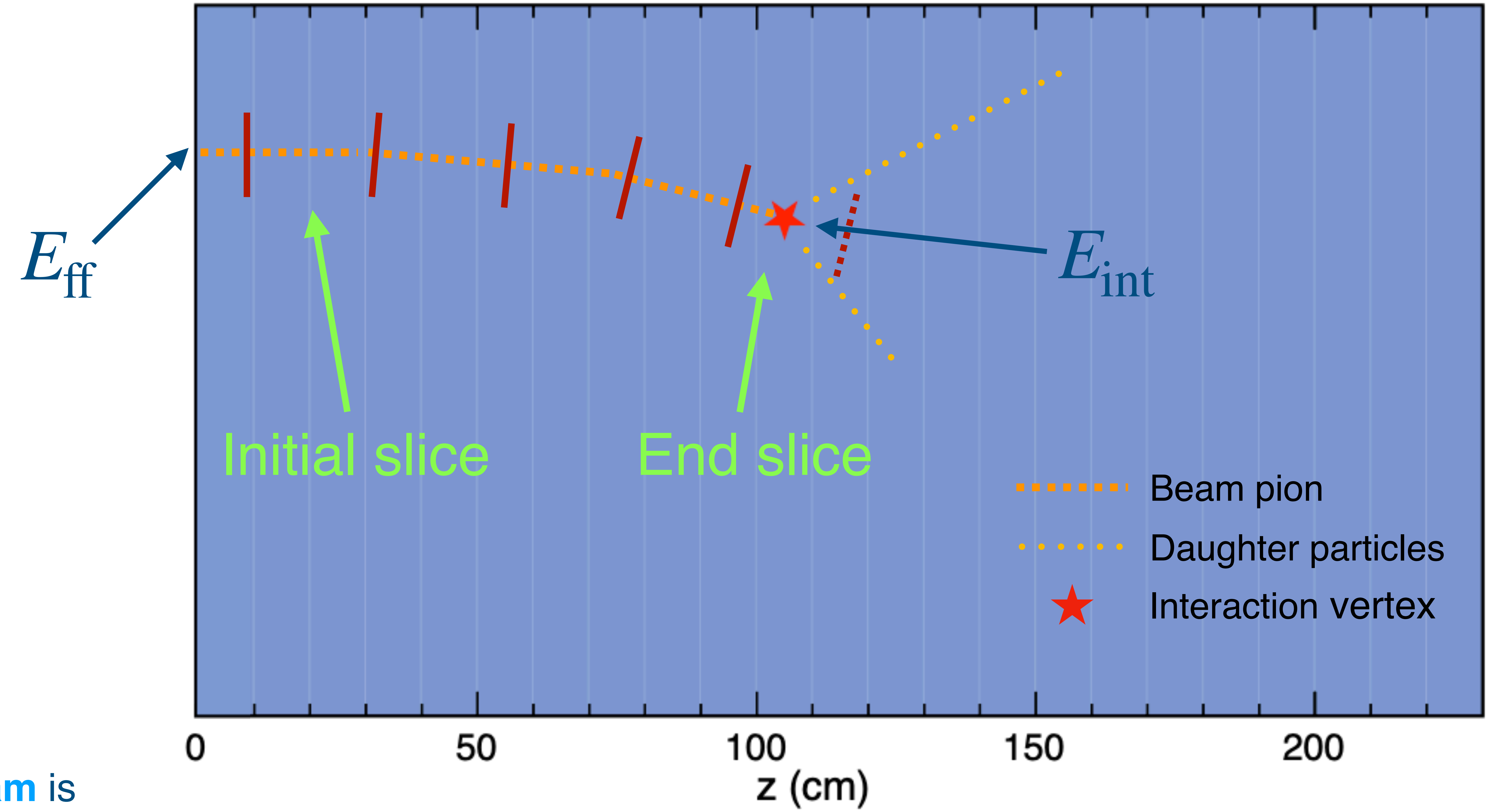
If there is pion inelastic scattering (signal) at the interaction vertex, then the **end slice** is also an **interaction slice**.

N_{incident}



The **incident histogram** is calculated by N_{initial} and N_{end}:

$$N_{\text{inc}}(i) = \sum_{j=i}^{N} N_{\text{end}}(j) - \sum_{j=i+1}^{N} N_{\text{ini}}(j)$$



The cross-section is calculated by N_{incident} and $N_{\text{interaction}}$

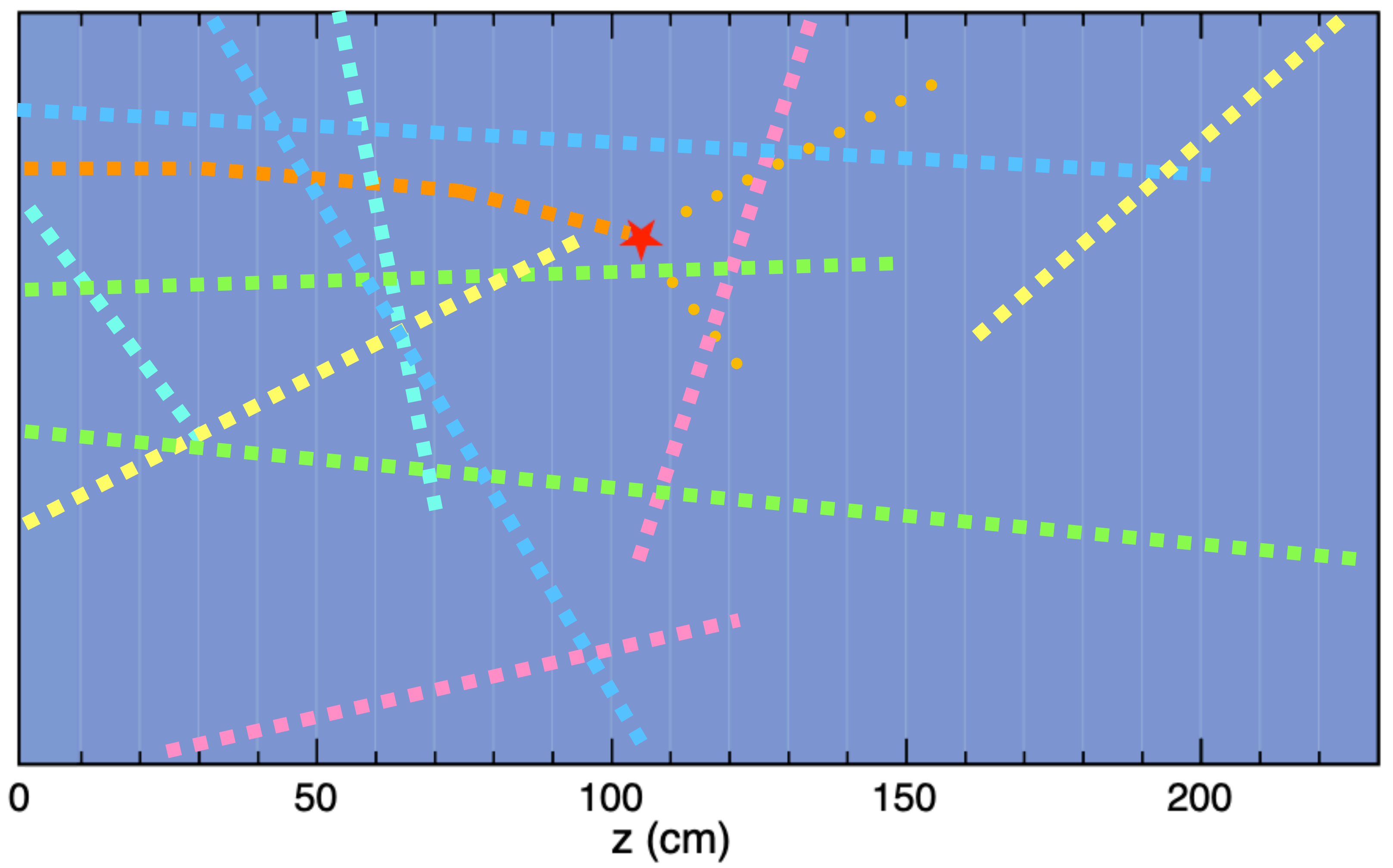
$$\sigma(E) = \frac{M_{\text{Ar}}}{\rho N_A \Delta E} \frac{dE}{dx}(E) \ln\left(\frac{N_{\text{inc}}(E)}{N_{\text{inc}}(E) - N_{\text{int}}(E)}\right)$$

Selections

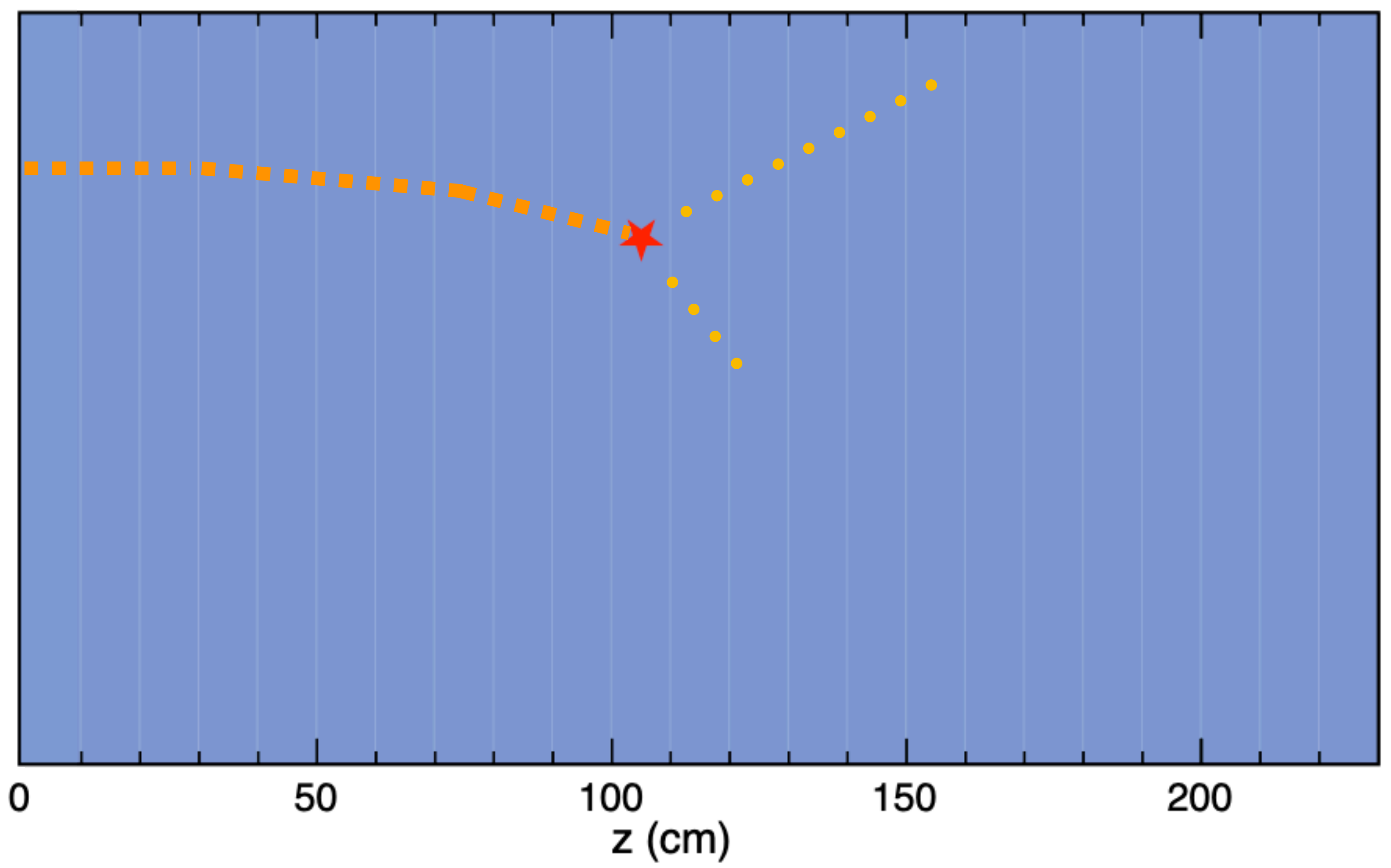
- For inclusive cross-section measurement, we need to select **pion beam events** (regardless of what daughter particles are).

  - Pandora identification
  - Precuts
  - Beam quality cut
  - Proton cut
  - Michel score cut
  - APA3 cut

Details in back-ups



Original event



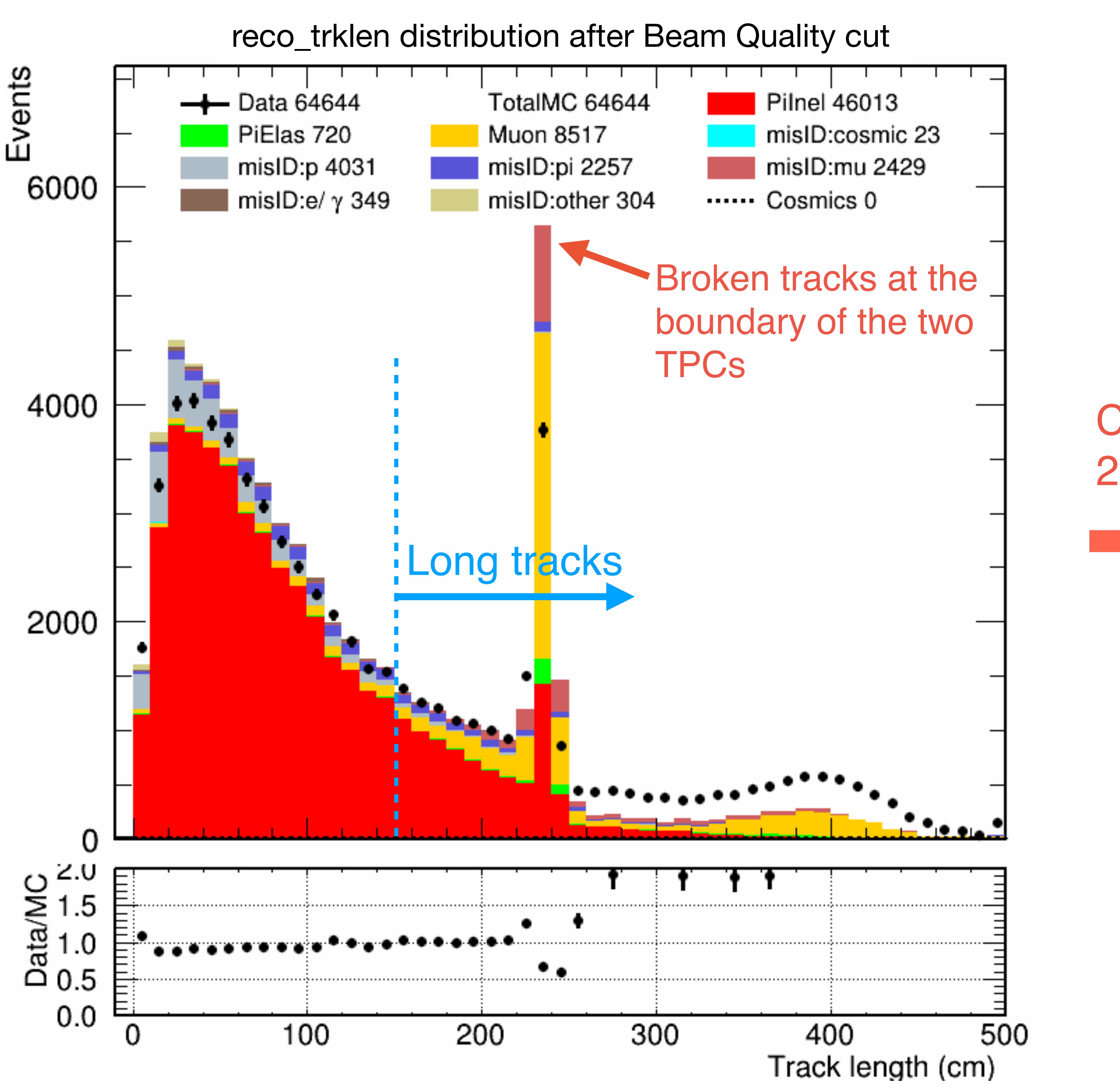
Pandora identified beam (pion) track

- After full selections, we have about **80%** pion inelastic events (signals)

Updates

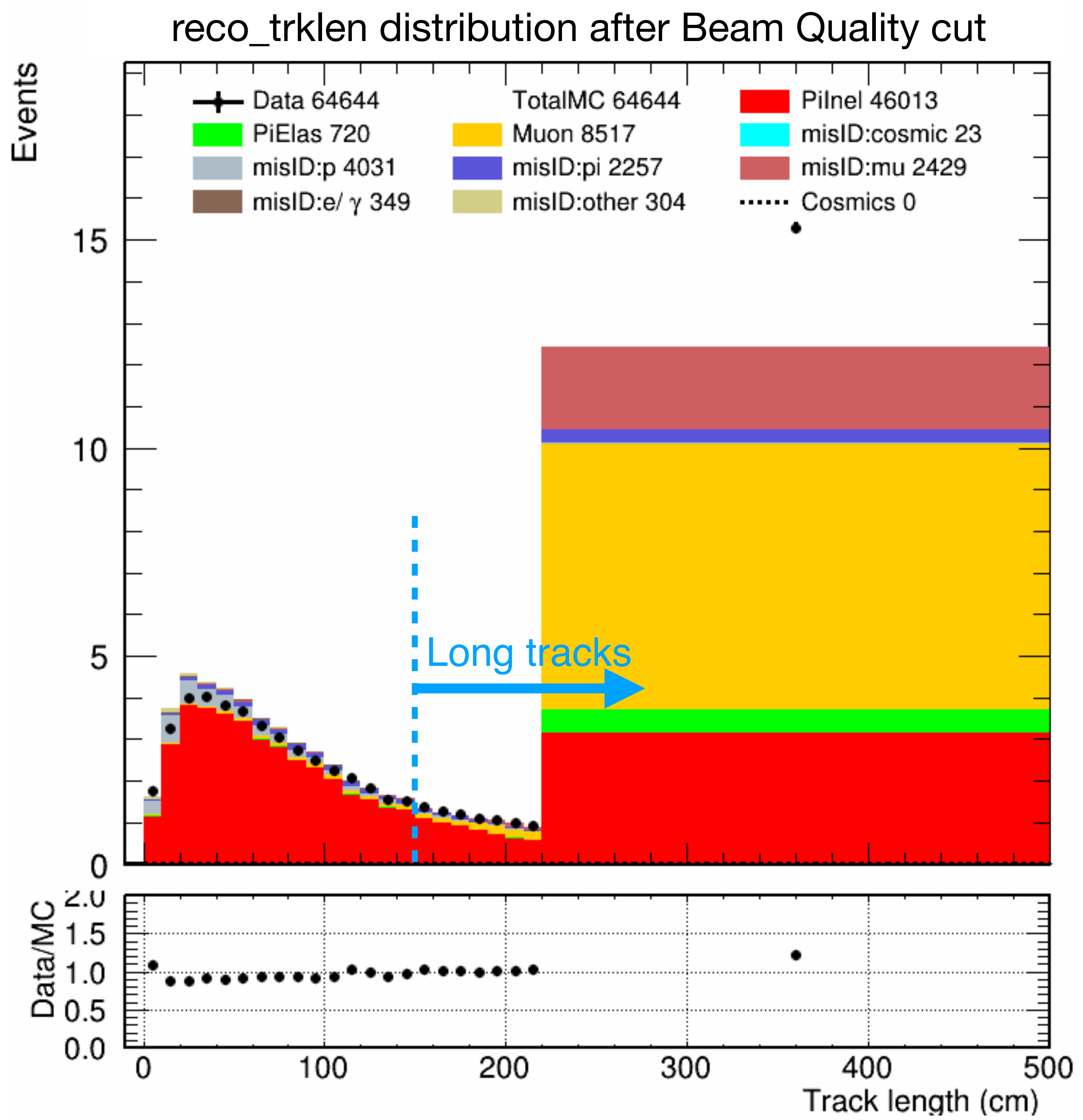
- **Reweighting MC**
  - Muon background reweight
  - Beam momentum reweight
- Background subtraction
- Unfolding and error propagation
- Results (mainly from fake data)

Muon bkg reweight



Combine tracks longer than 220 cm in a single bin

- We found there are more long tracks (length > 150 cm) in data than MC.



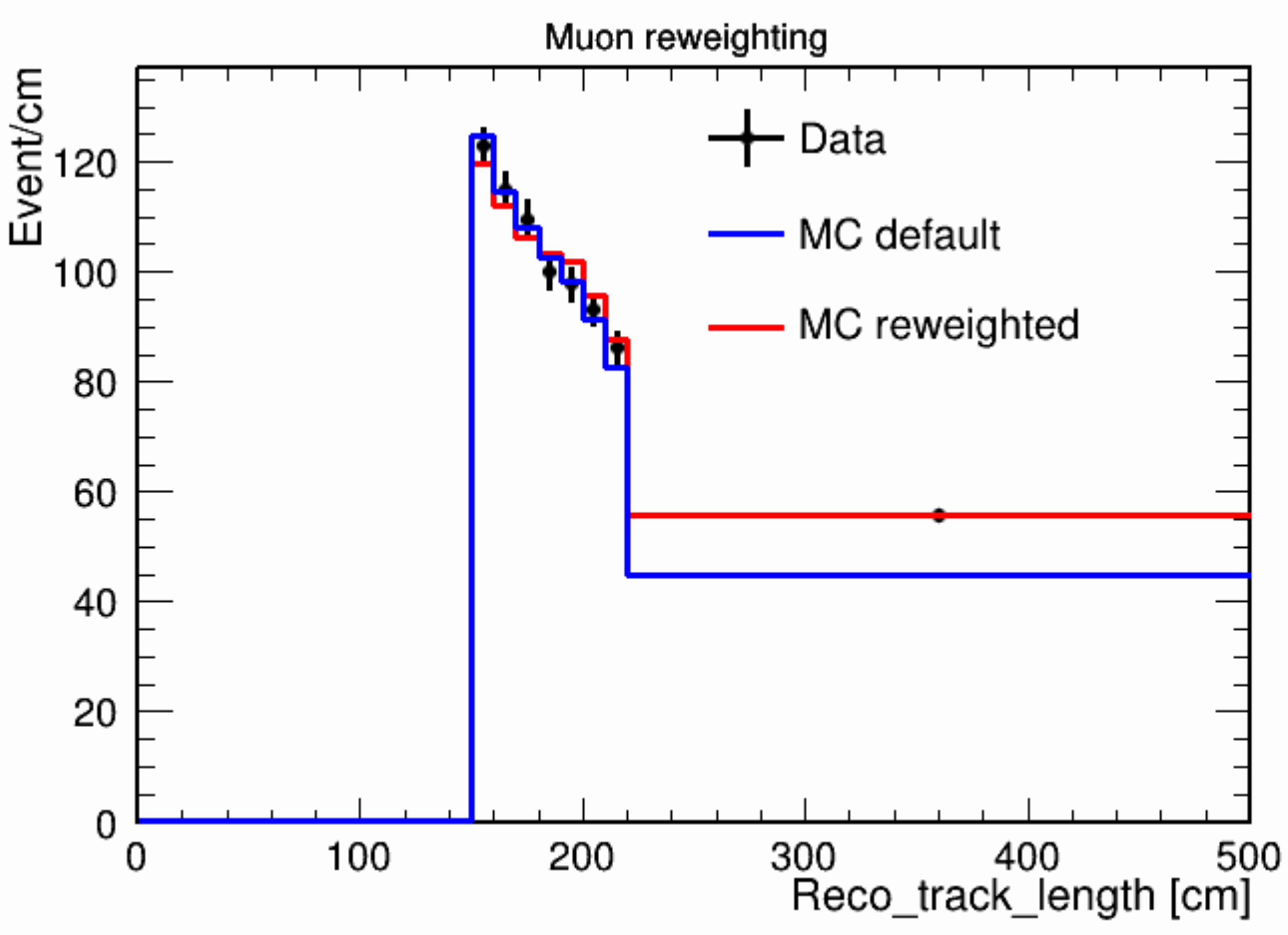
- This could be improved if we scale up the muon beam fraction in MC.

Muon bkg reweight

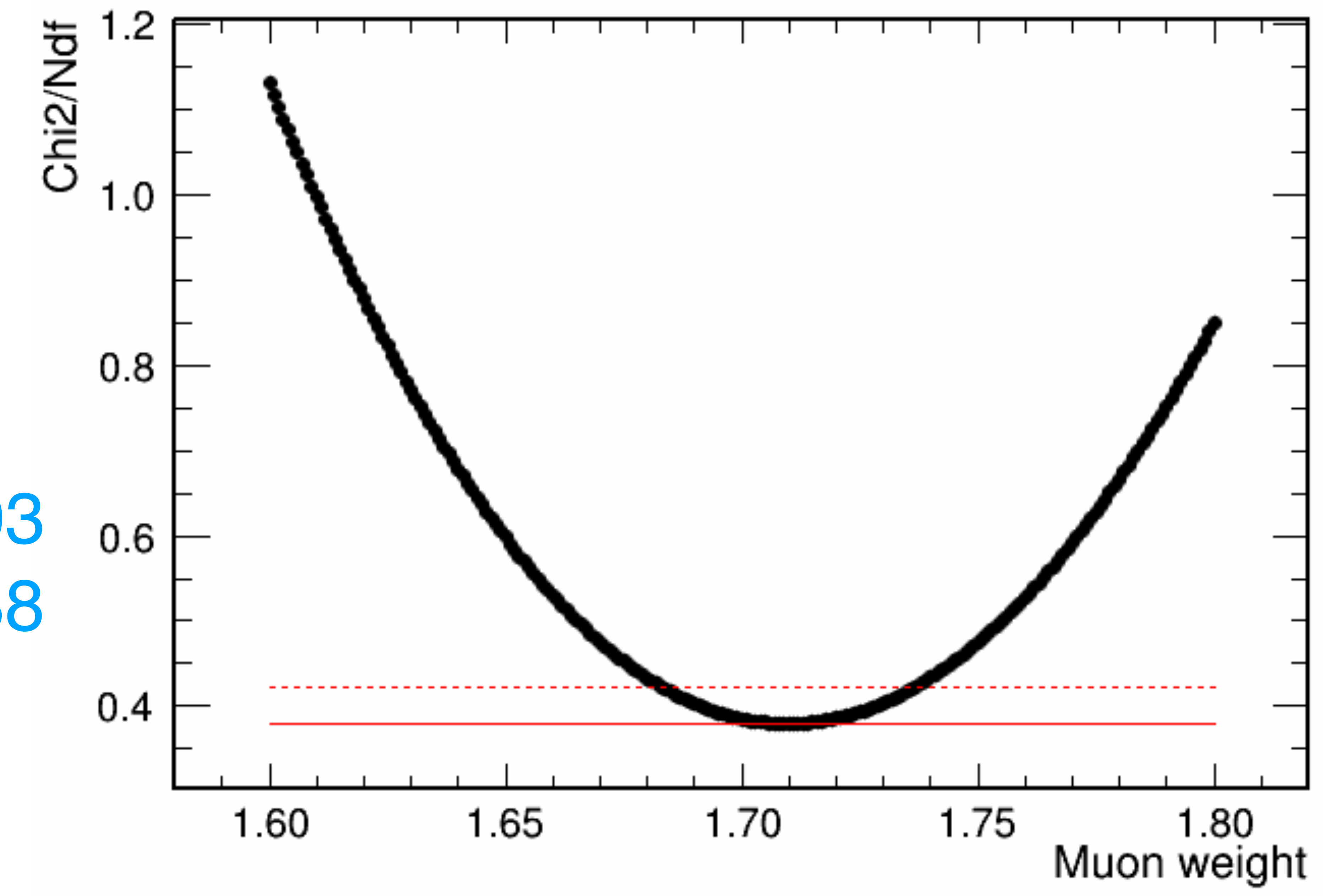
Compound Poisson distribution
https://www.sciencedirect.com/science/article/pii/S0168900214001776?ref=cra_js_challenge&fr=RR-1

- After all selections except muon-related cuts, we perform χ^2 fit in the long track region.

$$\chi^2/N_{df} = \frac{1}{N-1}\sum_{bin=1}^{N}\frac{(data - weightedMC)^2}{s\cdot data + weightedMC}, \text{ where } s = \frac{\sum(MCweight^2)}{\sum(MCweight)}$$

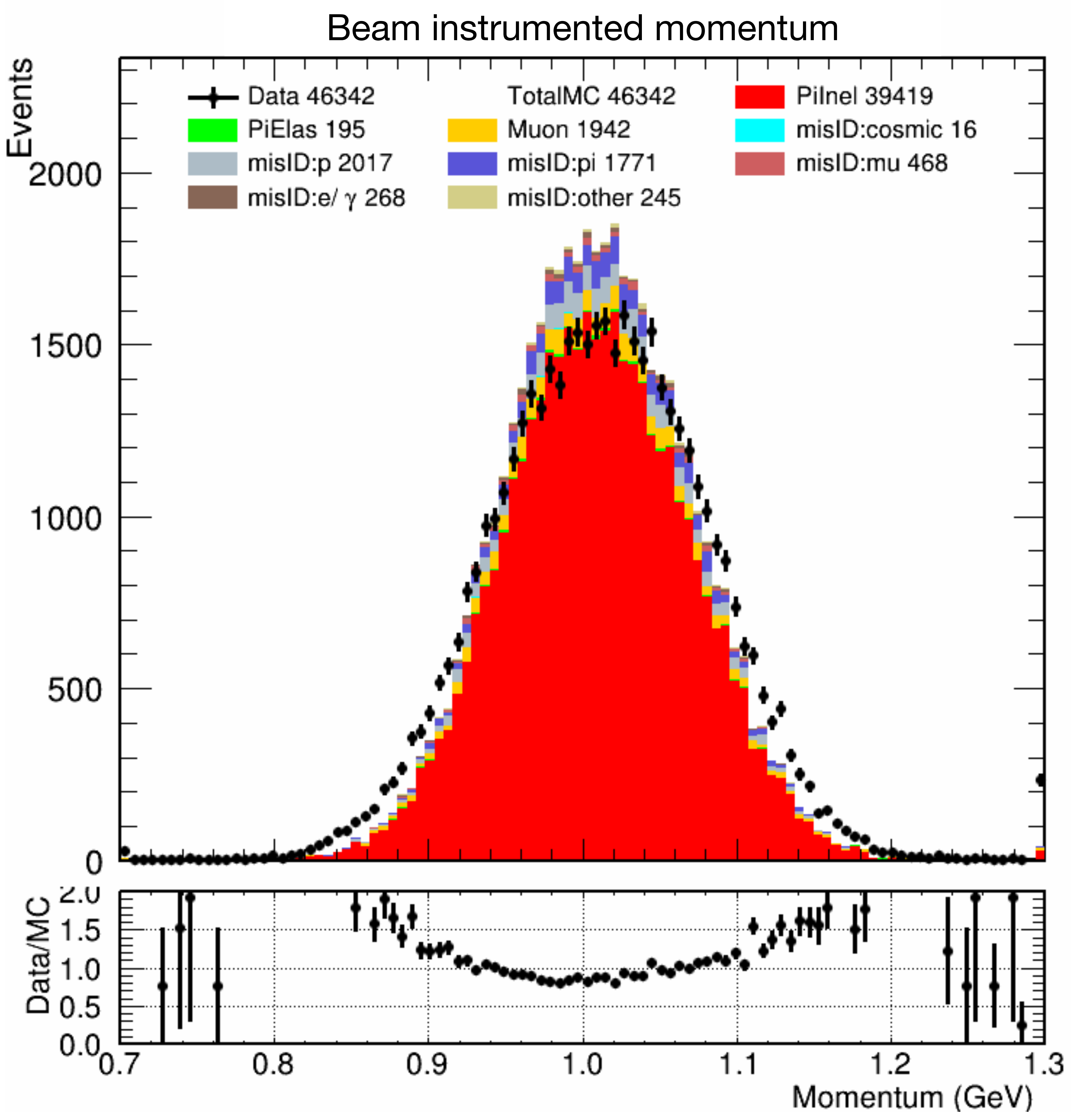


Fit result:
Weight = 1.71 ± 0.03
FOM = 3.03/8 = 0.38

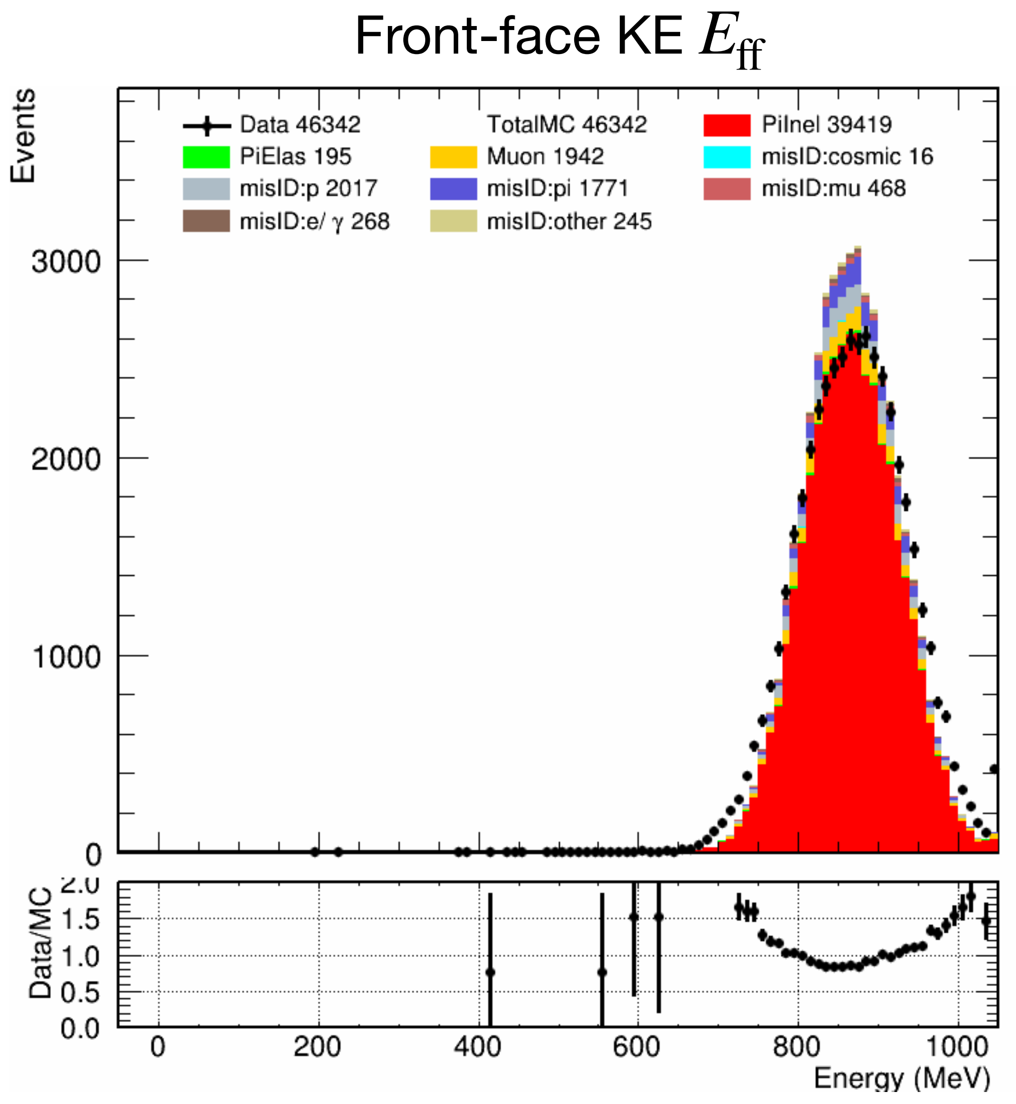


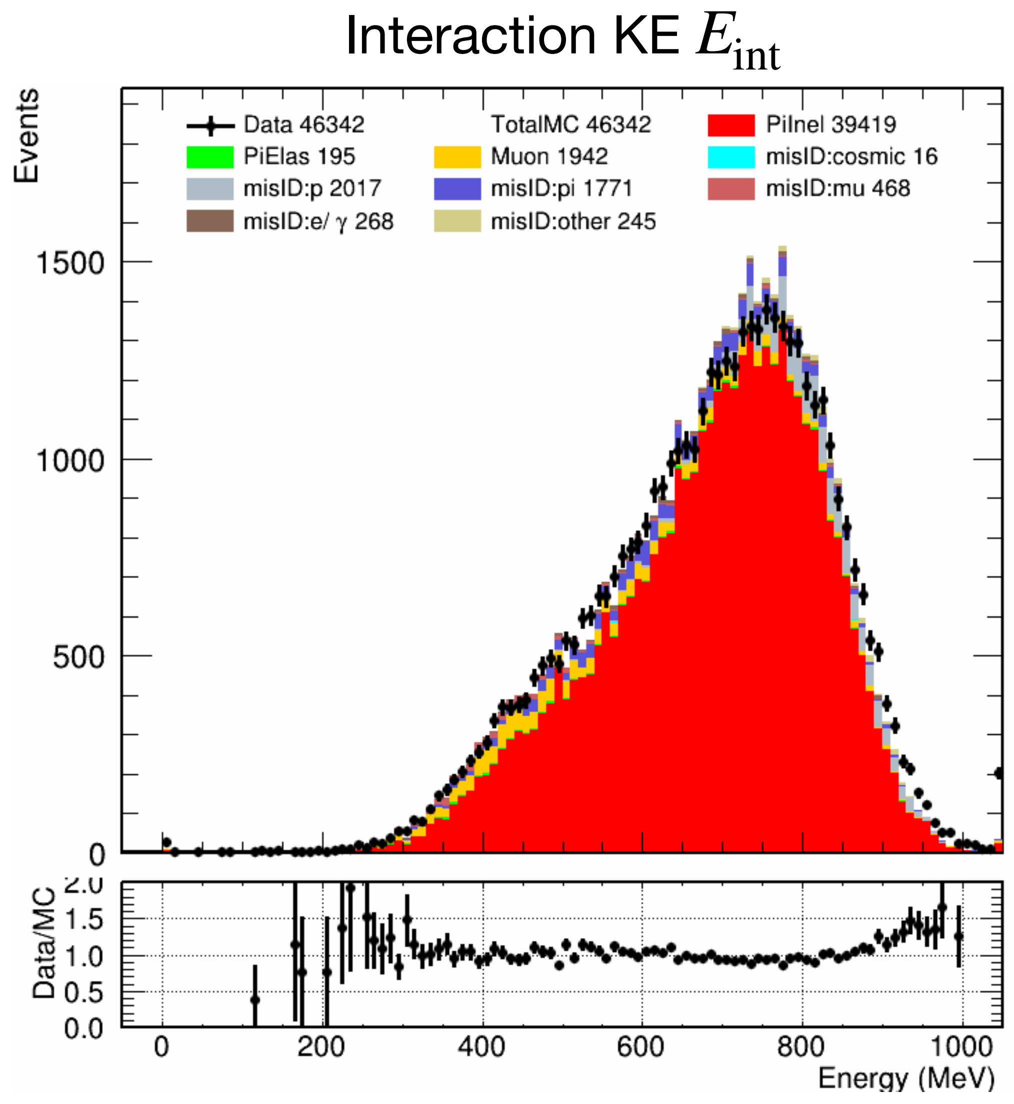
This mainly correct for long beam muons. For the remaining muons, we still use Michel score distribution after APA3 cut for sideband fit.

Beam momentum reweight



- The beam instrumented momentum distributions between data and MC are different.





- We use stopping muon sample as standard candle to calibrate the beam momentum, and applied the results to pions.

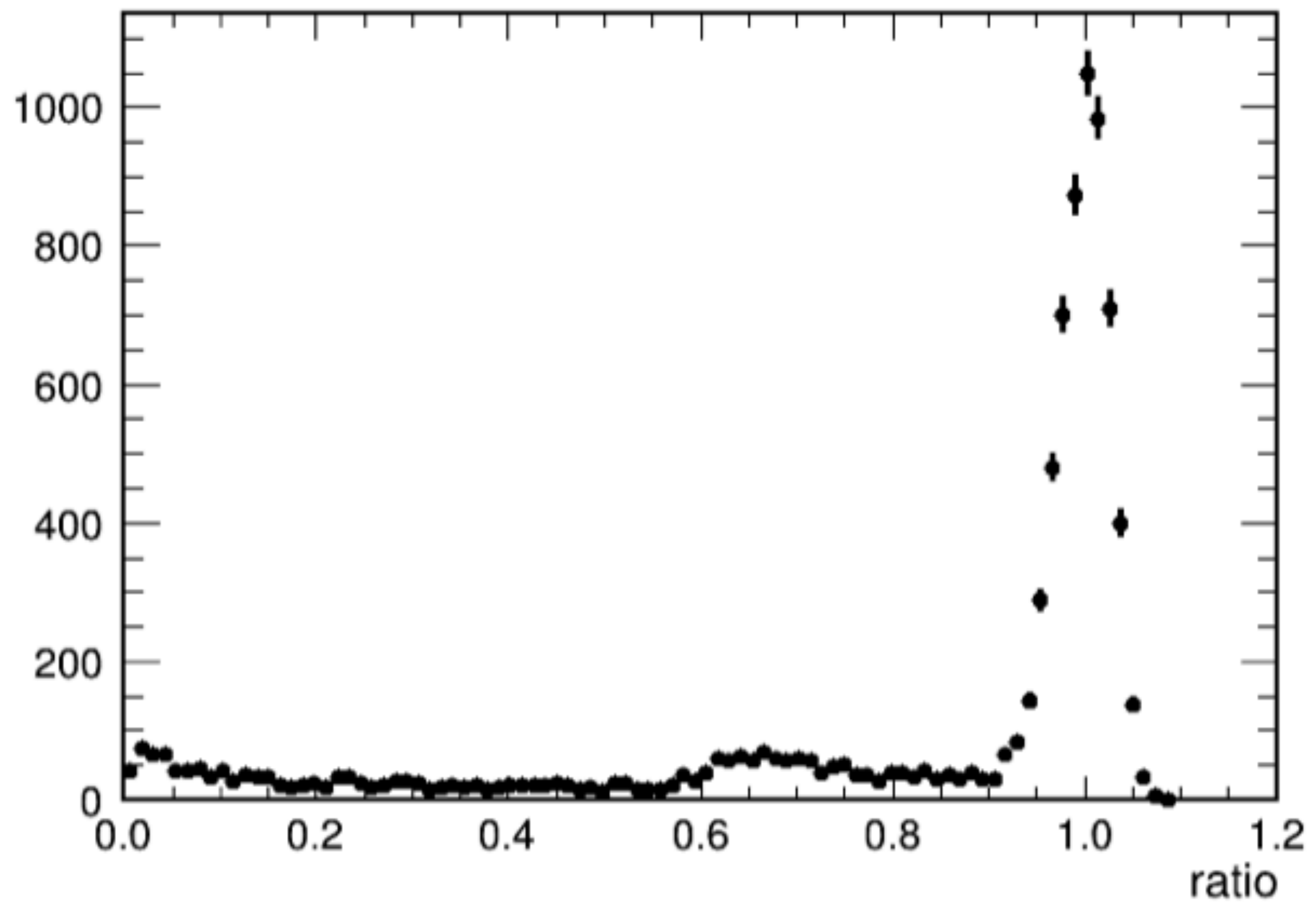
Select stopping beam muon

- Select beam muon: BeamQualityCut & Michel_score > 0.6
- By Bethe-Bloch formula, we can map between KE and residue range.

Thanks Heng-Ye and Tingjun for providing the method!

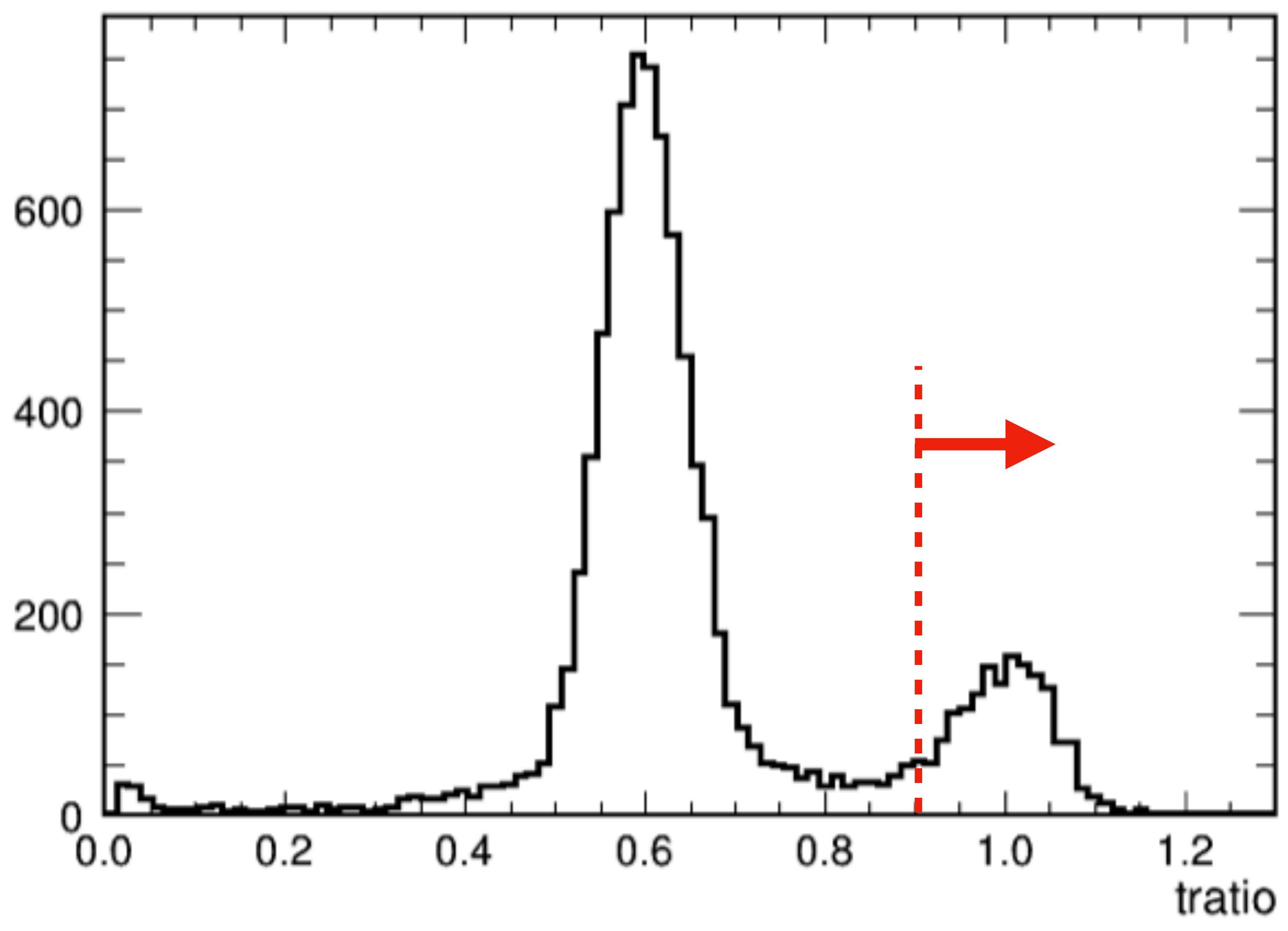
- Define Ratio = TrackLength / RangeFromKE($KE_{front\text{-}face}$)

MC true ratio



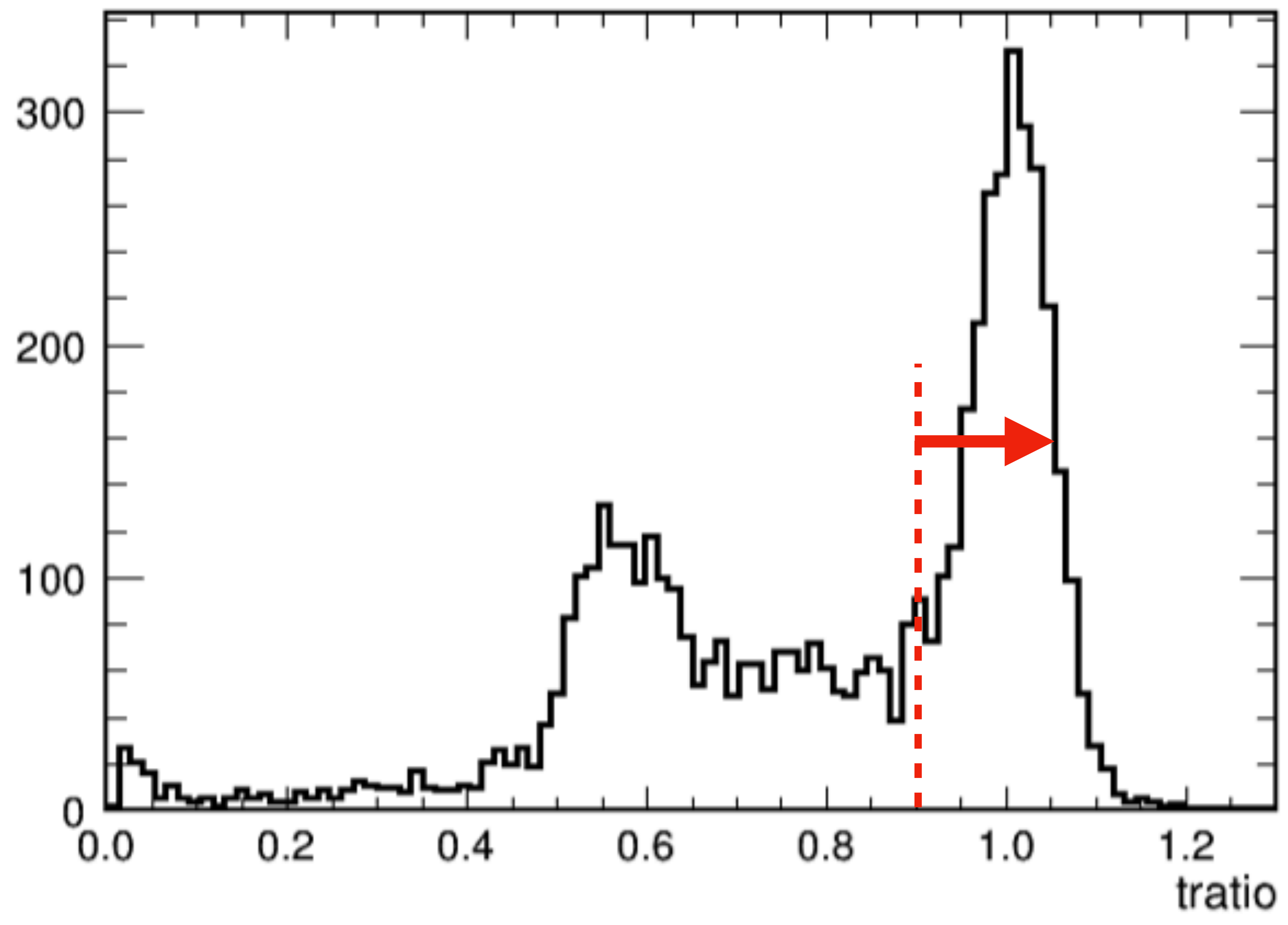
true_trklen / RangeFromKE(true_KEff)

MC reco ratio



reco_trklen / RangeFromKE(reco_KEff)

Data reco ratio



reco_trklen / RangeFromKE(reco_KEff)

Select Ratio > 0.9 as stopping muon sample

Beam momentum reweight

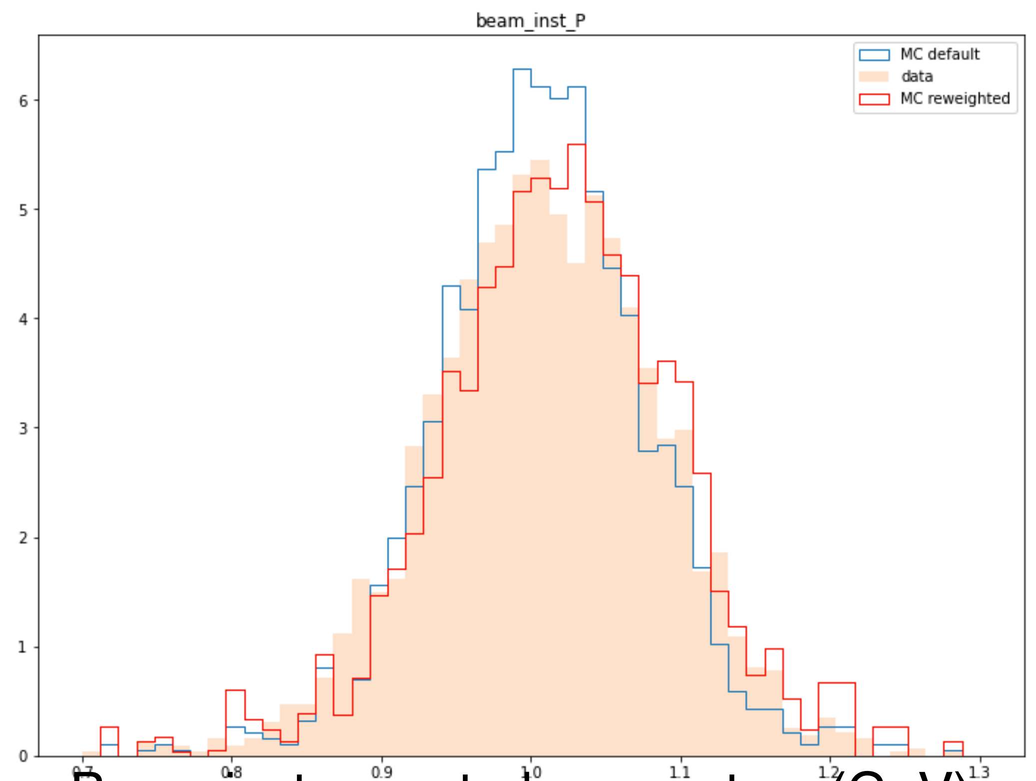
- A weight is assigned to each MC event

$$w = \frac{e^{-\frac{(p-\mu)^2}{2\sigma^2}}}{e^{-\frac{(p-\mu_0)^2}{2\sigma_0^2}}}$$

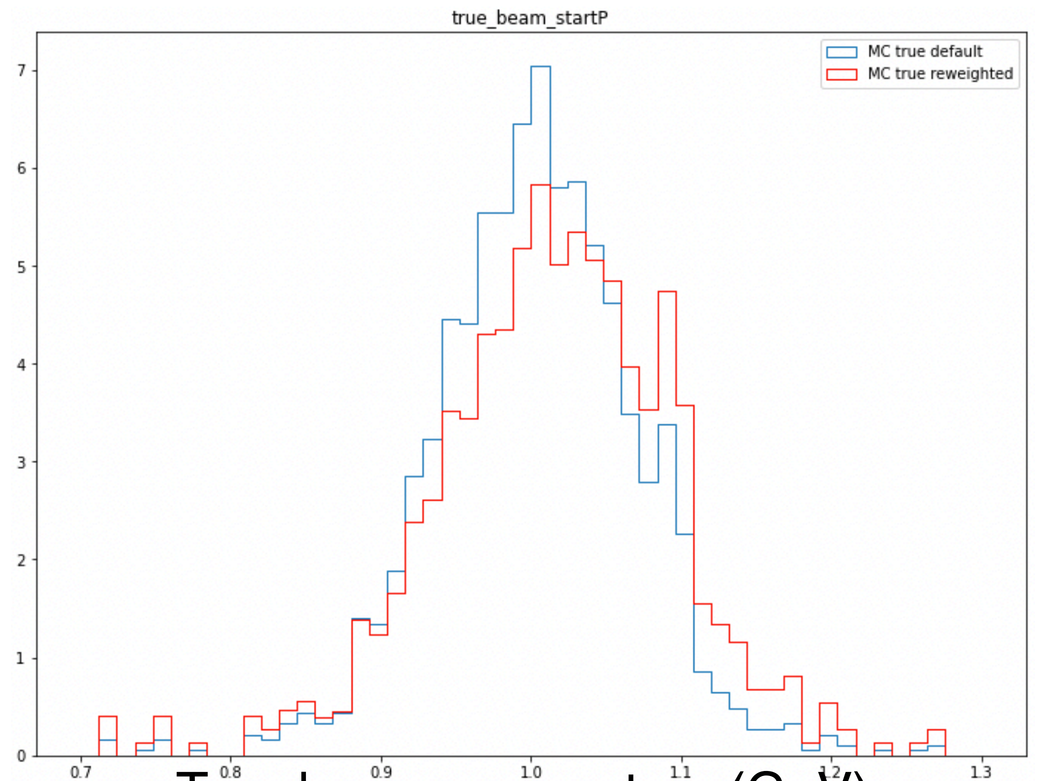
- p is the true momentum in each MC event
- μ_0 and σ_0 are fit to MC true momentum
- μ and σ are two **fit parameters**

Cutoff $w \leq 3$

- χ² fit is performed for the best agreement on KEff_from_range between data and MC stopping muons.

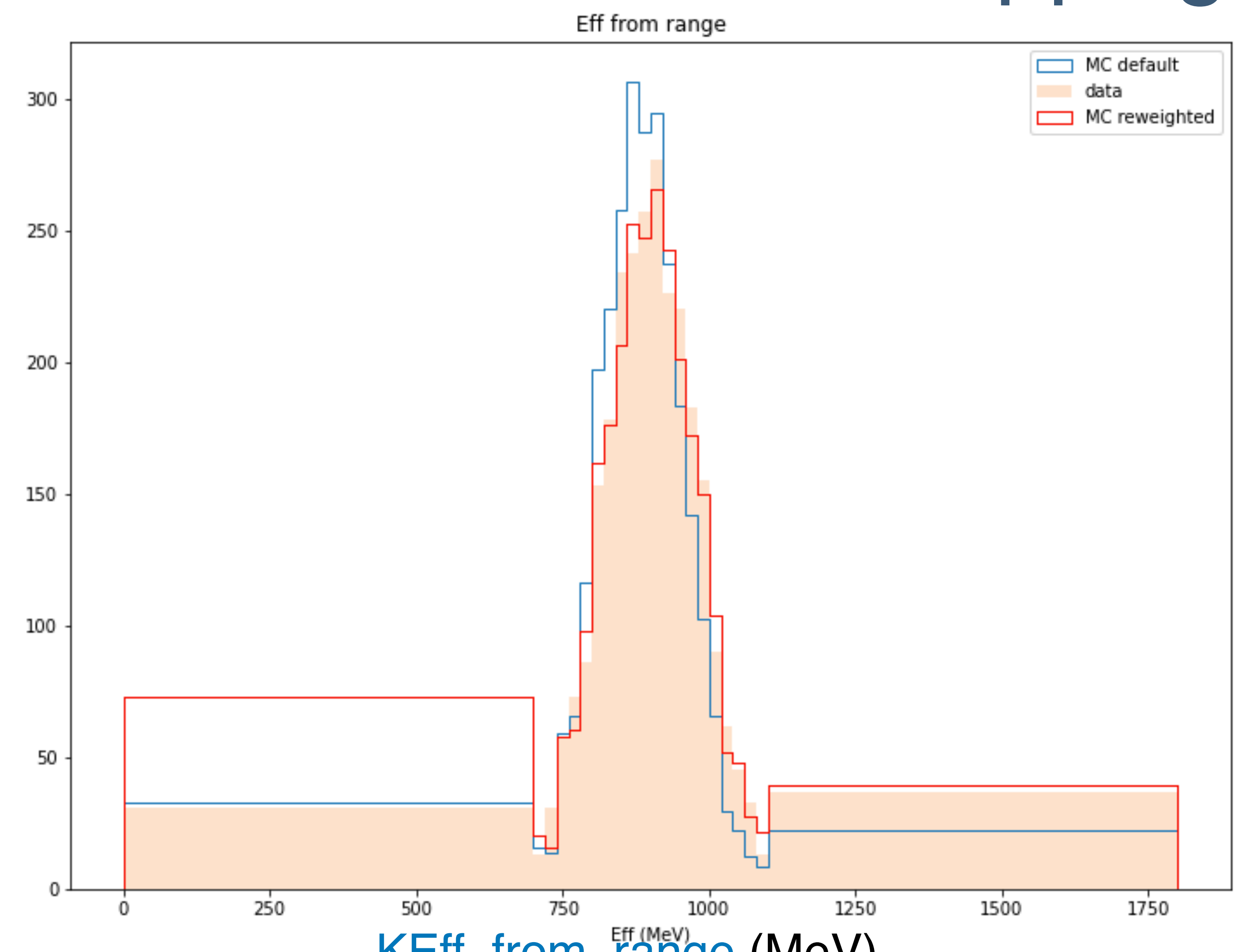


Beam instrumented momentum (GeV)

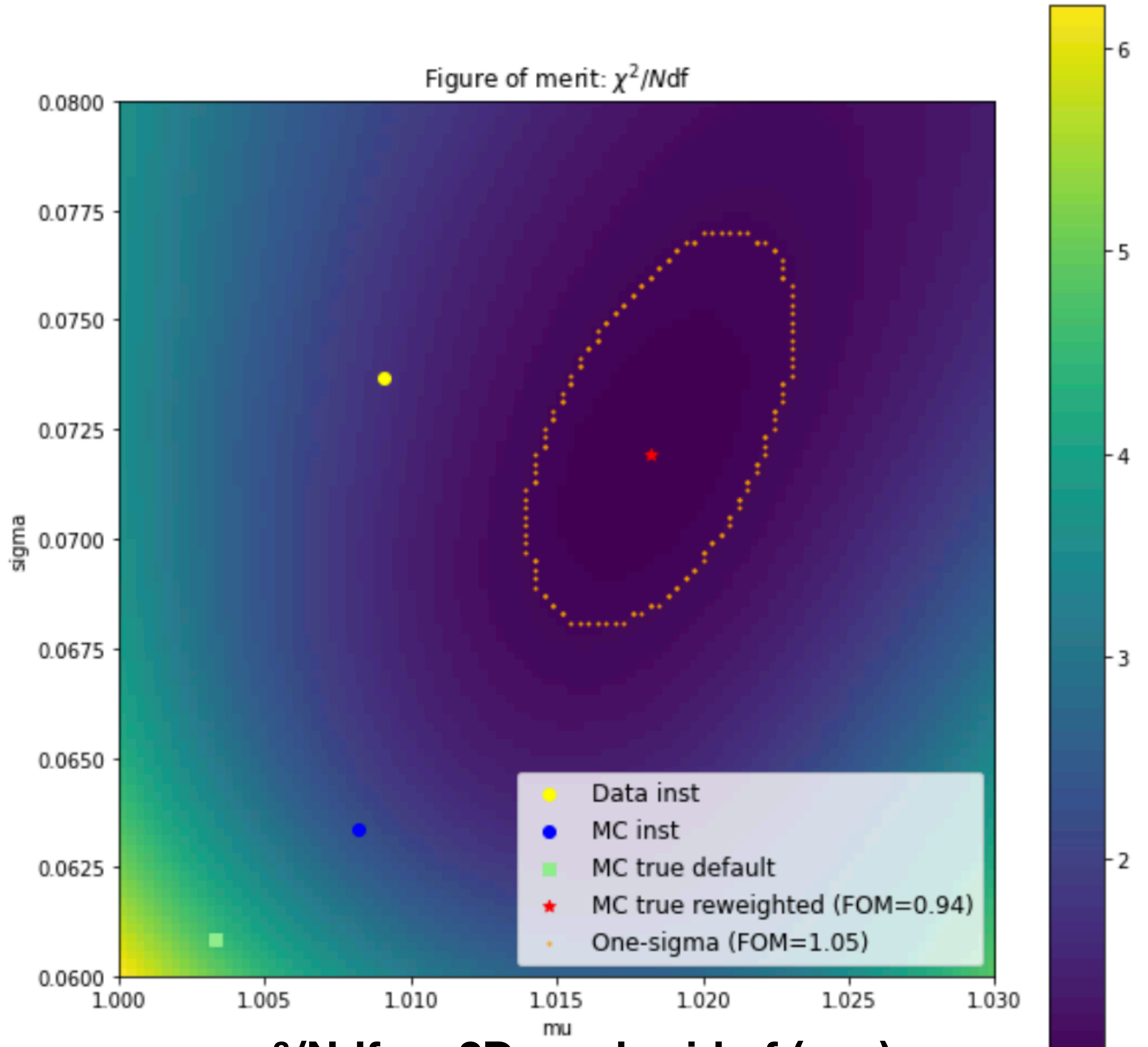


True beam momentum (GeV)

The front-face KE calculated by Bethe-Bloch formula given the reco track length (for a stopping muon).

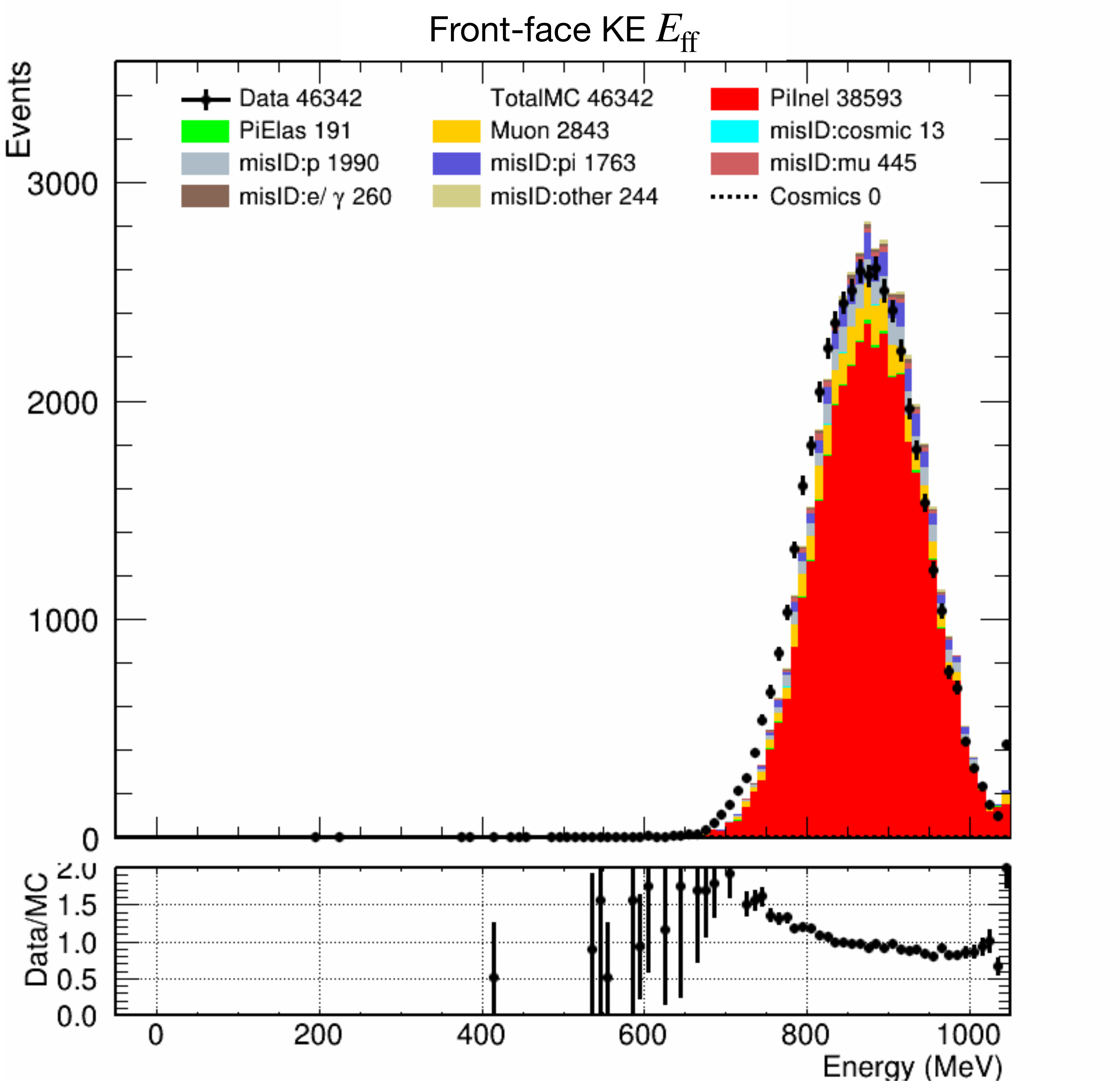


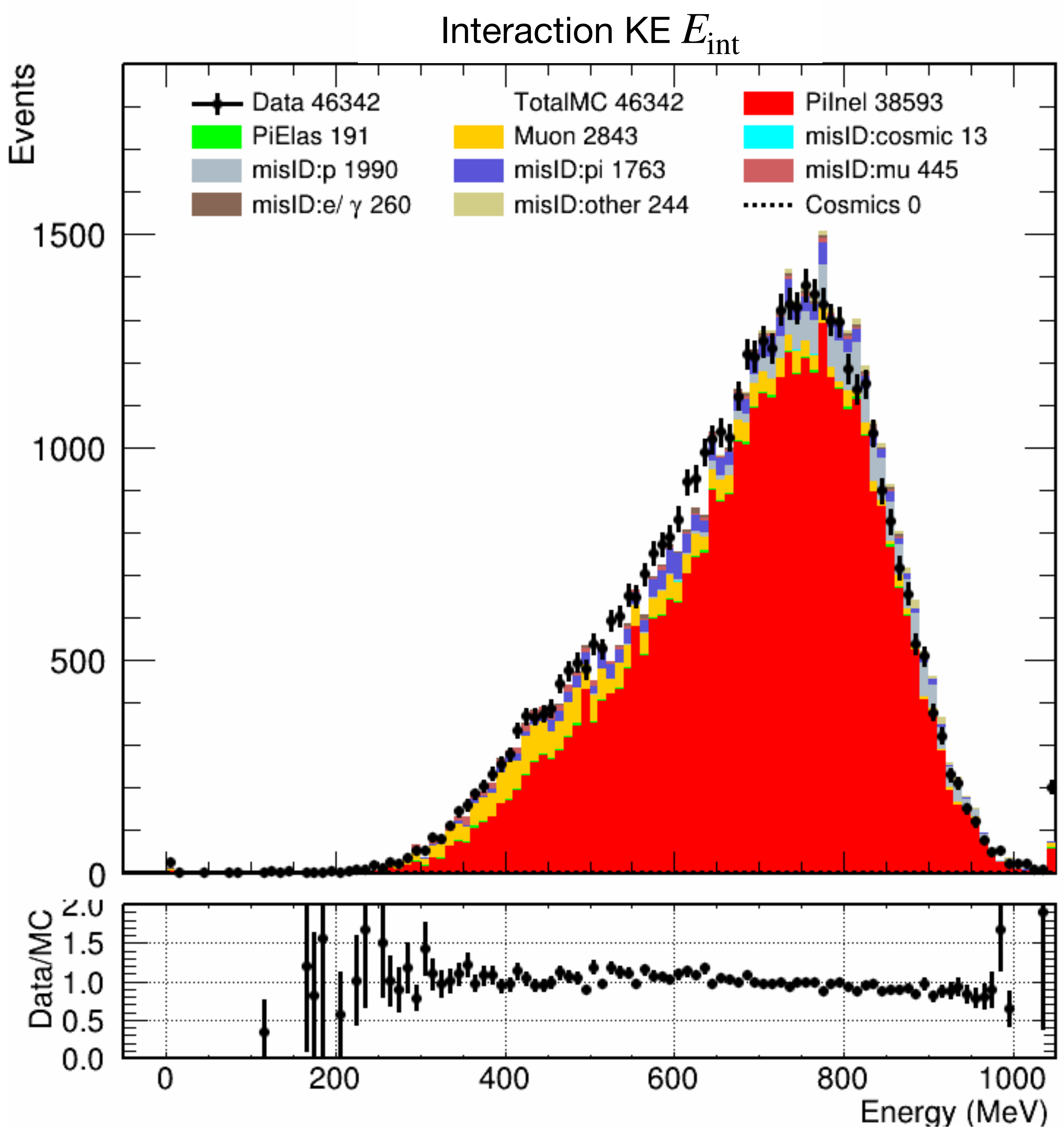
KEff_from_range (MeV)



χ²/Ndf on 2D meshgrid of (μ, σ)

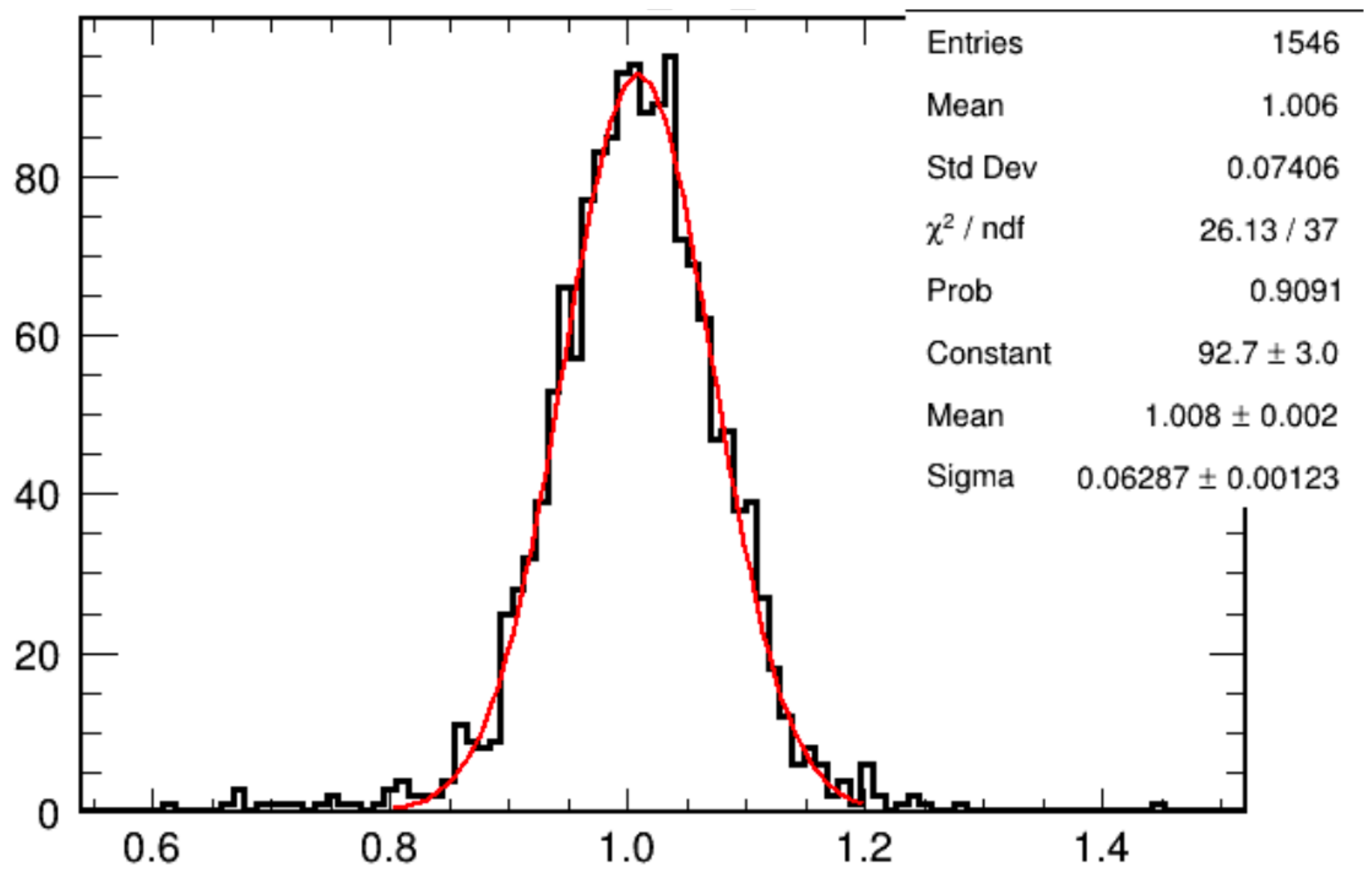
After reweighting



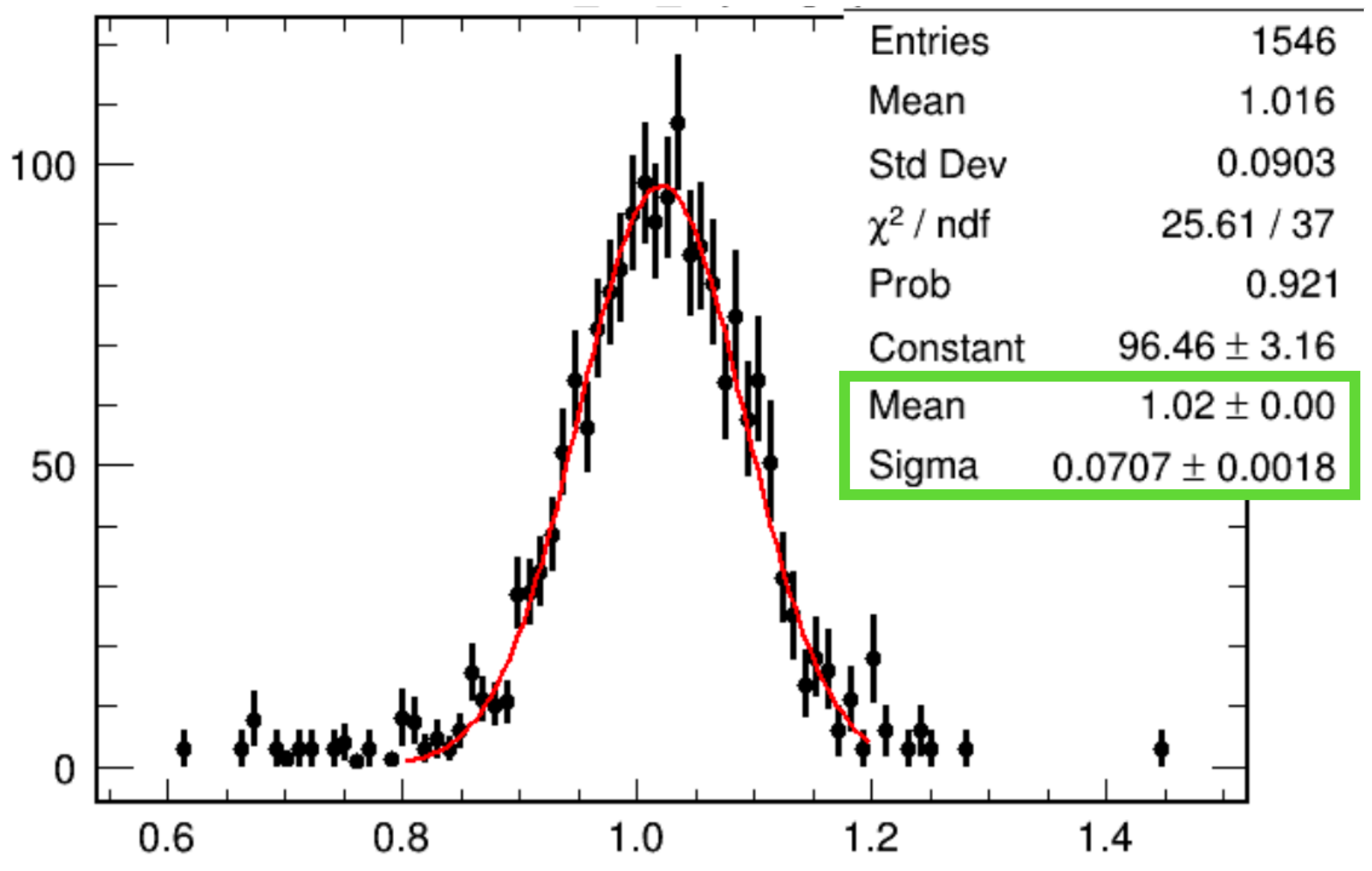


Beam momentum reweighting

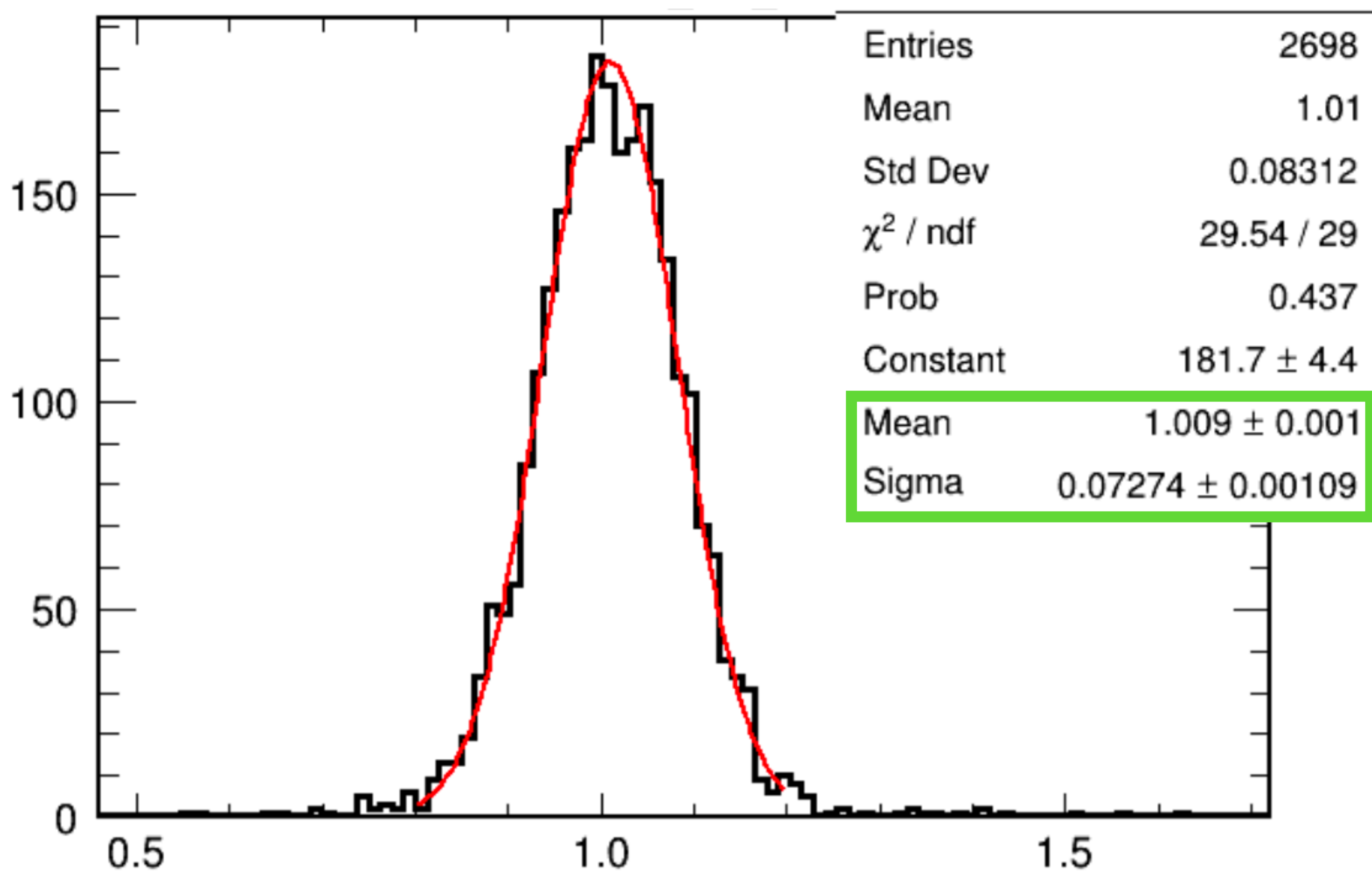
- After momentum reweighting, data and MC should have consistent **true momentum** distributions.
- However, the beam **instrumented momentum** distributions are still different.



MC beam momentum



MC reweighted beam momentum



Data beam momentum

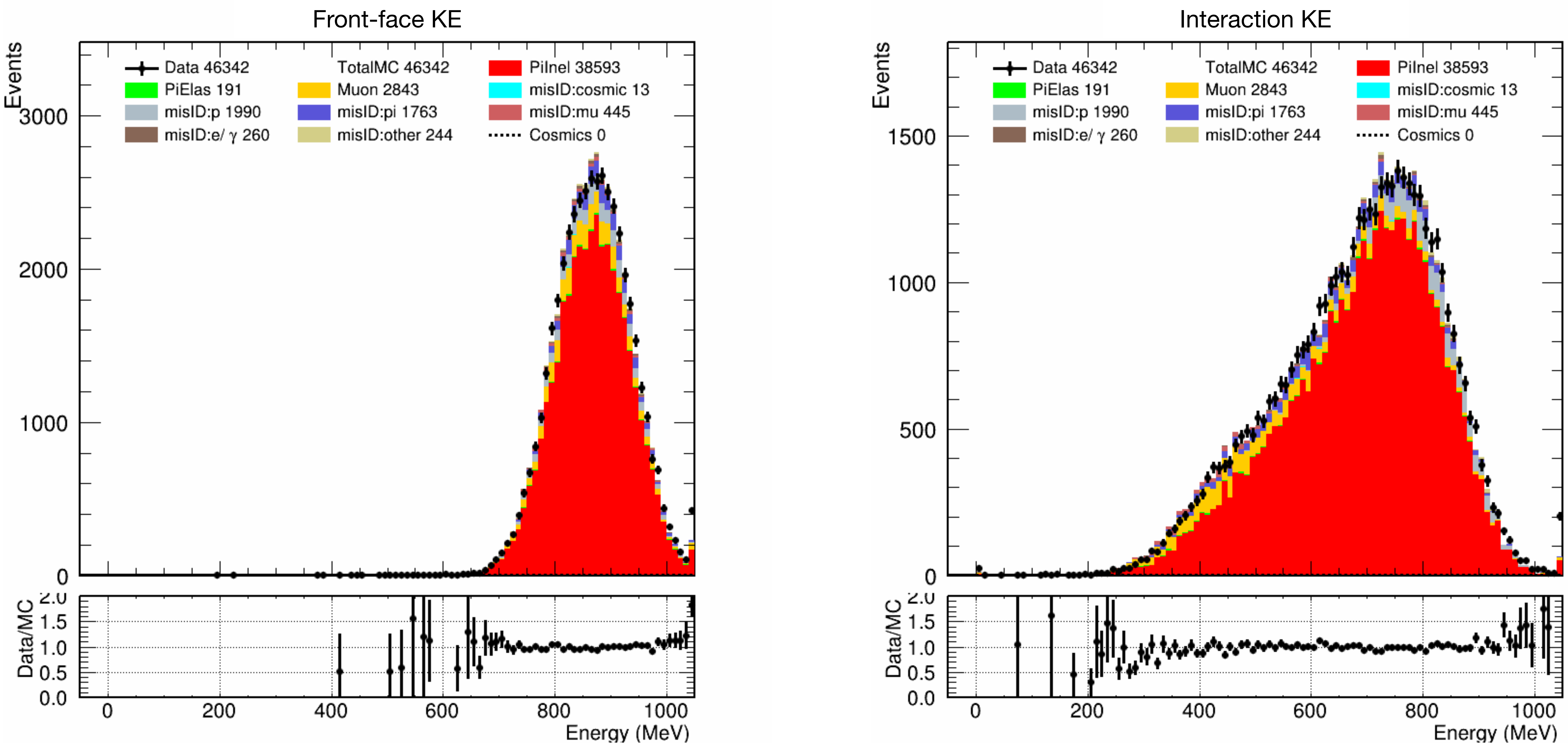
Difference in μ: upstream E loss. $\mu_{\text{data}} - \mu_{\text{MC}}$

- We add an extra random Gaus(-9.85, 17.76) MeV to each MC event.

Difference in σ: momentum resolution. Extra smearing: $\sqrt{\sigma^2_{\text{data}} - \sigma^2_{\text{MC}}}$

After reweighting and extra shifting/smearing

MC beam_inst_P add Gaus(-9.85, 17.76) MeV



Updates

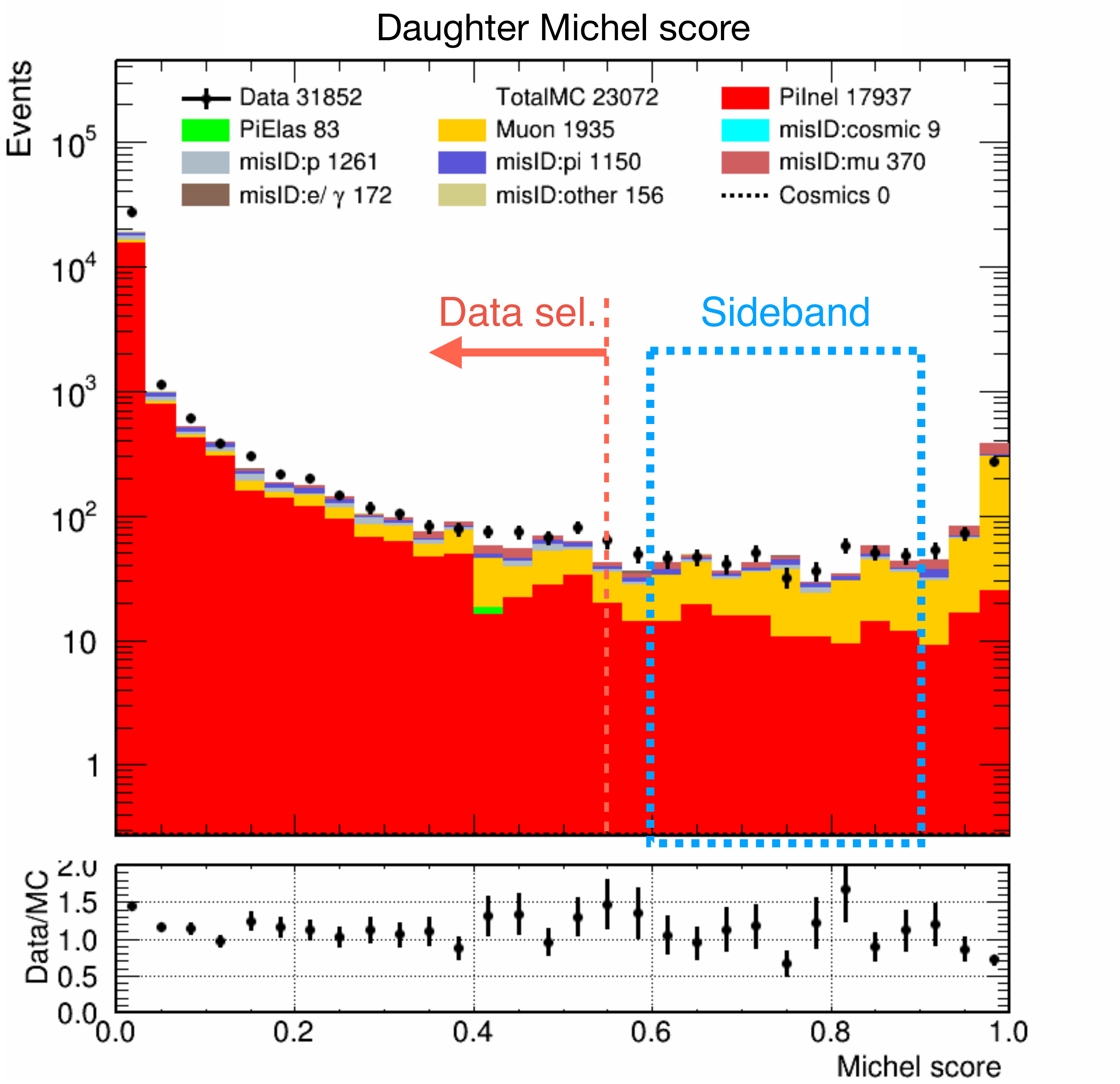
- Reweighting MC
- **Background constraints**
- Unfolding and error propagation
- Results (mainly from fake data)

Background constraints

- A data-driven method is used to account for difference of background fractions in data and MC
  - $$N_{\text{reco}}^{\text{sig}} = N_{\text{reco}} \cdot \left(1 - \sum_i f_i^{\text{data}}\right) = N_{\text{reco}} \cdot \left(1 - \sum_i f_i^{\text{MC}} \cdot \alpha_i\right)$$
- α_i is the scale factor for background i.
  - Three major backgrounds are considered.
    - Muon background
    - Proton background
    - Secondary pion background

Background constraints

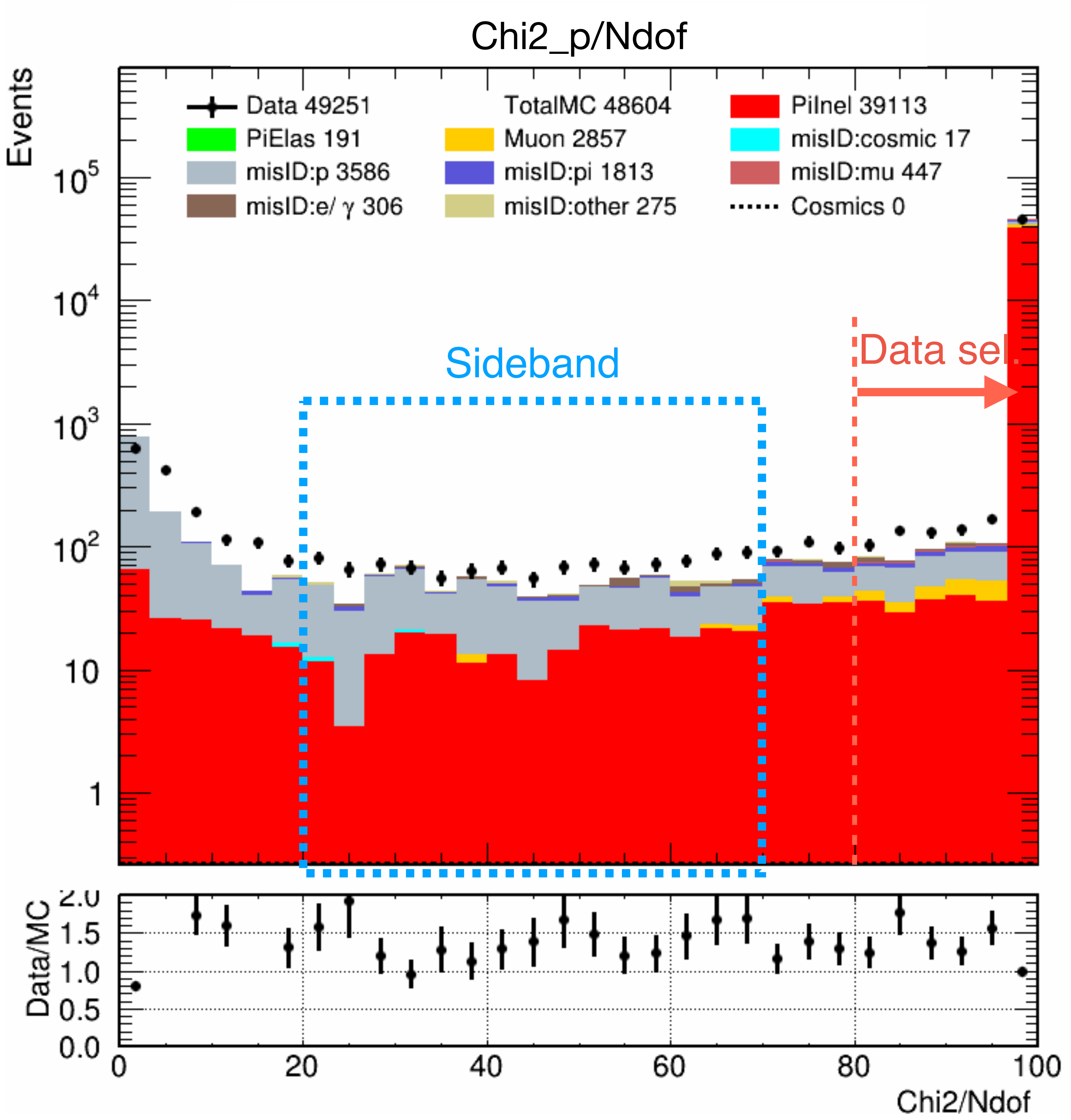
- α_i is fitted in the sideband of a distribution where background i dominates.
- E.g. the Michel score distribution is used for muon background constraint.



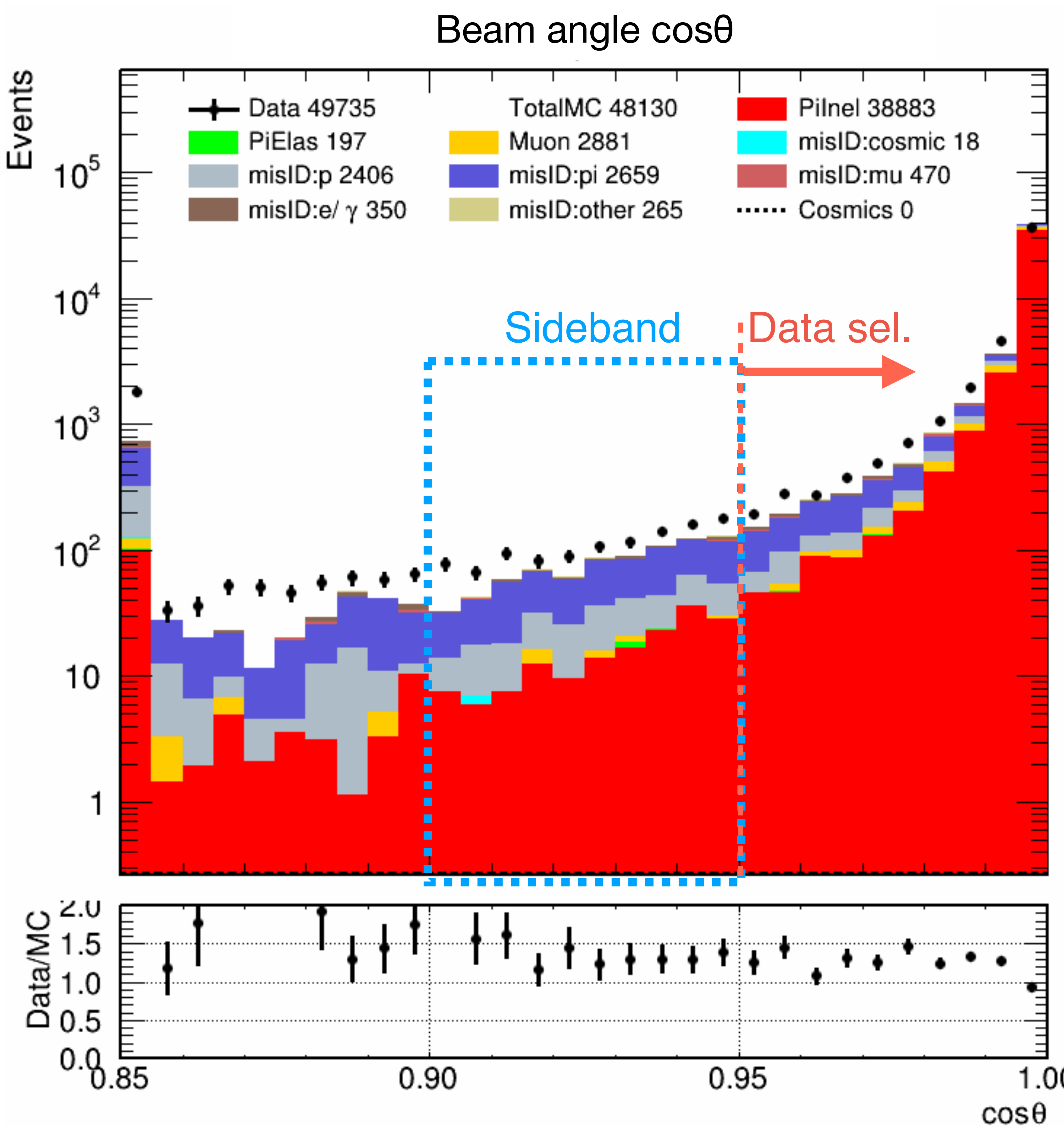
- The fit is done on the sample after all other selections except the Michel score cut.
- Fitted result: 0.65 ± 0.11

Background constraints

- Similarly for proton and secondary pion backgrounds, we use the Chi2_p/Ndof and the beam angle distributions respectively.



Fitted result:
1.65 ± 0.13



Fitted result:
1.47 ± 0.14

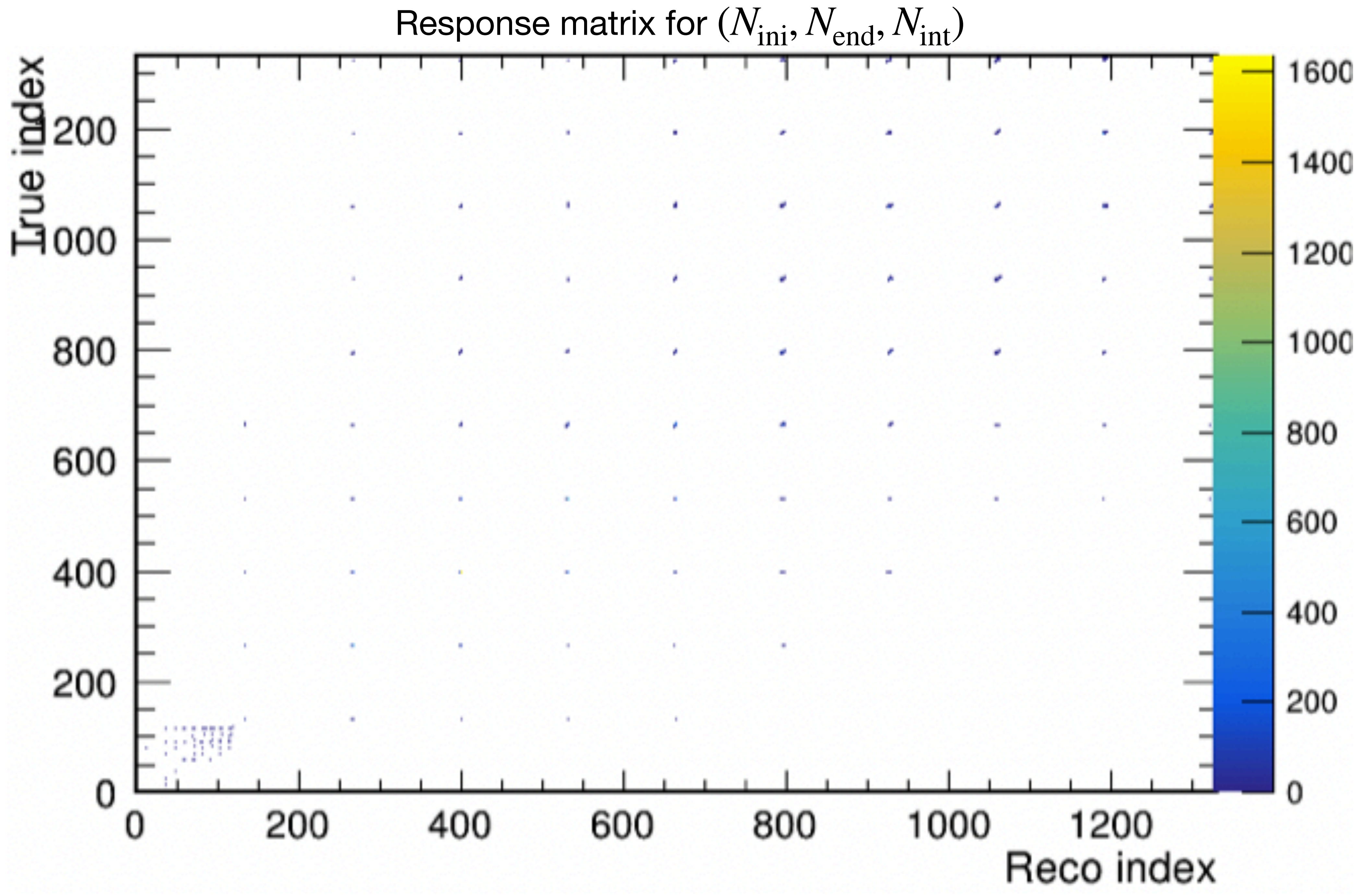
Updates

- Reweighting MC
- Background constraints
- **Unfolding and error propagation**
- Results (mainly from fake data)

THE UNIVERSITY OF CHICAGO DUNE

Multi-dimentional unfolding

- Because N_{ini}, N_{end} and N_{int} are related, we combine them as one variable $(N_{\text{ini}}, N_{\text{end}}, N_{\text{int}})$, and unfold them together.
  - Variable index = $N^2 \cdot \text{ID}_{\text{ini}} + N \cdot \text{ID}_{\text{end}} + \text{ID}_{\text{int}}$, where N is the number of slices
- The d'Agostini (iterative Bayesian) method is used to model the unfolding matrix
  - 10 iterations by default (need optimization)



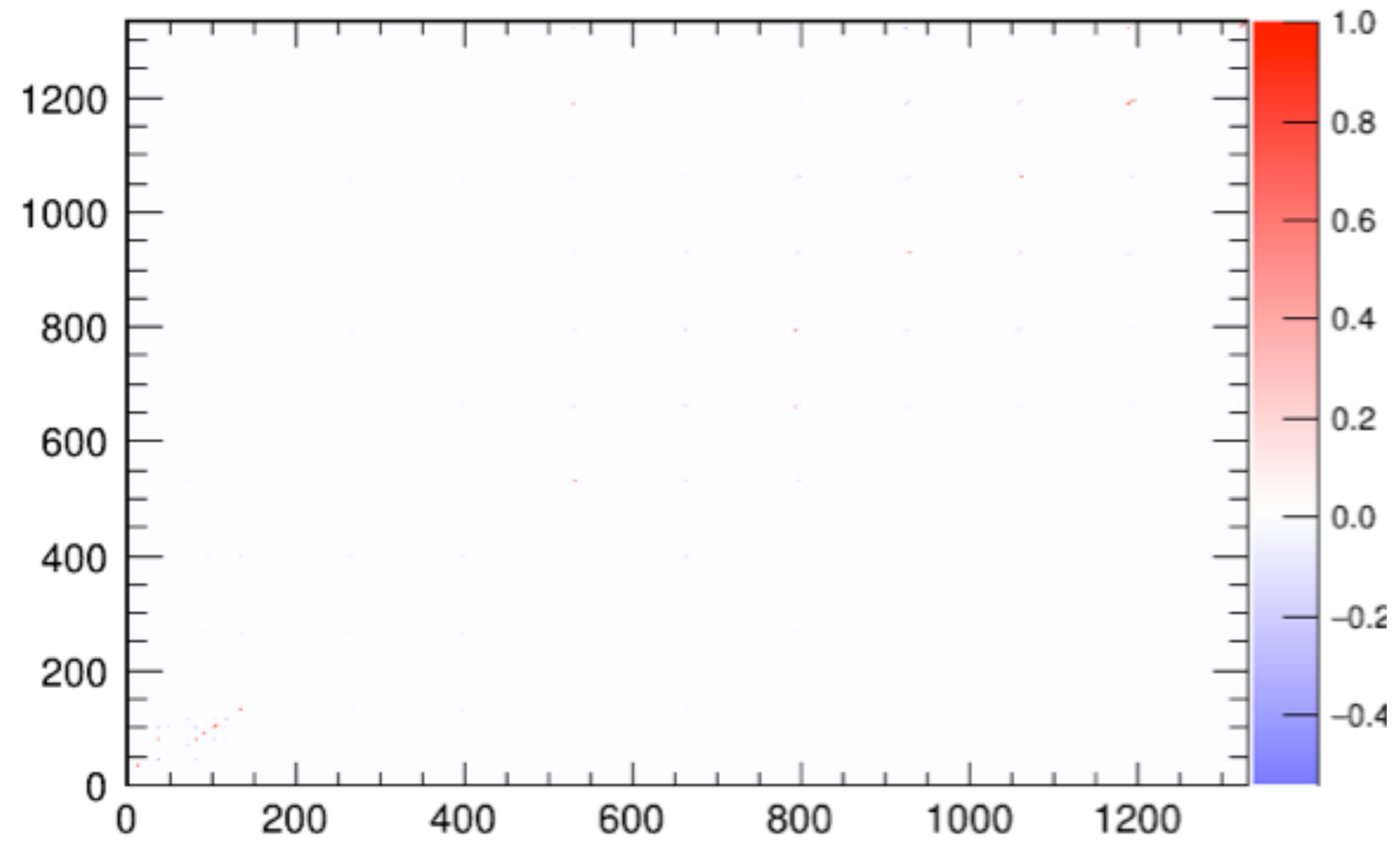
Error propagation

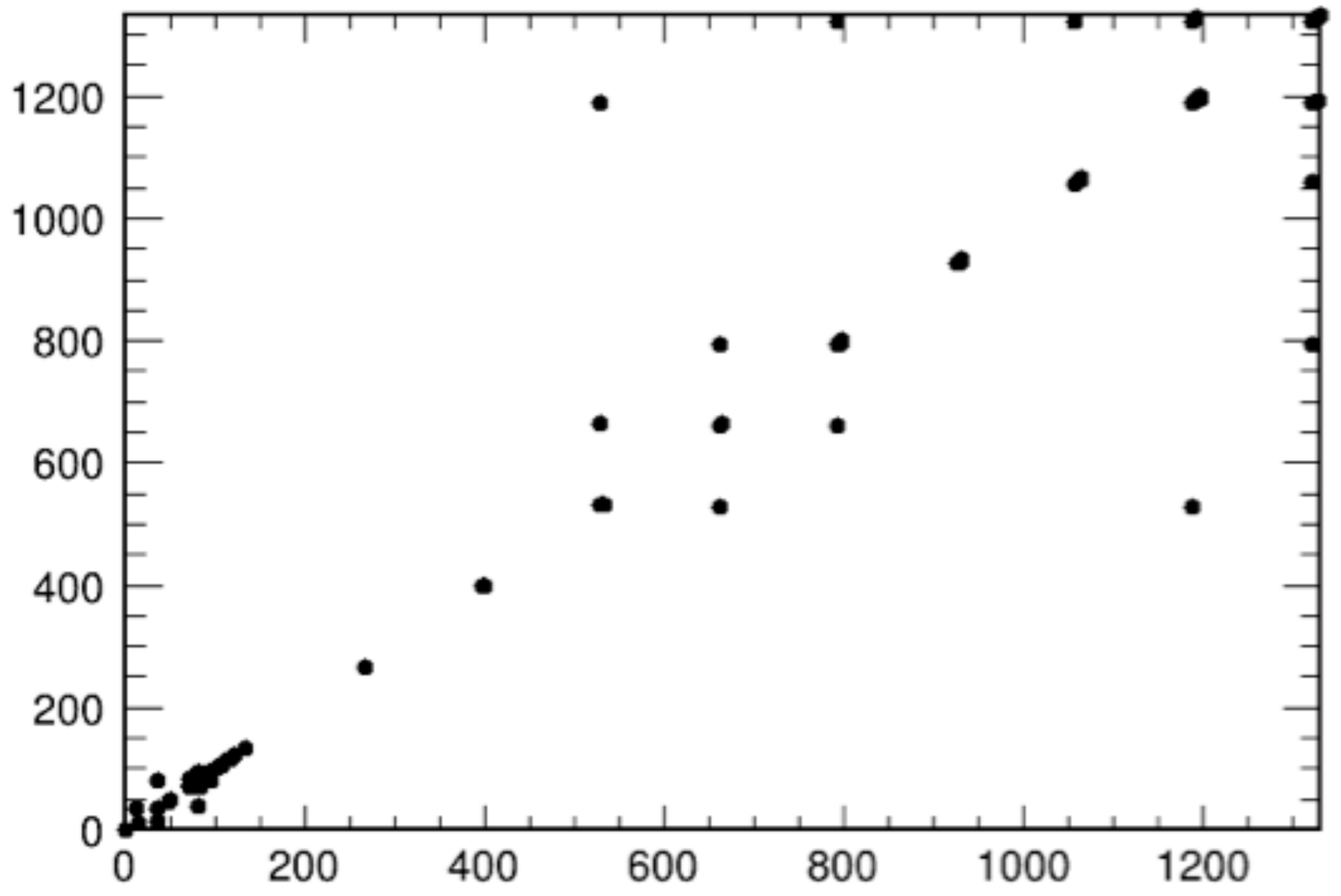
- Covariance matrix $V = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}$

- Jacobian matrix $J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$

- $V_f = J \cdot V_x \cdot J^T$

- In our case, x is the entry in each $(N_{\text{ini}}, N_{\text{end}}, N_{\text{int}})$ bin, and f is the cross-section.





Covariance matrix provided by RooUnfold
3D unfolding of $(N_{\text{ini}}, N_{\text{end}}, N_{\text{int}})$

Updates

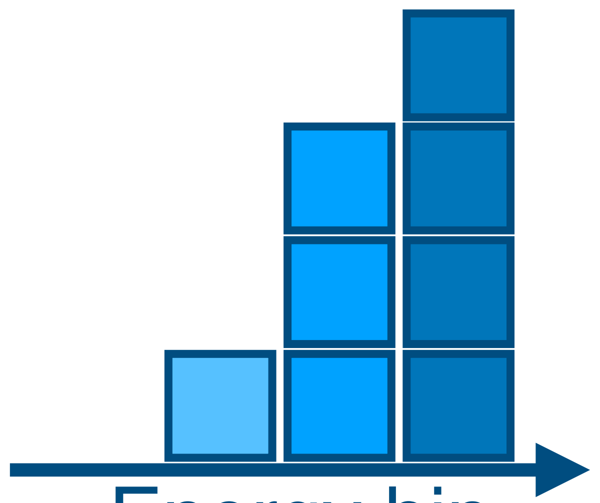
- Reweighting MC
- Background constraints
- Unfolding and error propagation
- **Results (mainly from fake data)**

Fake data

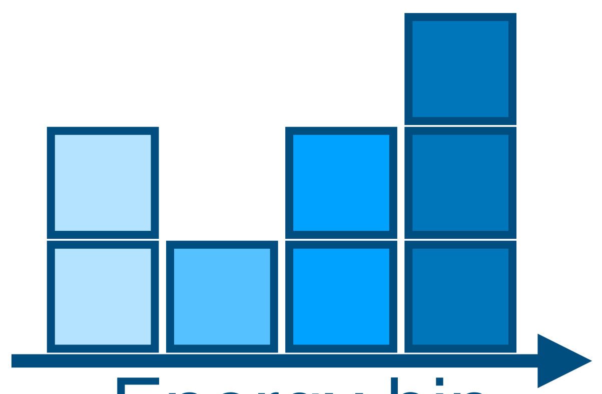
3D unfolding 10 iterations
(error bars are propagated from the covariance matrix provided by RooUnFold)

$$N_{\text{inc}}(i) = \sum_{j=i}^{N} N_{\text{end}}(j) - \sum_{j=i+1}^{N} N_{\text{ini}}(j)$$

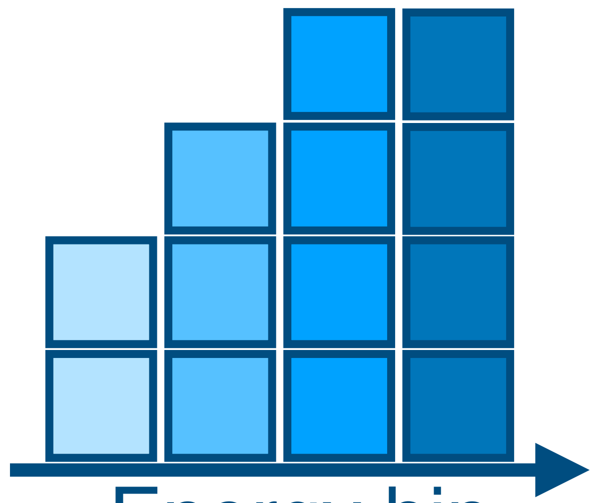
N_{initial}

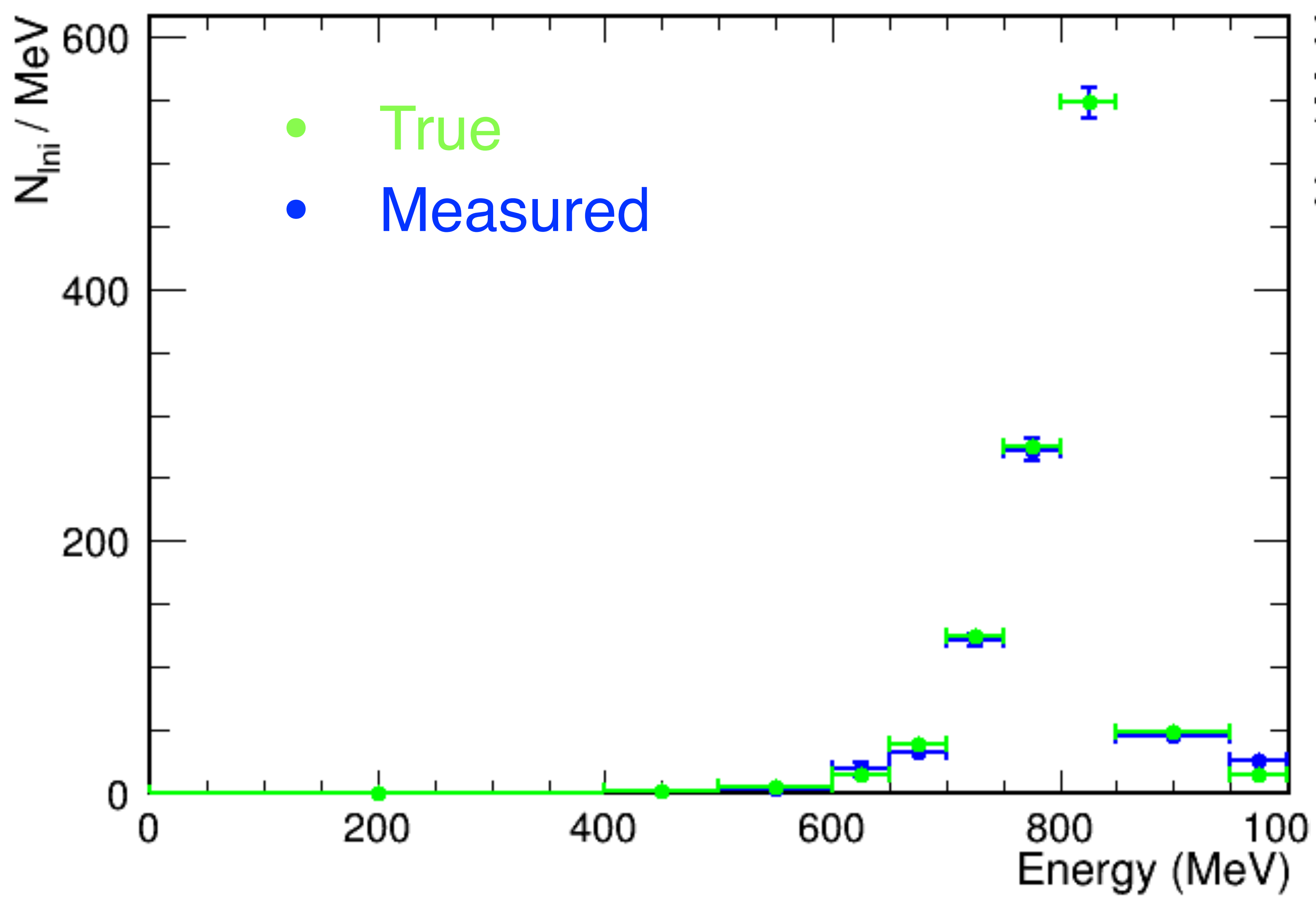


N_{end}

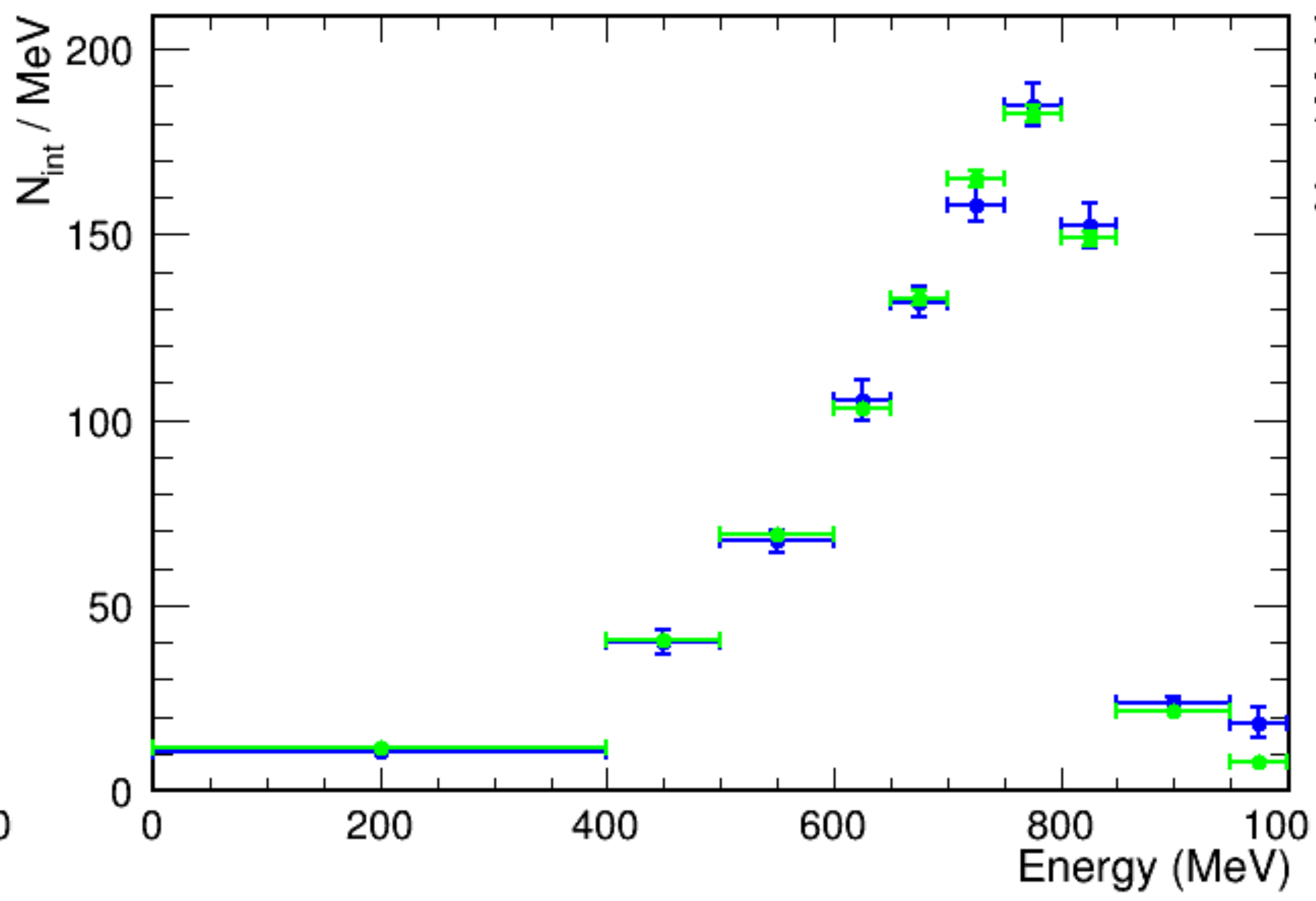


N_{incident}

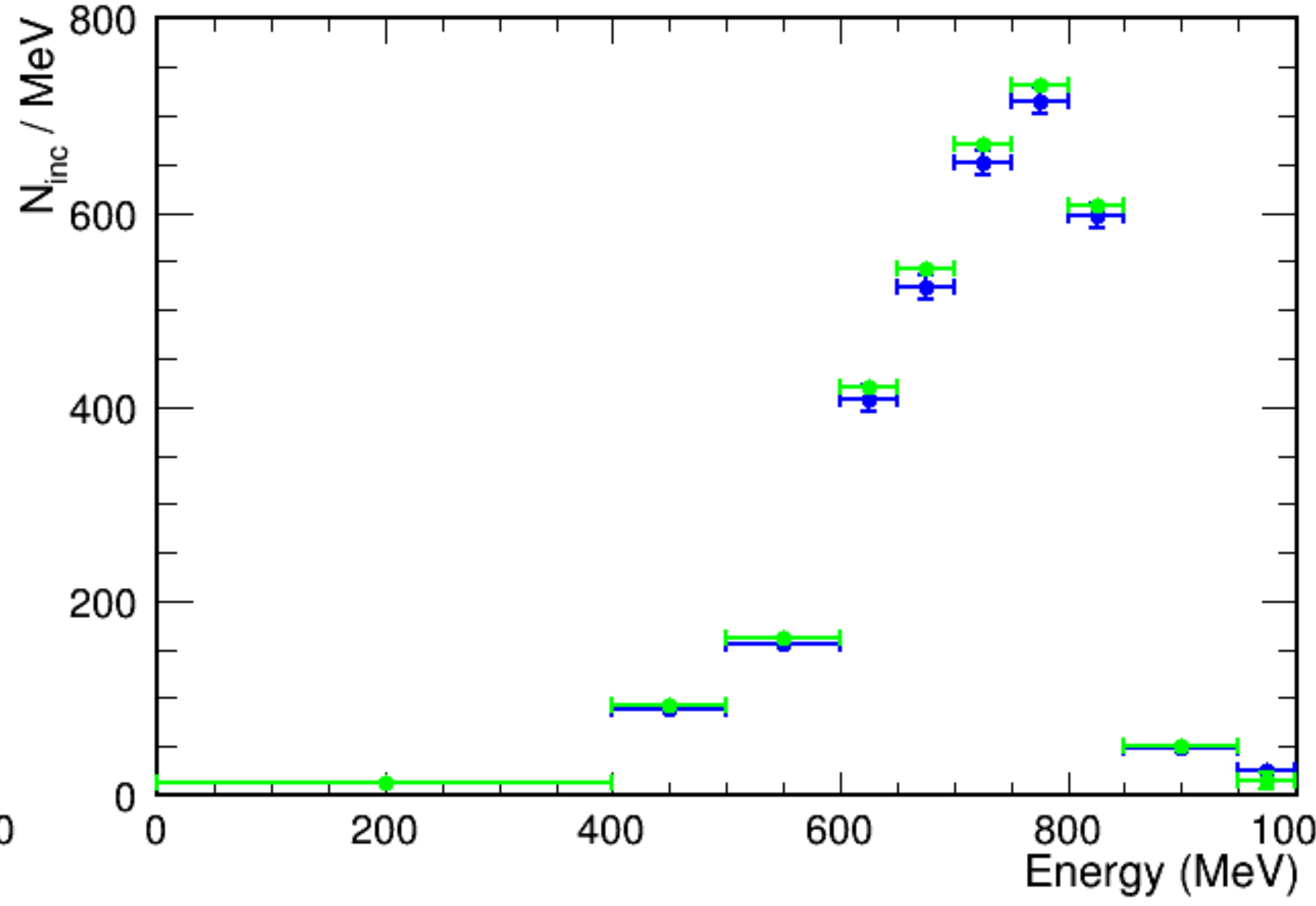




Initial histogram



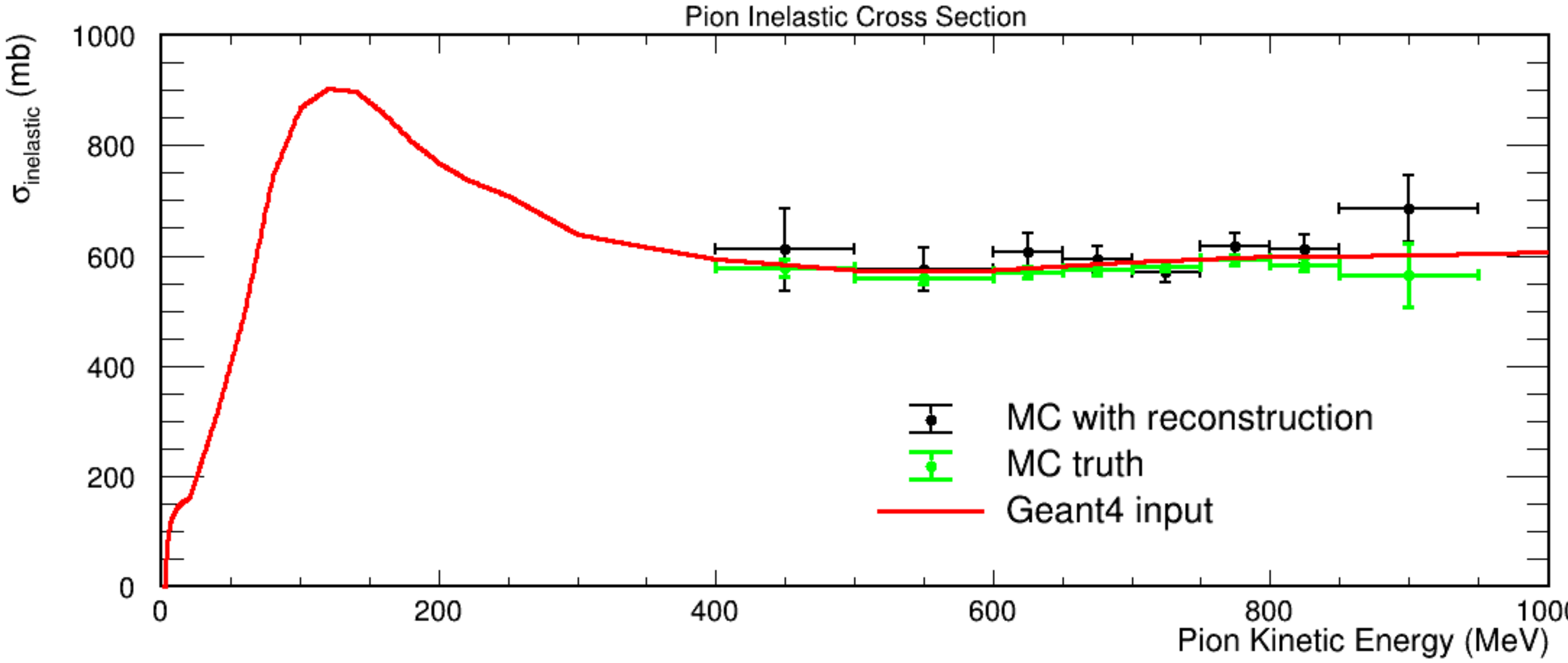
Interaction histogram

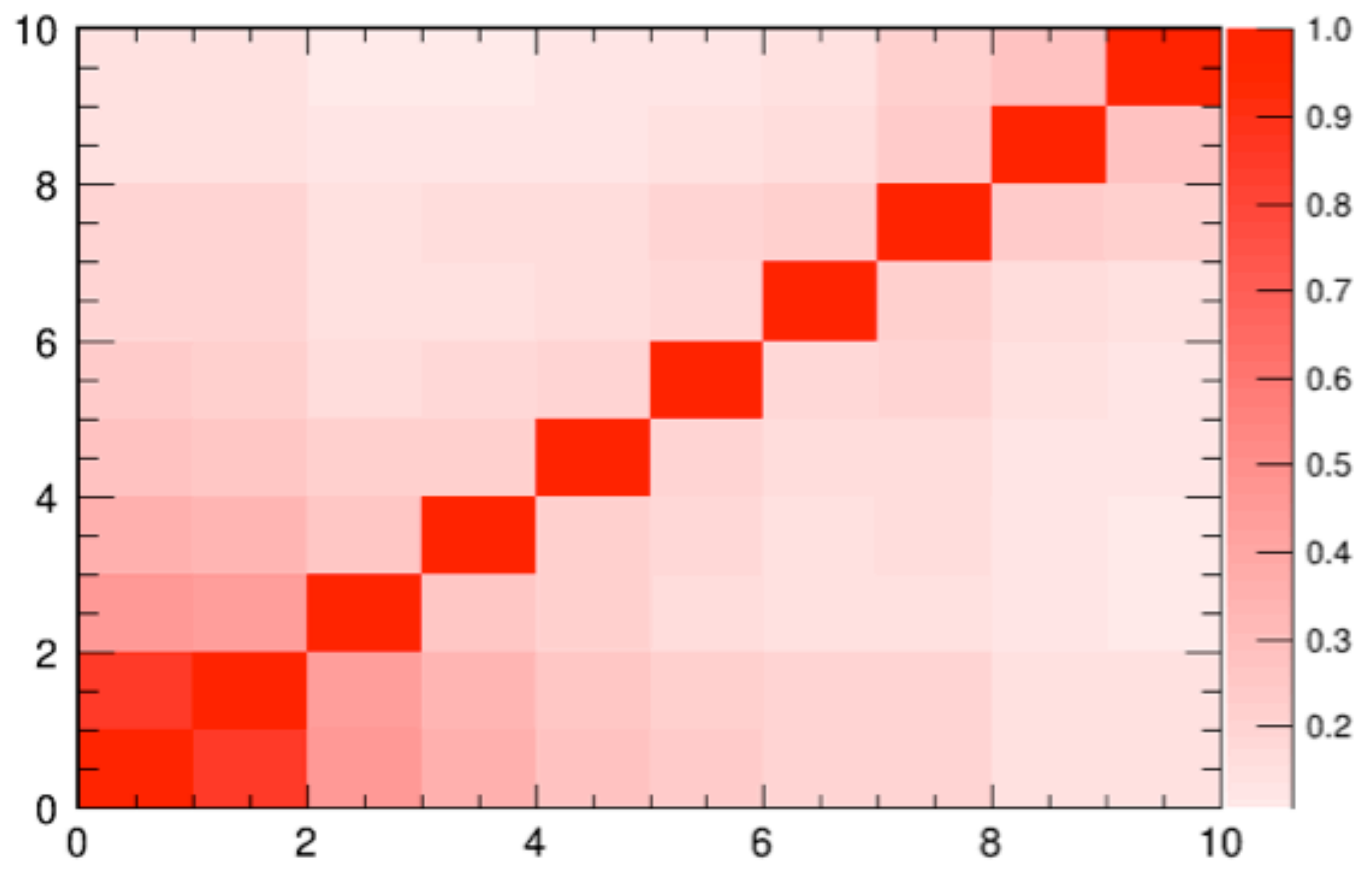


Incident histogram

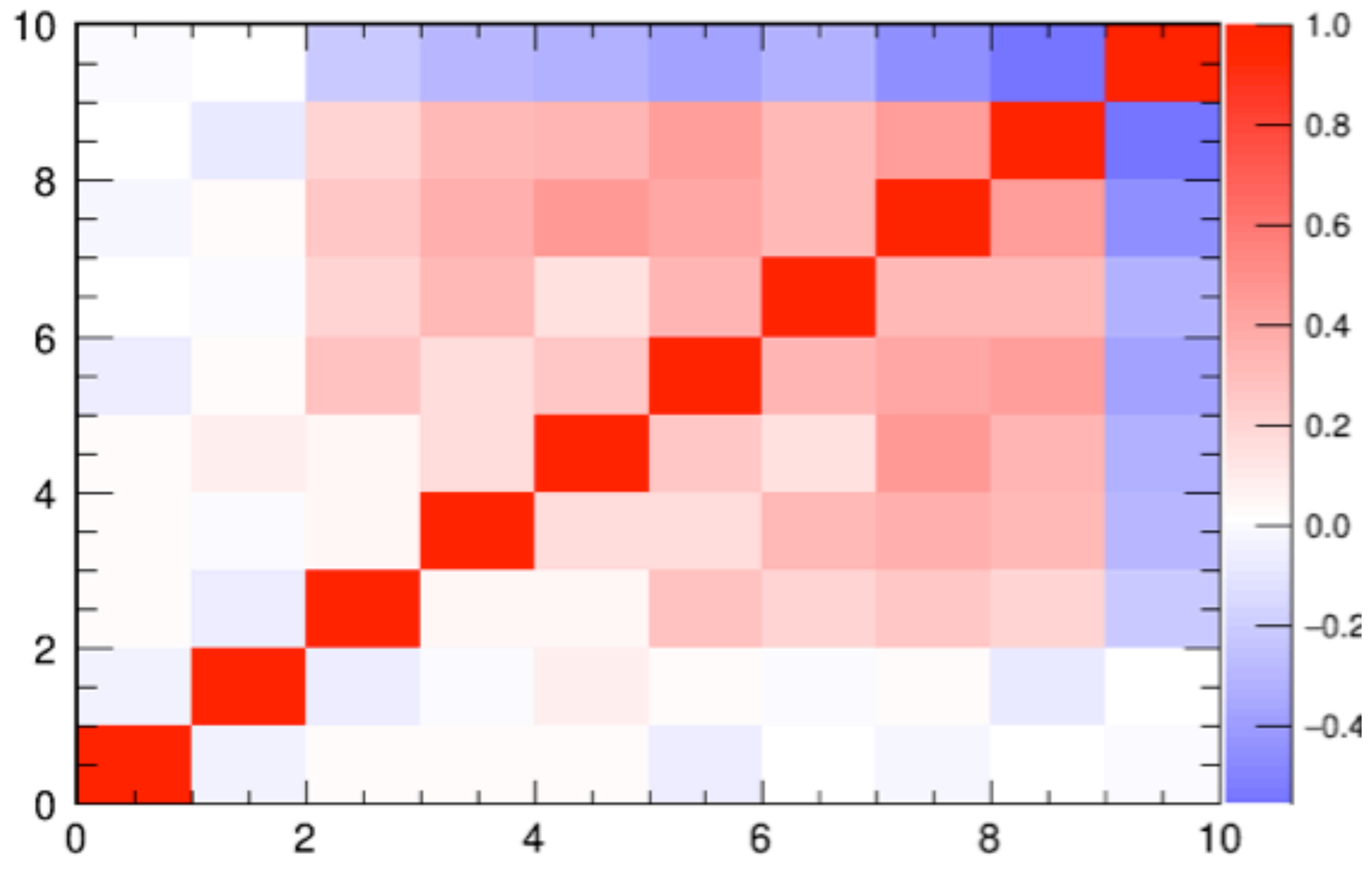
Fake data

3D unfolding 10 iterations
(error bars are propagated from the covariance matrix provided by RooUnFold)





Correlation matrix for true XS



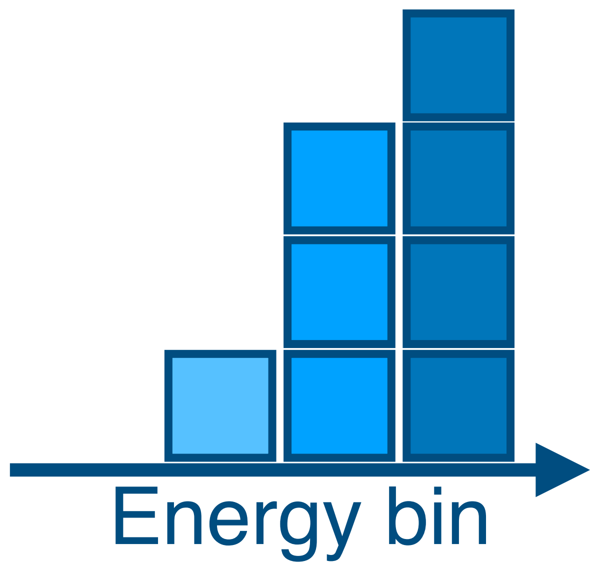
Correlation matrix for reco XS

Fake data

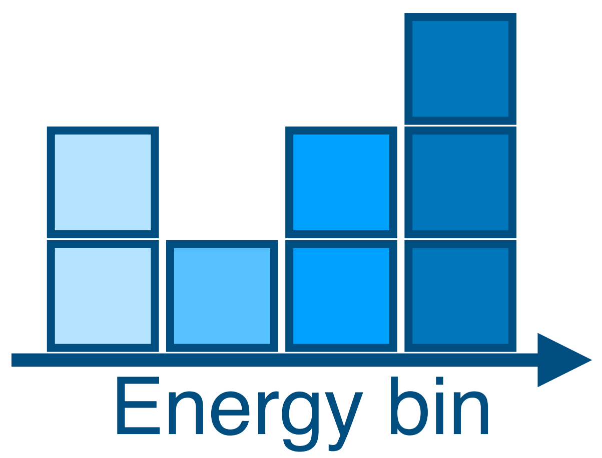
3D unfolding **200 iterations**
(error bars are propagated from the covariance matrix provided by RooUnFold)

$$N_{\mathrm{inc}}(i) = \sum_{j=i}^{N} N_{\mathrm{end}}(j) - \sum_{j=i+1}^{N} N_{\mathrm{ini}}(j)$$

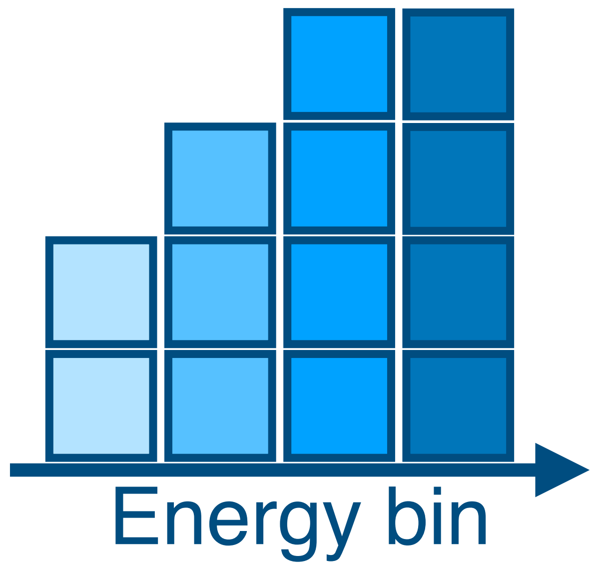
N_{initial}

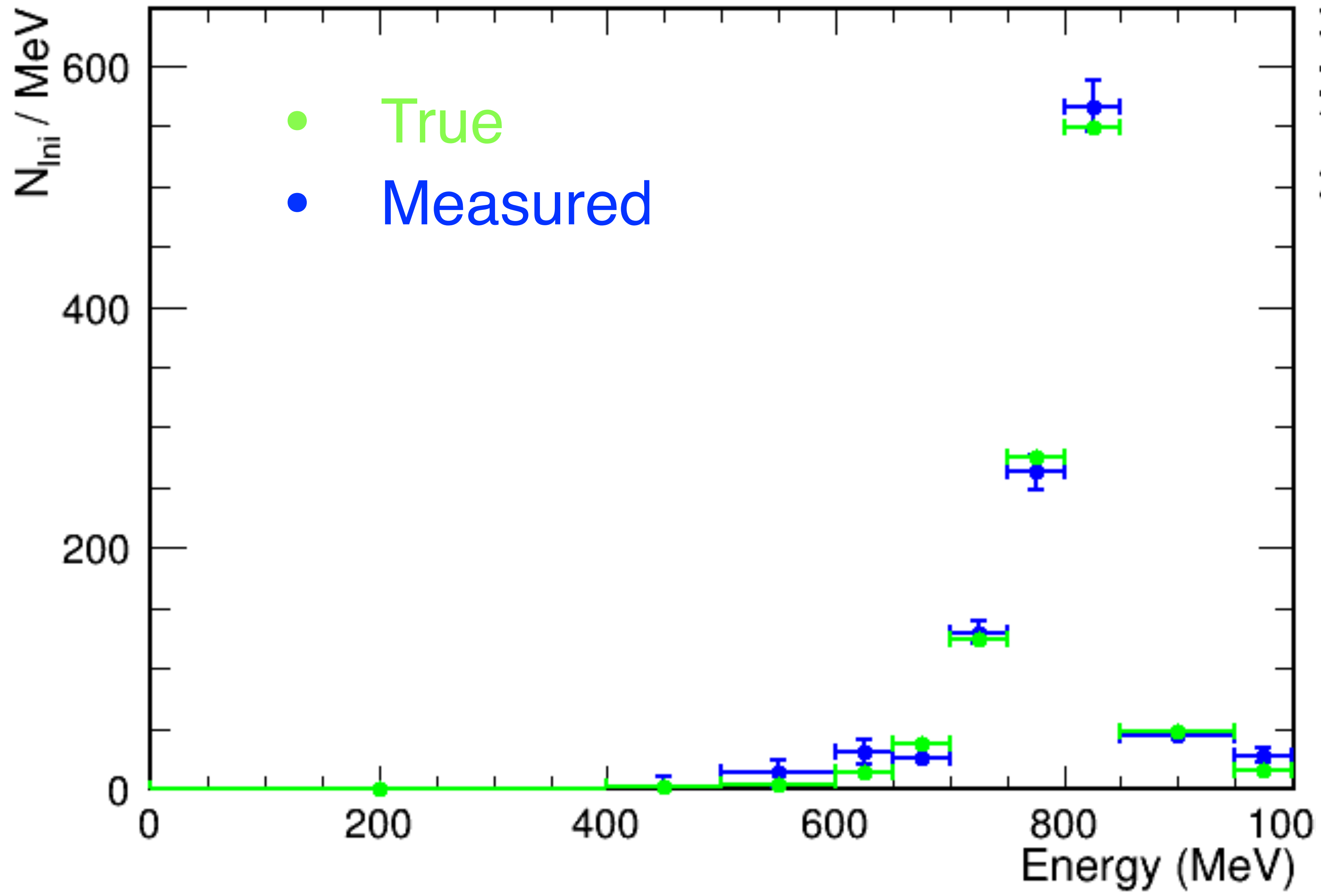


N_{end}

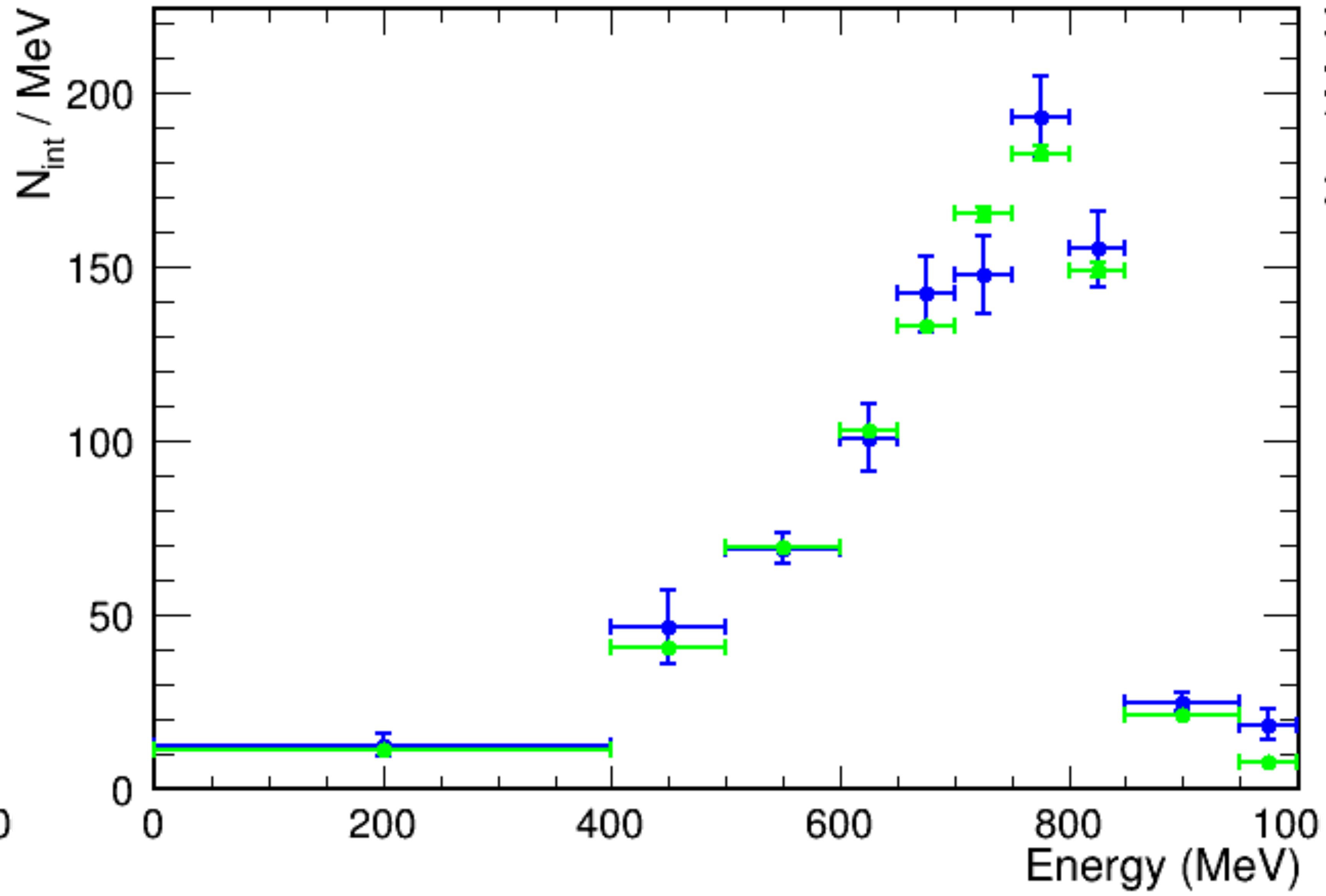


N_{incident}

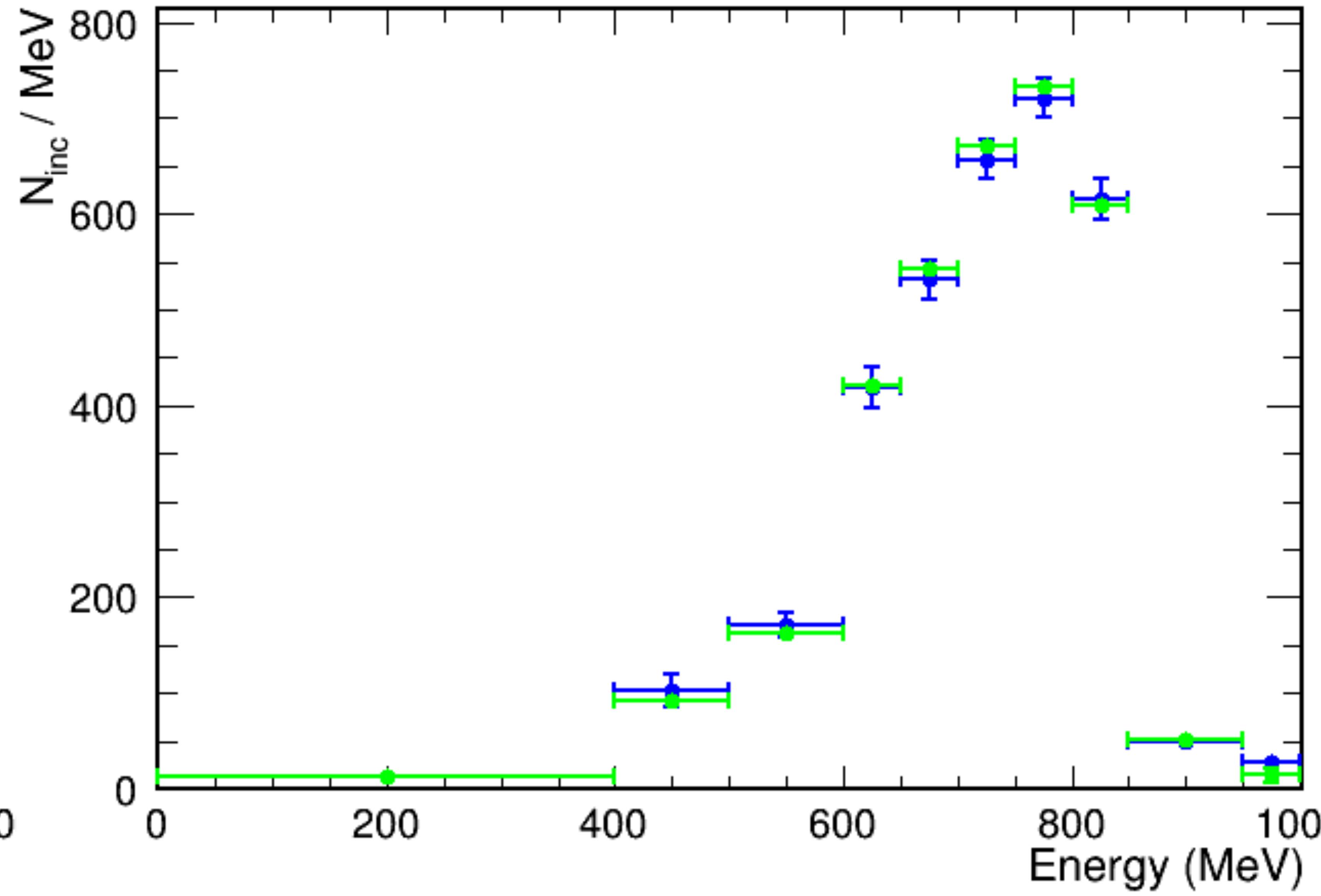




Initial histogram



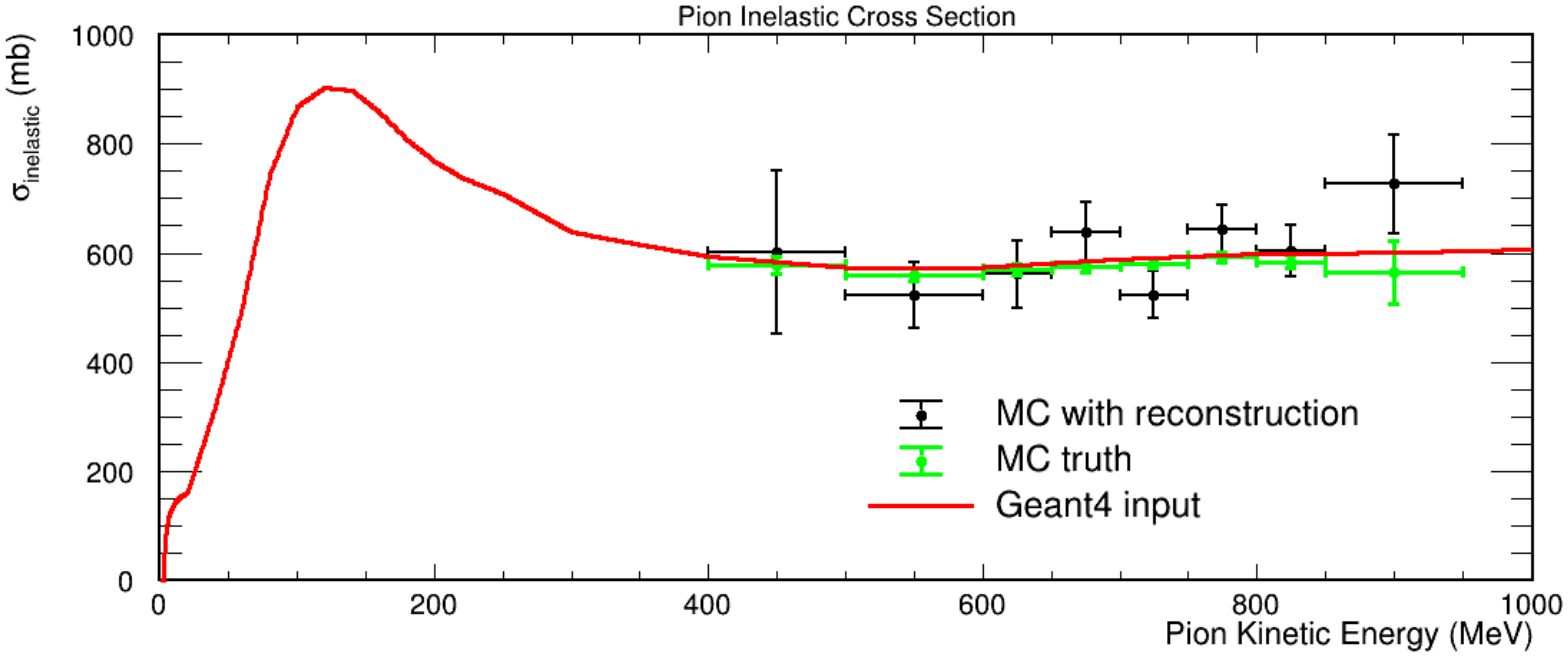
Interaction histogram

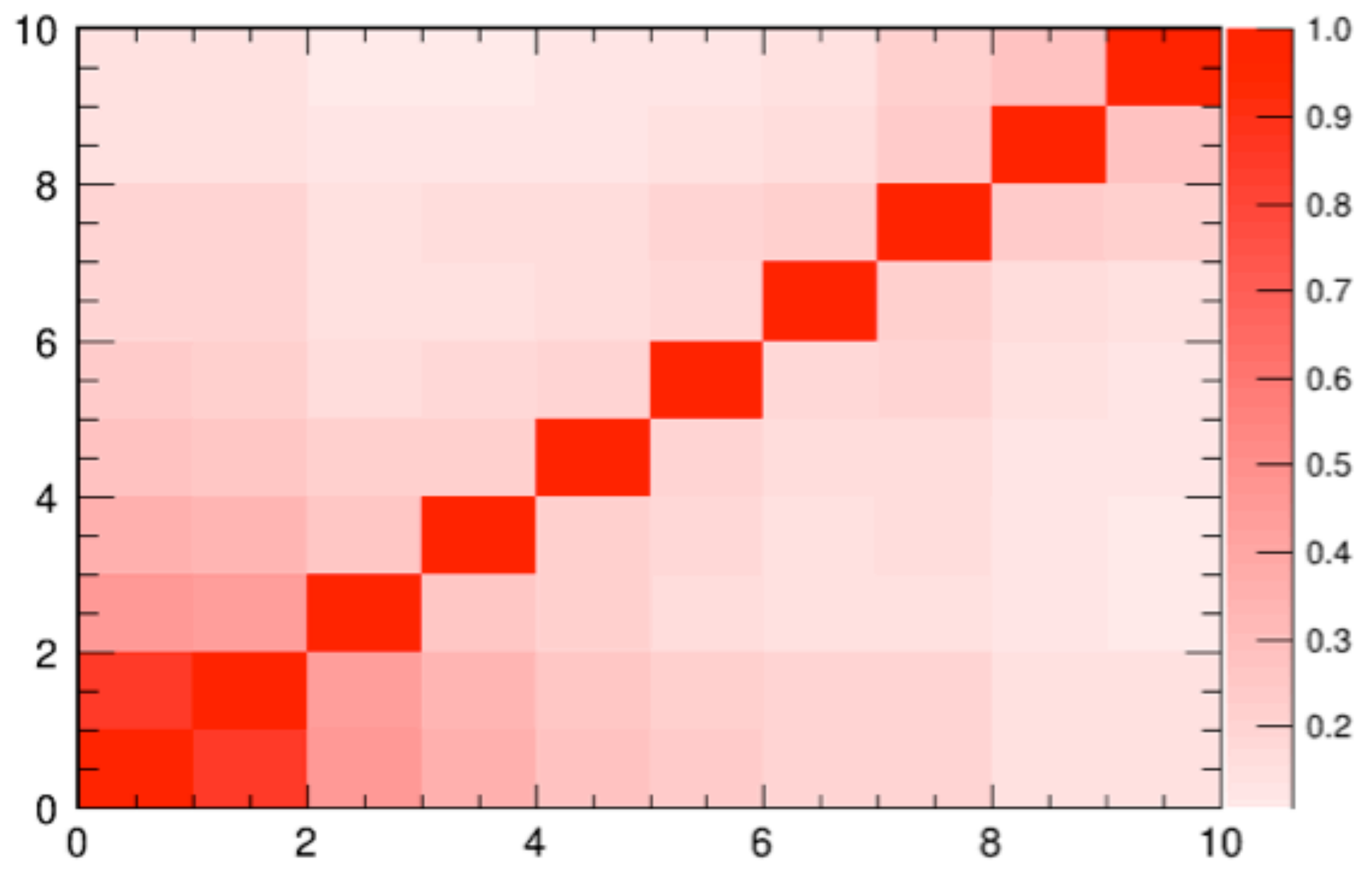


Incident histogram

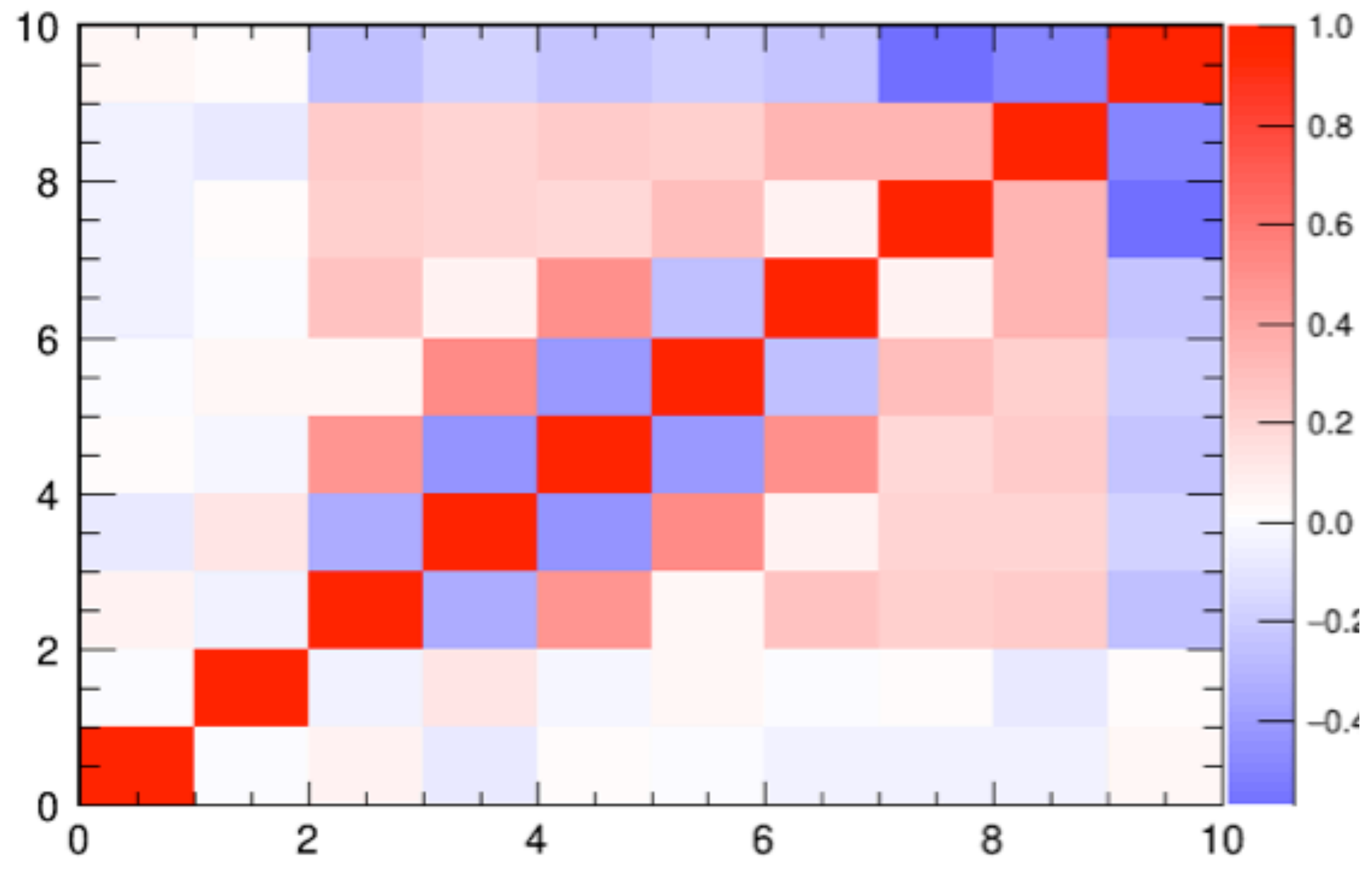
Fake data

3D unfolding **200 iterations**
(error bars are propagated from the covariance matrix provided by RooUnFold)





Correlation matrix for true XS

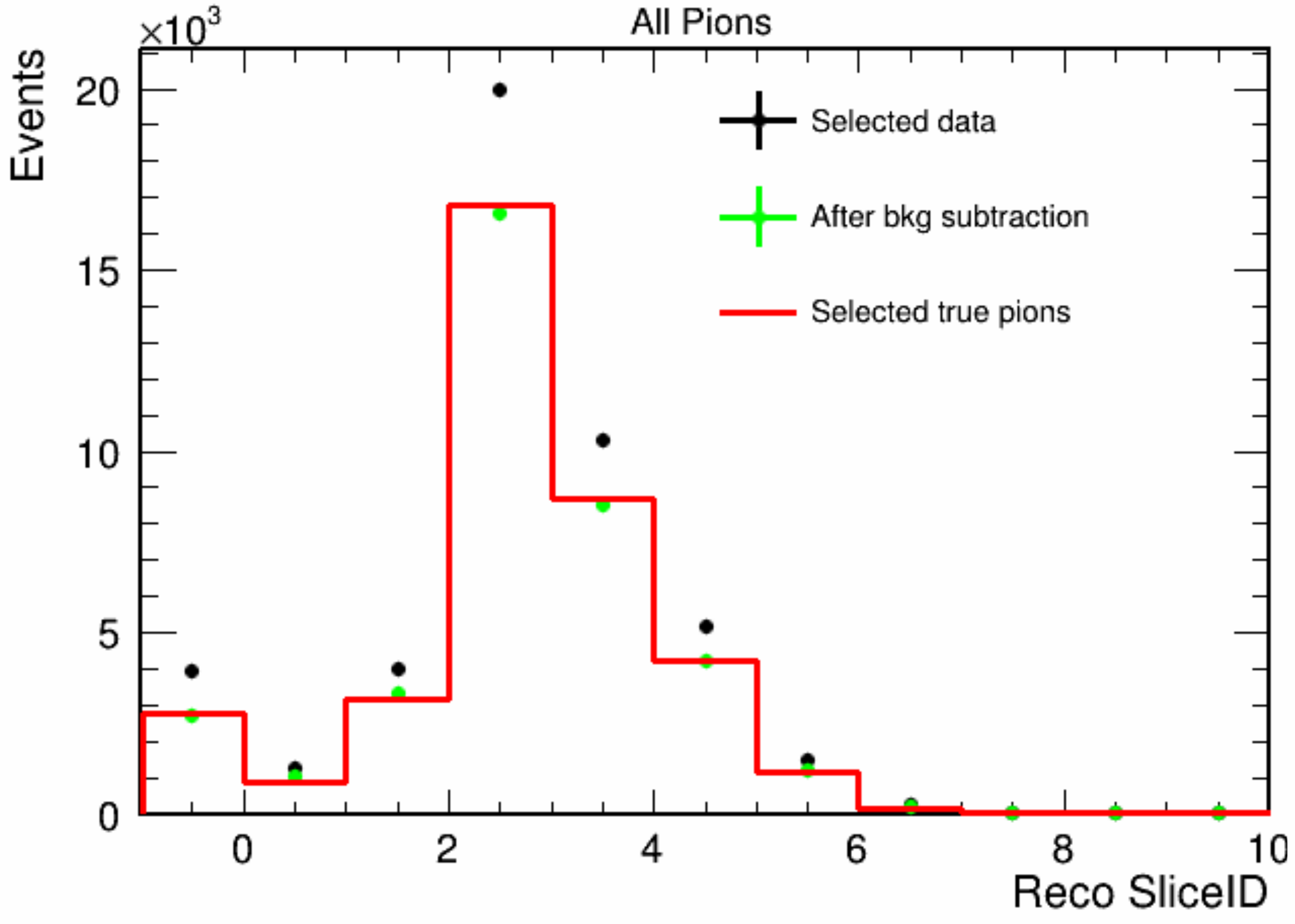


Correlation matrix for reco XS

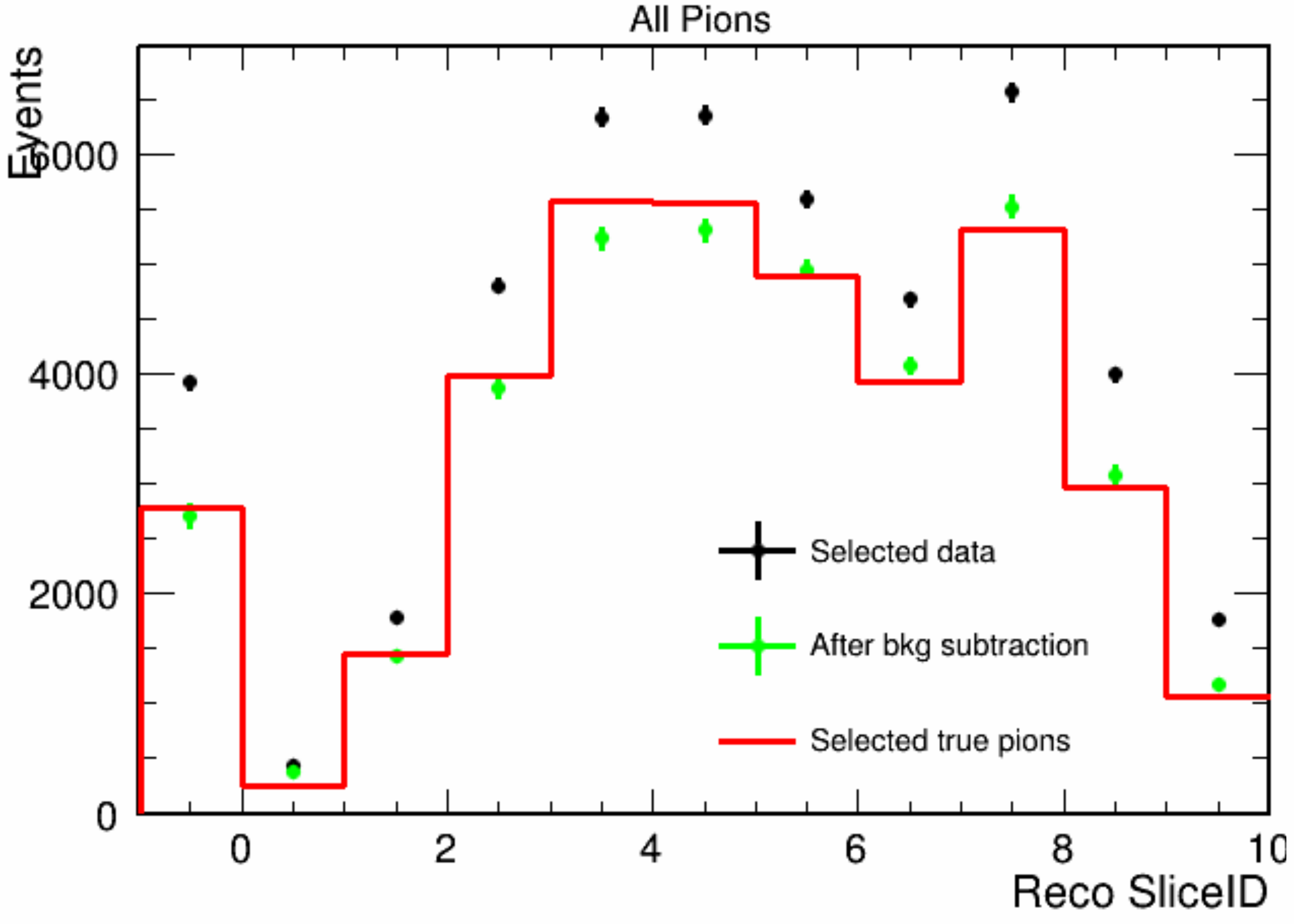
Real data

Muon scaling factor: 0.65 ± 0.11
Proton scaling factor: 1.65 ± 0.13
Pion scaling factor: 1.47 ± 0.14

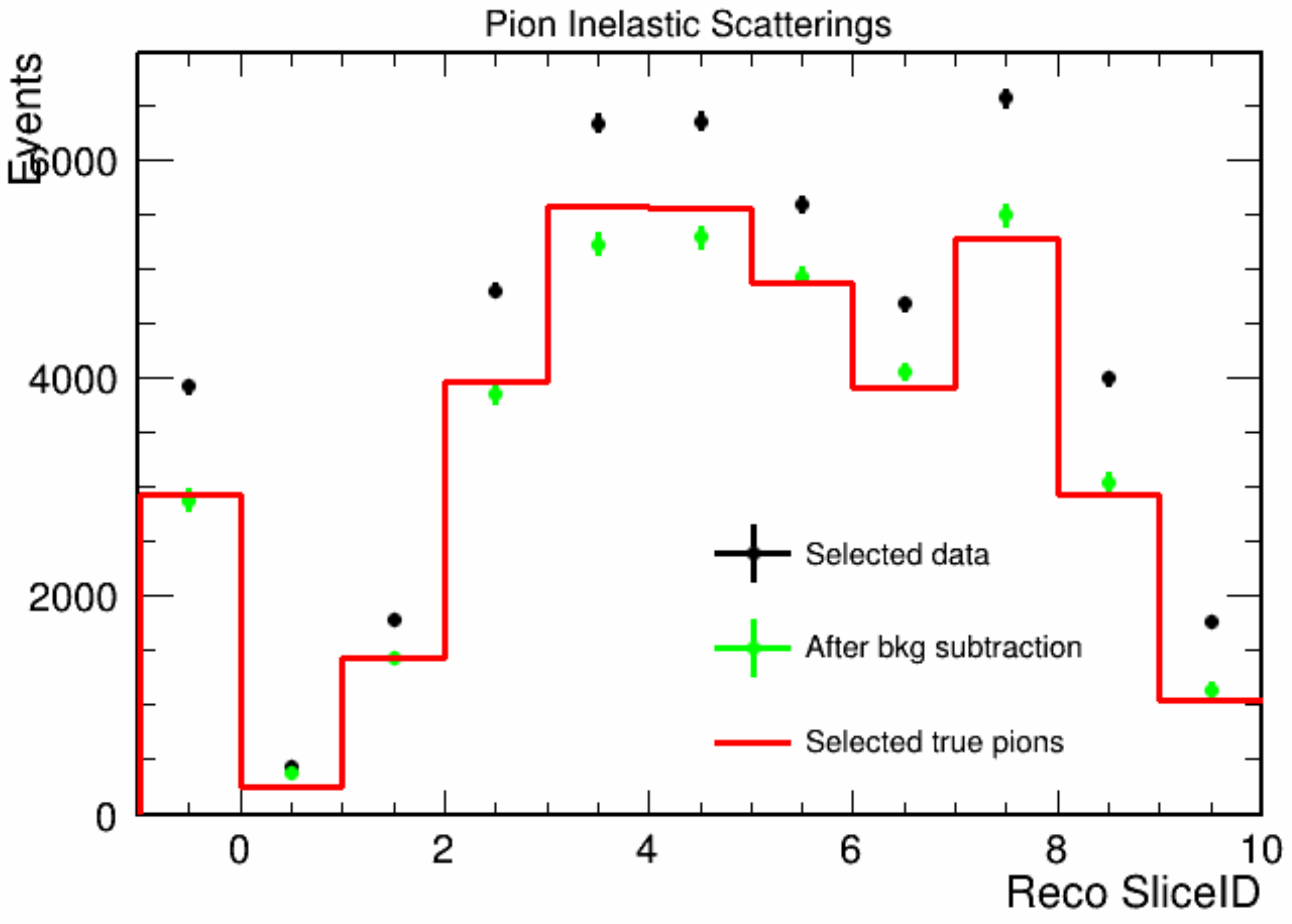
- After bkg subtraction



Initial sliceID



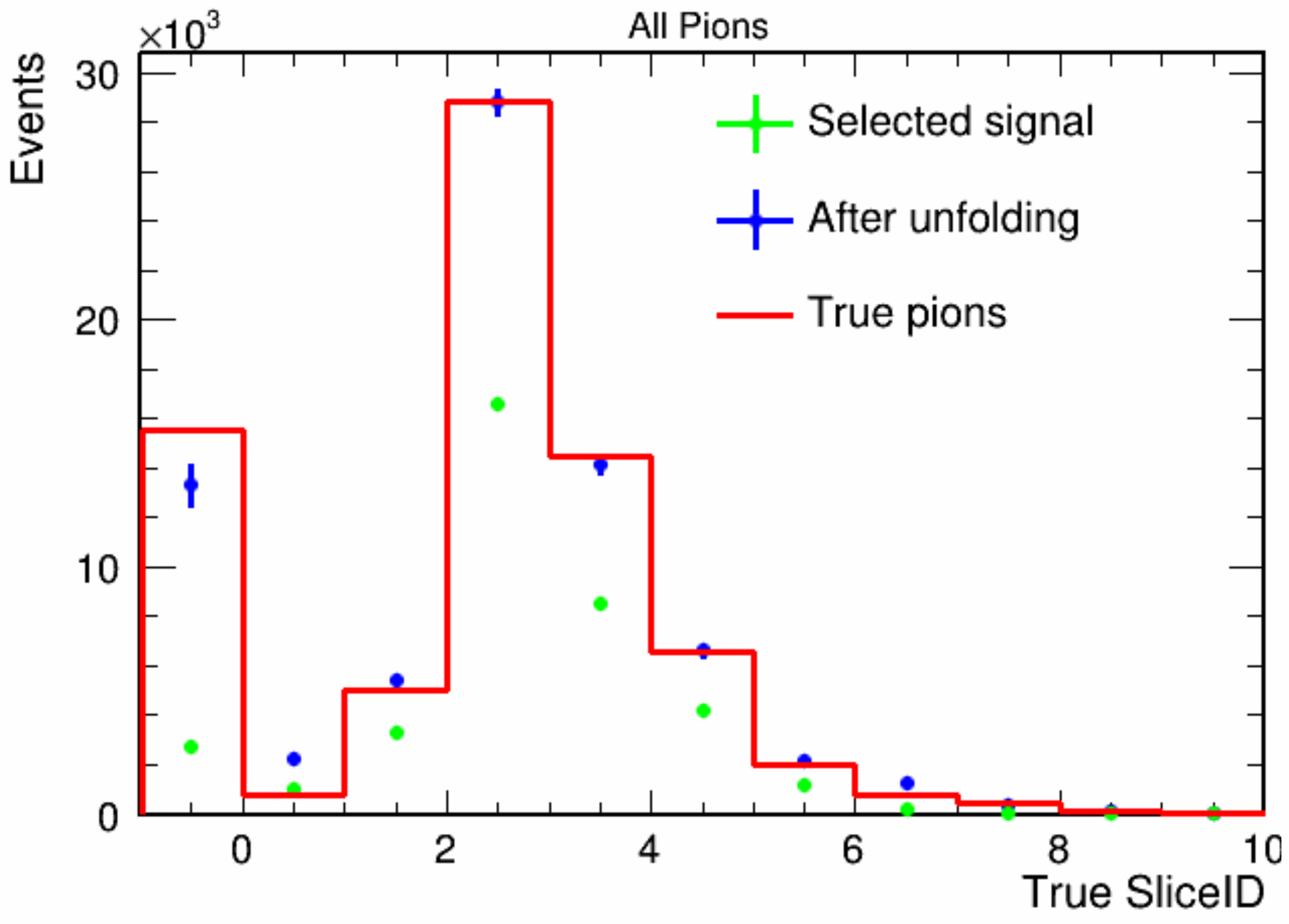
End sliceID



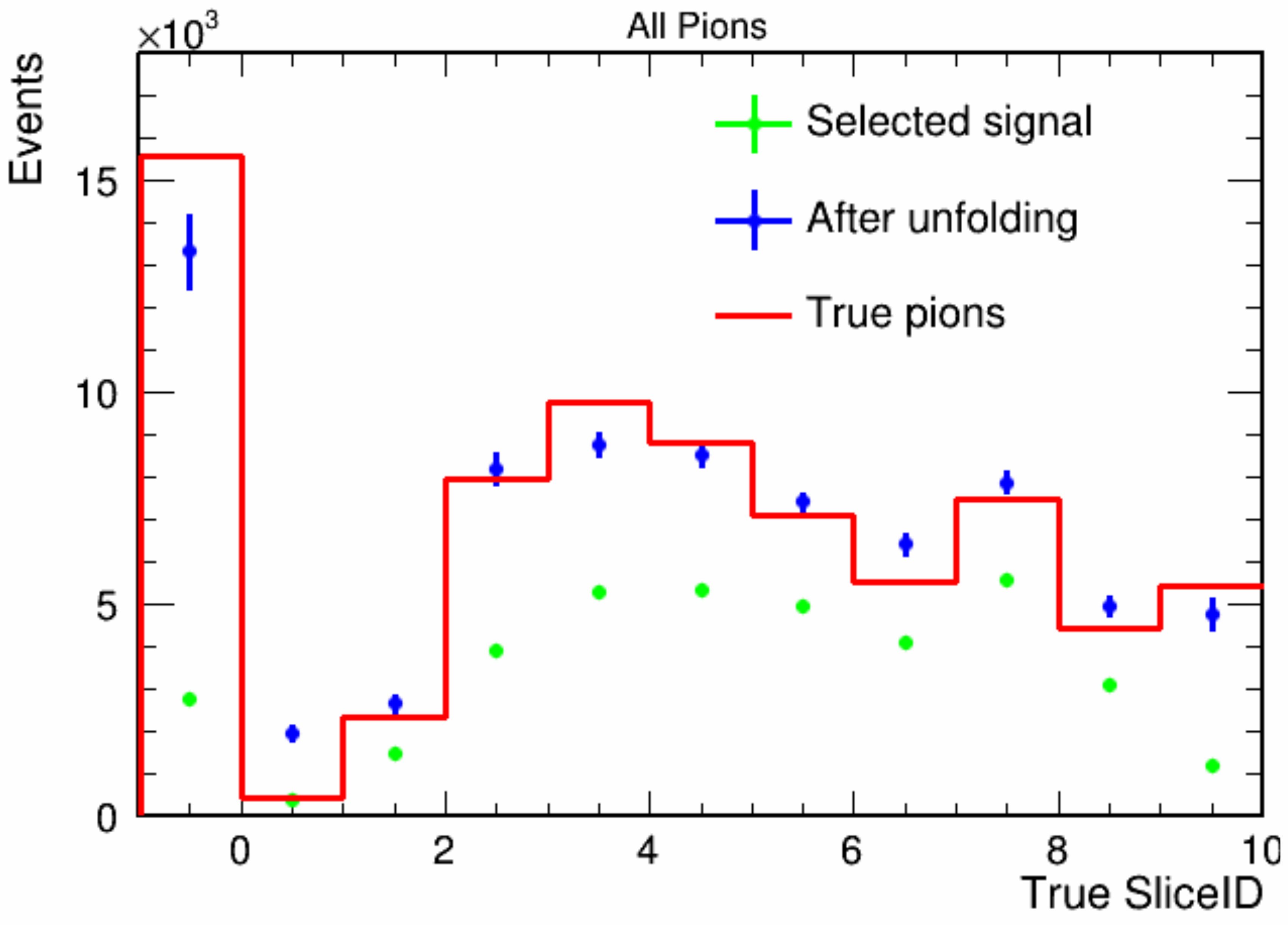
Interaction sliceID

Real data

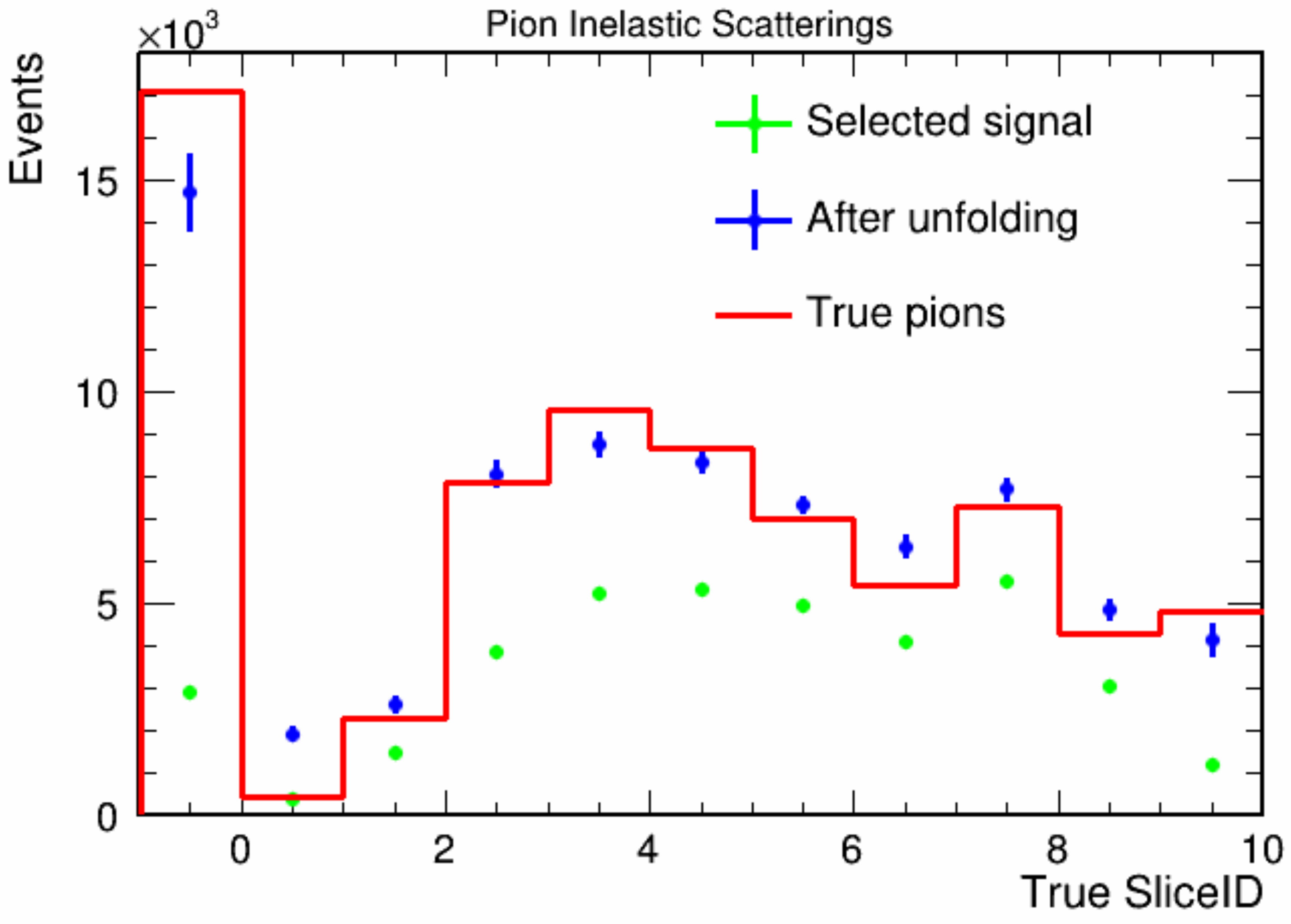
- After unfolding



Initial sliceID



End sliceID



Interaction sliceID

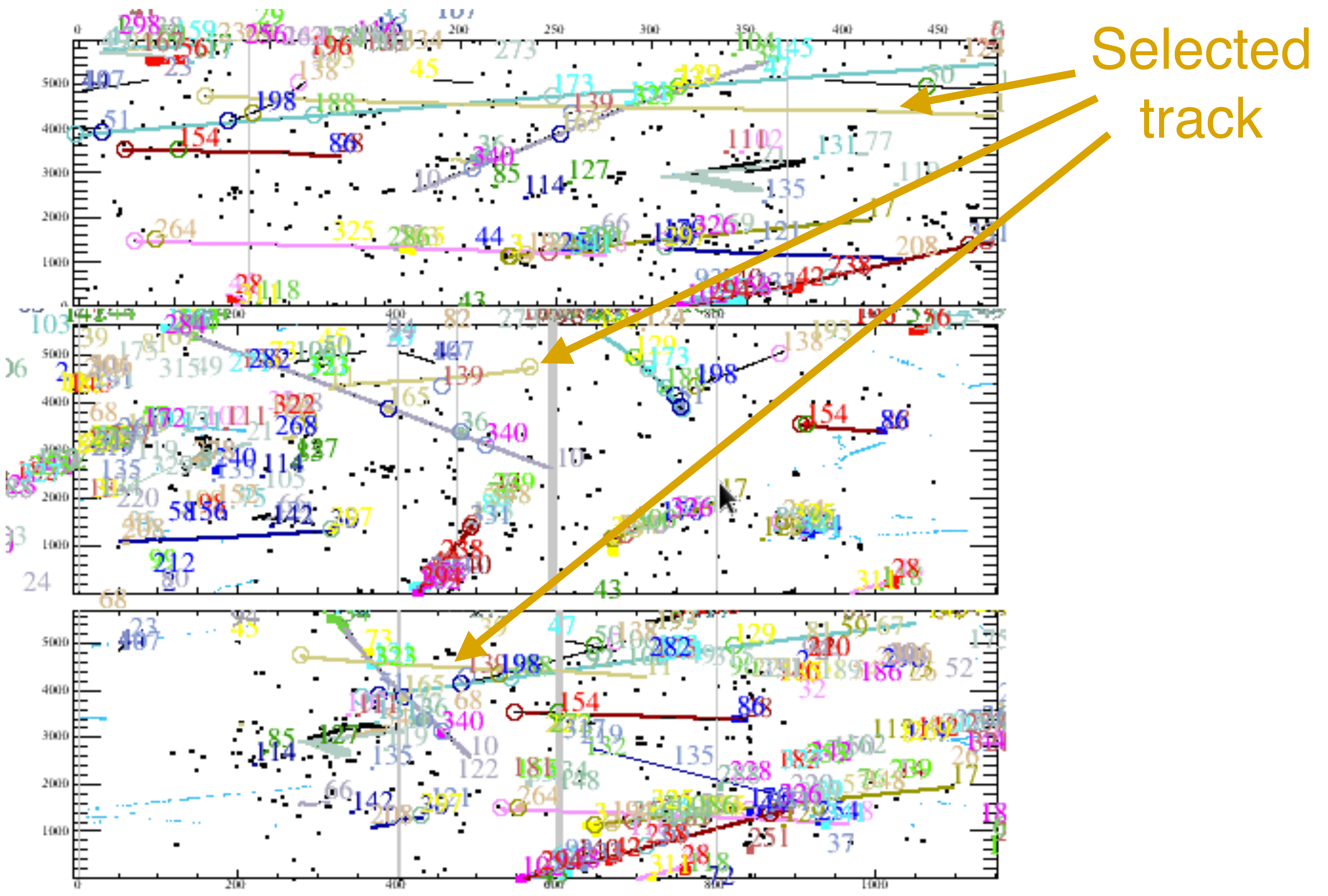
Proposal of systematic uncertainties

- We have propagated the systematic uncertainties from unfolding and the background scaling factors.
- Other systematics needed to be considered:
  - Reweighting factors
  - Energy reconstruction
  - …

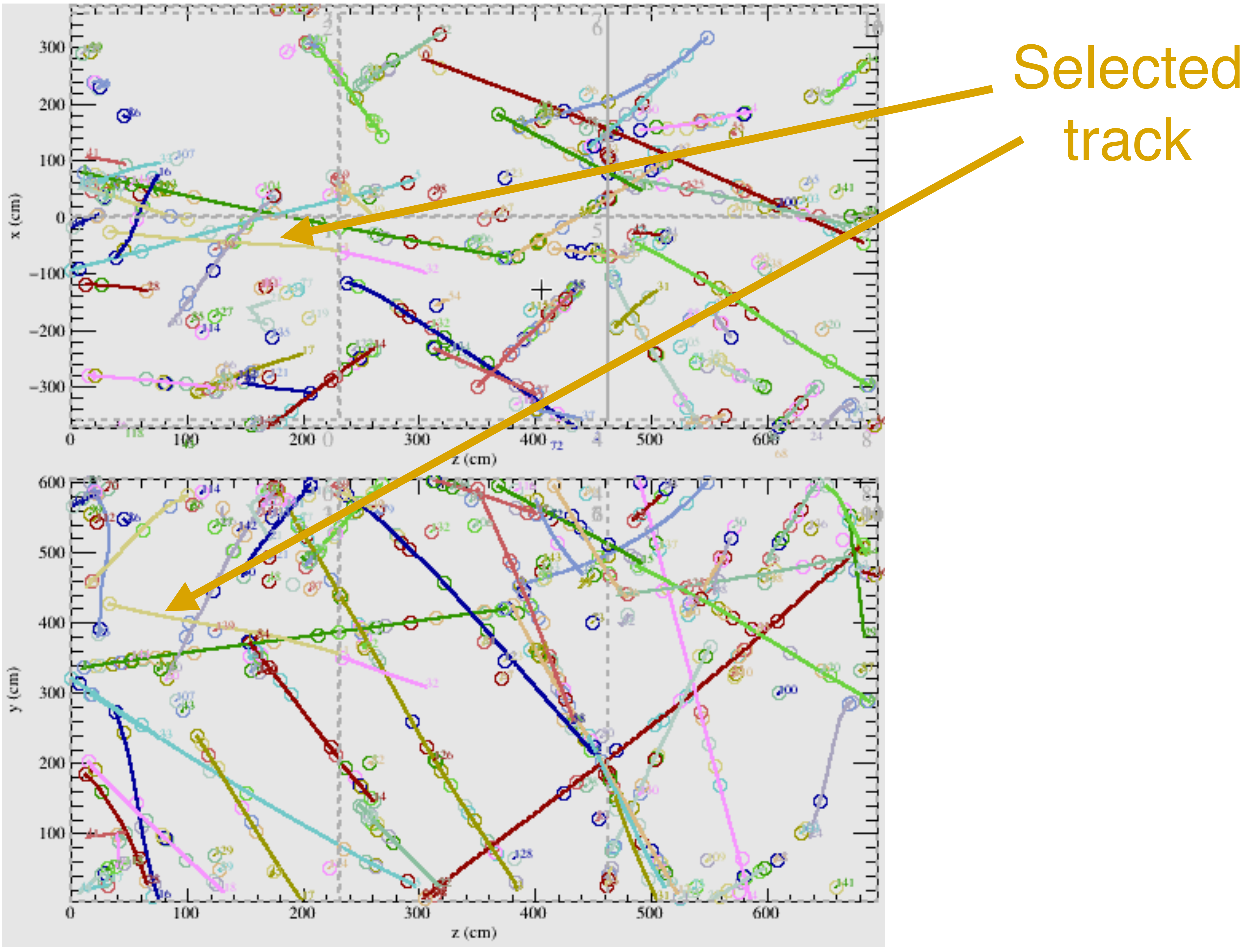
Back-ups

Pandora identification

- In each event, one track is selected as beam track by Pandora based on boosted decision tree (BDT) algorithm.



Wire view (from top to bottom: plane Y; U; V)



Ortho3D view (top: XZ view; bottom: YZ view)

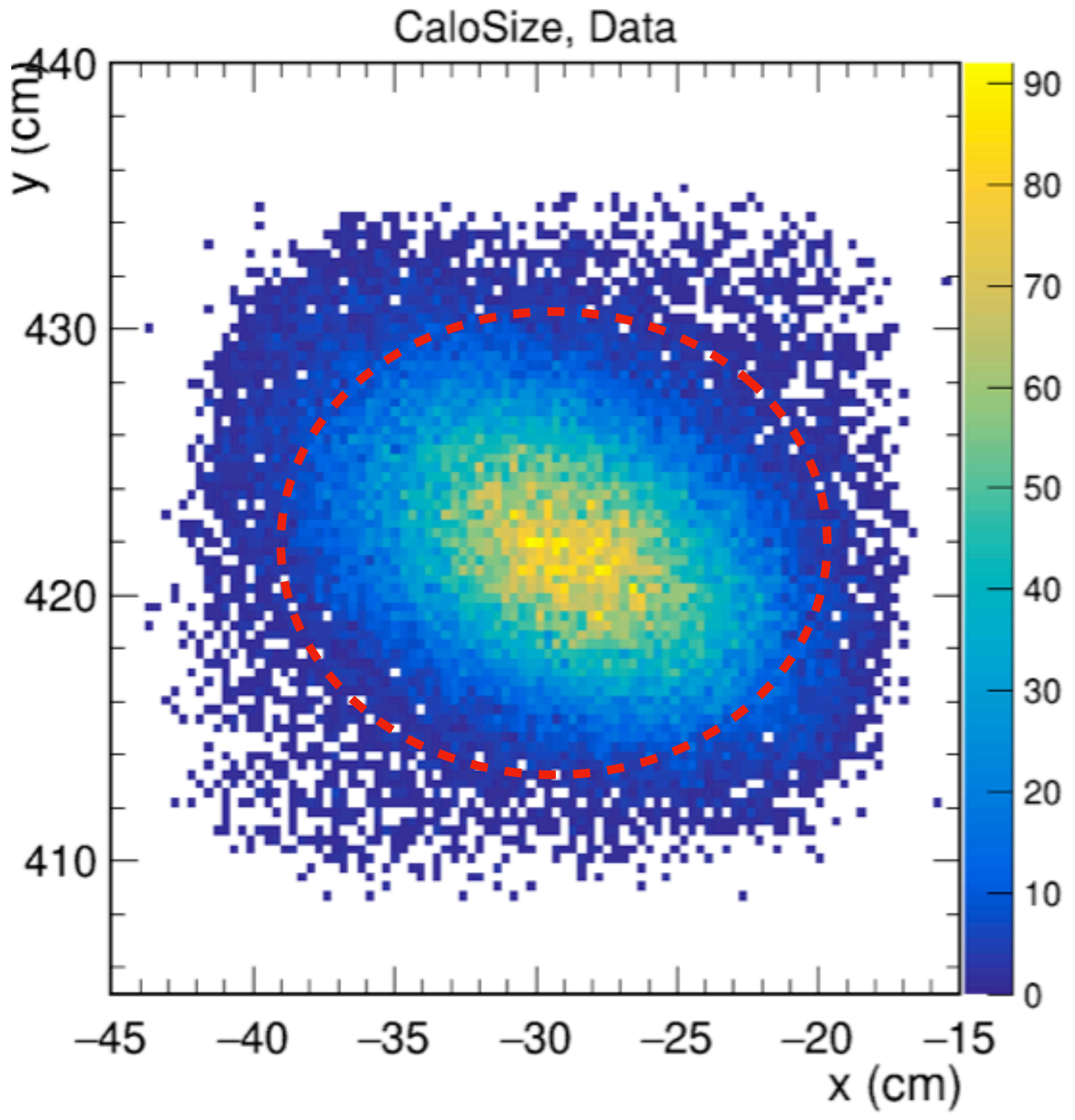
Precut

- Some technical cuts to ensure the beam track can be used.
  - **Upstream beam type selection**
    - MC   true_beam_PDG == -13 or 211
    - Data   beam_inst_trigger != 8
      beam_inst_nMomenta == 1 && evt.beam_inst_nTracks == 1
      beam_inst_PDG_candidates == -13 or 211
  - **Empty events removal**   reco_reconstructable_beam_event != 0
  - **Pandora Slice Cut** to ensure it is a track.   reco_beam_type == 13
  - **Calo Size Cut** require hit detected on collection plane.   ! (reco_beam_calo_wire->empty( ))

Variable definitions: https://wiki.dunescience.org/wiki/PDSPAnalyzer

Beam Quality Cut

- It consists of two parts. First, cuts on the position of instrumented beam particle projected to the front-face of the TPC.



- The events outside of the oval is more likely to be secondary particles produced by upstream interactions.

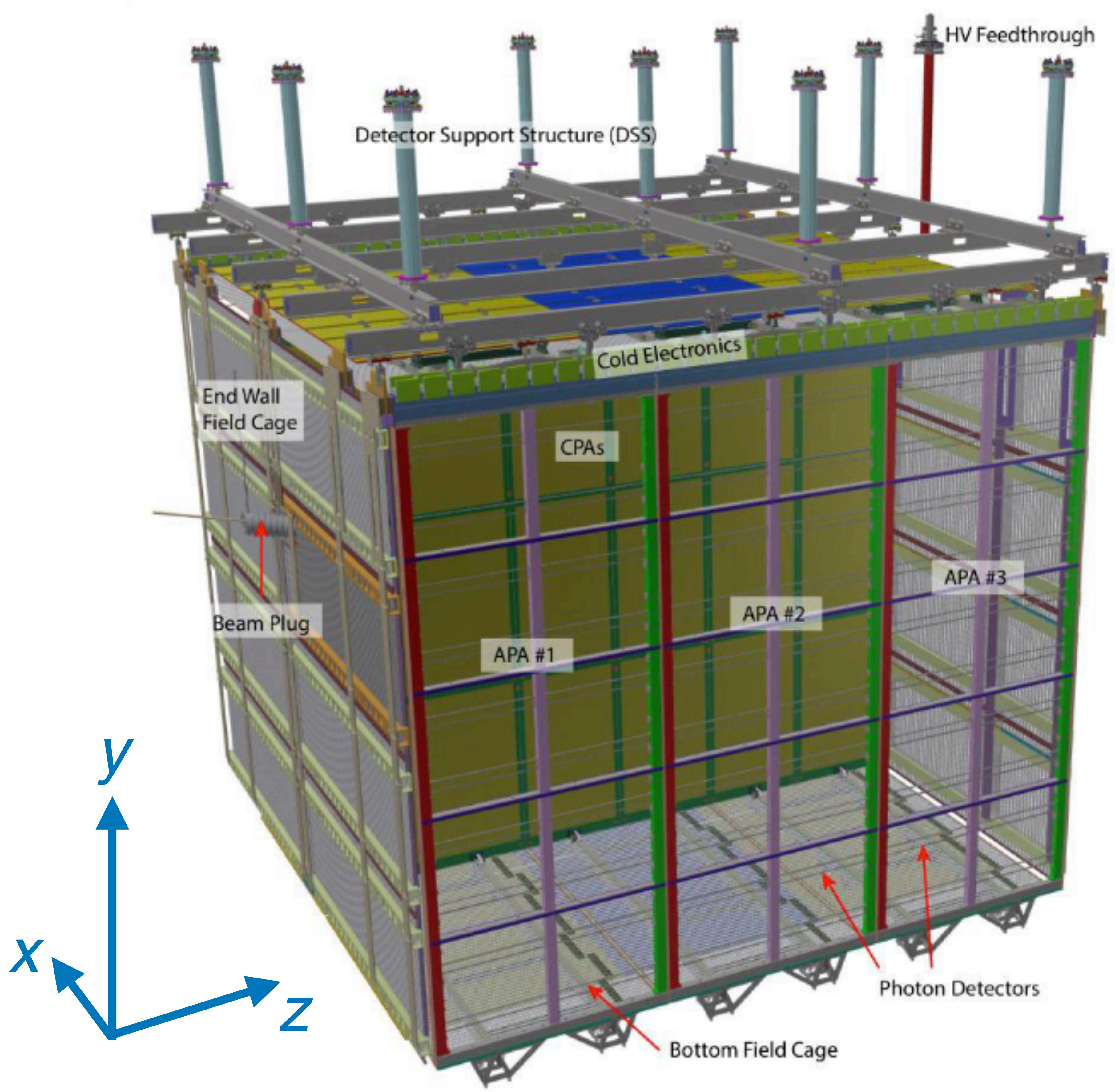
Selection: sqrt($\Delta x_{\mathrm{inst}}^2 + \Delta y_{\mathrm{inst}}^2$) < 4.5

- Δx_{inst} is $(x_{\mathrm{inst}} - \mu_{x_{\mathrm{inst}}})/\sigma_{x_{\mathrm{inst}}}$
- $\mu_{x_{\mathrm{inst}}}$ and $\sigma_{x_{\mathrm{inst}}}$ are derived before beam quality cut

Beam Quality Cut

- Second, cuts on beam entrance position and beam angle.
  - Entrance point on *xy* plane   sqrt(Δx^2+Δy^2) < 3
  - Start *z* position $|\Delta z| < 3$
  - Beam angle $\cos\theta > 0.95$

    - Δx is $(x - \mu_x)/\sigma_x$, where μ_x and σ_x are derived before beam quality cut. Δy and Δz are similar.
    - θ is the angle between the track and the mean direction μ_θ, derived before beam quality cut.

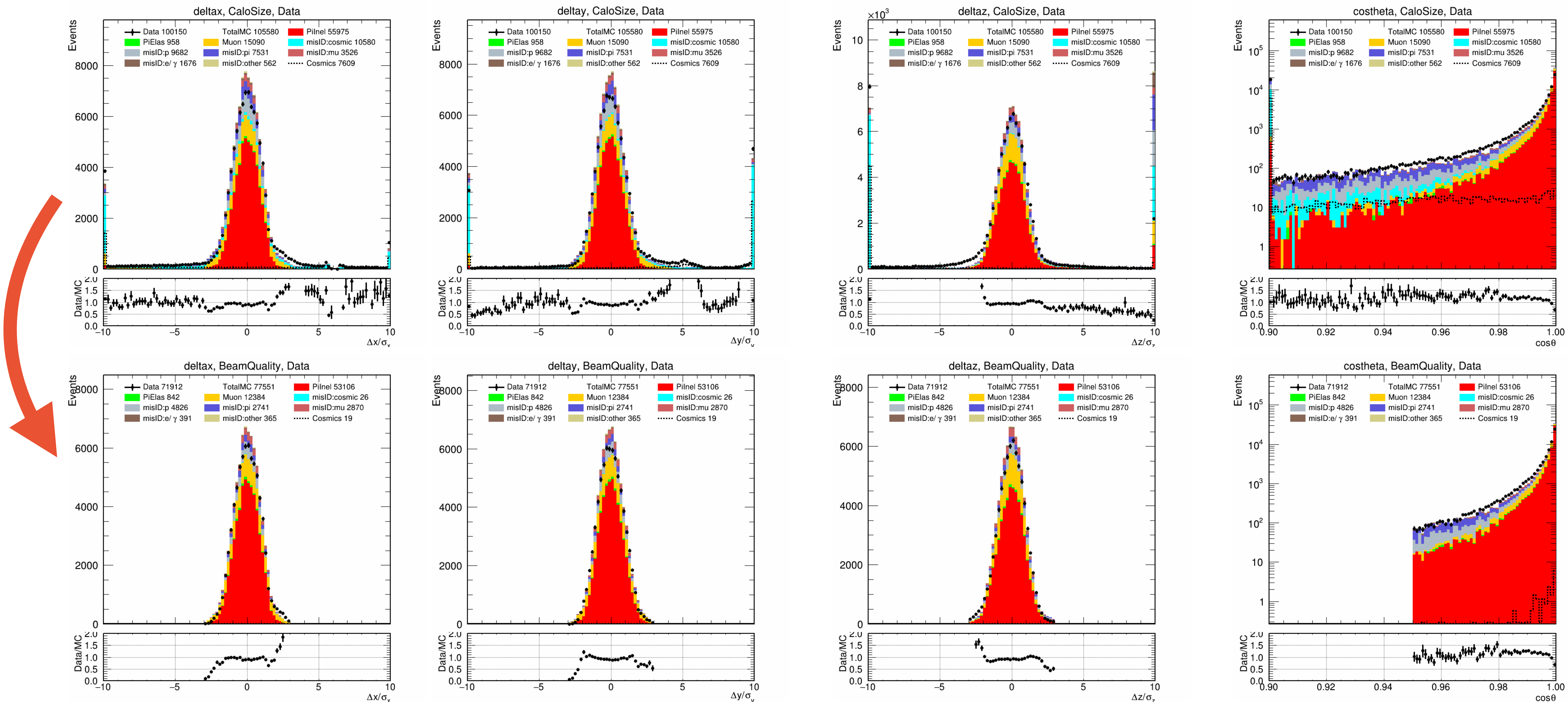


A view of ProtoDUNE-SP detector. The beam plug indicates the direction of beam track

Beam Quality Cut

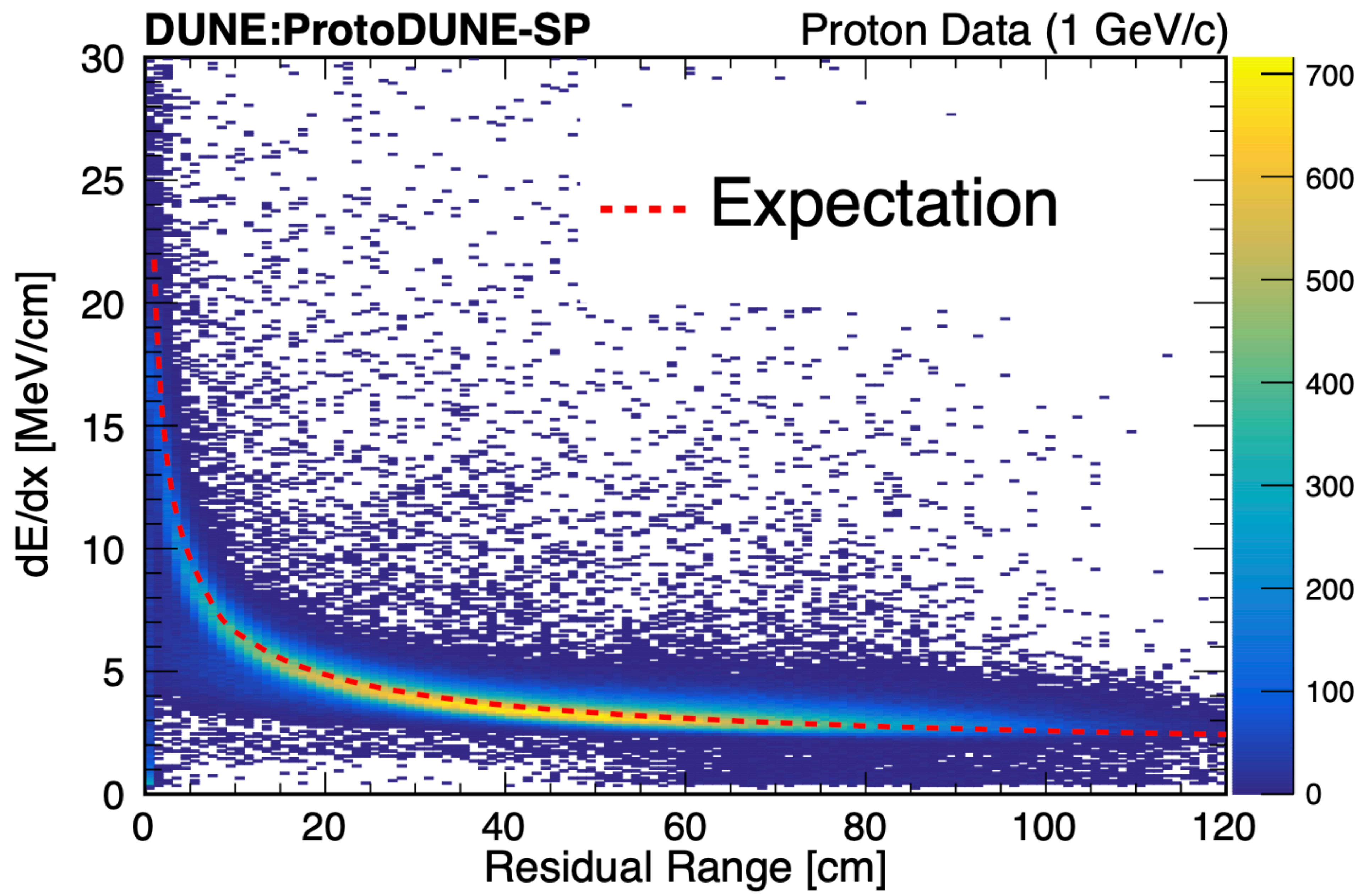
$\mathrm{sqrt}(\Delta x^2+\Delta y^2) < 3$ $|\Delta z| < 3$ $\cos\theta > 0.95$

After cut

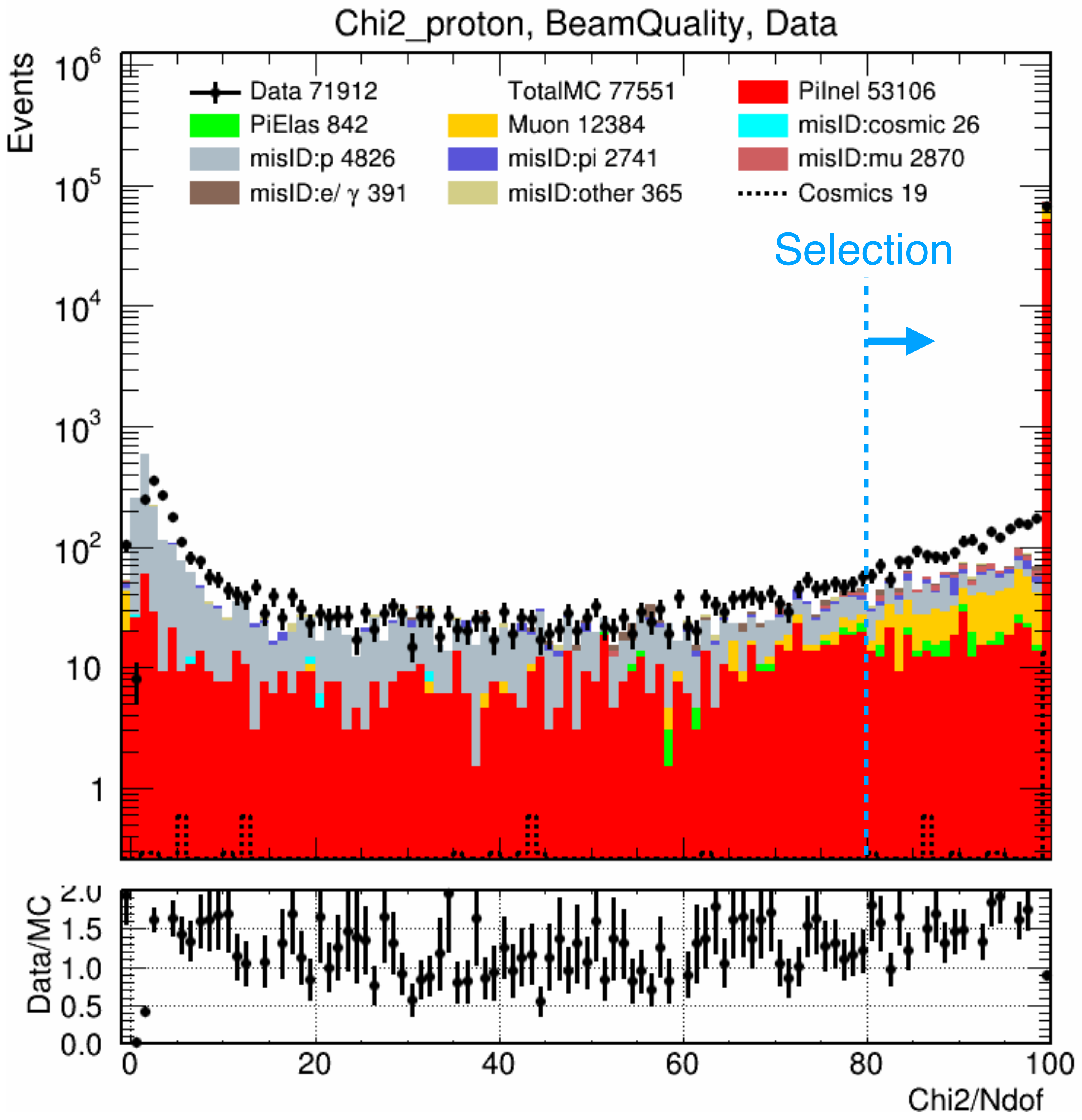


Proton Cut

- We use **Chi2_p/Ndof** to cut proton.
  - Assume it is a stopping proton, then fit dE/dx vs residue range to expectation.

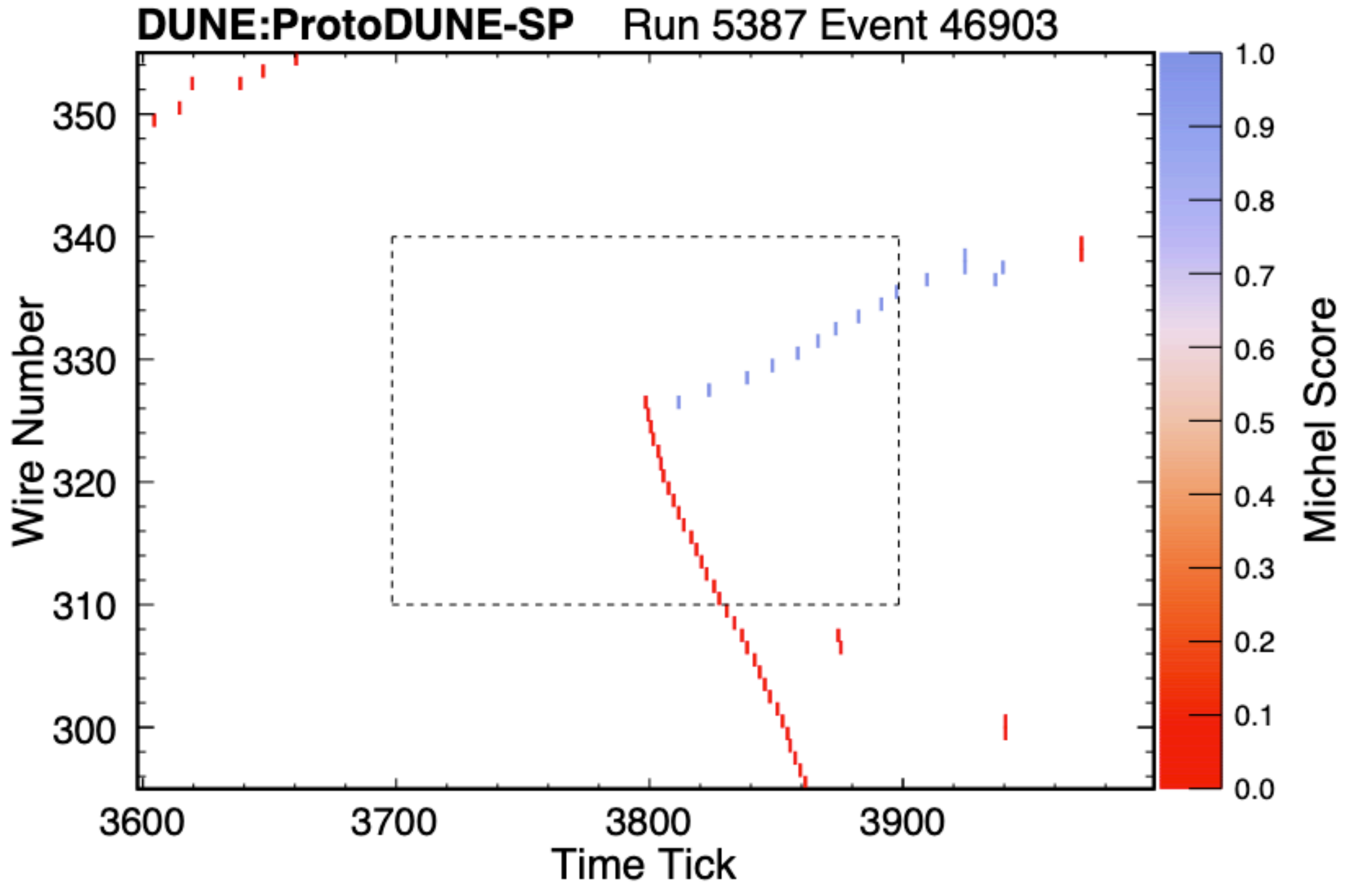


https://doi.org/10.1088/1748-0221/15/12/P12004

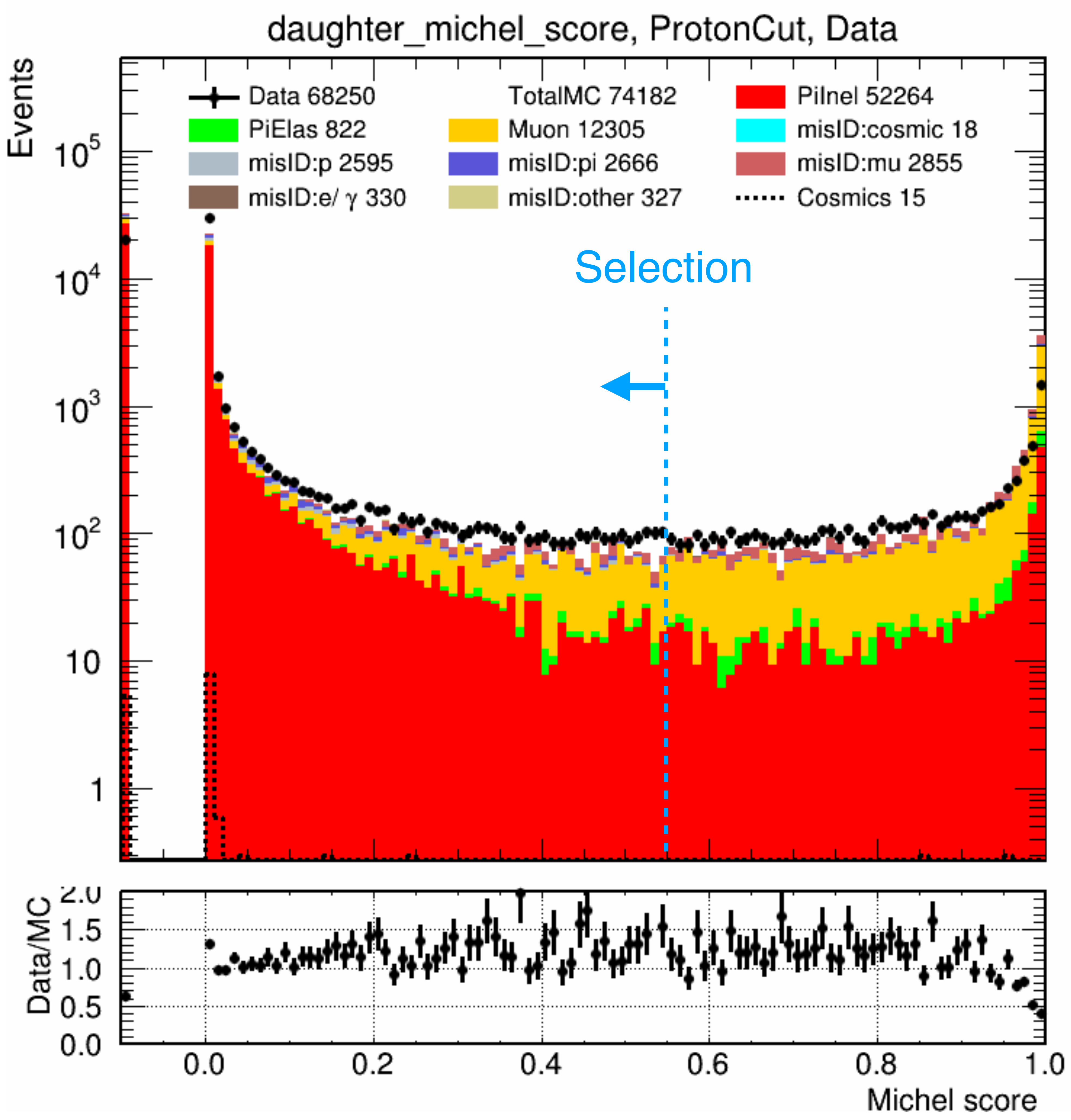


Michel Score Cut

- Michel electron shows some features which can be detected by pattern recognition, thus can be used to cut muon.

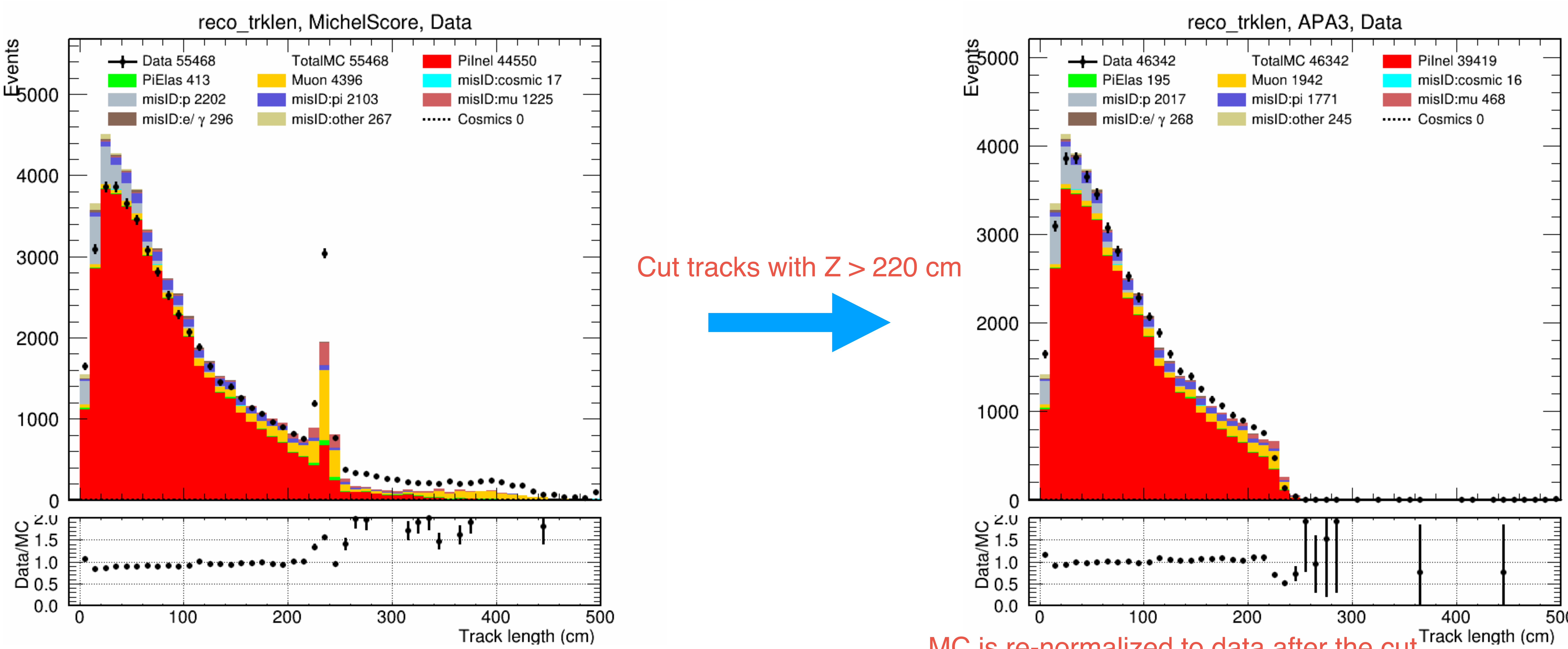


A Michel candidate
https://docs.dunescience.org/cgi-bin/private/ShowDocument?docid=23125

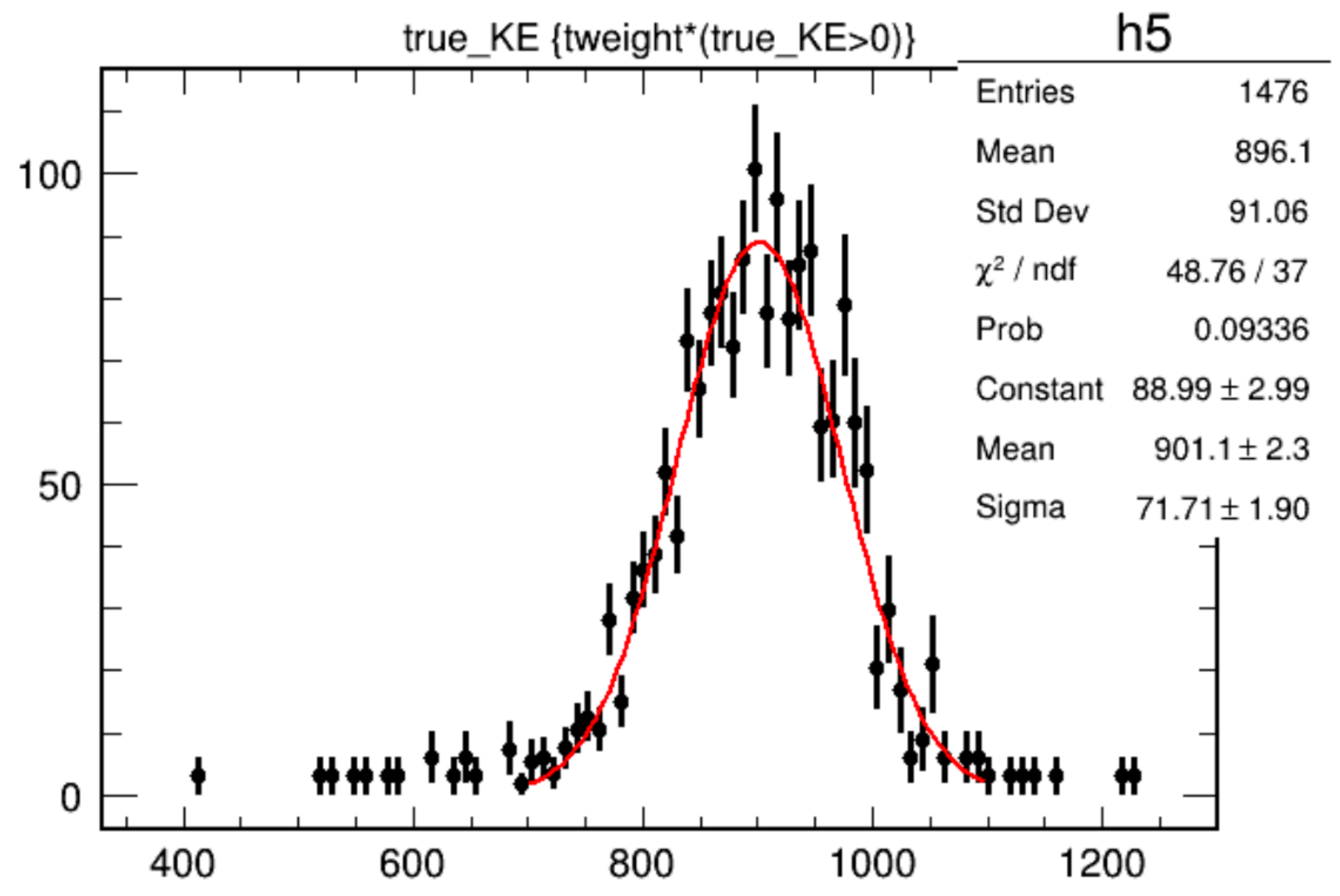


APA3 cut

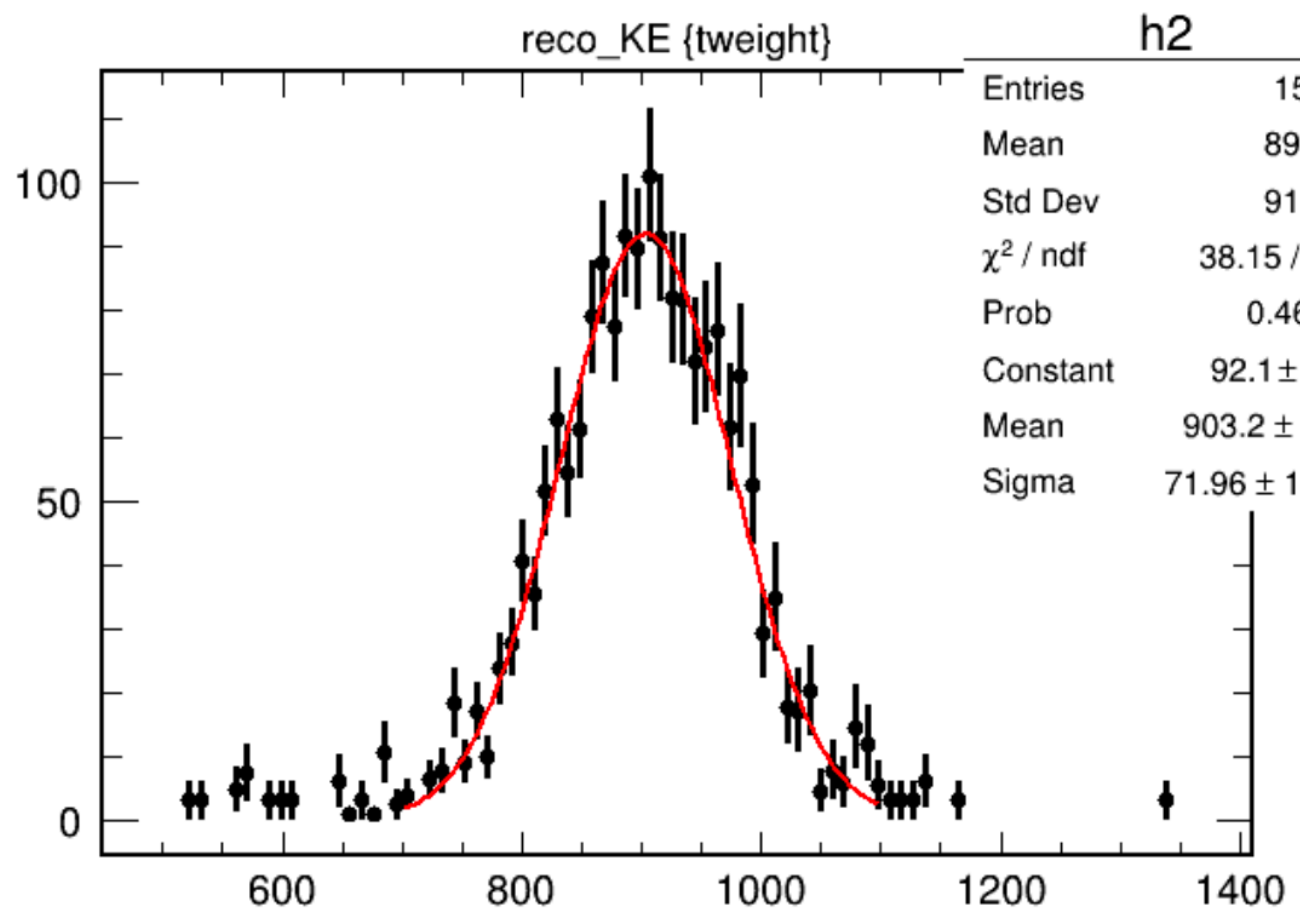
- We only use tracks in the first TPC, which can further cut long muon tracks.



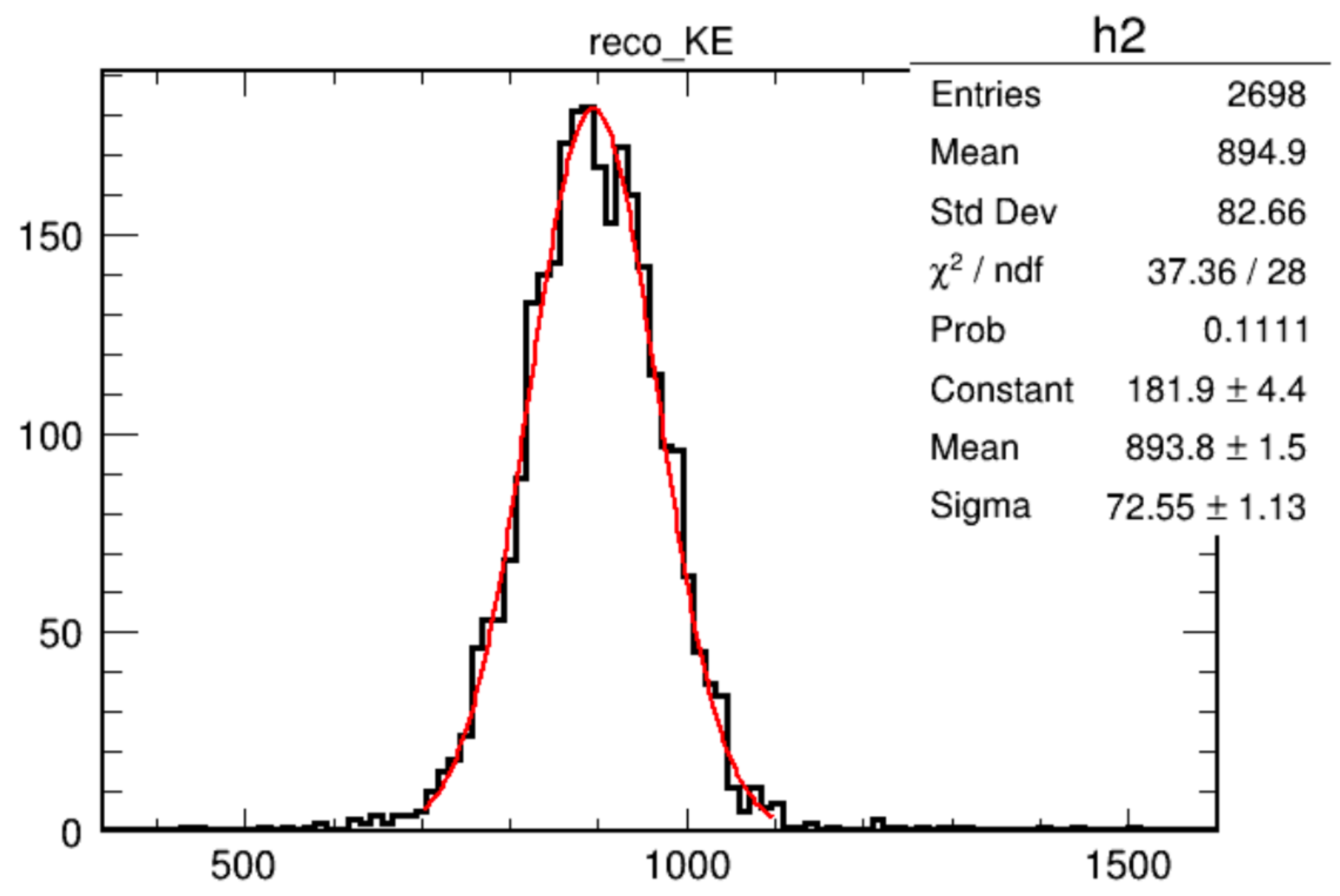
MC is re-normalized to data after the cut



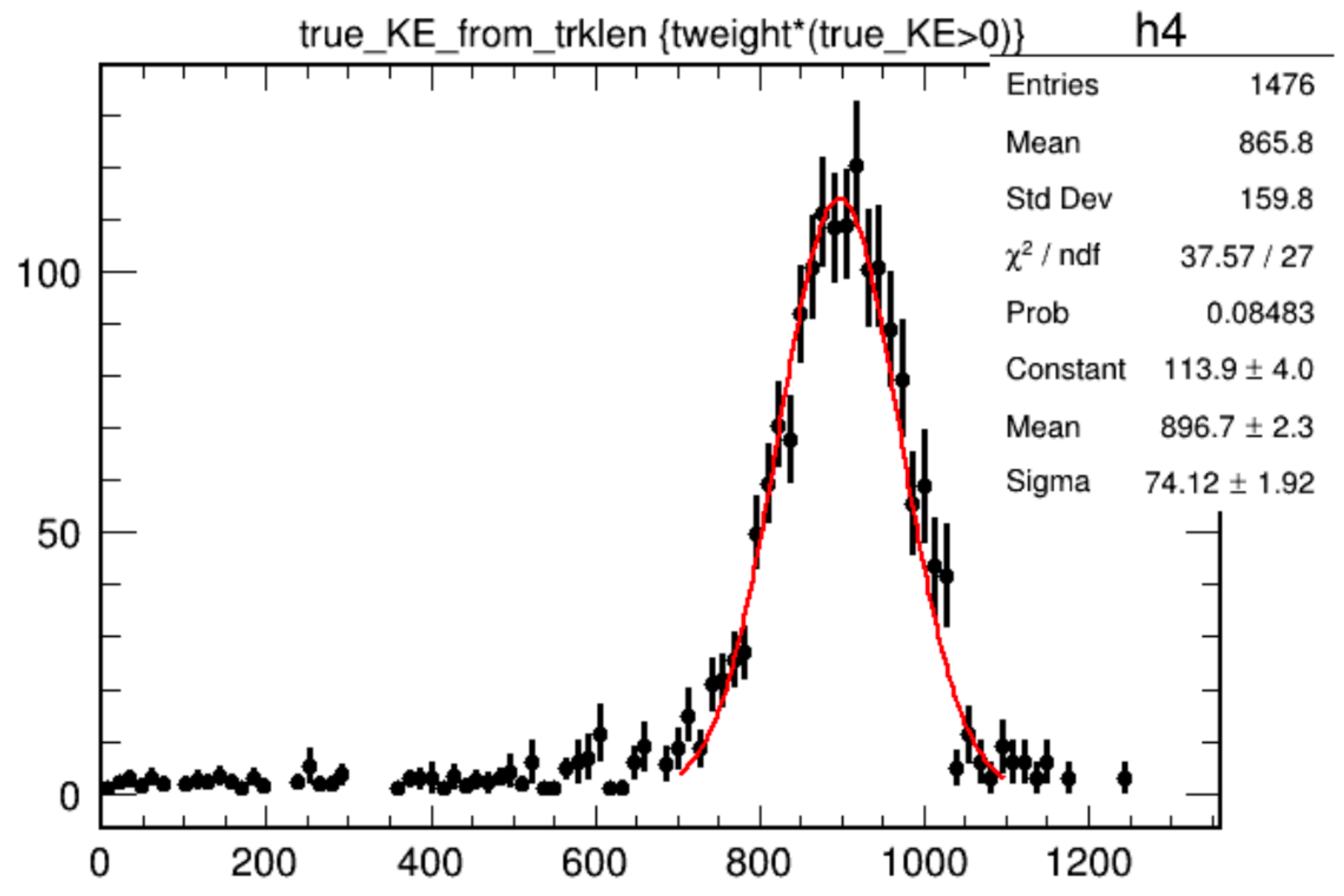
MC true KE



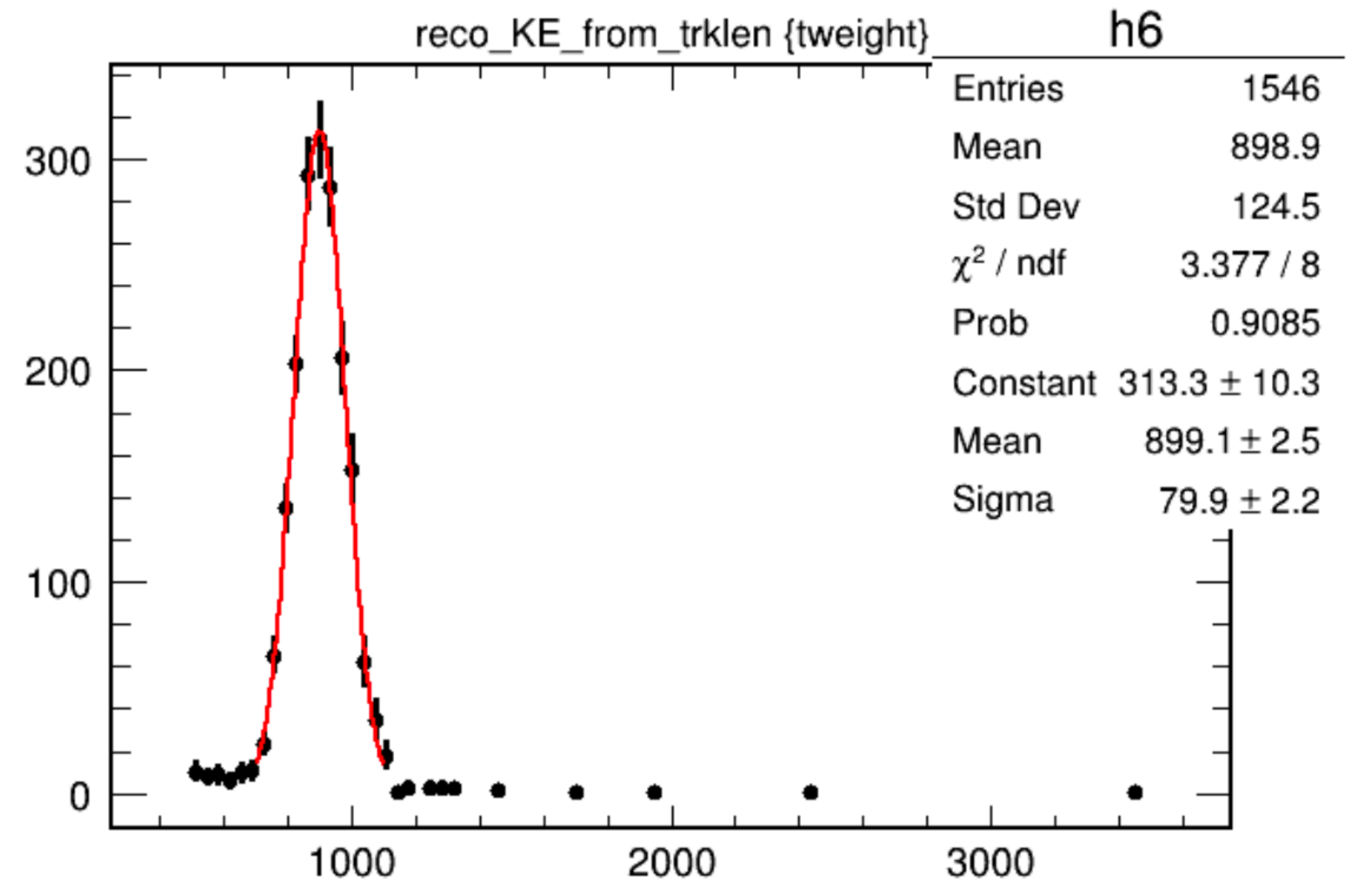
MC reco KE



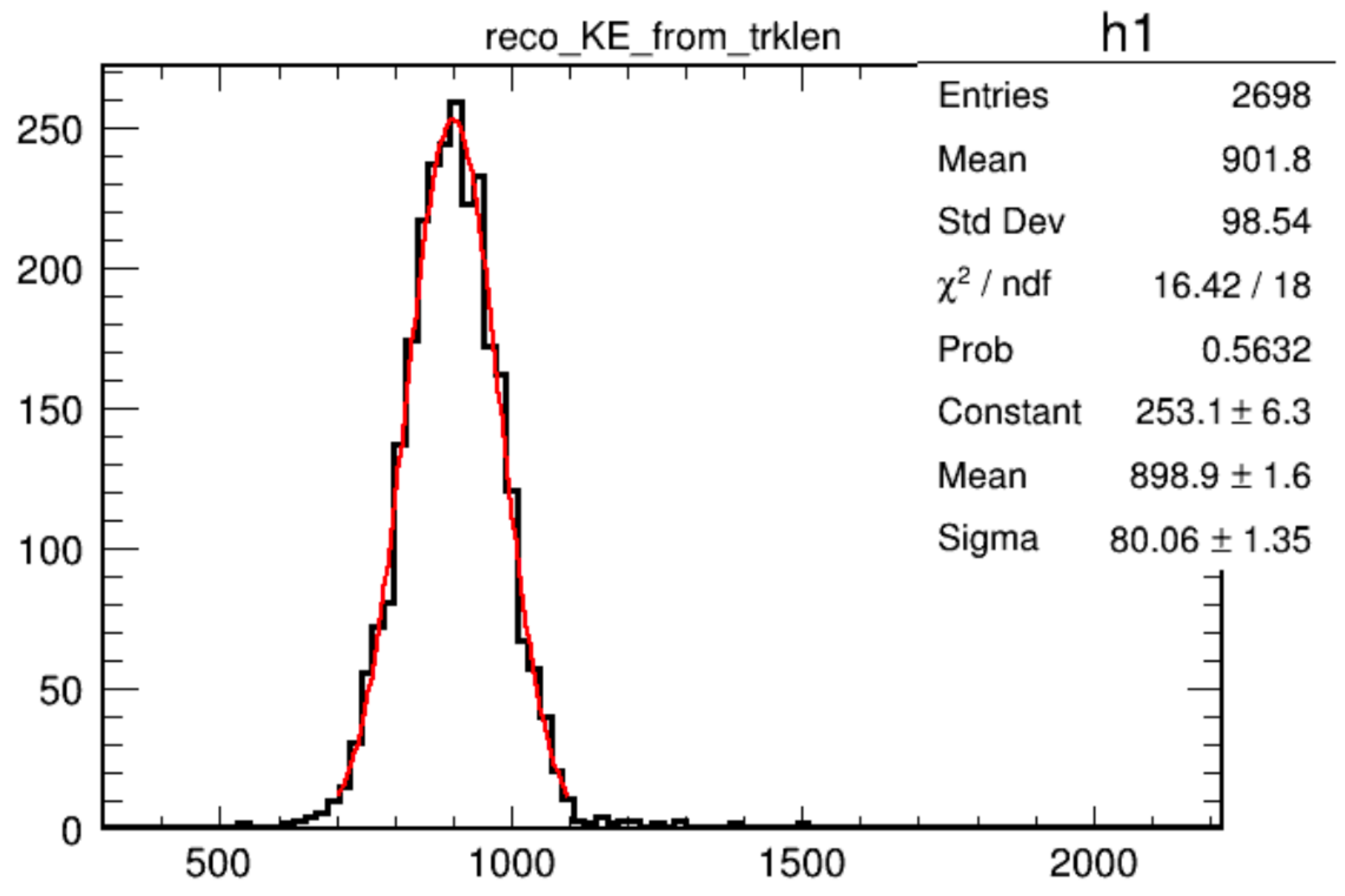
Data reco KE



MC KE calculated by true length



MC KE calculated by reco length



Data KE calculated by reco length