MINERvA Cross Sections
The Nitty Gritty Details

Daniel Ruterbories
On Behalf of the MINERvA Collaboration

June 23rd , 2016

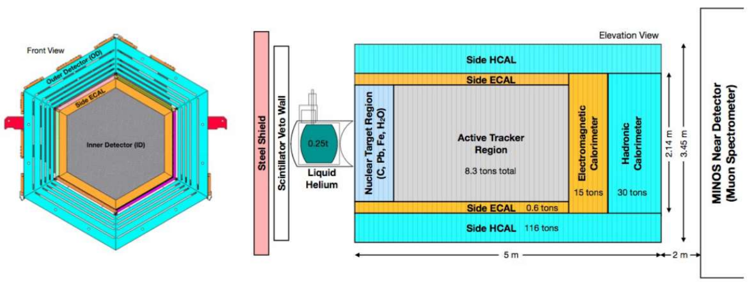
Outline

- MINERvA *very* briefly
- General approach to cross sections
- Analysis Framework
- Specific case – CC0π Double Differential

The MINERvA Experiment

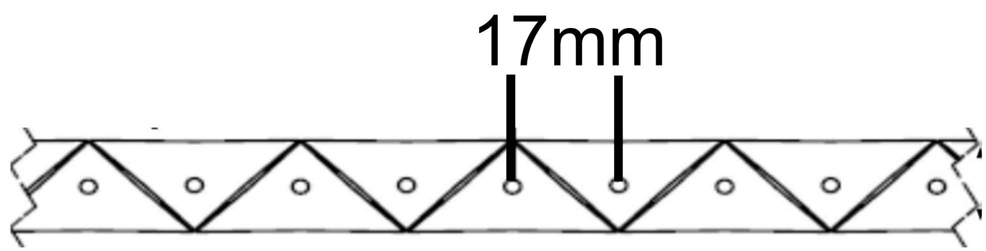
- Study neutrino-nucleus scattering at a few GeV
  - Measure the effects of the nuclear environment on neutrino scattering
  - Improve understanding of neutrino-nucleus cross section model by working with generators
  - Benefits current and future neutrino oscillation experiments
  - Measure A-dependence using the same detector in the same beam simultaneously

Experimental Apparatus



Beam Direction →

Three views:
X: Vertical
U,V: ±60



Spatial resolution ~3mm
Timing resolution ~3ns

Nucl. Inst. and Meth. A743 (2014) 130
arXiv:1305.5199

General Approach

Wish list – What's the right order?!

- Select as many events of the type you are interested in
- Reject as many of the events which are not the type you are interested in
- Minimize your systematic errors – critical with large exposures where you will not have statistics issues
- Report something which is both interesting and useful to the community
- Minimize your reliance on the model in the simulation

Design your analysis

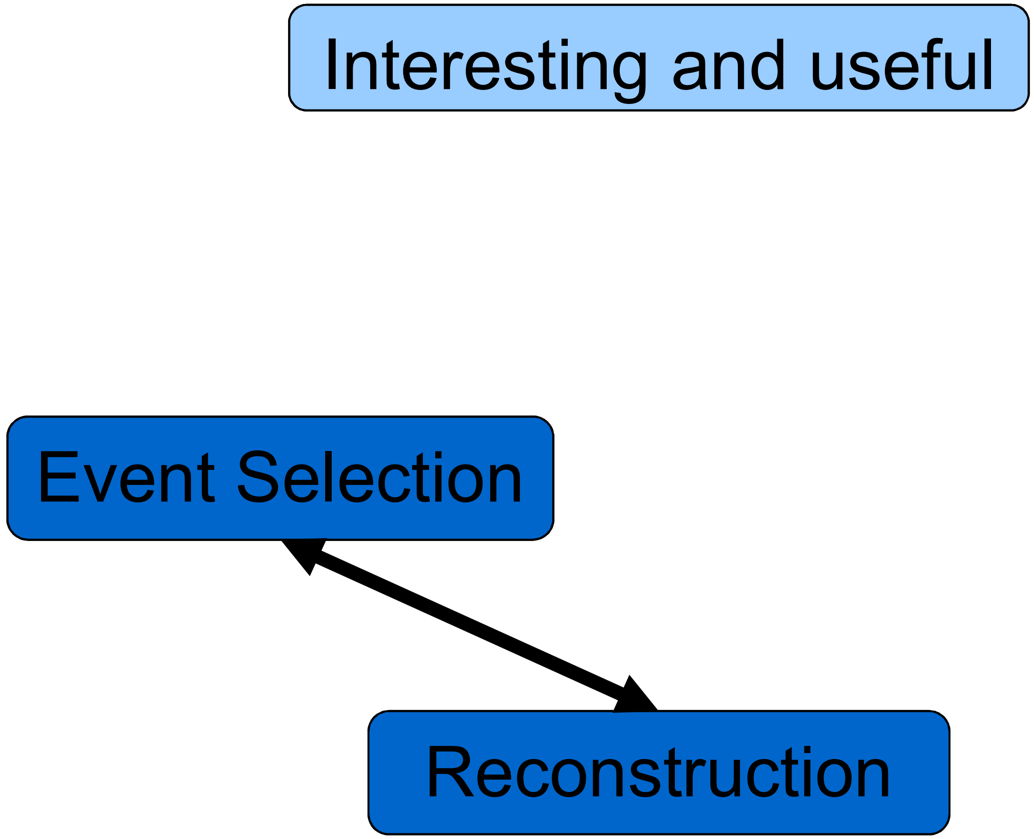
Interesting and useful

Design your analysis

Interesting and useful

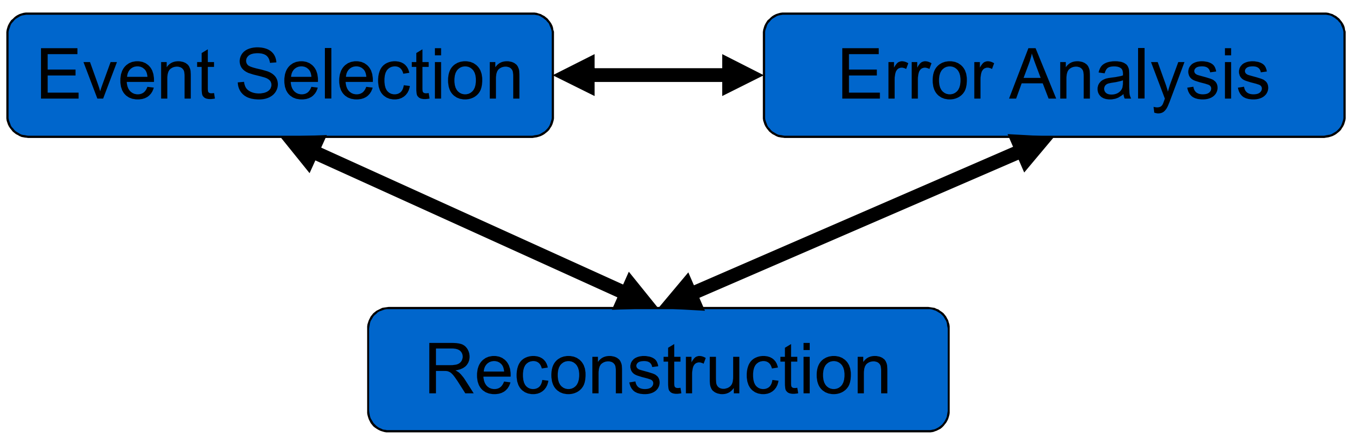
Event Selection

Design your analysis



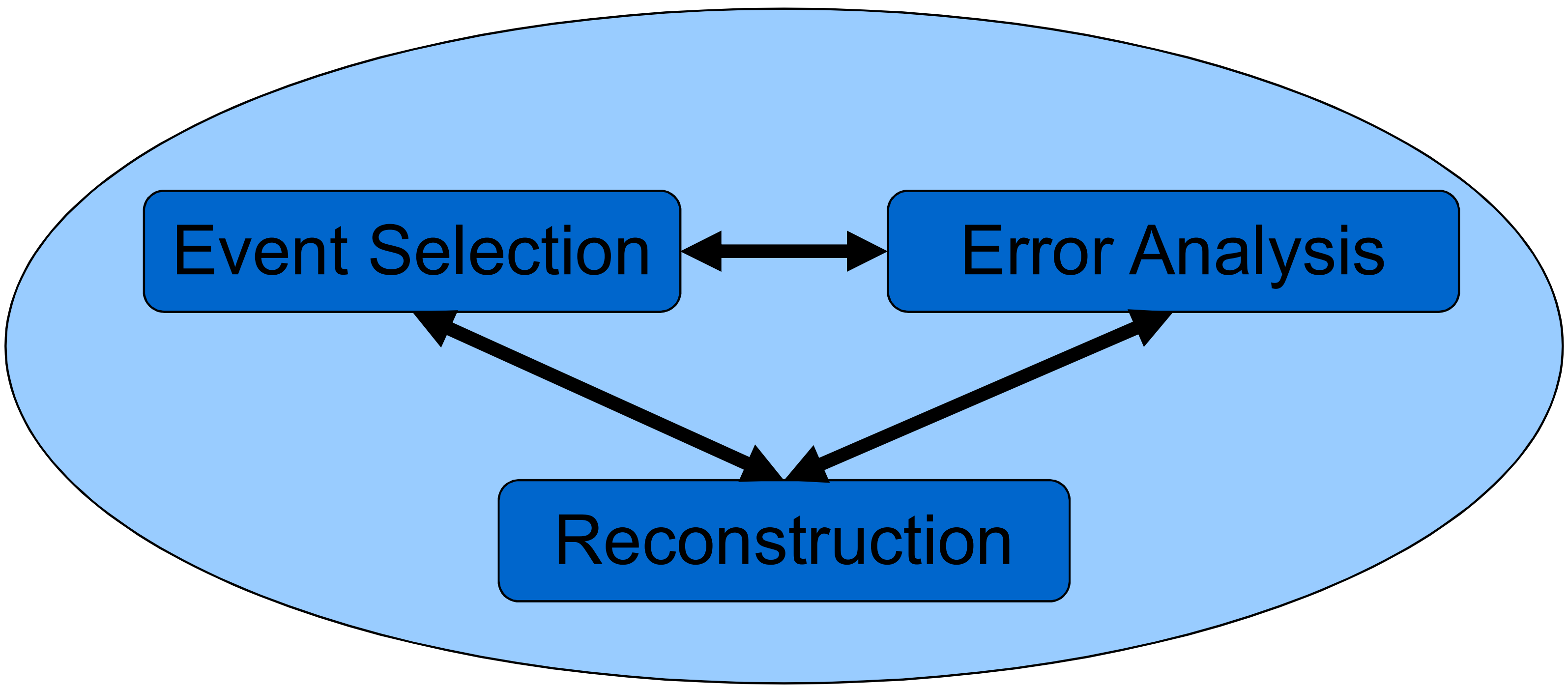
Design your analysis

Interesting and useful



Design your analysis

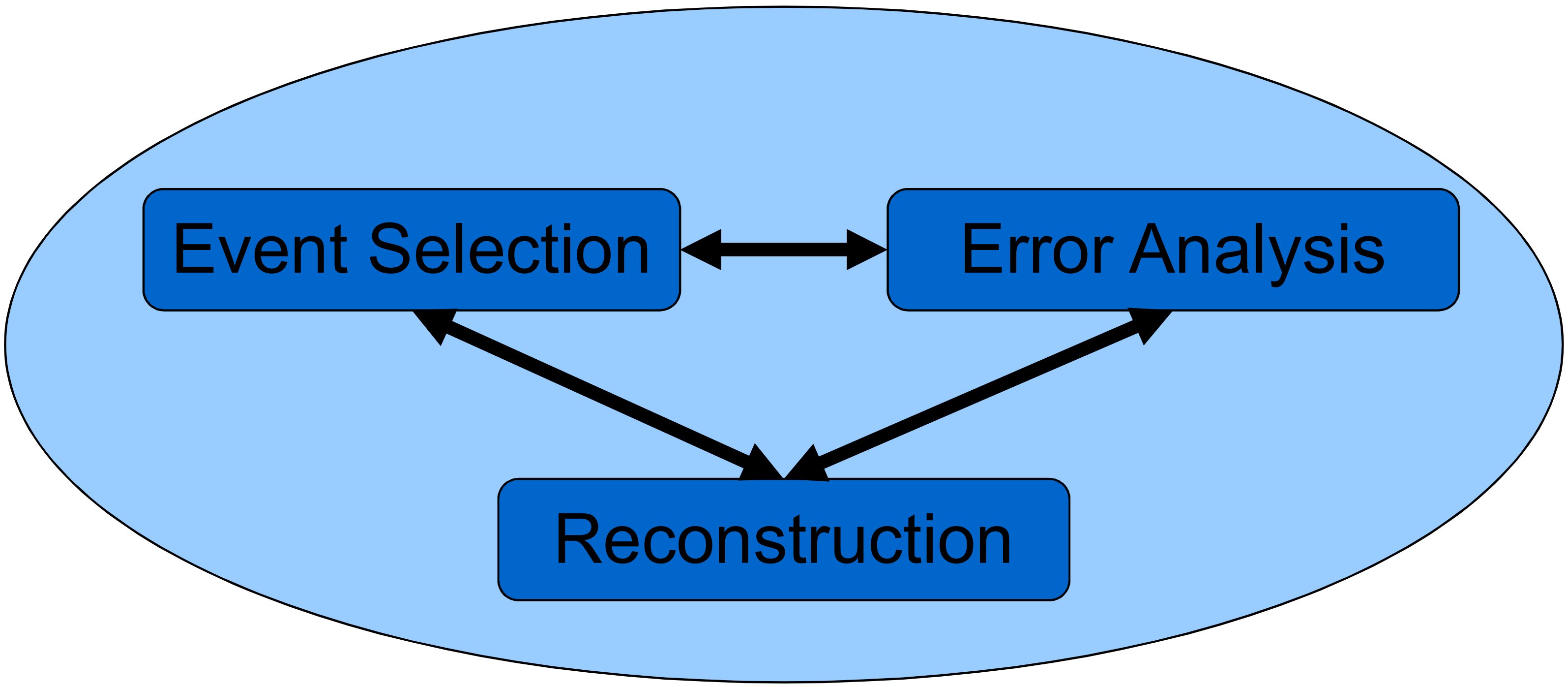
Interesting and useful



This loop is where, as an analyst, we spend almost all our time.

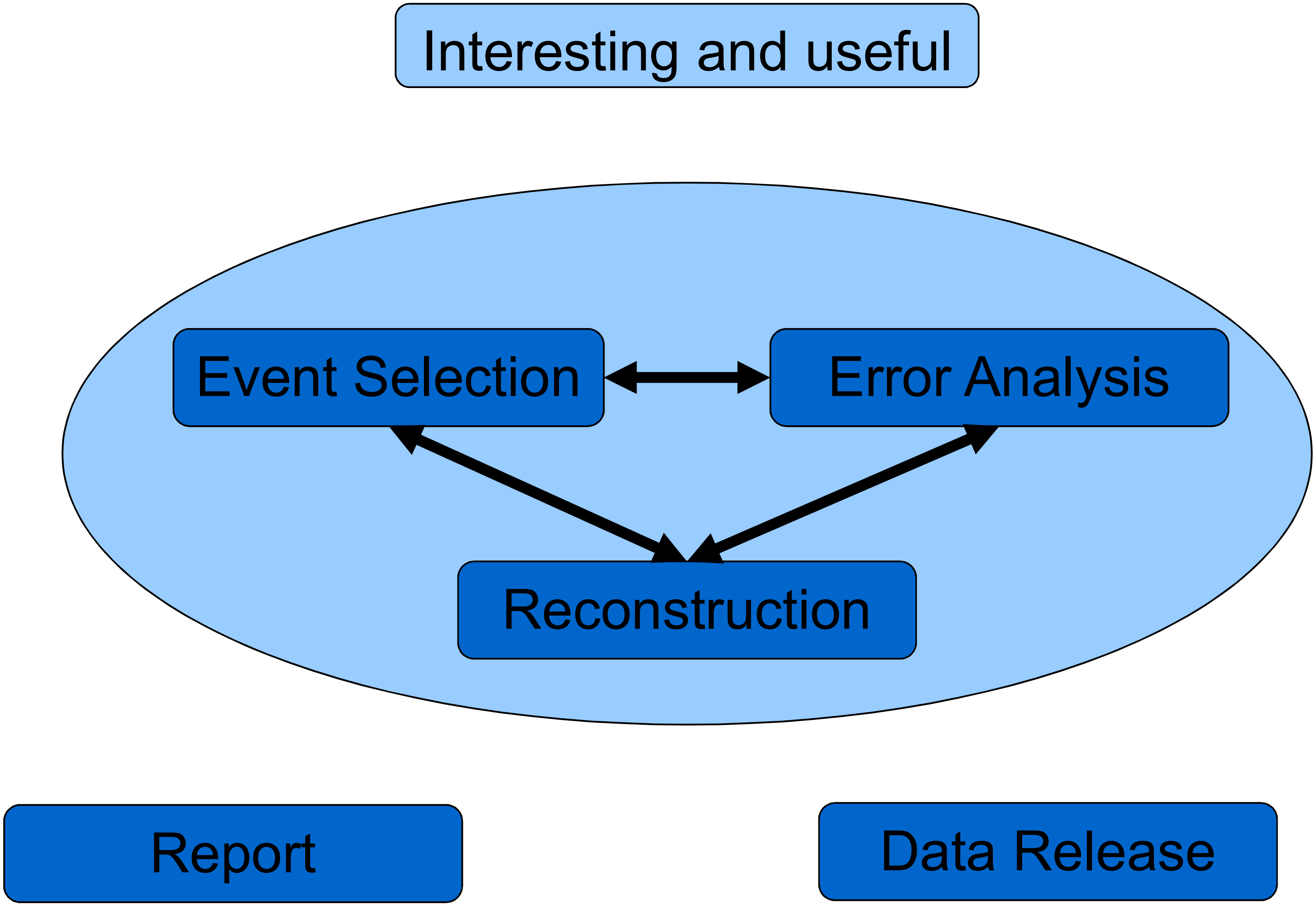
Design your analysis

Interesting and useful



Report

Design your analysis



So, How?

$$\left(\frac{d\sigma}{dx}\right)_\alpha = \frac{\sum_j U_{j\alpha}\left(N_{data,j} - N_{data,j}^{bkgd}\right)}{A_\alpha (\Phi T)(\Delta x)}$$

- Unfolding: $U_{j\alpha}$
- Selected data: $N_{data,j}$
- Data constrained background: $N_{data,j}^{bkgd}$
- Acceptance: A_α
- Flux and Targets: ΦT
- Bin width: Δx

Do not trust the simulated background!

$$\left(\frac{d\sigma}{dx}\right)_\alpha = \frac{\sum_j \boxed{} \left(N_{data,j} - N^{bkgd}_{data,j}\right)}{\boxed{}\left(\boxed{}\right)\left(\boxed{}\right)}$$

- Large uncertainties (typically) on the backgrounds in the model
- We have all sorts of data to constrain these – use it.
- Sidebands, shape analysis, anything to help understand your background

Unfolding

$$\left(\frac{d\sigma}{dx}\right)_\alpha = \frac{\sum_j U_{j\alpha}\left(\quad - \quad\right)}{\quad\left(\quad\right)\left(\quad\right)}$$

- MINERvA's goal in analysis design is to make the unfolding handle detector smearing, but not model effects.
  - Unfold in observable variables, not model variables
- Framework used is iterative unfolding implemented in RooUnfold
- MINERvA does extensive testing to understand an appropriate number of iterations
- Unfold in all systematic universes when possible
  - Need to watch for statistical fluctuations inflating systematic uncertainties

Acceptance

$$\left(\frac{d\sigma}{dx}\right)_\alpha = \frac{\sum_j \square \left(\square - \square \right)}{A_\alpha (\square)(\square)}$$

- To do full phase space or not that is the question
  - Fiducial Cross Sections (measure what you see) are more appealing
- <u>Design your signal to match what you reconstruct.</u>

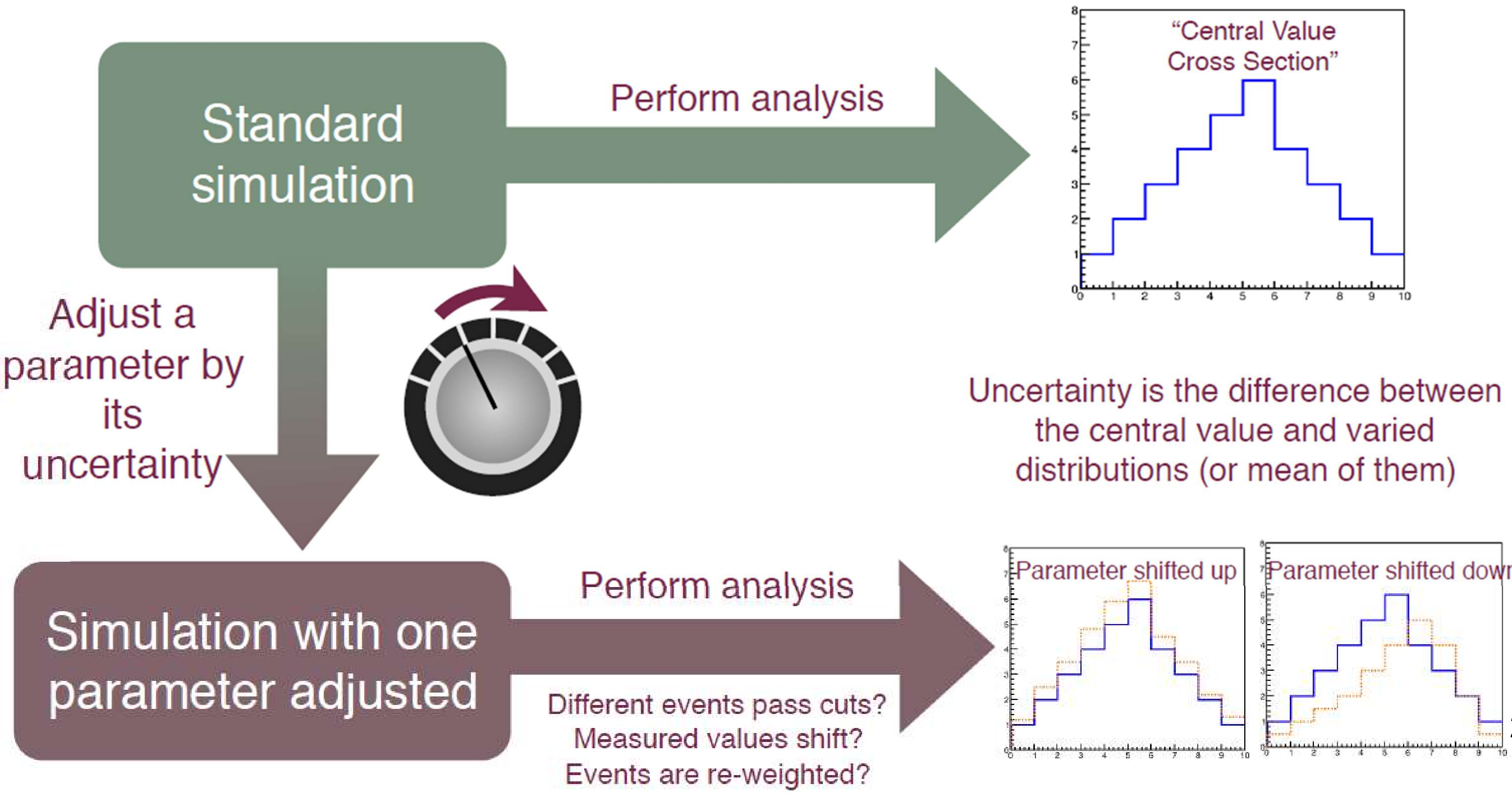
Framework

2-stages of analysis

- General reconstruction is run – slice time, make long tracks, match to MINOS
- (1) Every analysis designs an analysis tool
  - Can run short trackers, shower reconstruction, finer time slicing, Michel taggers, etc.
  - Can, in principle, do a completely different set of reconstruction
  - Output = Anatuples
- (2) Macros use anatuples to do all the steps I described earlier – I focus here today

MINERvA in the Multi-verse

- We use the "many universe" method to evaluate systematics
  - That means, LOTS of histograms
  - Do you like bookkeeping?



MINERvA's Swiss Army Knife

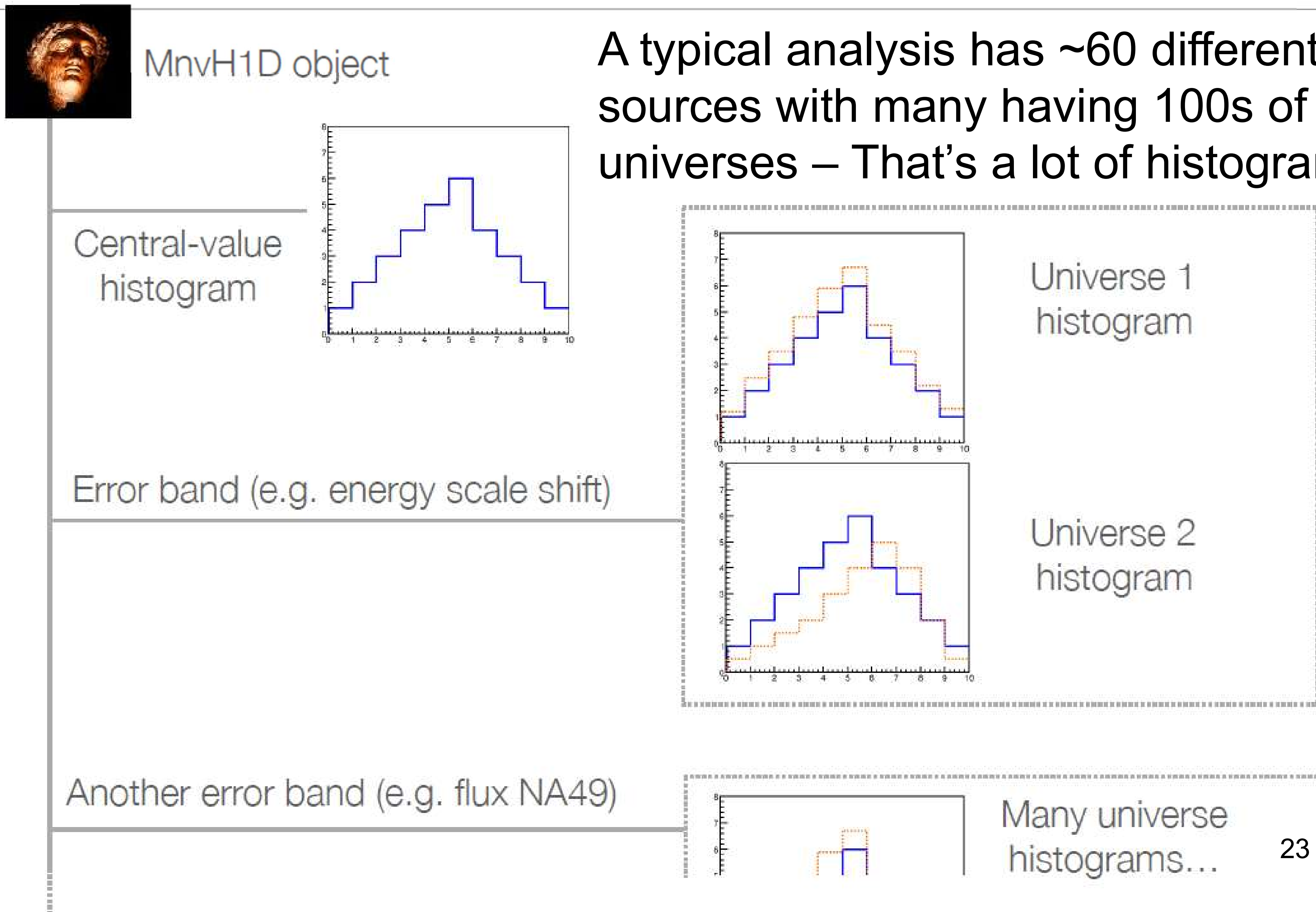
- MnvH1D and MnvH2D are the general tools and container for our analyses
- Supported with generalized tools to provide various systematic universes
  - Secondary interaction in the detector
  - Flux uncertainties

What’s a MnvH1D?

- MINERvA histogram object which is an extension to the ROOT TH1 object.
- It does the bookkeeping of all the systematic universes
- It handles all the error propagation and calculation
- If given an MnvH1D you have all the components you’d like to report a cross section

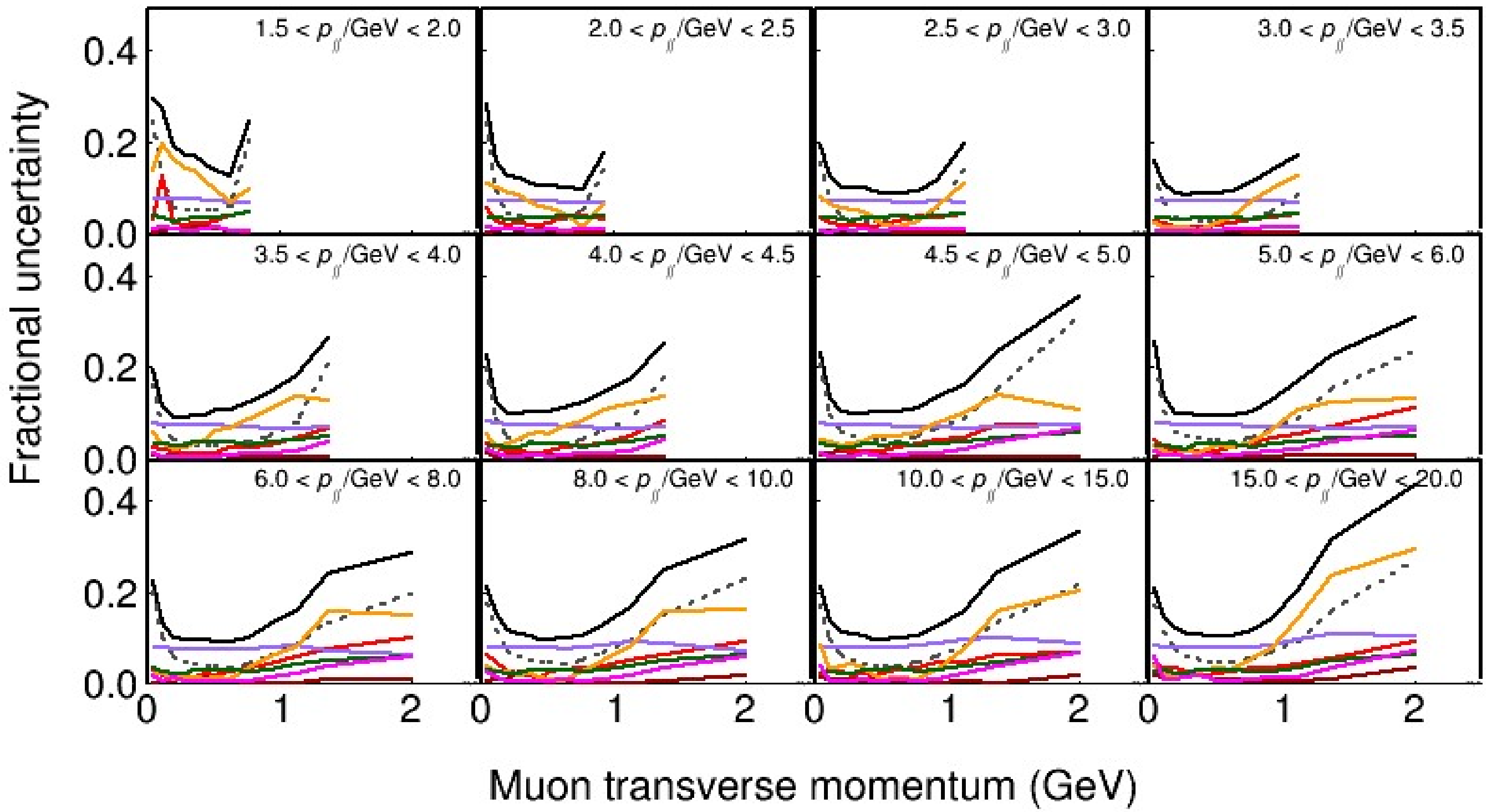
MnvH1D

A typical analysis has ~60 different sources with many having 100s of universes – That's a lot of histograms!



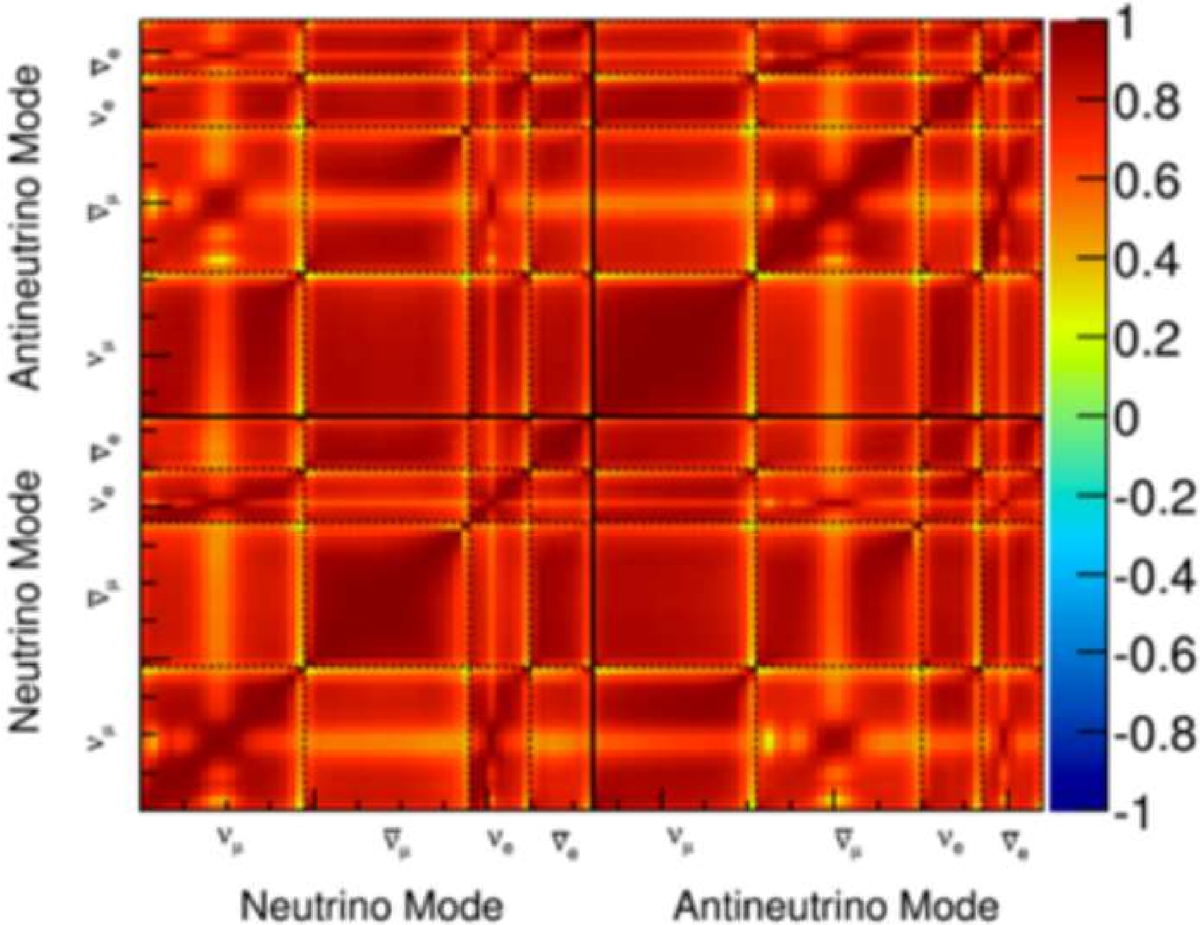
This allows….

- Error summaries plots with ease –



This allows….

- Correlation matrices with ease – Also, individual sources if you want

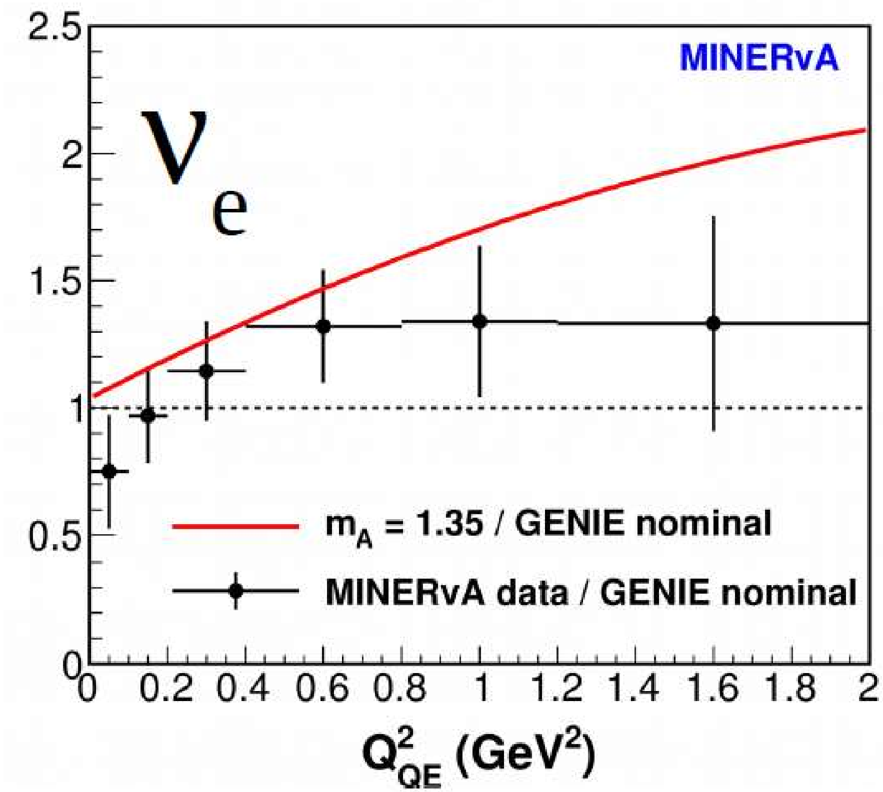


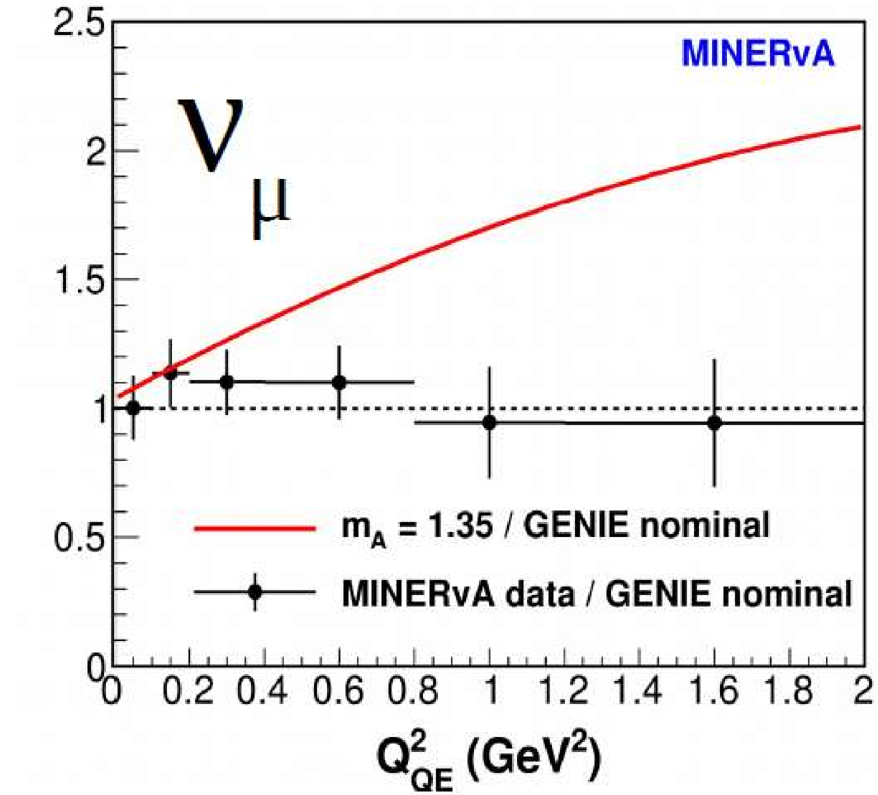
Correlation between flux bins in energy and species provided to DUNE

This allows….

- Ratio analyses with ease

J. Wolcott JTEP Seminar

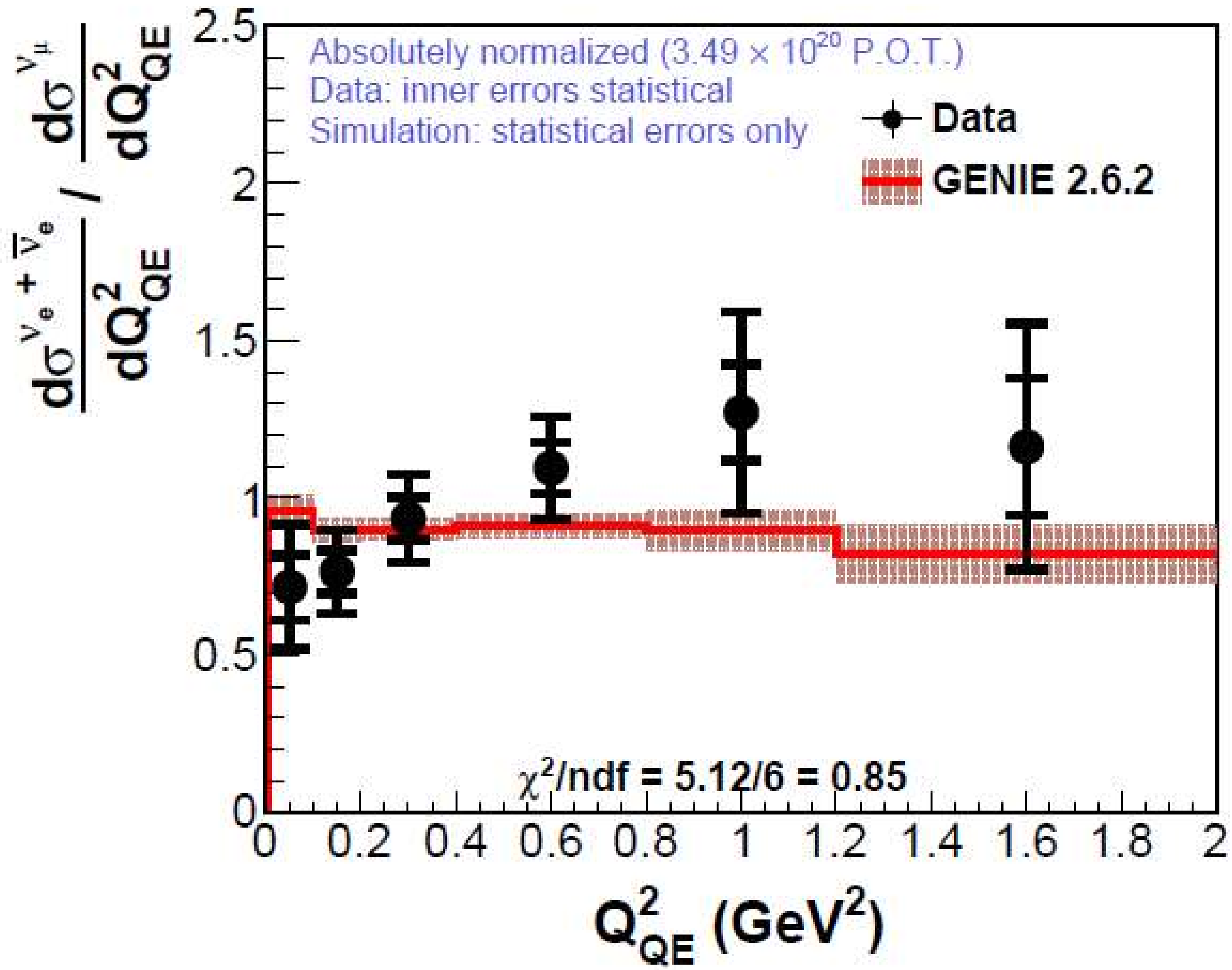




This allows….

- Ratio analyses with ease

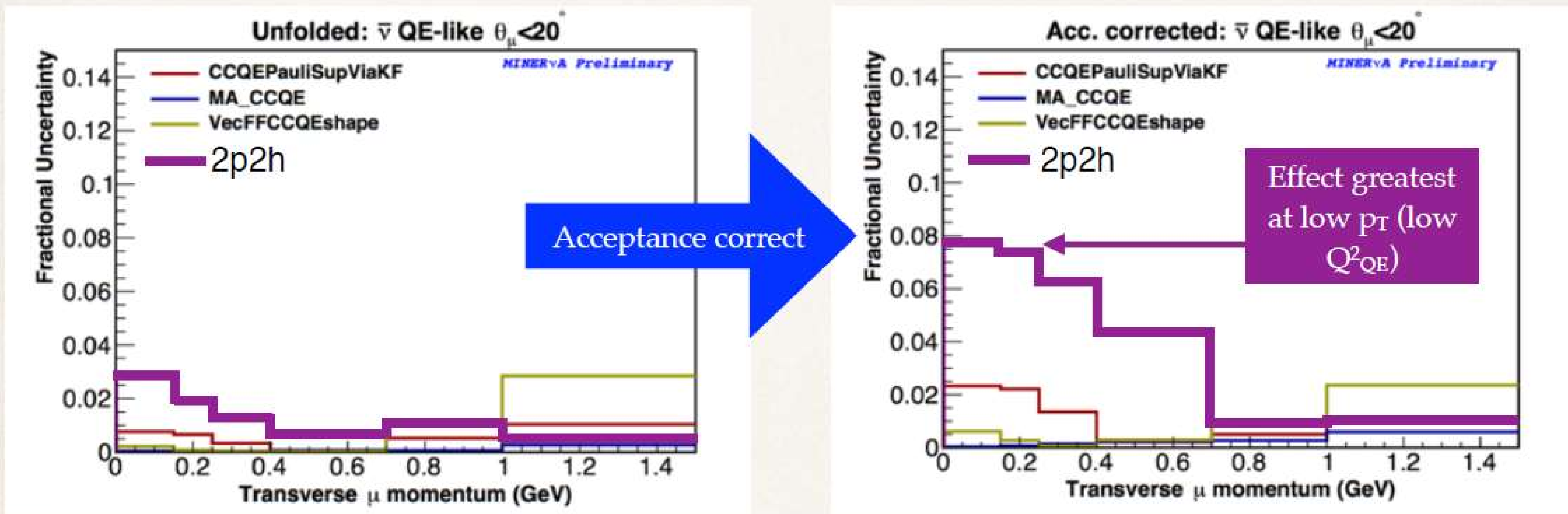
Phys. Rev. Lett. 116, 081802 (2016)



This allows….

- Understand how systematics change with each step of the extraction process since you start off with MnvH1D's from the start.
- This allows an analyst to improve selection and background constraint methods and quantify quickly how effective they are.

Example where this is useful!



Background subtraction:
small uncertainty due to differing background fractions in the 2 samples

Acceptance correction:
large increase in uncertainty as acceptance is very different for 2p2h events

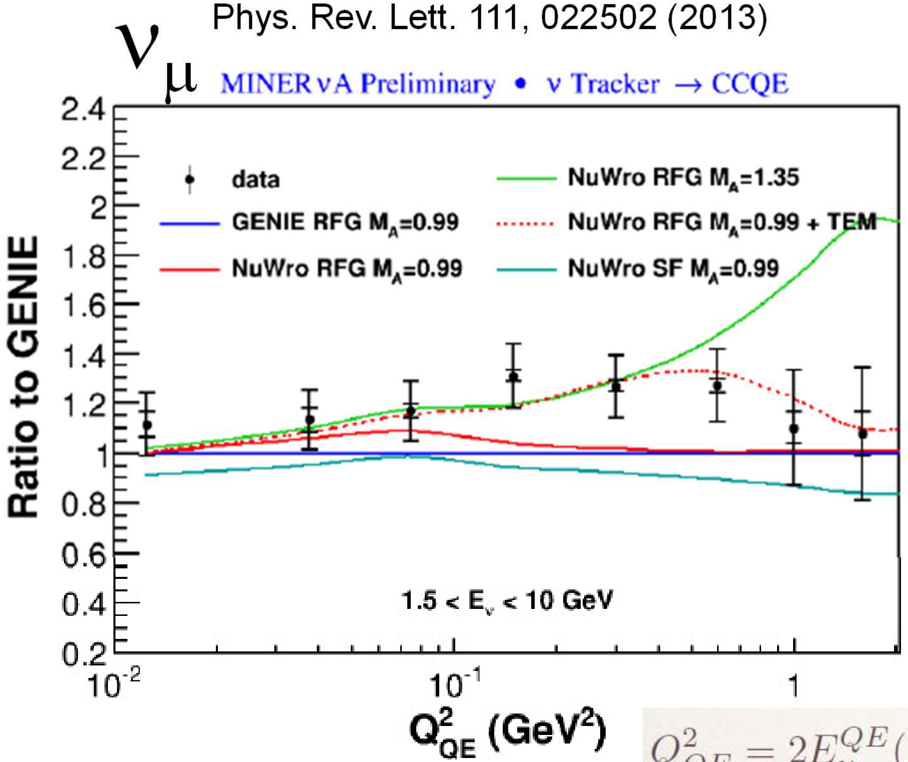
C. Patrick FNAL W&C Seminar, 17 June 2016

Specific Case: CC0π

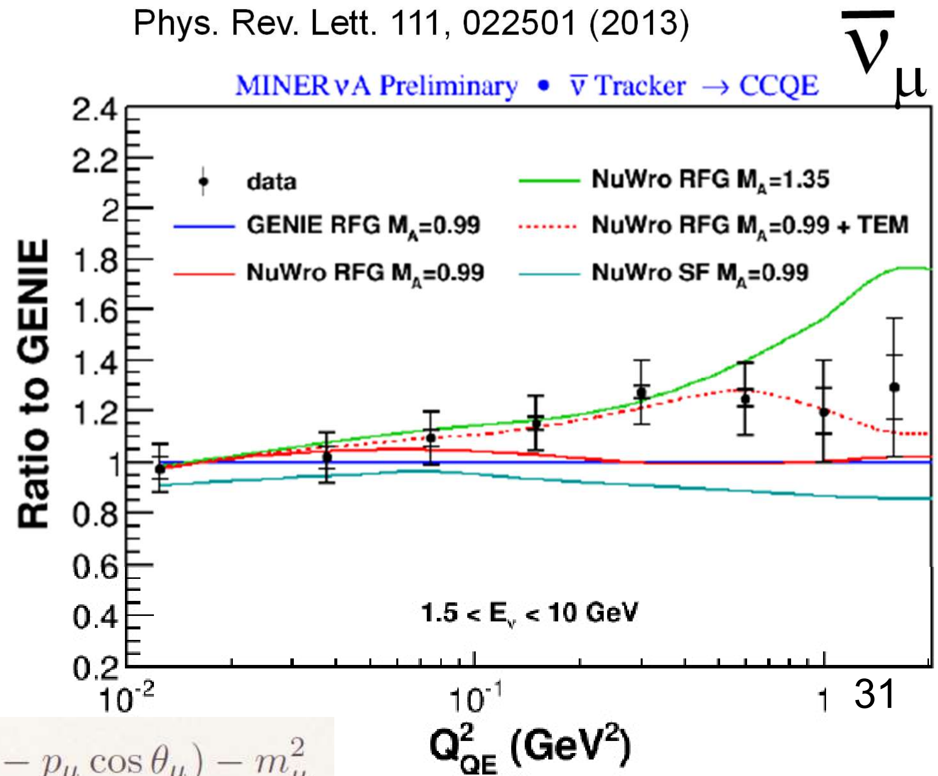
What and Why CC0π

- The explosion of models and generator improvements provide an expanded world to compare data to
- Original MINERvA CCQE measurements indicated our data preferred 2p2h-like effects

Phys. Rev. Lett. 111, 022502 (2013)



Phys. Rev. Lett. 111, 022501 (2013)



$$Q^2_{QE} = 2E_\nu^{QE}(E_\mu - p_\mu \cos\theta_\mu) - m_\mu^2$$

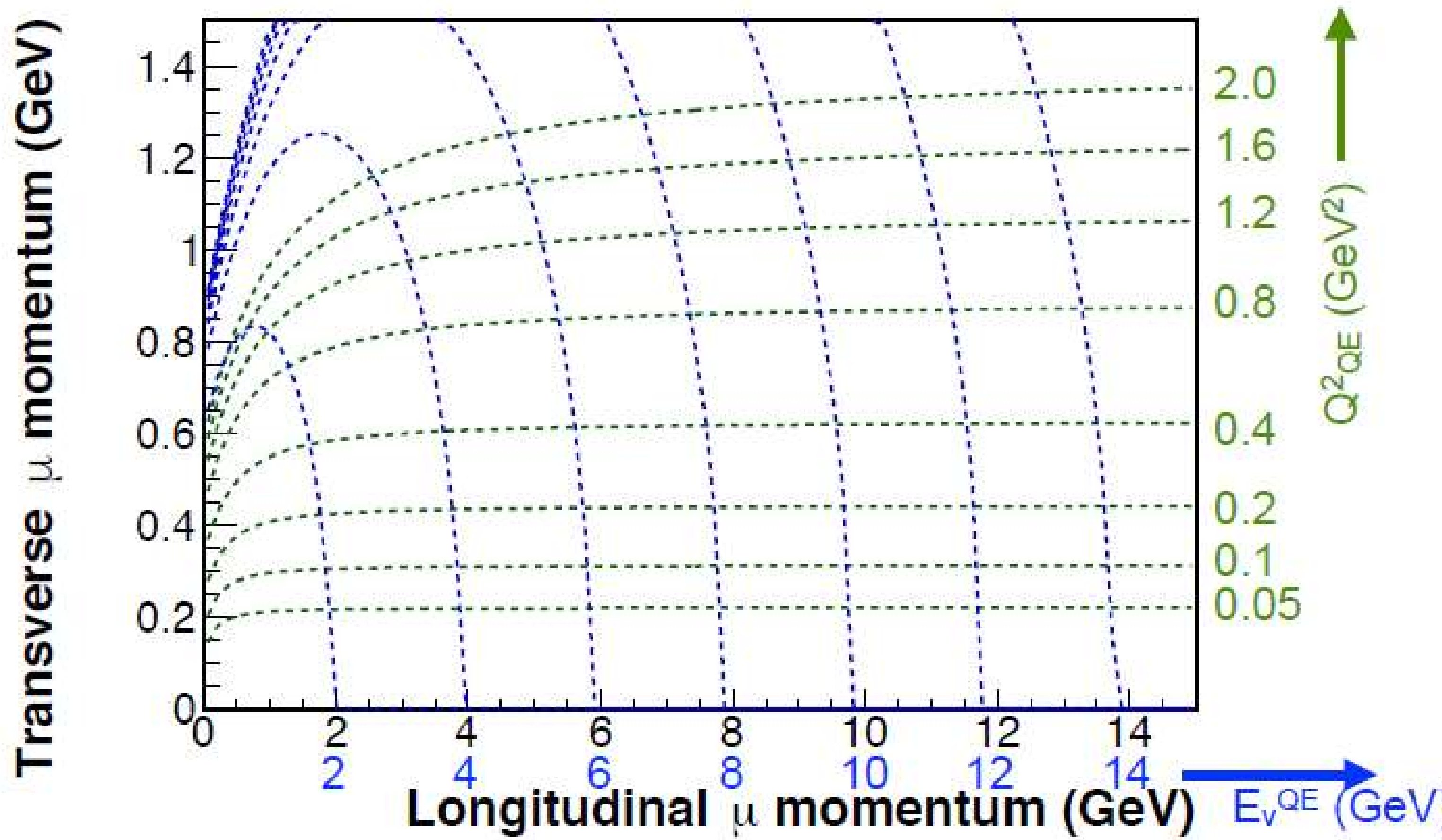
What and Why CC0π

- Lots learned about techniques and new modeling necessitates redirection of the analysis
- Advancement of reconstruction techniques allows for a different type of analysis
- General idea: Select events with/without visible extra tracks, reject pions, constrain pions, and report what's left, CC0π

Signal Definition and Deliverable

- CCQE-like, CC0π
  - **<u>ANY</u>** number of nucleons, energy doesn't matter
    - Not rejecting events based on reconstructed multiplicity
  - **<u>NO</u>** pions, heavy baryons
  - **<u>NO</u>** gammas > 10MeV
    - Data has de-excitation gammas, GENIE simulates this on oxygen, FV has a few % oxygen
  - Muon angle < 20 degrees
    - Geometric acceptance of MINERvA+MINOS
- Output: Muon 2D differential in $P_t P_{||}$

Why P_t $P_{||}$?



These are the variables we directly measure experimentally
They align, mostly, with interesting “QE” variables

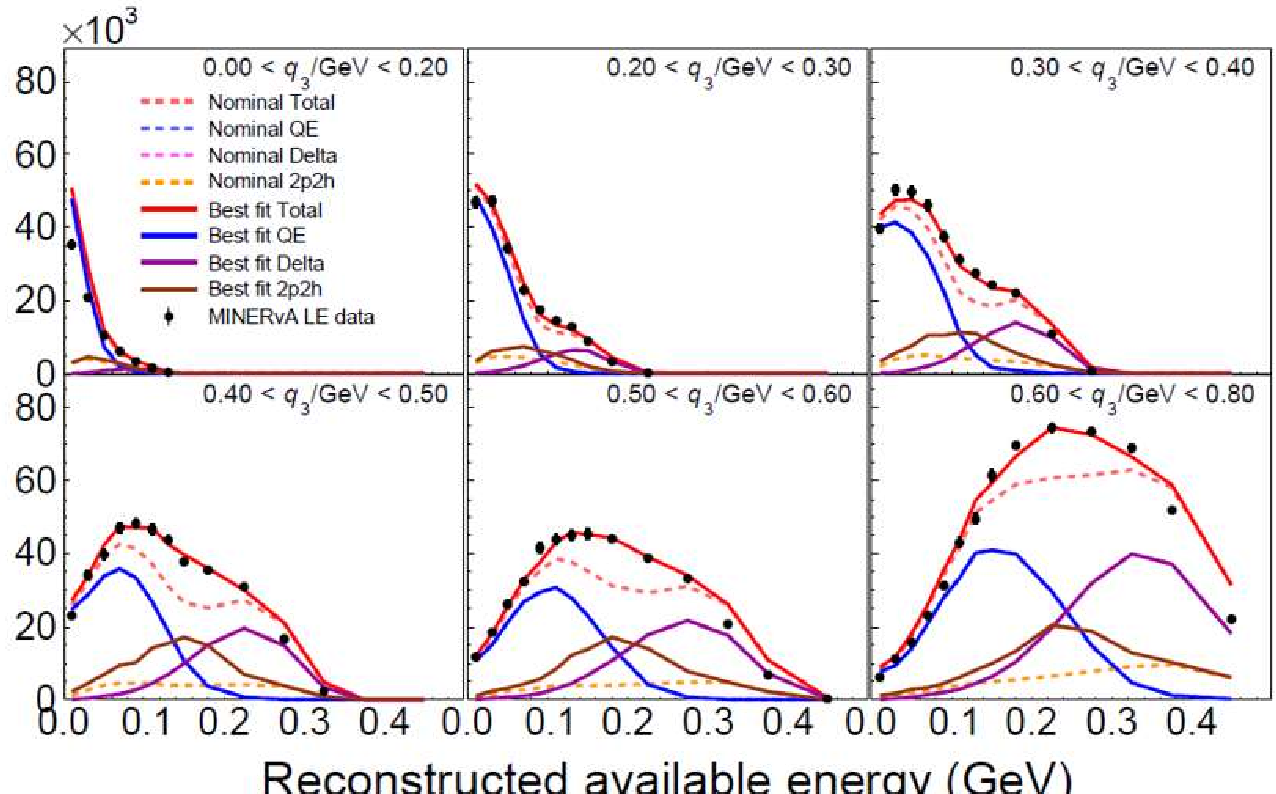
What MC do we use?

- GENIE 2.8.4[1] is the foundation
- Latest flux [2]
- Non-resonant pion production reduced by 57%[3]
- Valencia RPA suppression applied to CCQE [4]
- Valencia 2p2h[5]
- Low recoil analysis fit based on [6]

[1] Nucl.Instrum.Meth.A614 (2010) 87-104
[2] Phys. Rev. D 94, 092005 (2016)
[3] Phys. Rev. D 90, 112017 (2014)
[4] PRC 70, 055503 (2004); PRC 83, 045501 (2011)
[5] PRC 70, 055503 (2004); PRD 88, 113007 (2013)
[6] Phys. Rev. Lett. 116, 071802 (2016)

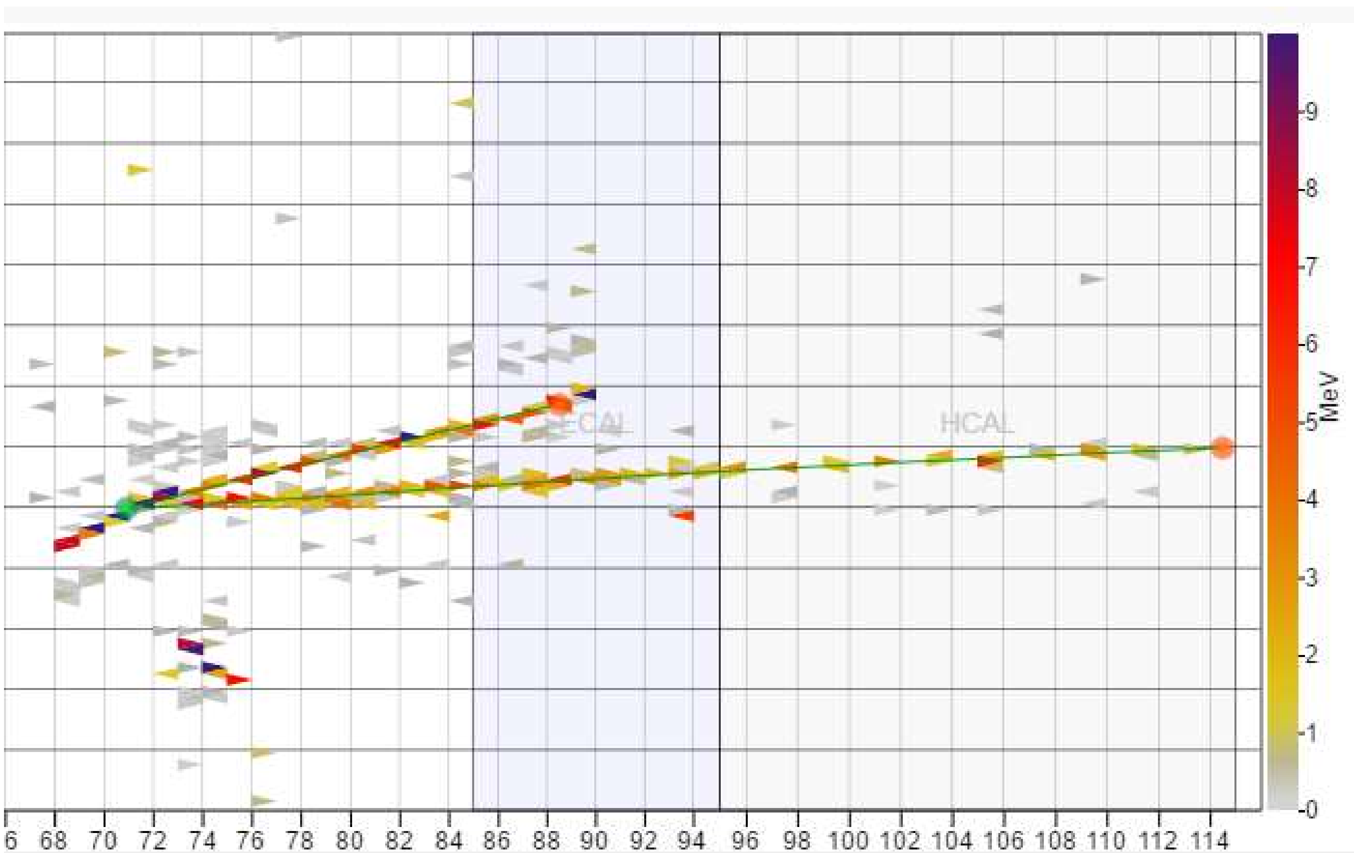
The low recoil fit

- Fit a 2D Gaussian in true (q_0,q_3) as a reweighting function to the 2p2h contributions to get the best agreement
- Does not scale true QE or resonant production.
- More on how we treat uncertainty later.



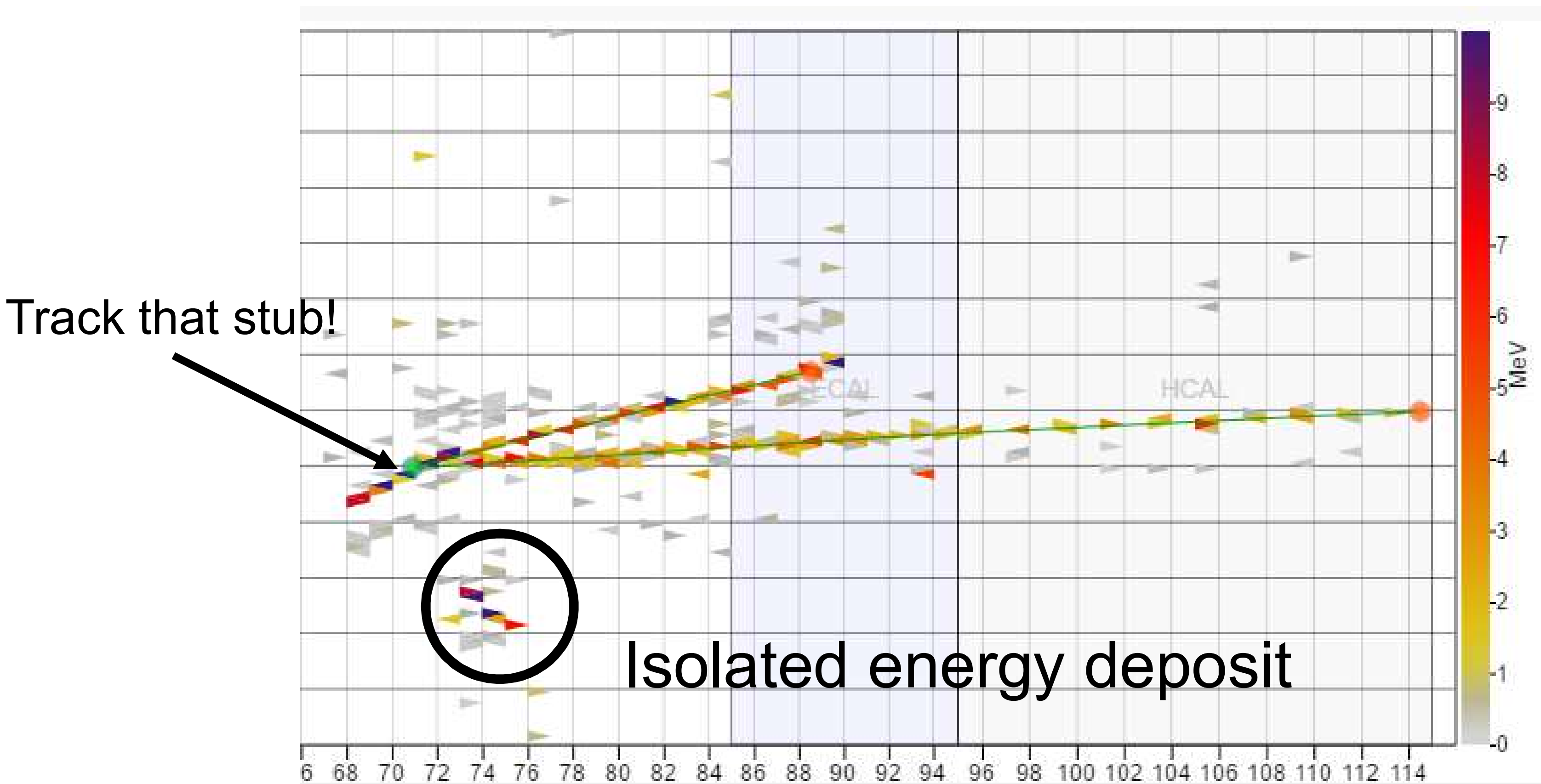
Event Selection In Pictures

- Strategy –Track pions and protons
- Select events based on particle identification
- Constrain pion background using side band fits



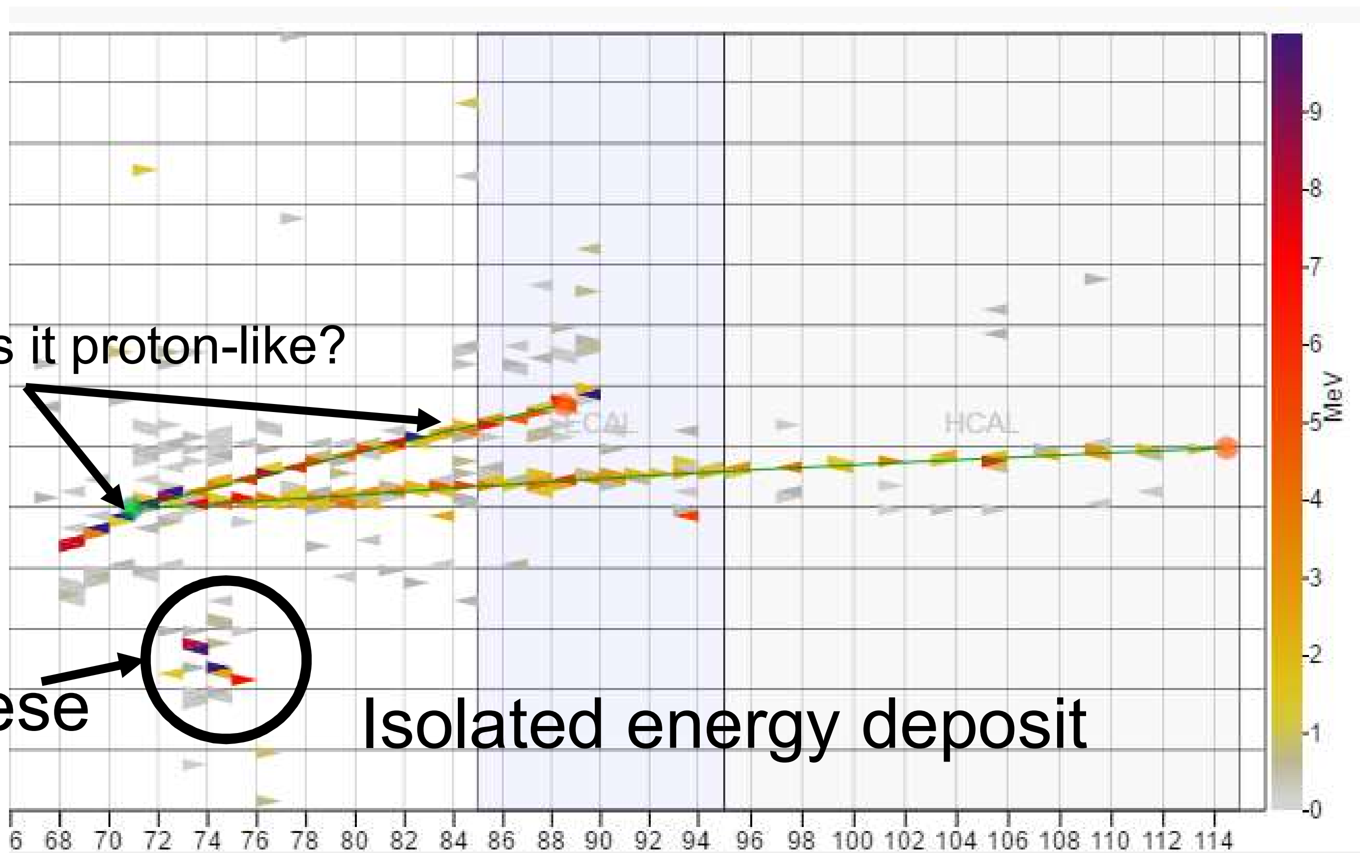
Event Selection In Pictures

- Strategy –Track pions and protons
- Select events based on particle identification
- Constrain pion background using side band fits



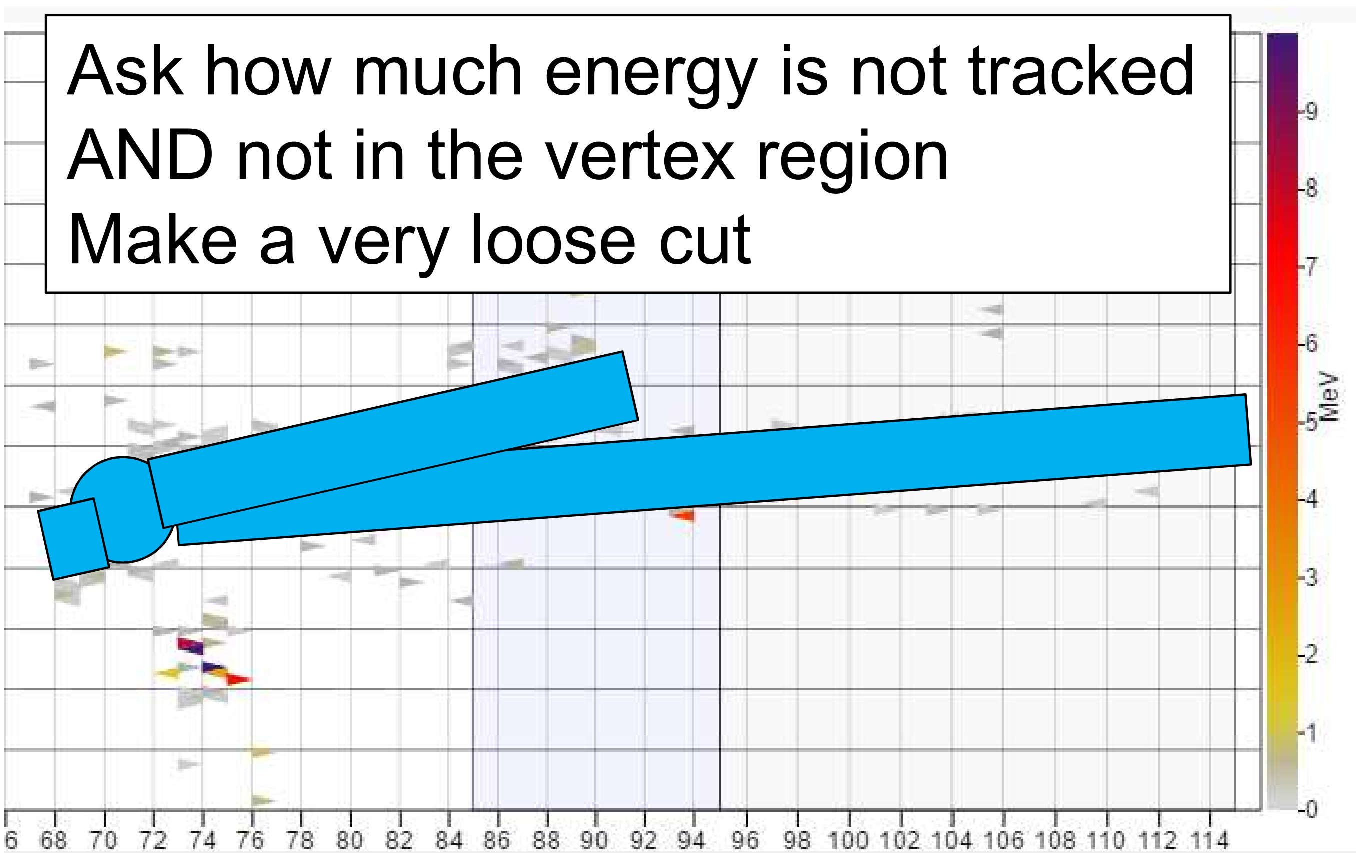
Event Selection In Pictures

- Strategy –Track pions and protons
- Select events based on particle identification
- Constrain pion background using side band fits



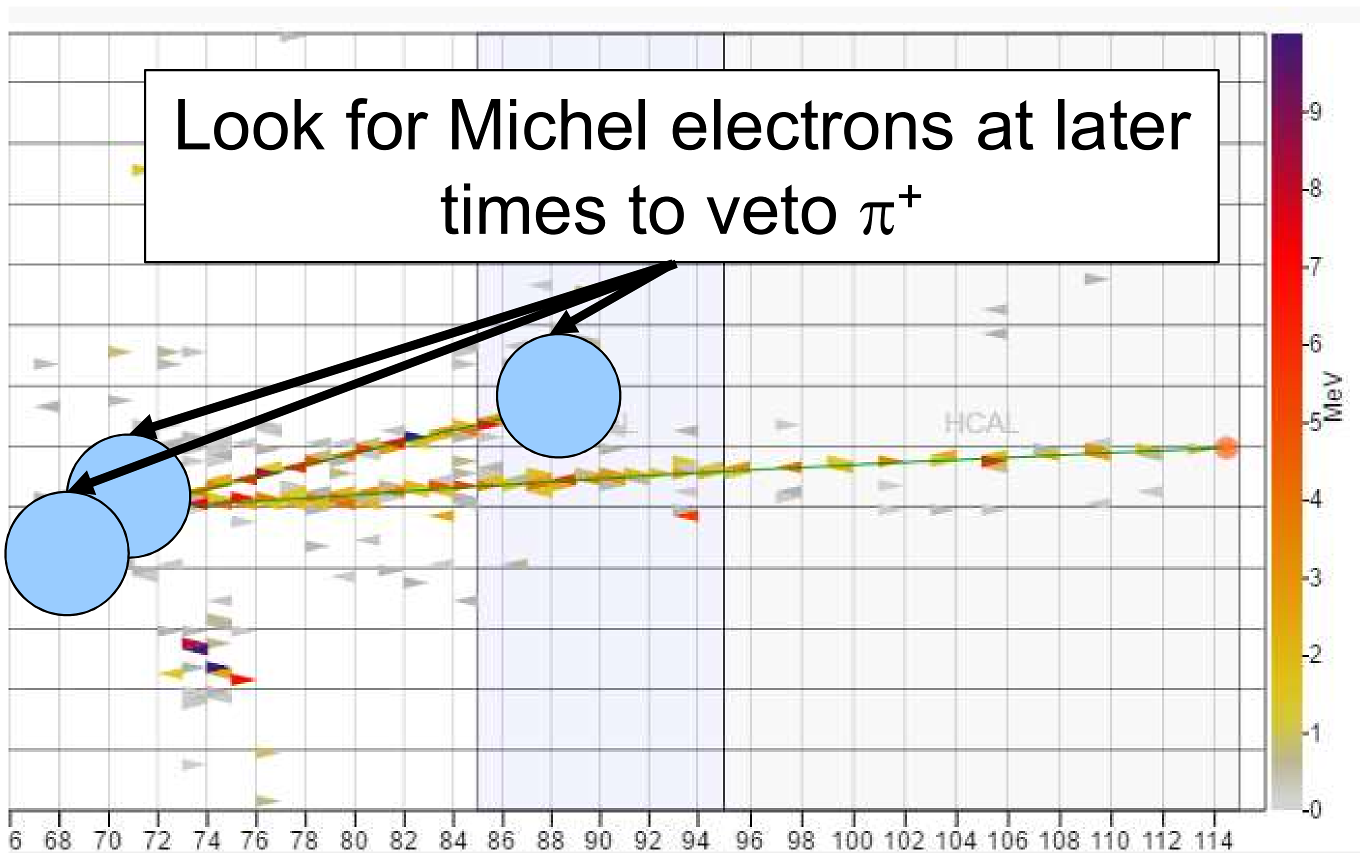
Event Selection In Pictures

- Strategy –Track pions and protons
- Select events based on particle identification
- Constrain pion background using side band fits



Event Selection In Pictures

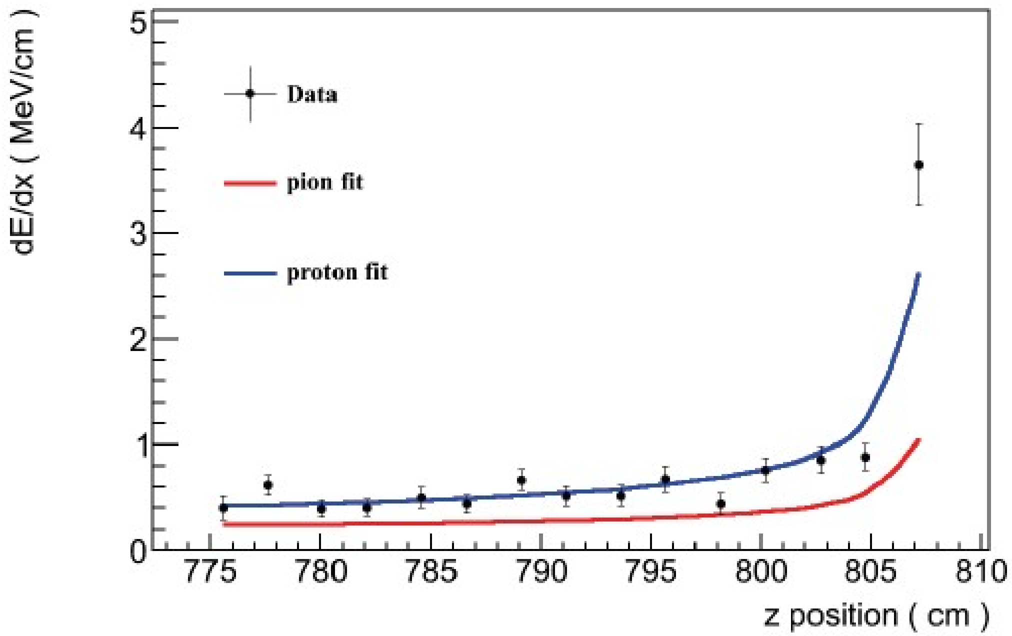
- Strategy –Track pions and protons
- Select events based on particle identification
- Constrain pion background using side band fits



Summary of Cuts

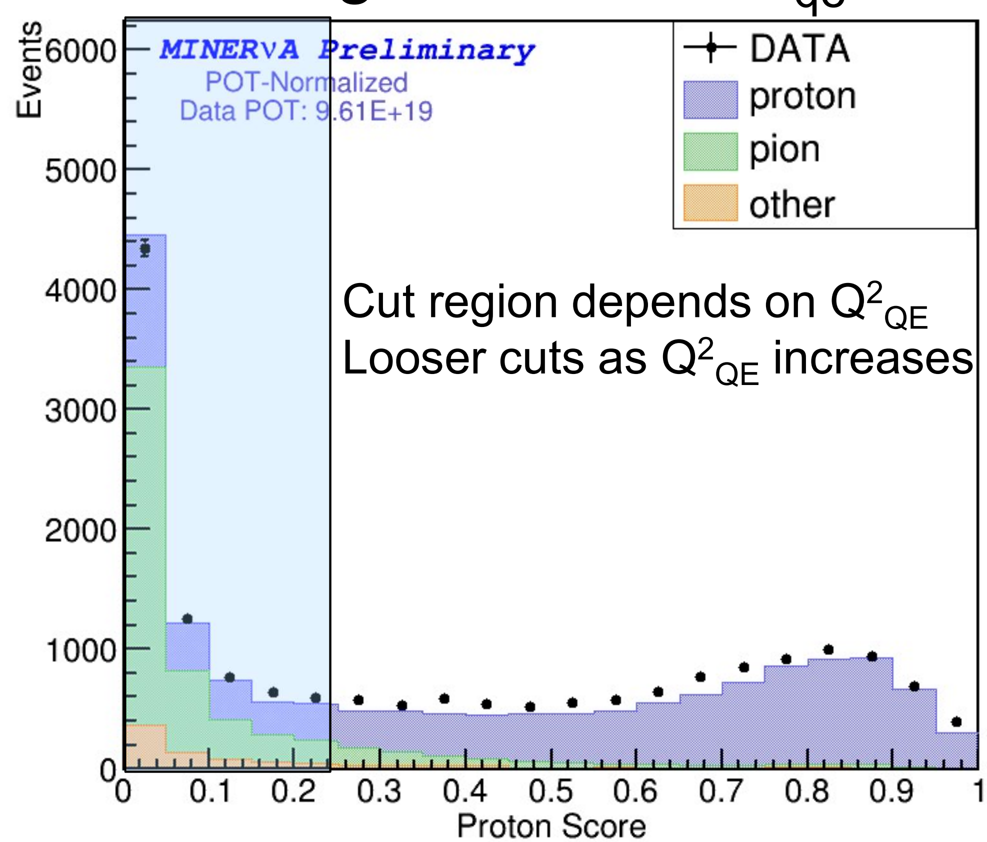
- dE/dX based π/p rejection – Targets $\pi^{+/-}$
- Isolated energy clusters – Targets π^0
- Michel tagging – Targets π^+
- Loose recoil cut – Targets inelastic events

PID broken down by particle



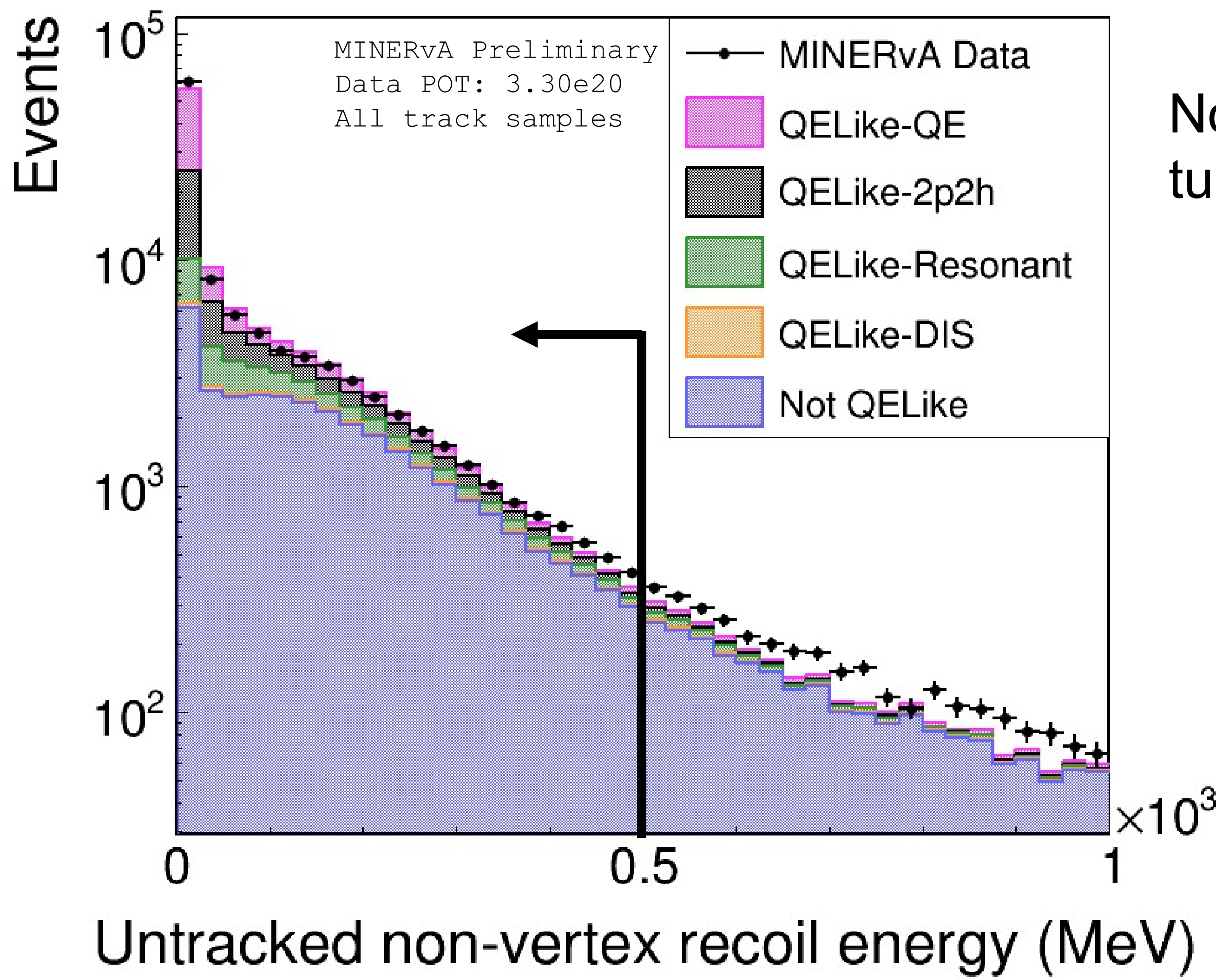
- This is applied to all tracks which are not the muon
- Loosen cut as Q^2_{QE} increases because protons are harder and interact more

Integrated over Q^2_{qe}



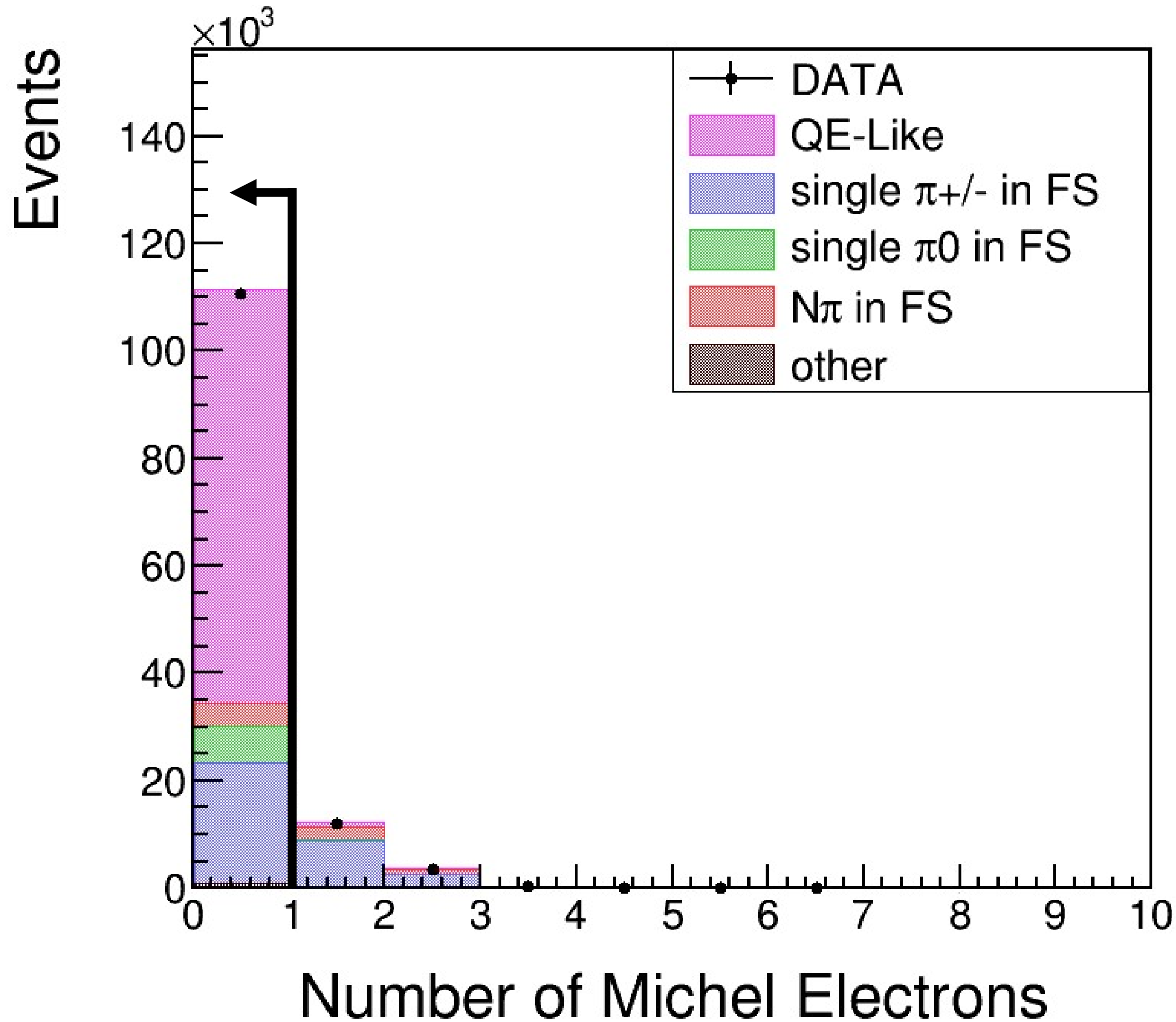
Hadronic Recoil

- Very loose cut on the untracked energy outside the vertex region.
- Sample here passes the rest of the selection.

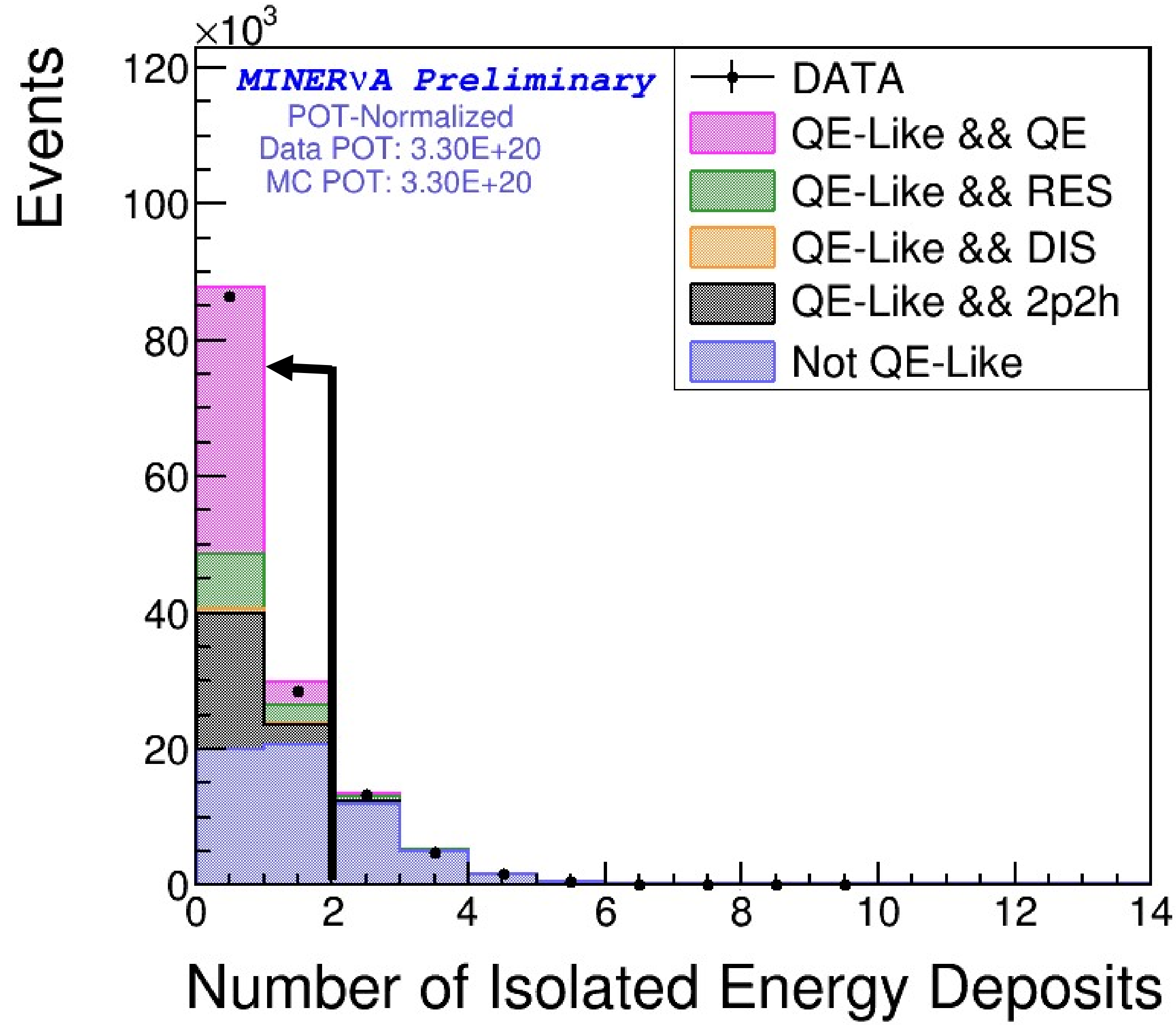


No background tuning applied

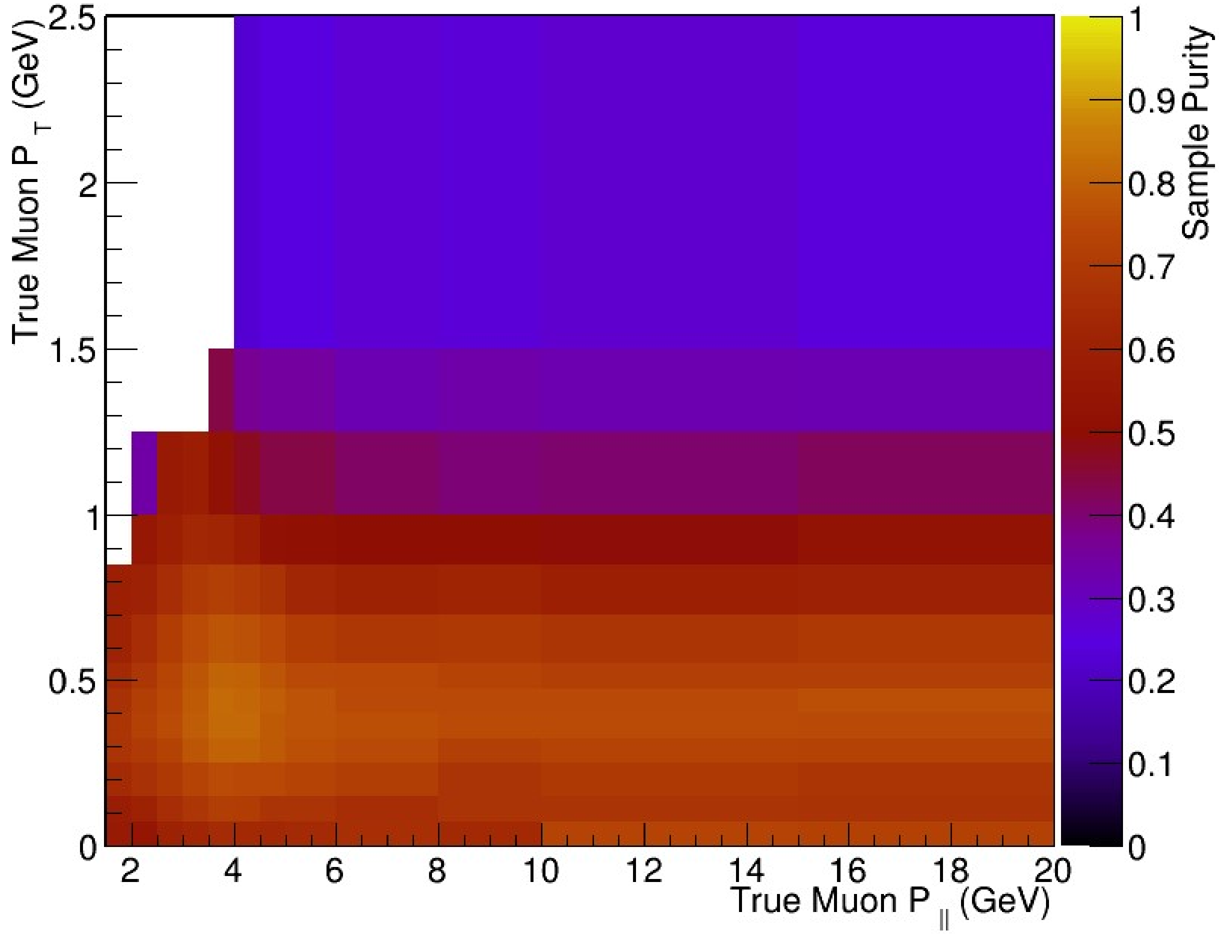
Number of Michel Electrons



Isolated energy deposits



Purity

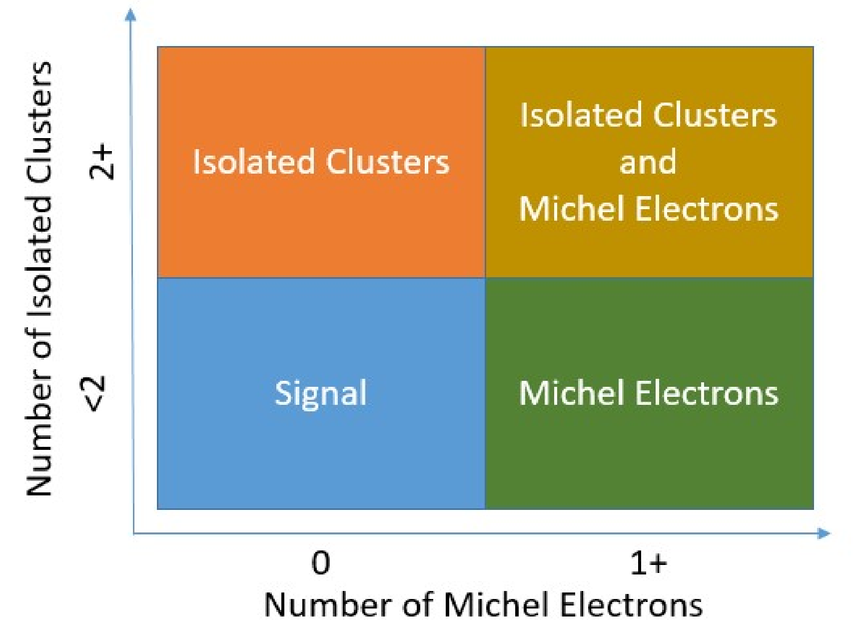


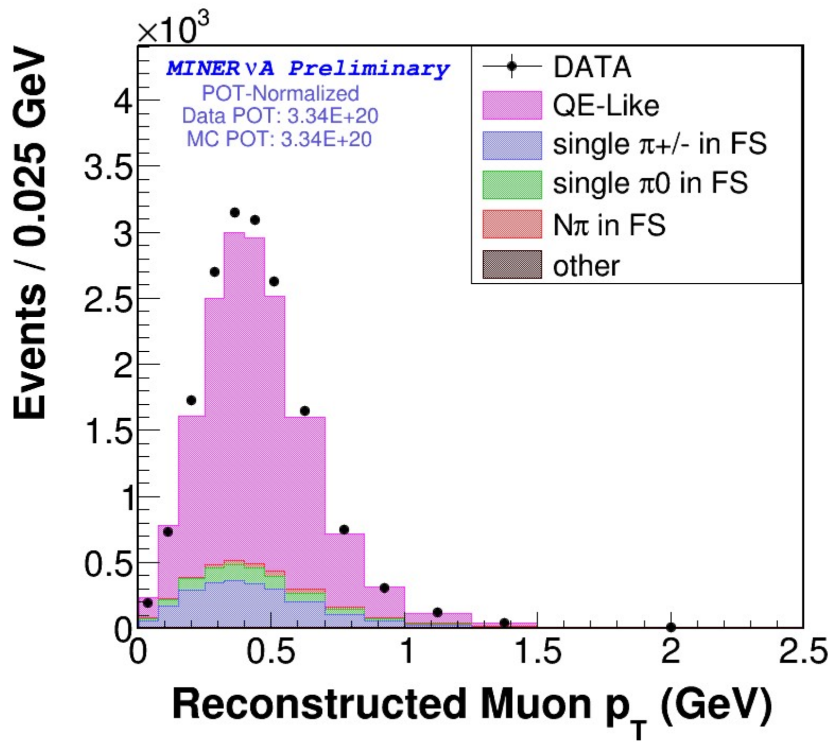
True Muon P T (GeV)
2.5
2
1.5
1
0.5
0
2
4
6
8
10
12
14
16
18
20
True Muon P || (GeV)
Sample Purity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

Constrain the Background

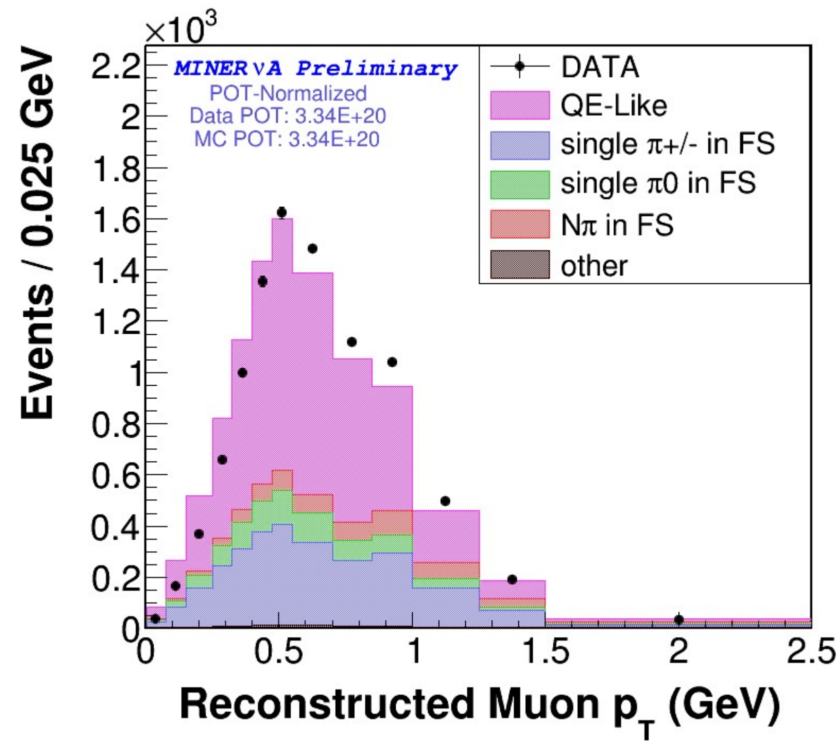
- Extract scaling factors to control single pion events (charged or neutral) and multi-pion events.
- 3 sidebands used
  - Michel electron(s)
  - >1 isolated cluster of energy
  - Michel electron(s) AND >1 isolated cluster of energy
- Simultaneous fit in P_t bins (may combine)
- Muon only and Muon + N tracks treated separately

Signal



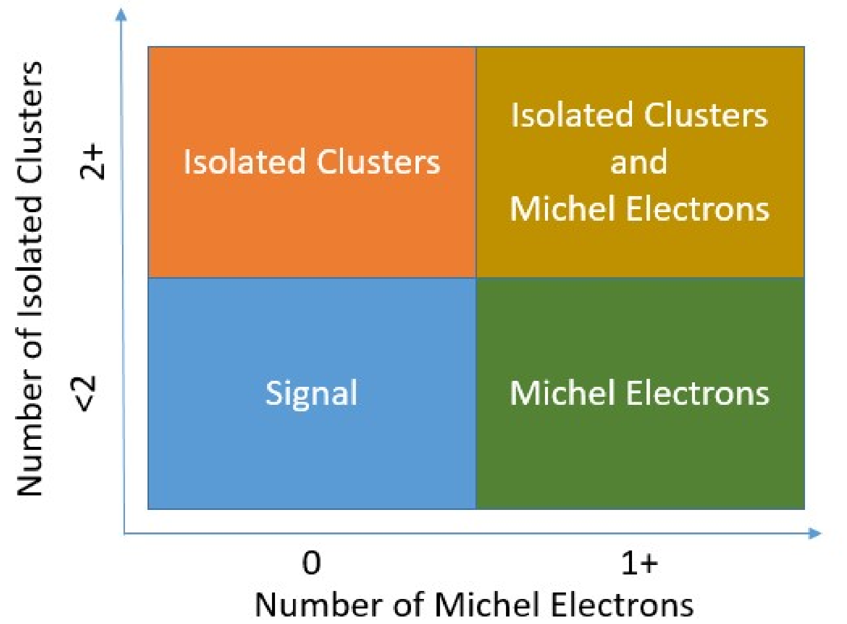


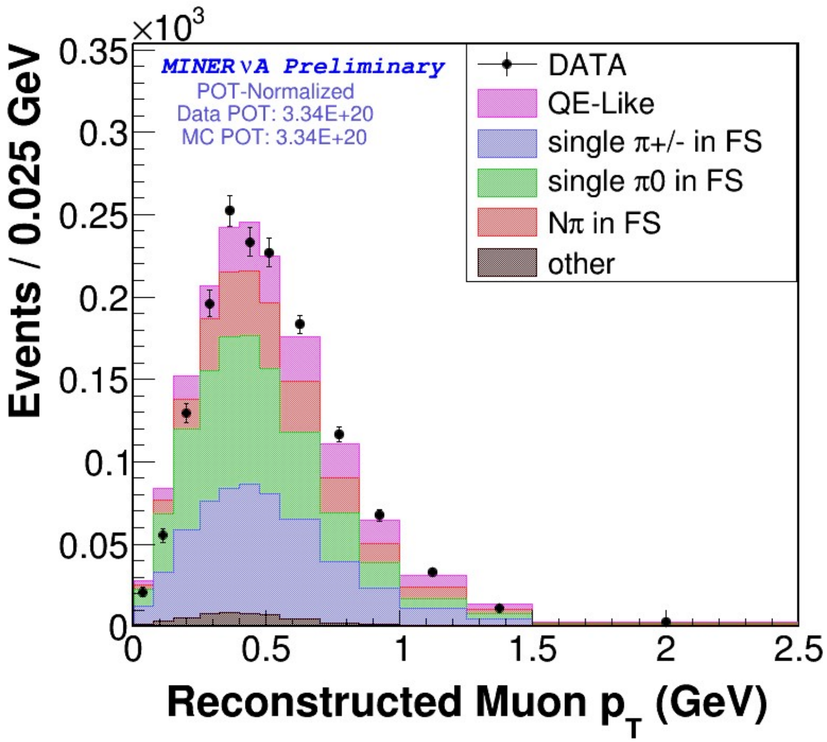
1-track



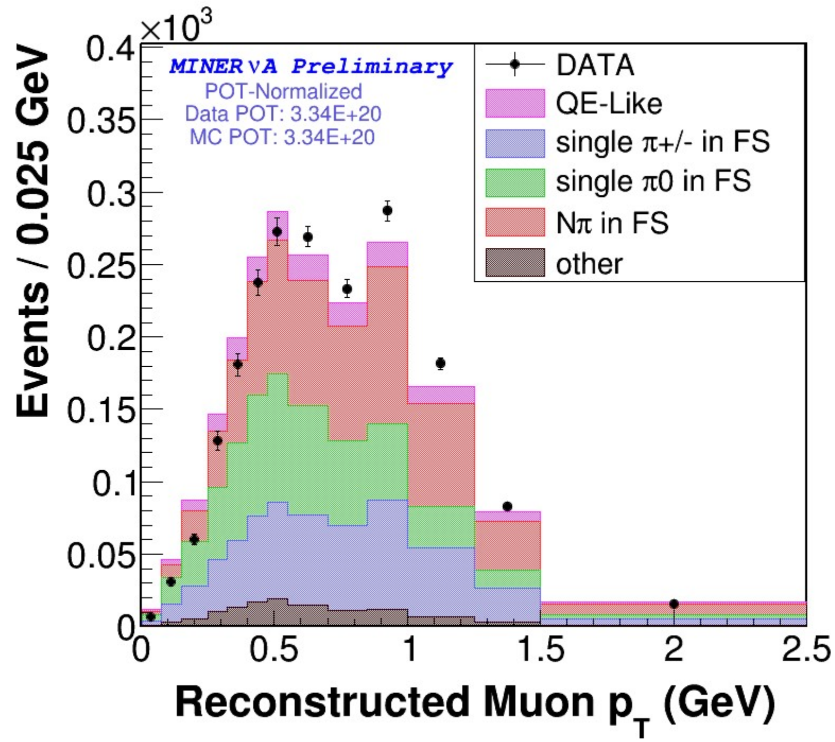
2-track

Isolated Clusters



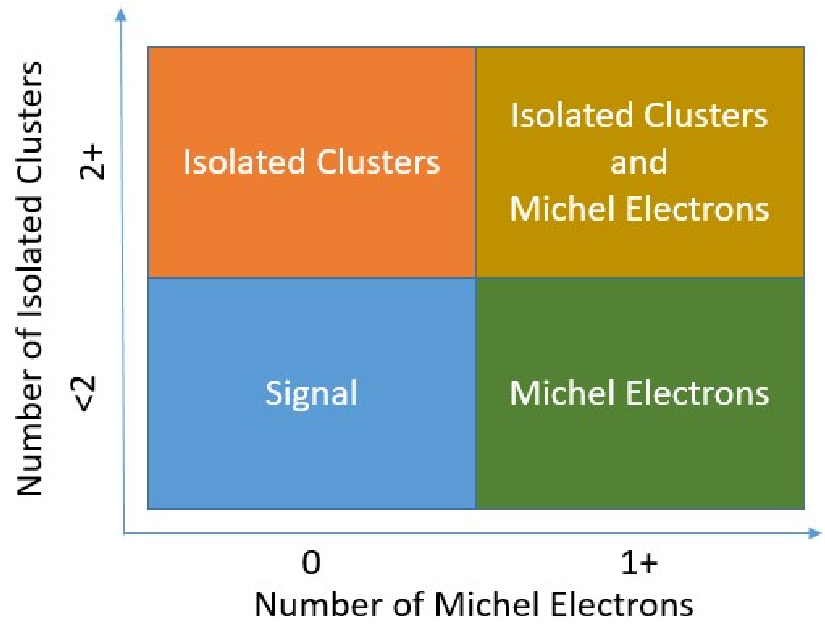


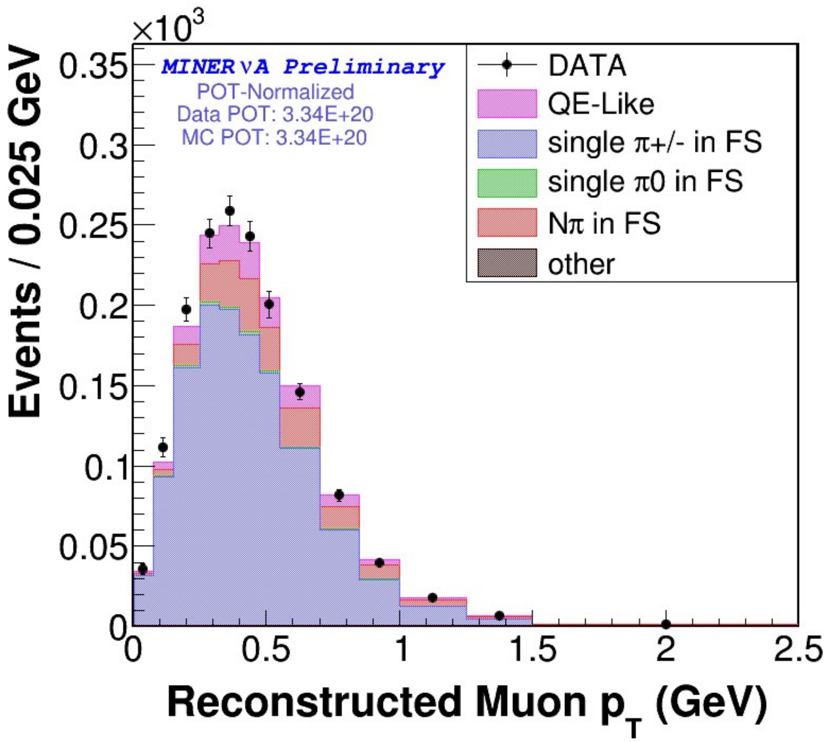
1-track



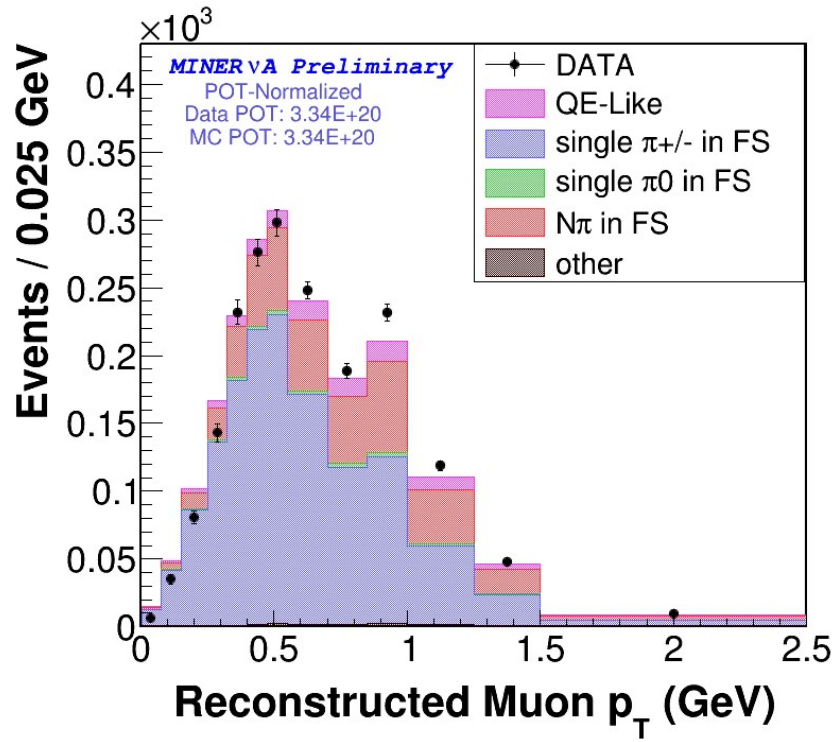
2-track

Michel Electrons



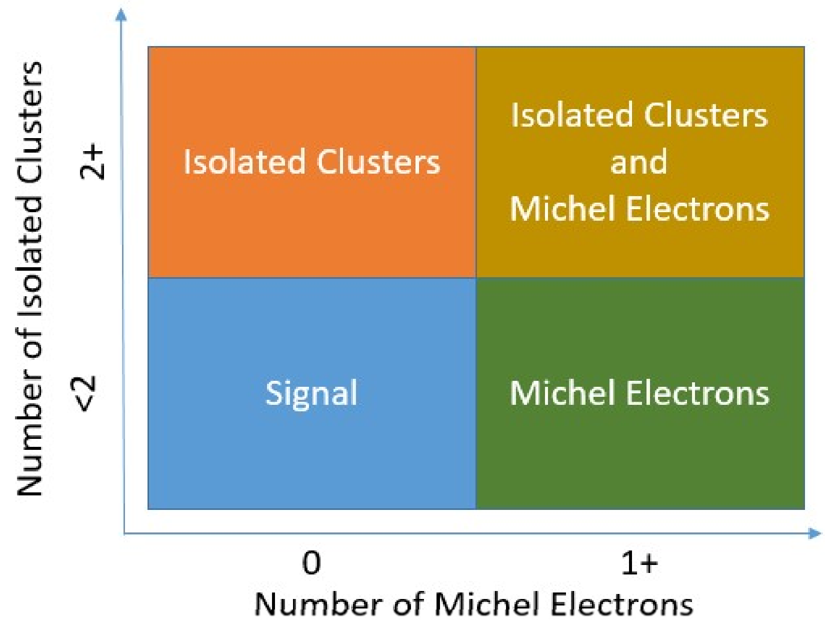


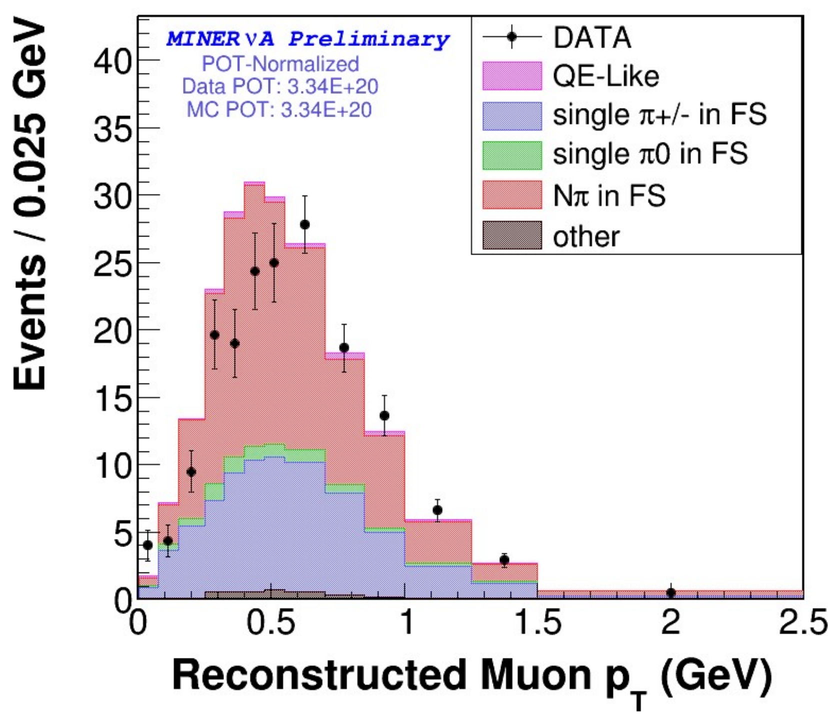
1-track



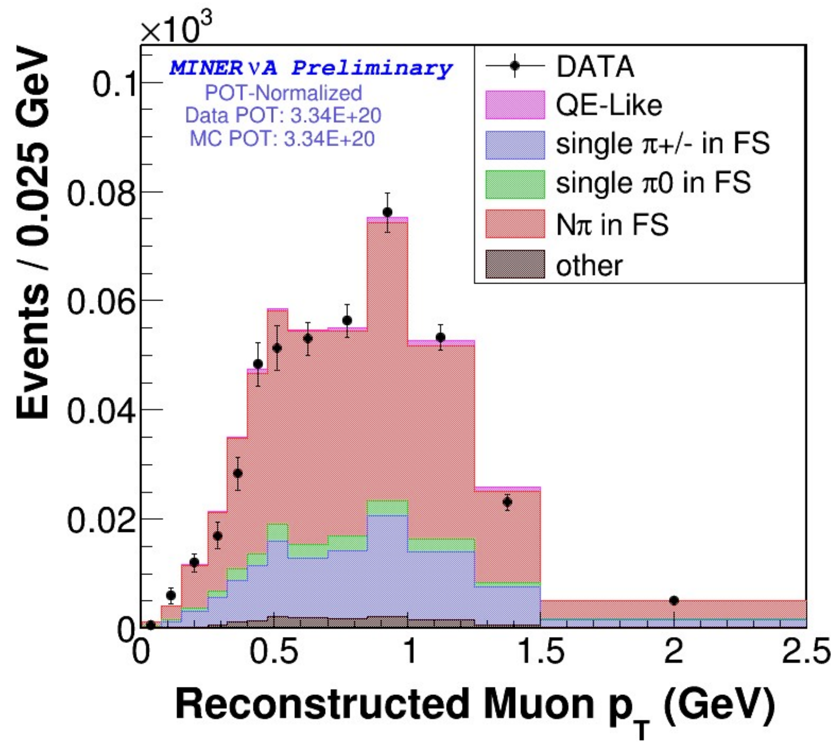
2-track

Isolated Clusters and Michel Electrons



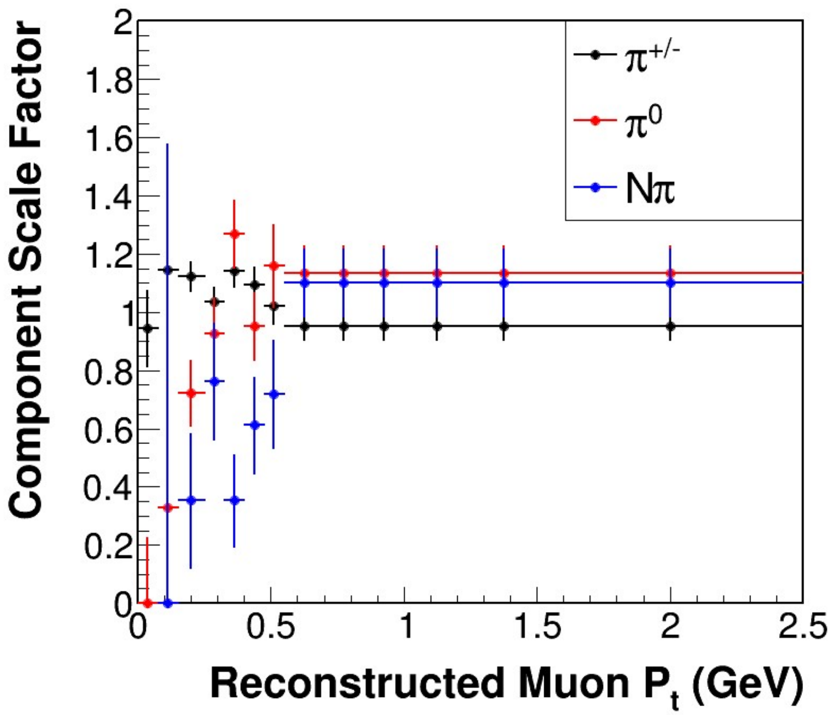


1-track

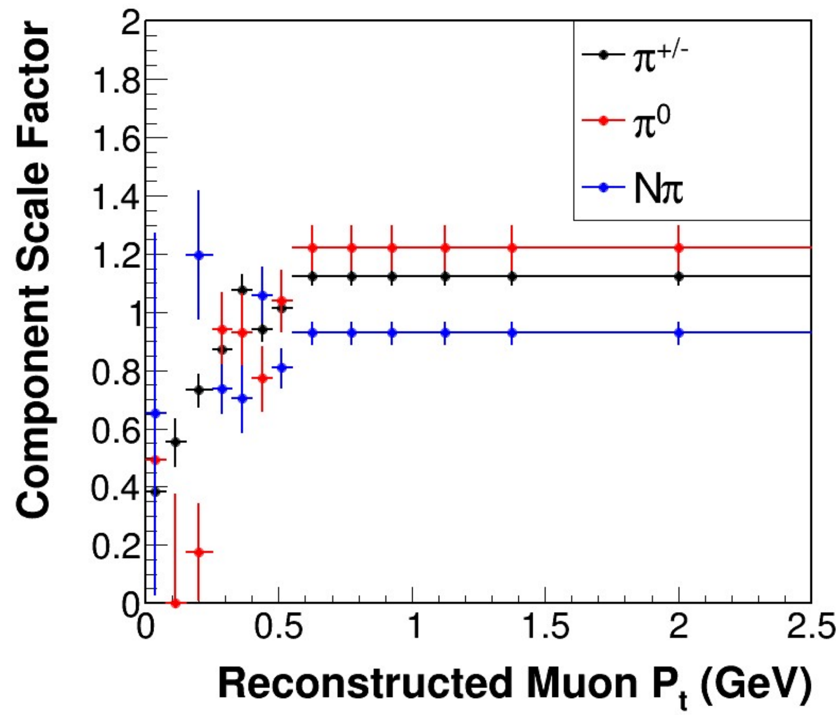


2-track

Component Scales

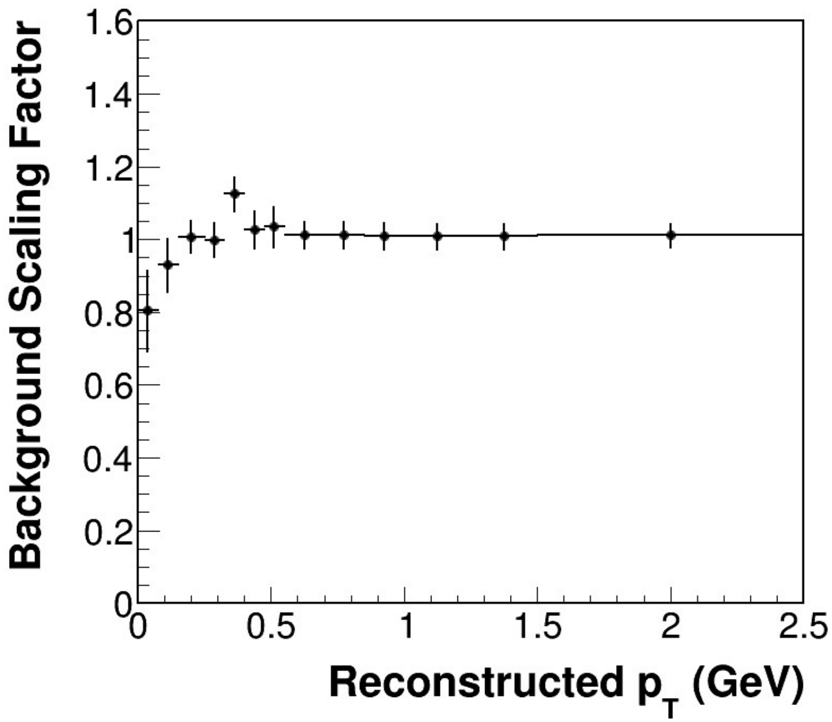


1-track

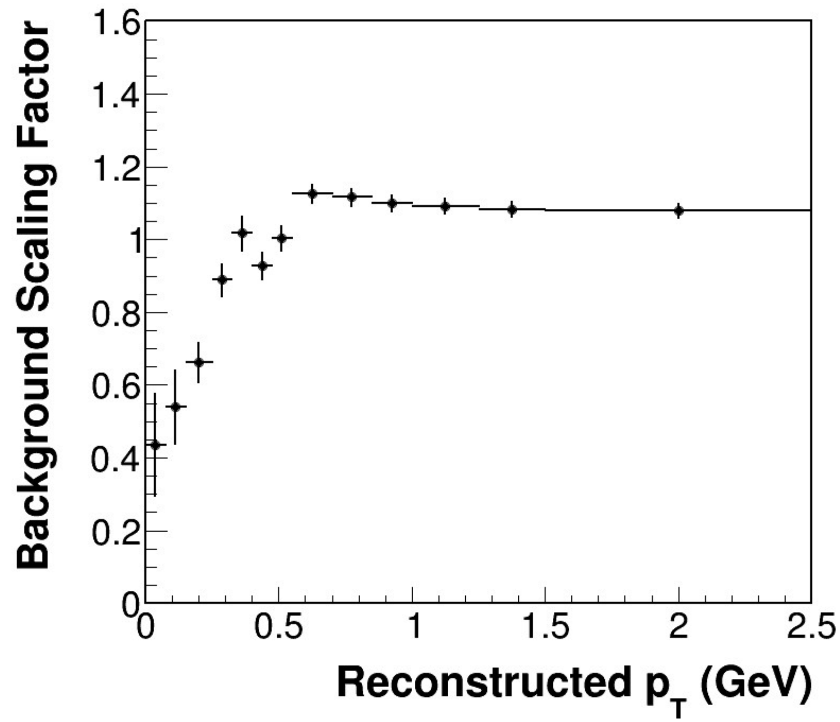


2-track

Overall scale factors

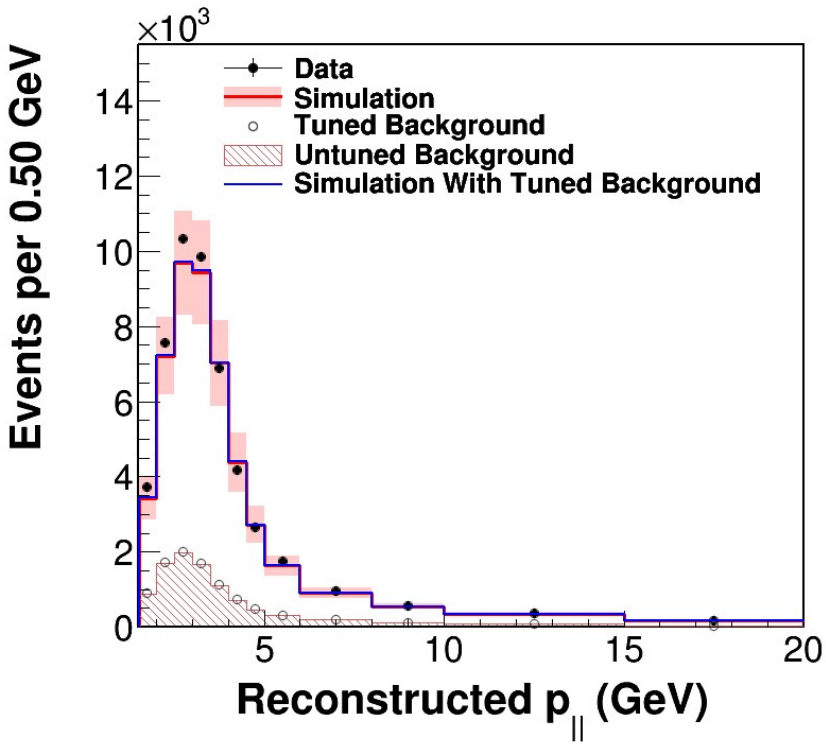


1-track

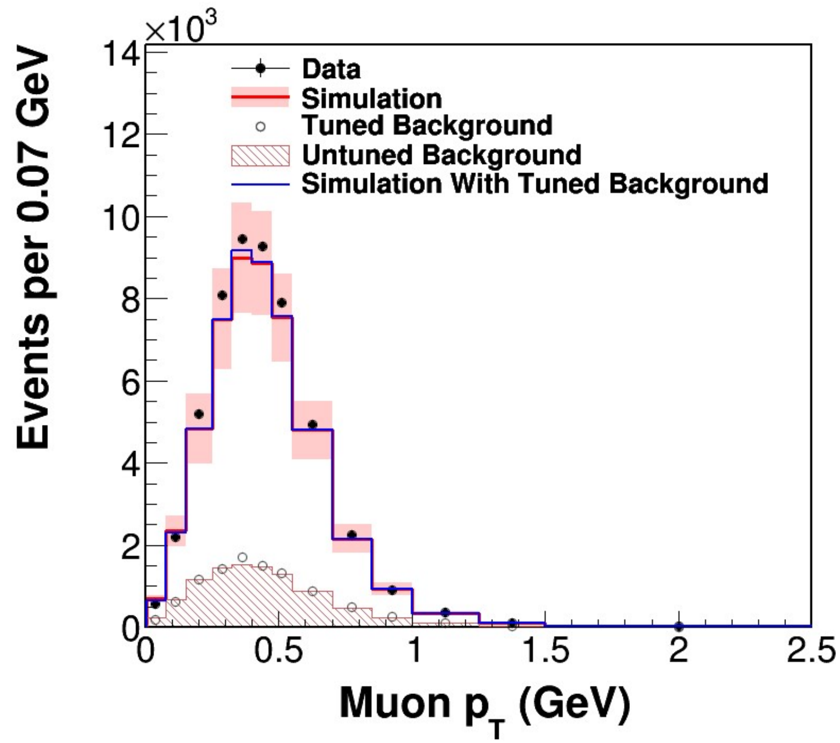


2-track

Signal 1 track

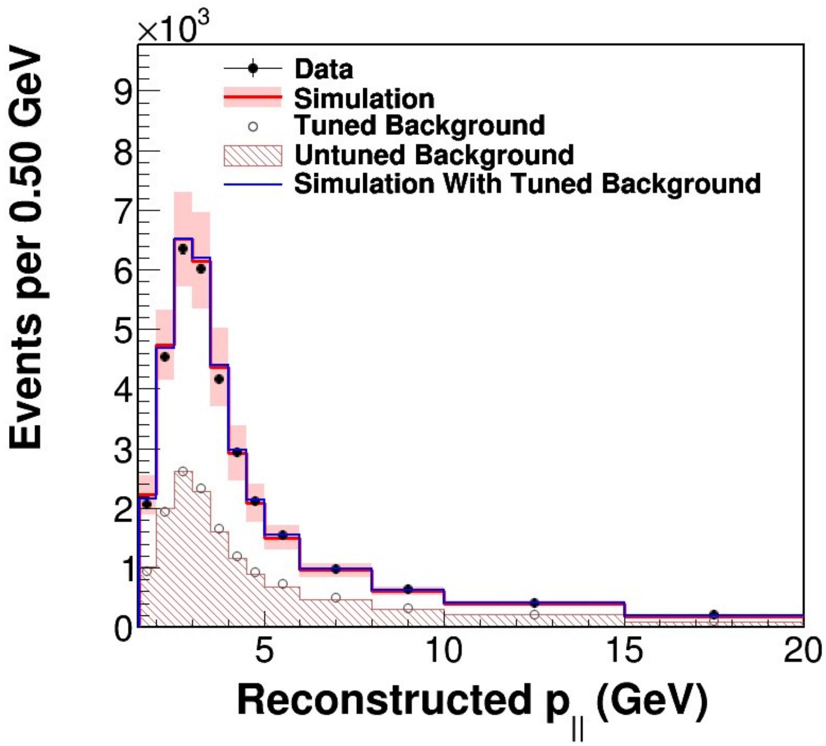


1-track

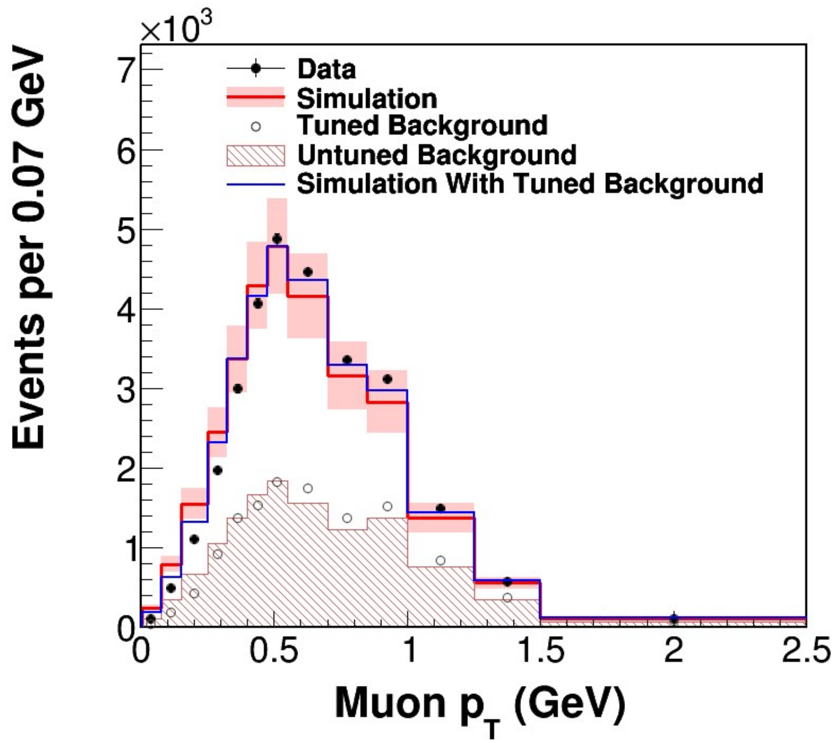


2-track

Signal 2 track



1-track

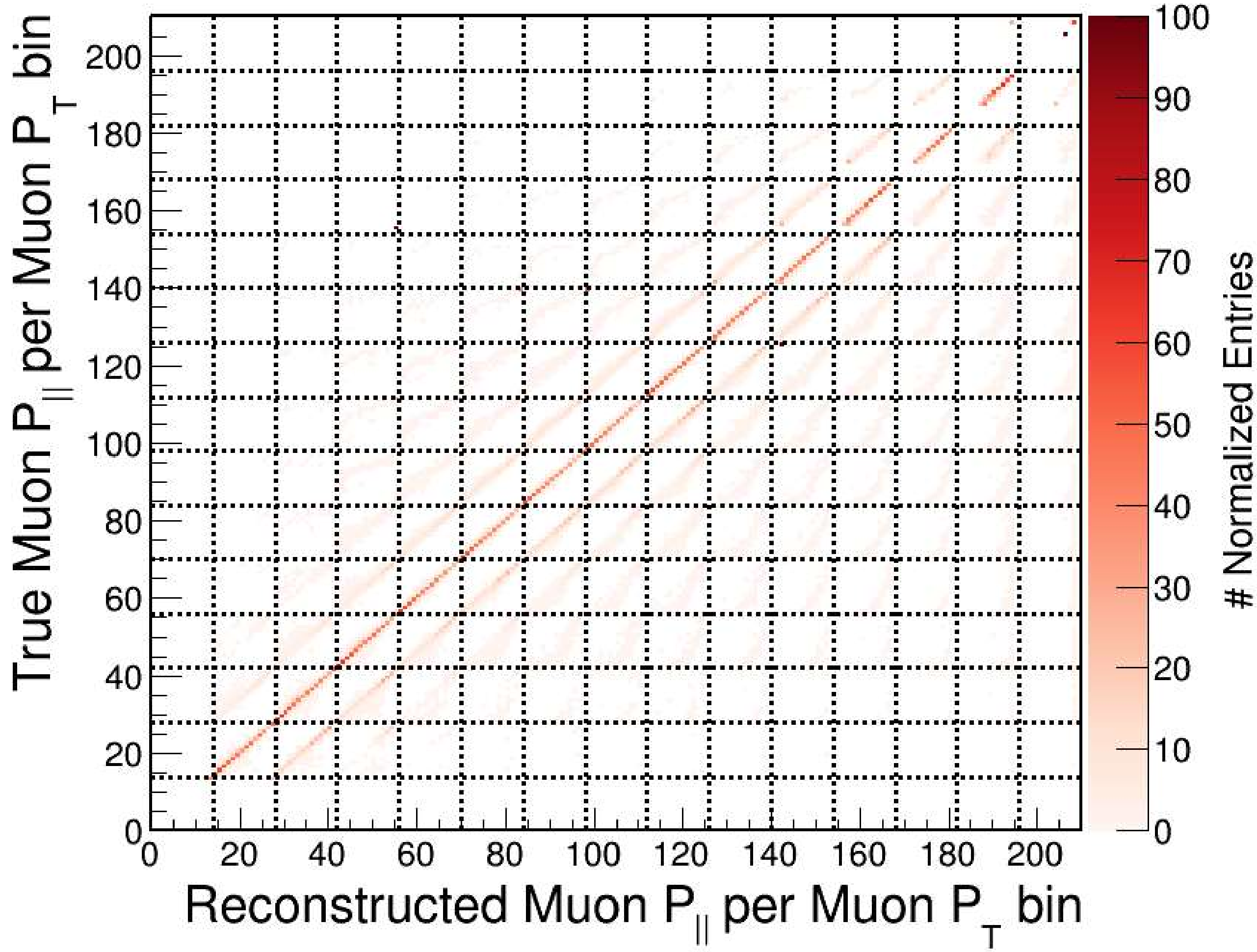


2-track

Migration Matrix

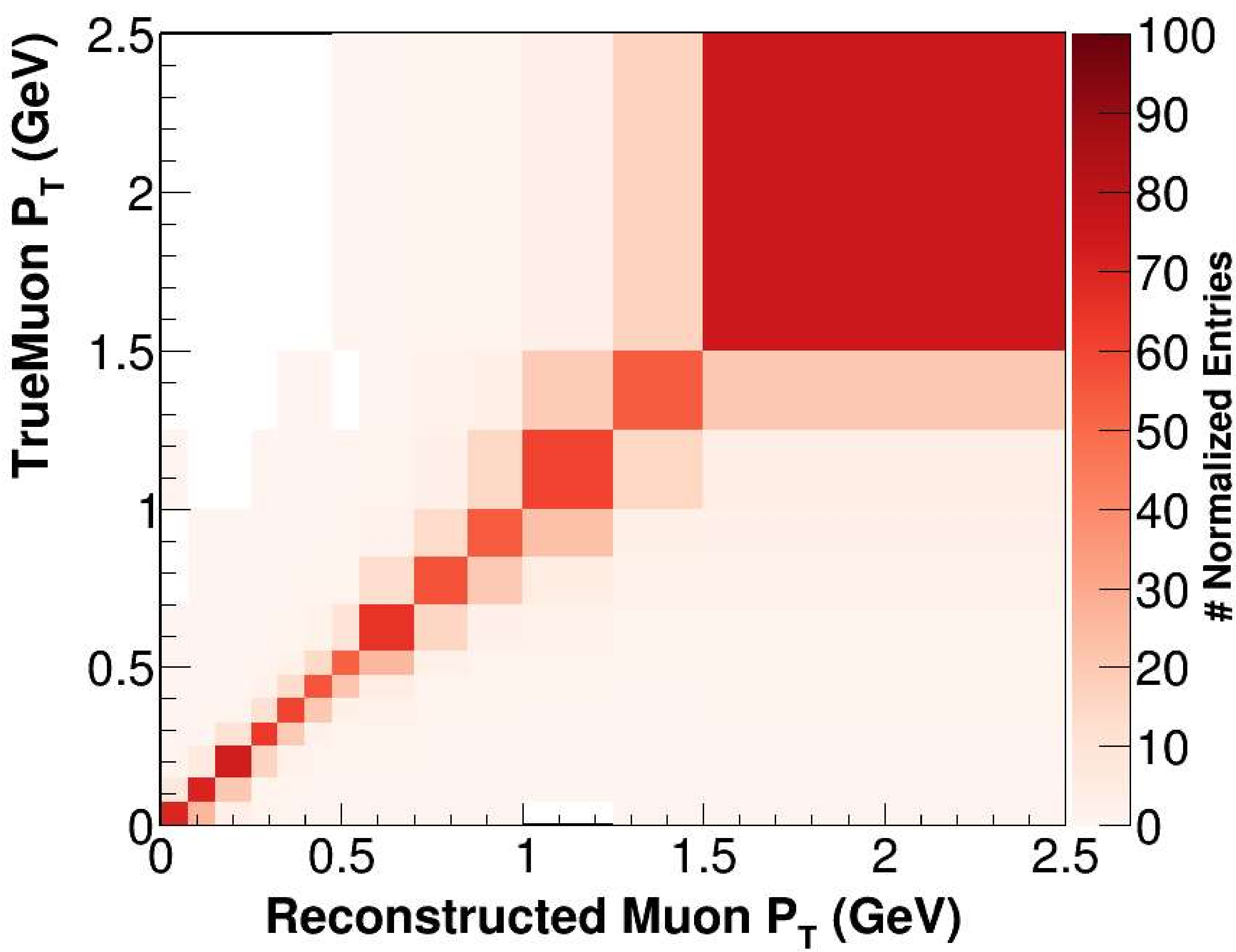
- Analysis uses the D'Agosti unfolding method implemented in RooUnfold.
- Based on bias studies, the necessary number of iterations is 2
- Mostly diagonal, with most elements in the 60-70% or more on the diagonal.

PtP||



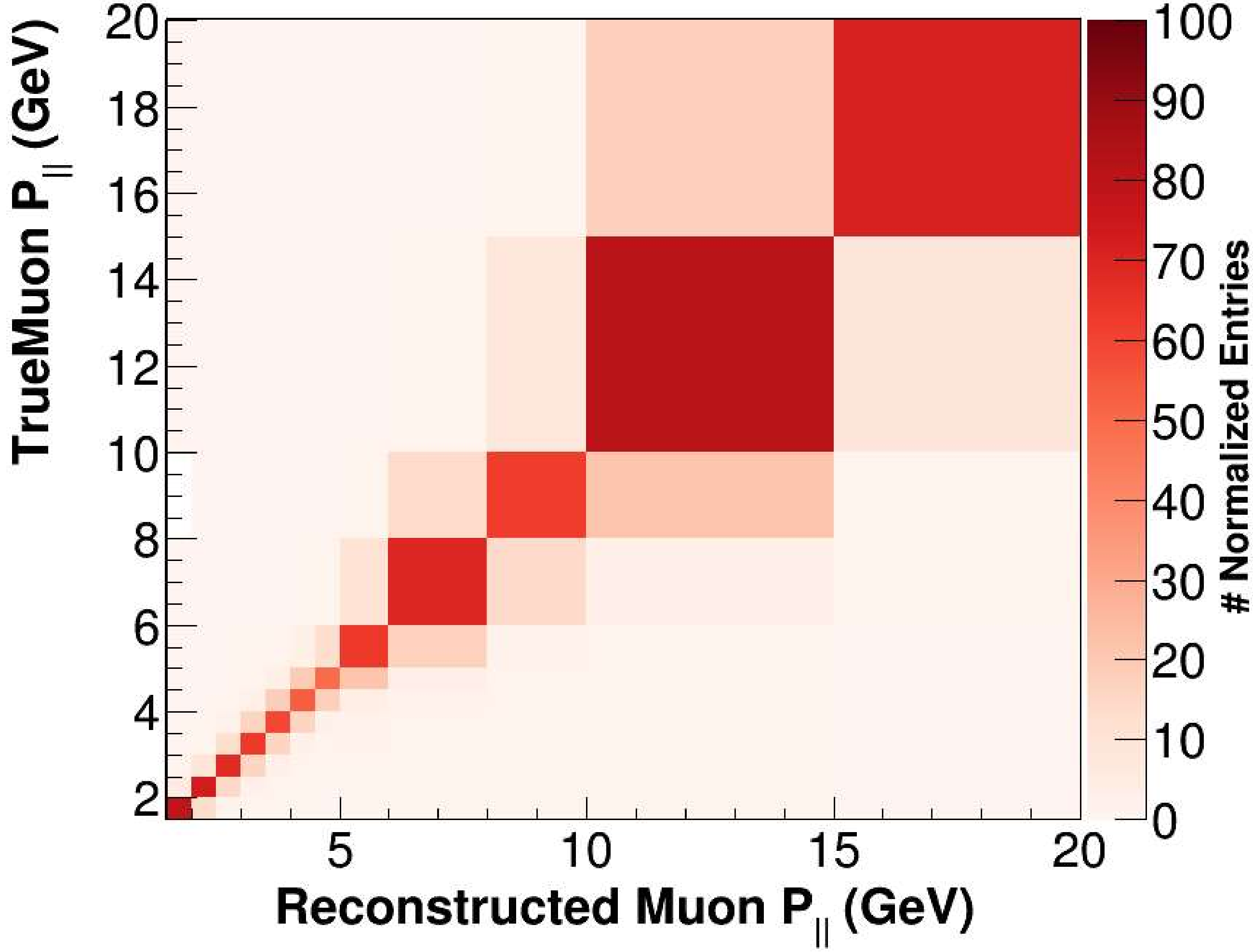
True Muon $P_{||}$ per Muon P_T bin
Reconstructed Muon $P_{||}$ per Muon P_T bin
Normalized Entries
0
20
40
60
80
100
120
140
160
180
200
10
30
50
70
90

Pt



TrueMuon P_T (GeV)
Reconstructed Muon P_T (GeV)
Normalized Entries
0
0.5
1
1.5
2
2.5
0
10
20
30
40
50
60
70
80
90
100

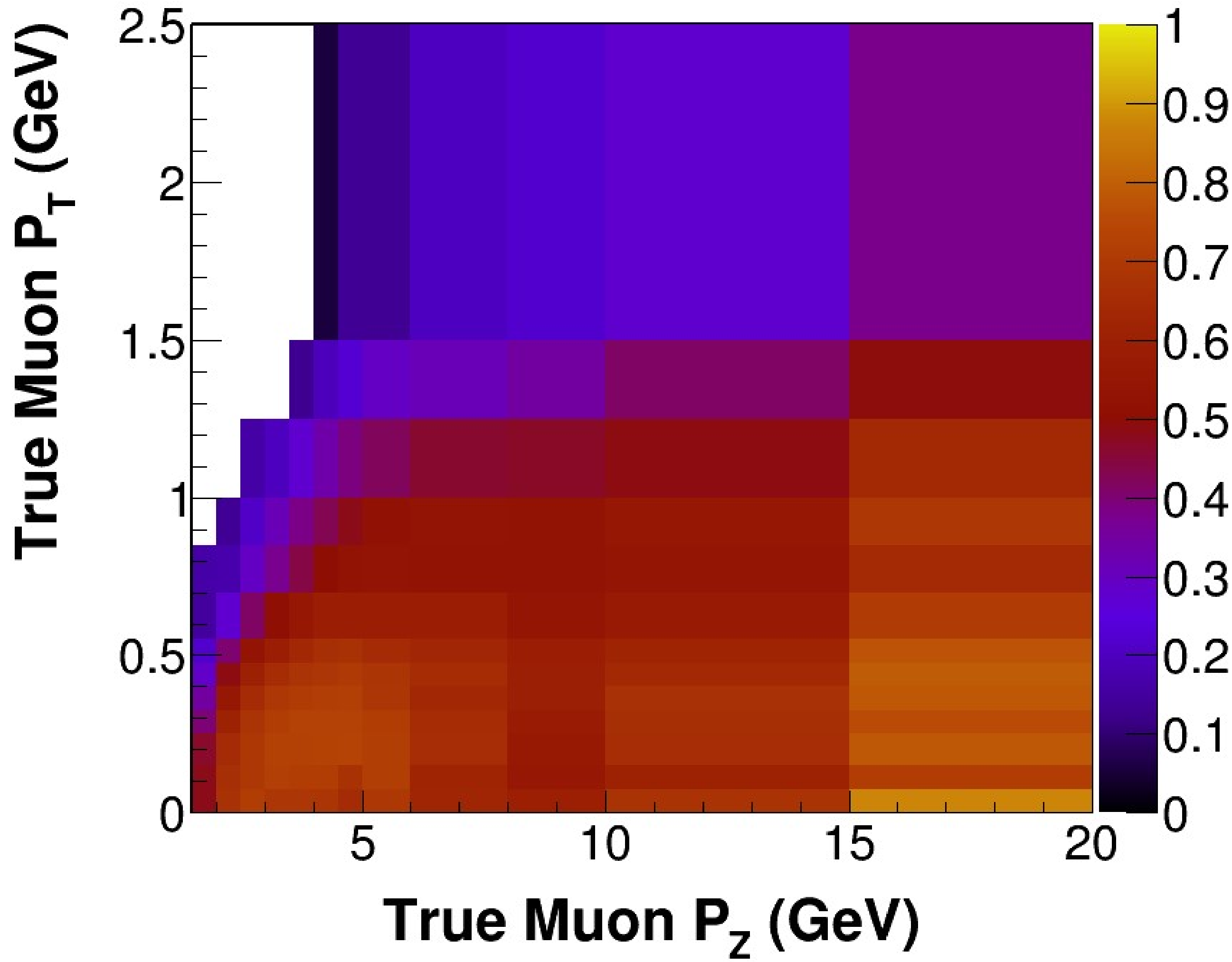
$P_{||}$


TrueMuon P|| (GeV)
20
18
16
14
12
10
8
6
4
2
5
10
15
20
Reconstructed Muon P|| (GeV)
100
90
80
70
60
50
40
30
20
10
0
Normalized Entries

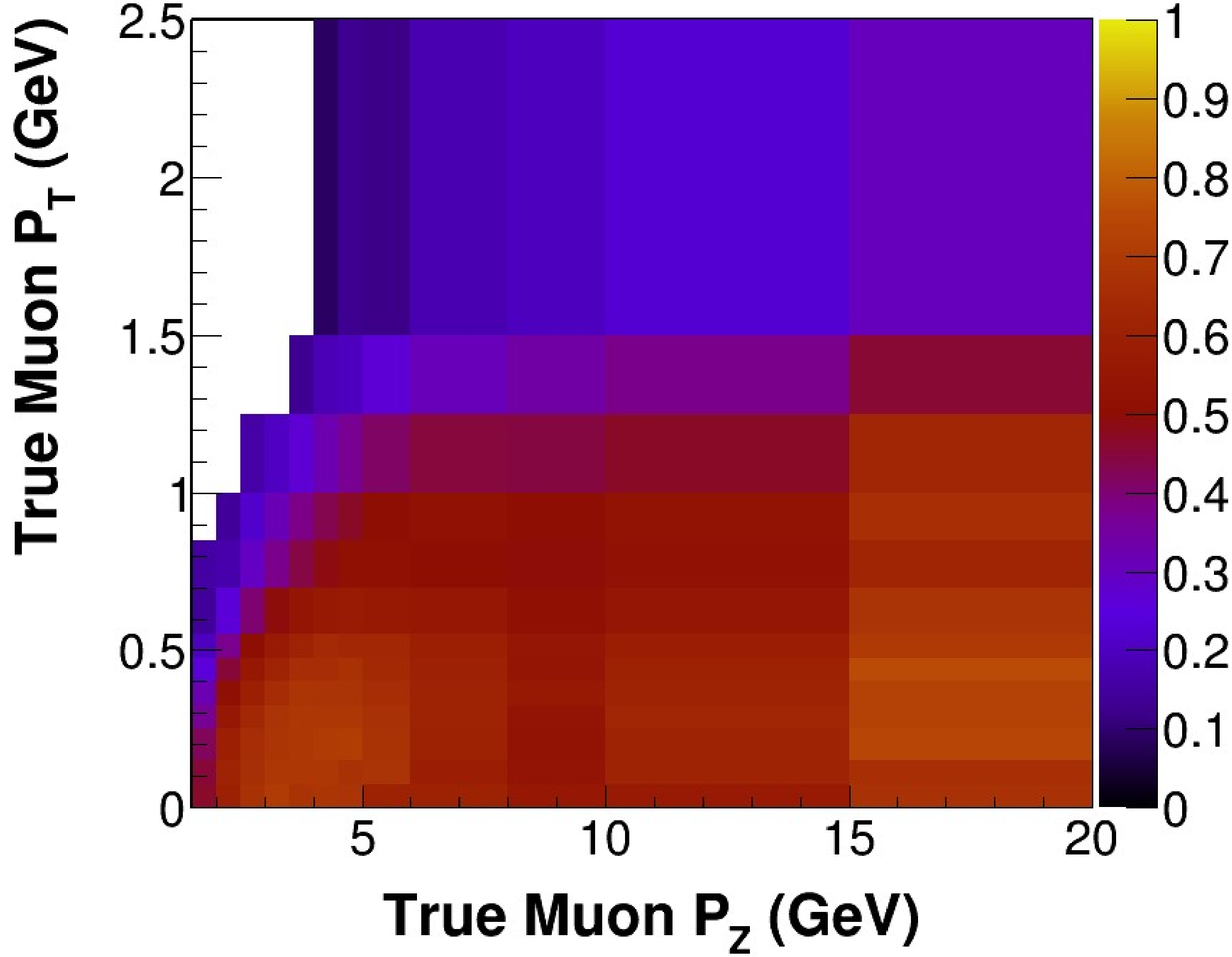
Efficiency

- One of biggest sources of model dependence.
- If selection picks on features of the underlying model you depend on that model
  - Ex. Recoil system energy cuts – QE type cuts great for QE, bad for QELike – See Minerba's talk
- So, how different are the various efficiencies of components of the sample
- Also, what fraction of the signal is coming from that sample

True QE Efficiency



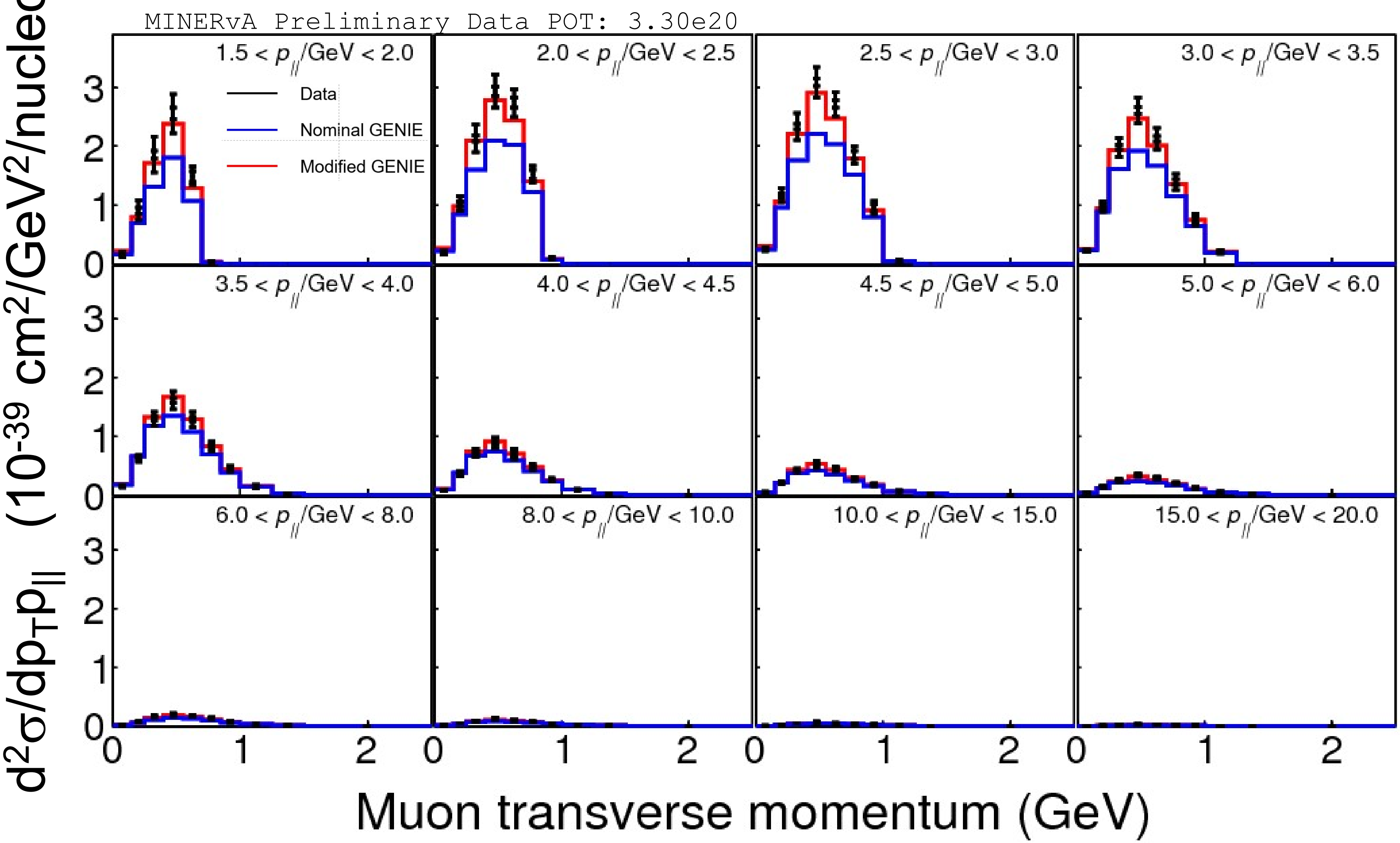
Overall CC0π Efficiency



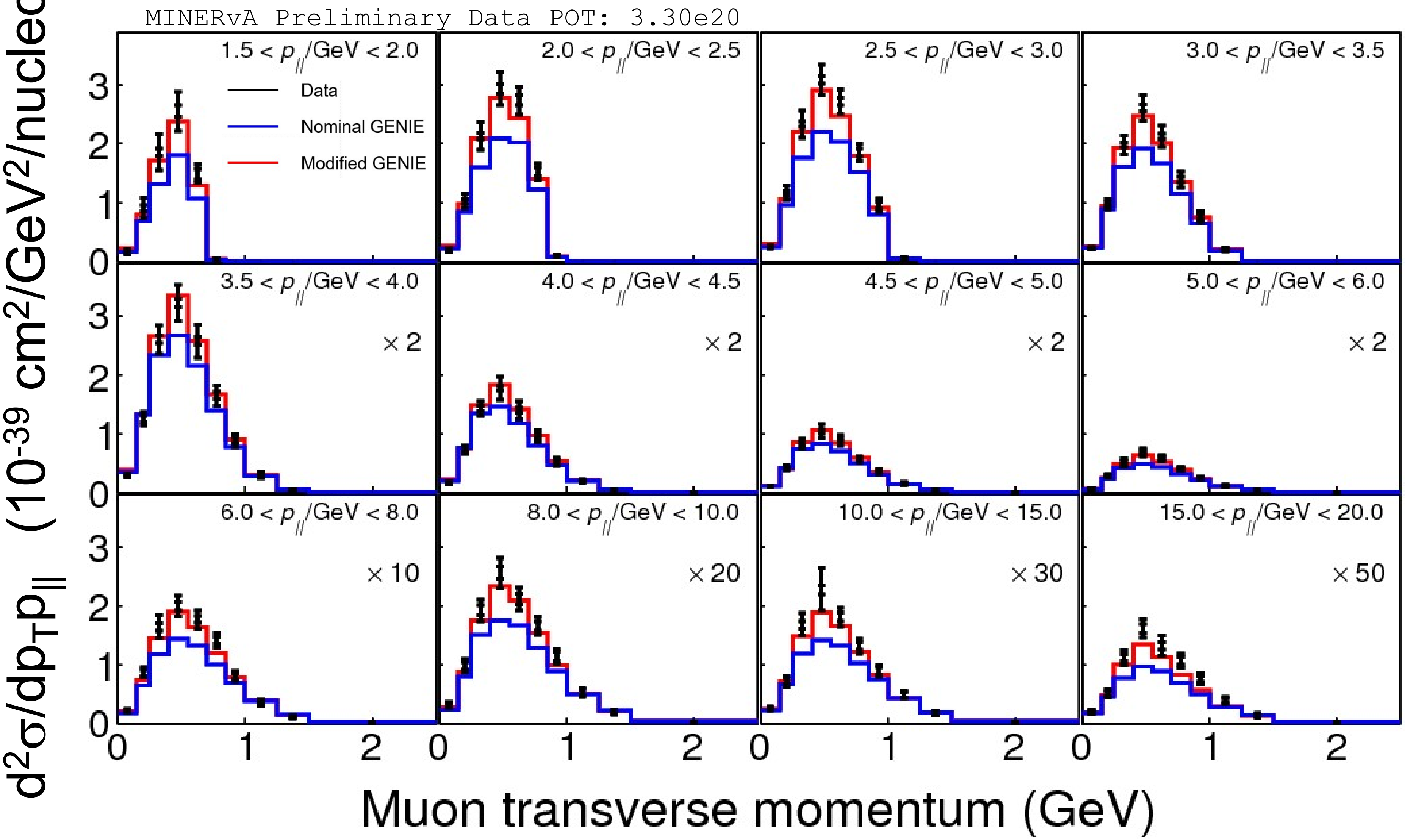
$$\left(\frac{d\sigma}{dx}\right)_\alpha = \frac{\sum_j U_{j\alpha}\left(N_{data,j} - N^{bkgd}_{data,j}\right)}{A_\alpha(\Phi T)(\Delta x)}$$

Put it all together

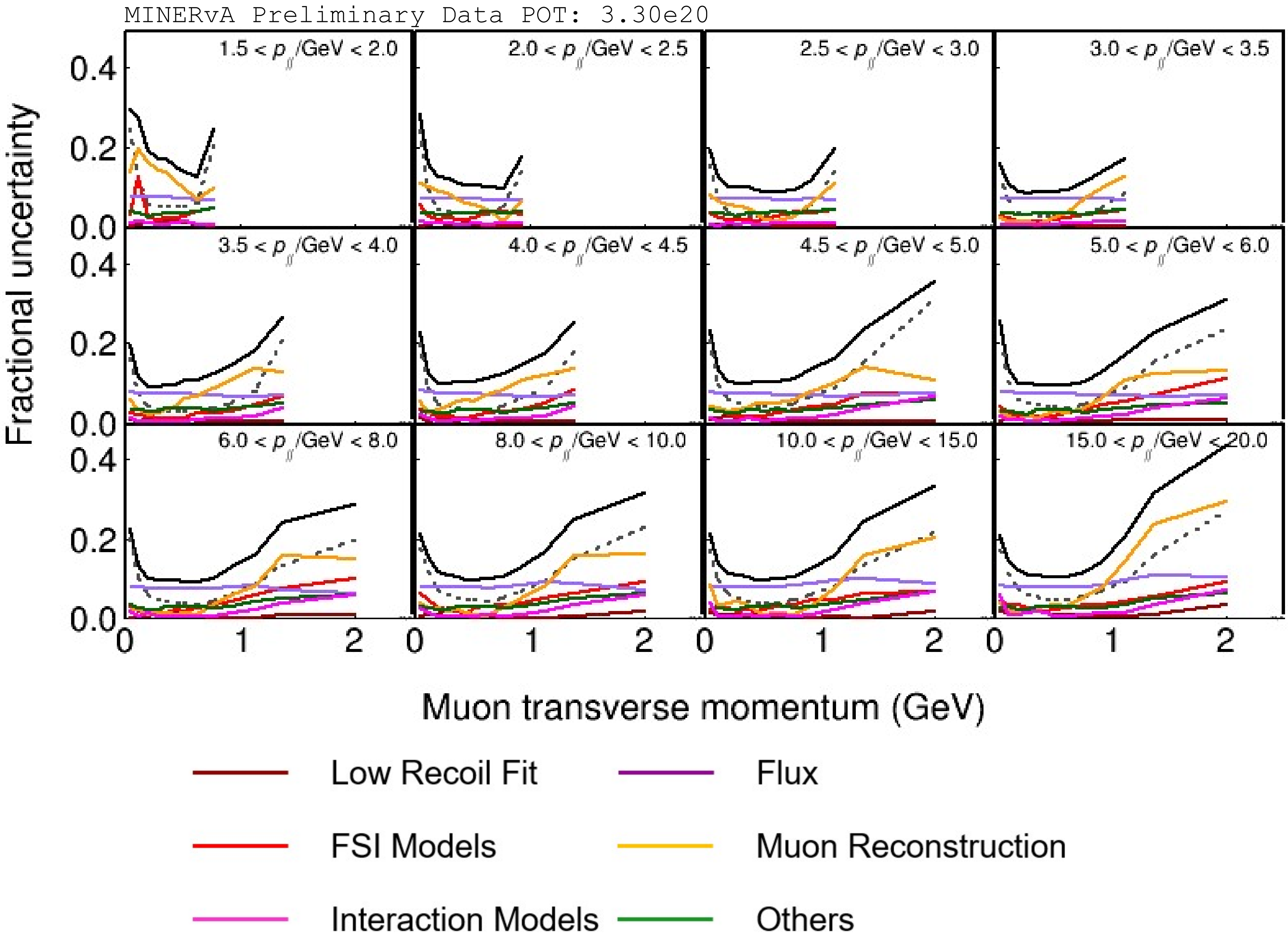
$P_t P_{||}$ Cross Section



$P_tP_{||}$ Cross Section



Systematic Uncertainties



Low Recoil Fit

- Taking the Low recoil $q_0 q_3$ analysis fit 4 fits are performed
  - Allow nn+np 2p2h modes vary
  - Allow only nn 2p2h modes vary
  - Allow only np 2p2h modes vary
  - Allow only QE modes vary
- Despite the very different inputs, the results on the CC0pi analysis are very similar

FSI Models

- Dominated by pion absorption which causes a signal<->background migration.

Interaction Models

- All cross section related GENIE knobs
- Small in most of the measurement except very low P_t and high P_t
- Low P_t dominated by QE model, Pauli Suppression, RPA @few % for high $P_{||}$
- High P_t dominated by Pion/DIS knobs and RPA

Muon Reconstruction

- 11MeV shift from MINERvA material assay
- 30 MeV shift from energy deposition per cm
- 2% for energy by range MINOS
- 0.6% > 1GeV or 2.5% <1GeV if measured by curvature
- Added in quadrature

Others

- Includes – **particle response in detector**, energy of hits, number of targets, matching efficiencies between MINOS and MINERvA, Bethe-Bloch.
- Notables – Proton efficiency, Bethe-Bloch at high P_t is at ~3%

χ^2 Reporting

- MINERvA compares to various models, and reports χ^2 compared to the data
- Recently been discussing the effect of highly correlated data and calculation of the χ^2
  - Can lead to χ^2 which don't follow what your eye says has to be right
- Known as "Peelle's Pertinent Puzzle" to nuclear physicists
  - International evaluation of neutron cross-section standards", IAEA 2007
  - "Box-Cox transformation for resolving the Peelle's Pertinent Puzzle in curve fitting", Oh and Seo 2004
- Cross section typically have at least one highly correlated uncertainty - Flux

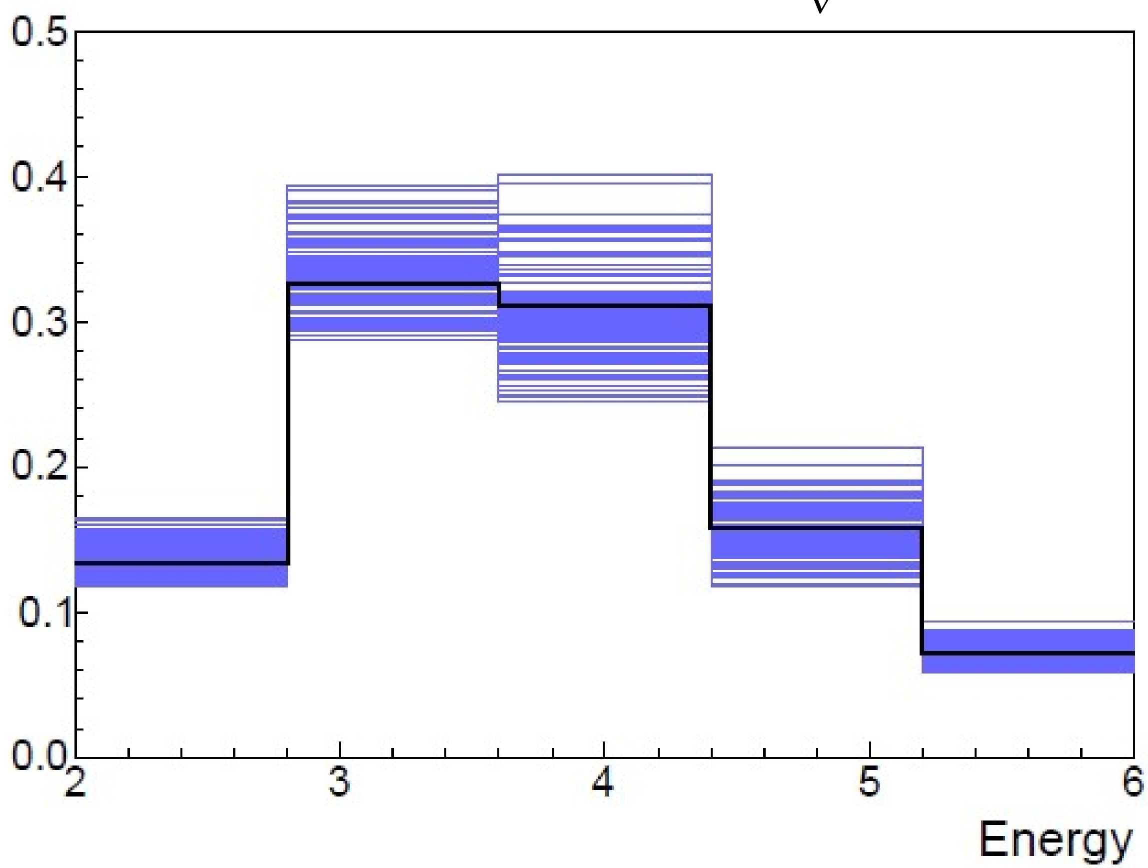
Uncertainties

- Given a central value of 1 with a 1σ value of 0.8. What is 2σ?
- Additive uncertainties: 1-2*0.8
  - This results in $-\infty$ to ∞ for an arbitrary number of deviations
  - A Gaussian distribution has this property
- Multiplicative uncertainties: $1-0.8^2$
  - This results in 0 to ∞ for an arbitrary number of deviations
  - A Log-Normal distribution has this property

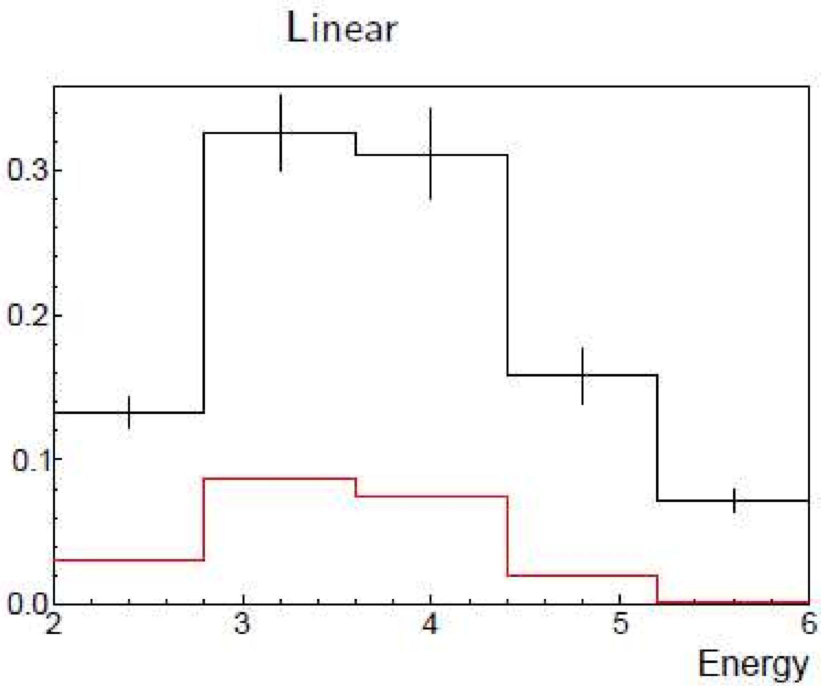
Example

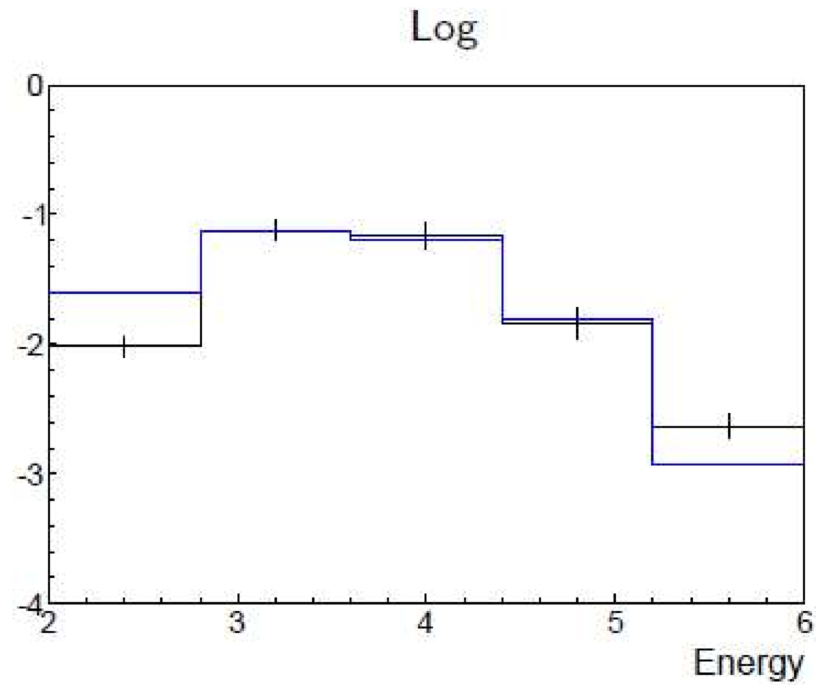
- Fit with Gaussian in standard way
- Log transform and fit and transform back

Flux errors on E_ν

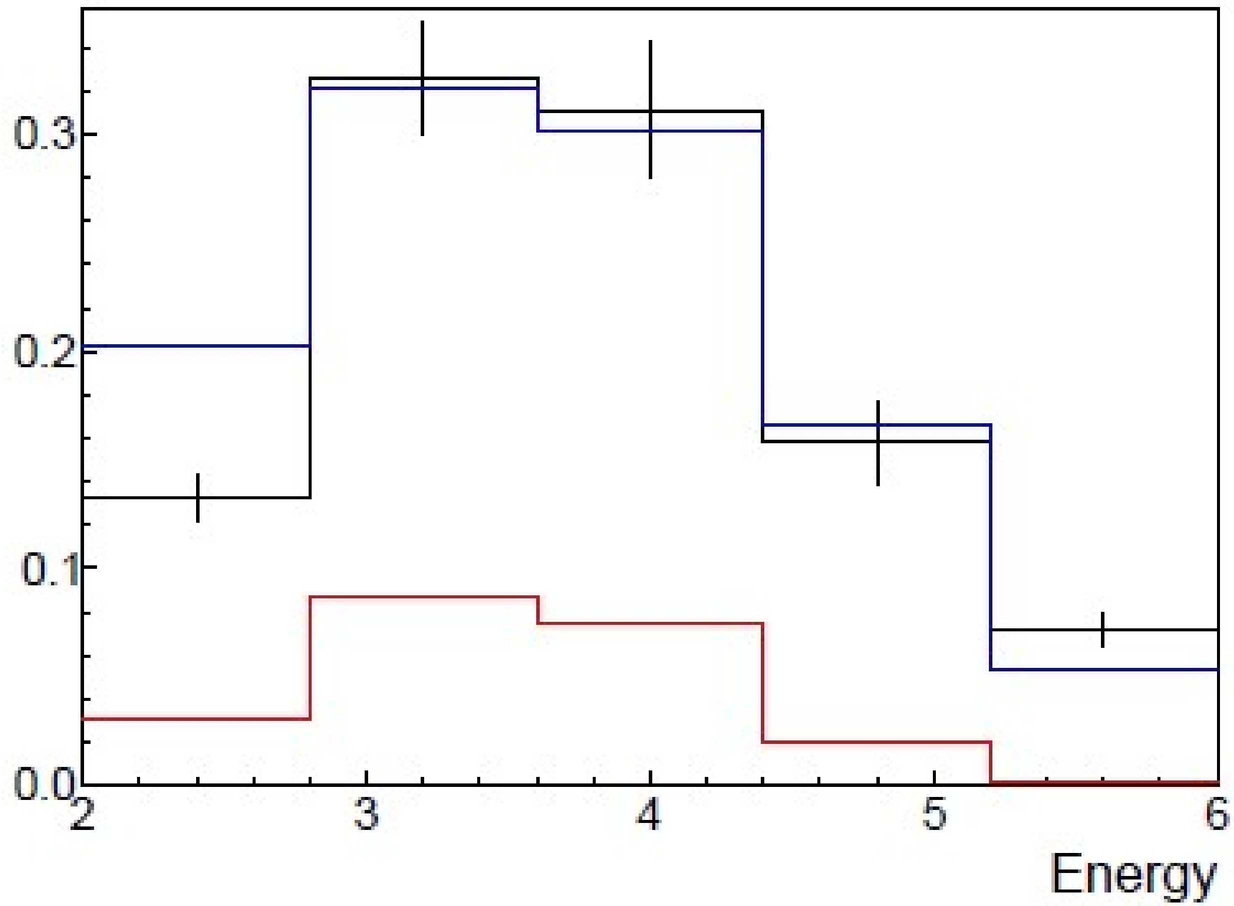


Comparison





Transform back



Future

- Model comparisons with large correlations dominated by scale errors (flux!!) can return χ^2 we don't expect
- Application of log transformations improves this
- Of course the errors on cross sections are both multiplicative and additive
  - Literature suggests solution transformation
  - G.E.P. Box and D.R. Cox
    Journal of the Royal Statistical Society. Series B (Methodological)
    Vol. 26, No. 2 (1964), pp. 211-252

Conclusion

- MINERvA has a mature cross section program with the goal of model independent, and interesting/useful results
- MnvH1D provide a useful tool to handle the complex process of extracting a cross section
  - Discussion if this is a viable way to release our data
- Signal definitions and what you reconstruct should align
  - Fiducial cross sections!
- Cross sections are difficult, complex, and have many internal tensions which you need to describe clearly