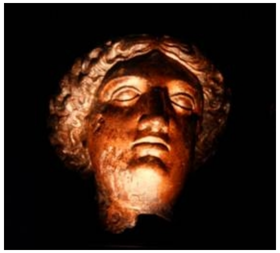
User analysis computing

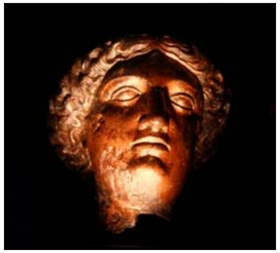
Trung Le
Tufts University
17 October 2016

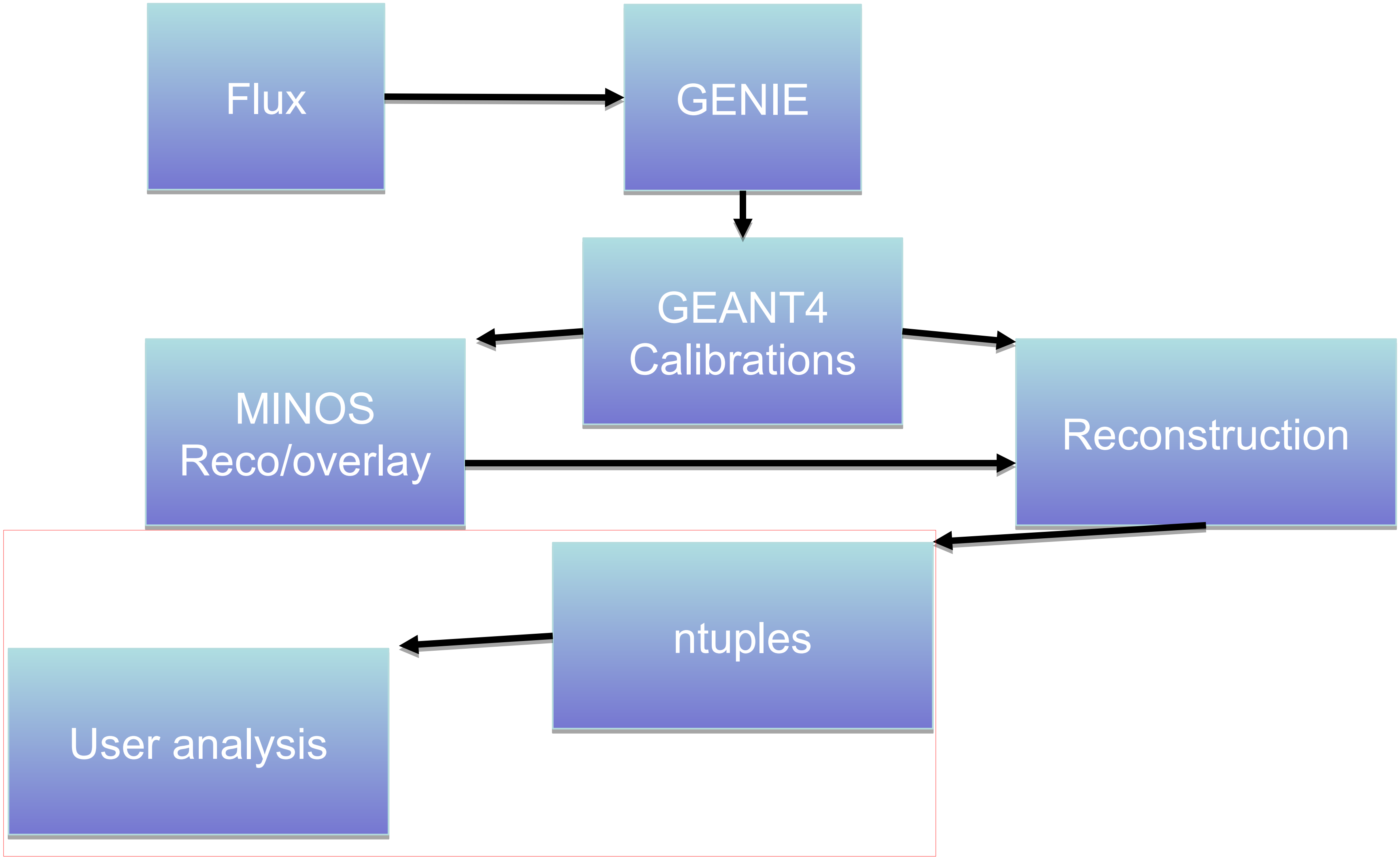
Outline



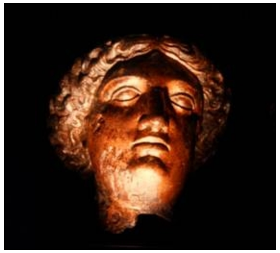
- Making analysis ntuples
- Physics analysis using the ntuples

Data flow



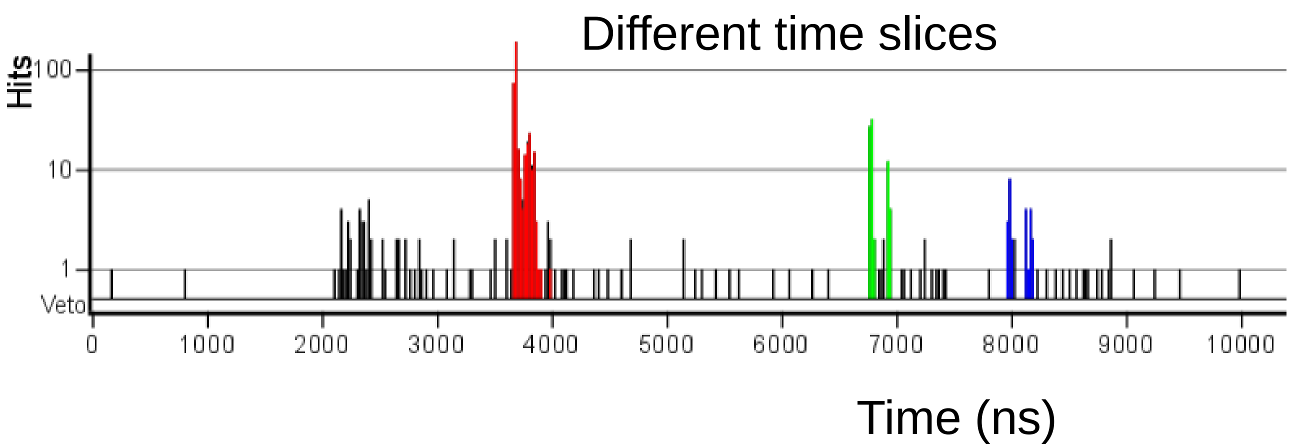


Flux
GENIE
GEANT4
Calibrations
MINOS
Reco/overlay
Reconstruction
ntuples
User analysis



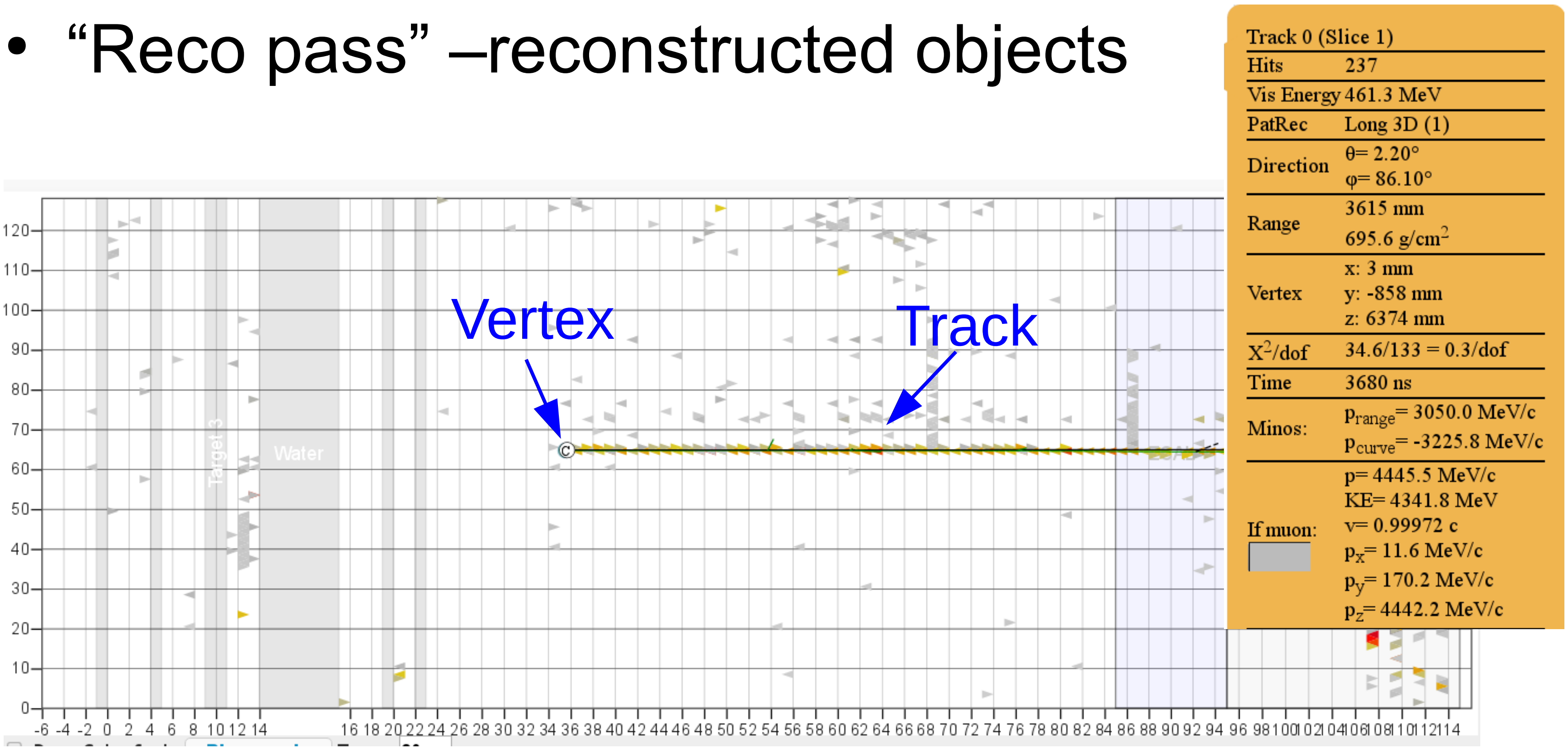
Data processing stages

- "Reco pass" –reconstructed objects
  - Data gates are divided into time slices or physics events
  - Each event has an interaction vertex and tracks

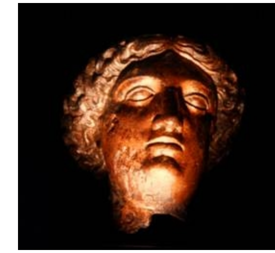


Data processing stages

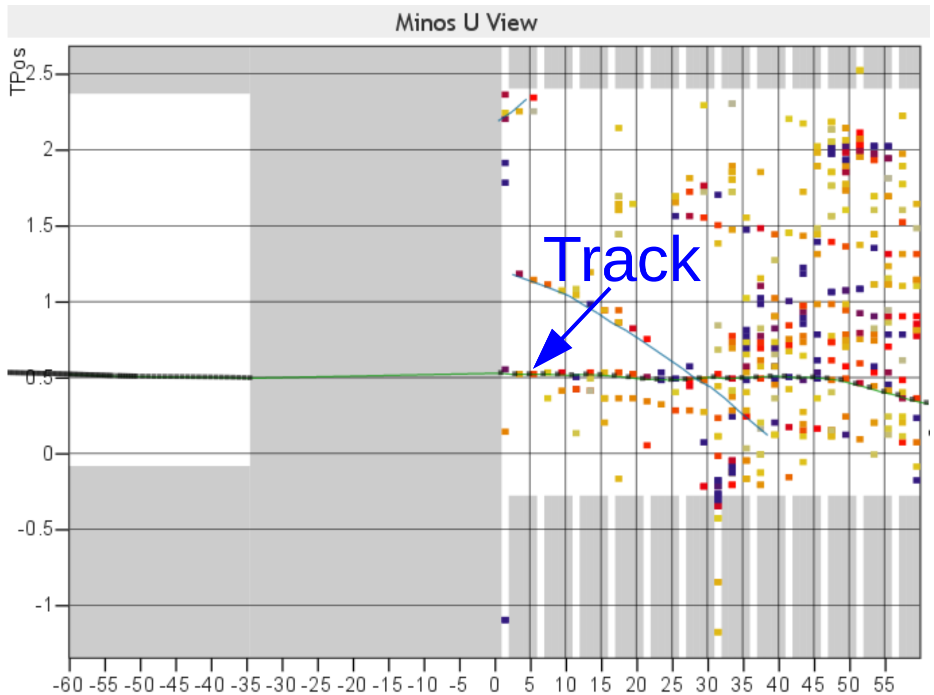
- “Reco pass” –reconstructed objects



Data processing stages



- "Reco pass" –reconstructed objects



Track 0 (Slice 1)	
Hits	237
Vis Energy	461.3 MeV
PatRec	Long 3D (1)
Direction	θ= 2.20° φ= 86.10°
Range	3615 mm 695.6 g/cm^2
Vertex	x: 3 mm y: -858 mm z: 6374 mm
X^2/dof	34.6/133 = 0.3/dof
Time	3680 ns
Minos:	p_{range}= 3050.0 MeV/c p_{curve}= -3225.8 MeV/c
If muon:	p= 4445.5 MeV/c KE= 4341.8 MeV v= 0.99972 c p_x= 11.6 MeV/c p_y= 170.2 MeV/c p_z= 4442.2 MeV/c

The muon is reconstructed in MINOS and matched to MINERvA

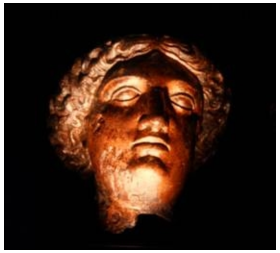
Data processing stages

- “Ana pass” –make analysis ntuples
- Make custom reconstruction
- Calculate reconstructed physics variables
- Fill variables in analysis ntuple

User analysis tool

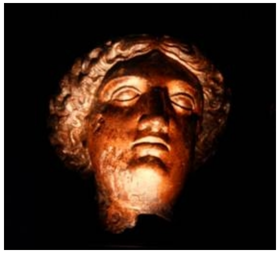
- Each physics analysis has a corresponding analysis tool derived from MinervaAnalysisTool
- User analysis tool must implement these methods:
  - initialize()
  - reconstructEvent()
  - interpretEvent()
  - tagTruth()

initialize() –Example

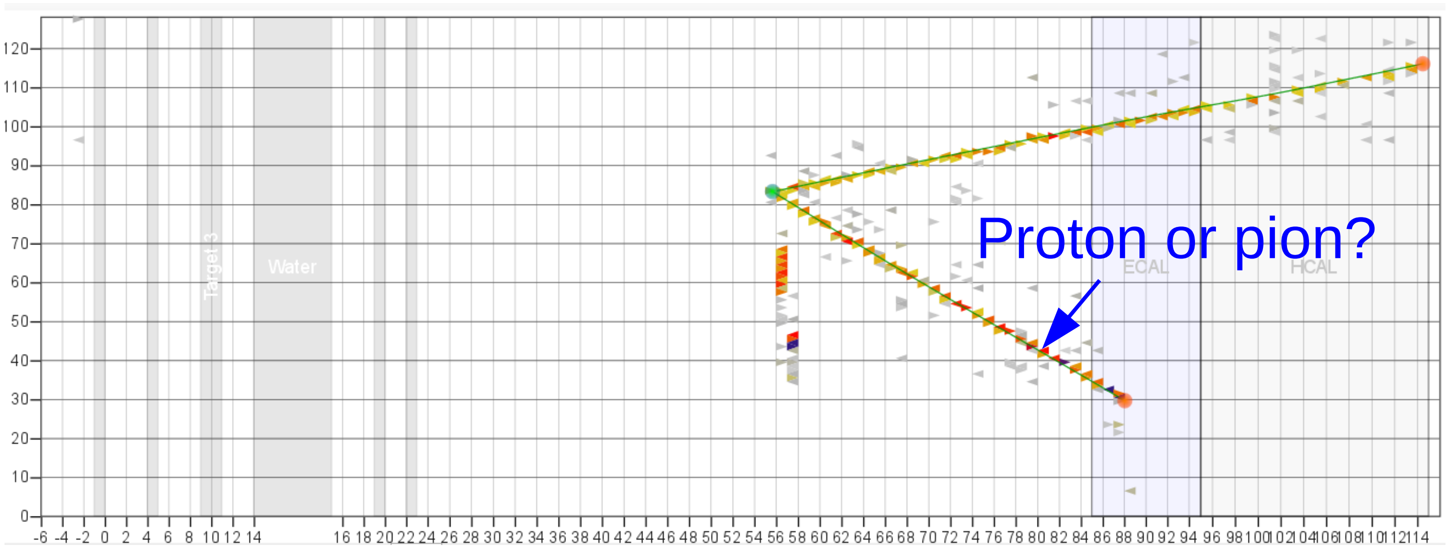


- Declare ntuple branches
  - declareIntEventBranch(“ntrack”, -1)
  - declareDoubleEventBranch(“evis”, 0.0)
  - declareContainerIntEventBranch(“track_pdg_vec”)
  - declareContainerDoubleEventBranch(“track_length_vec”)

reconstructEvent() –Example



- Additional custom reconstruction (e.g., pi0 reconstruction, particle identification)
- Advantage: fast develop-test cycle



interpretEvent() –Example

- Assign hypothesis to the event, e.g., CCQE or RES
- Calculate reconstructed physics variables
- Decide whether to save the event to ntuple

tagTruth() –Example

- While the purity can be calculated from the selected sample, we need the sample before any event selection cuts to calculate the efficiency

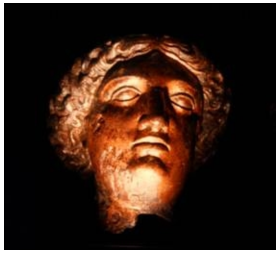
$$\epsilon = N_{\text{selected}}/N_{\text{total}}$$

Ana pass –data processing

- *python ProcessAna.py --ana_tool ACCPionMinus --data --playlist minerva5 --inv v10r8p9 --max_inputs 10 --outdir /pnfs/minerva/persistent/users/ltrung/acc1pi-*

- The output of this command are analysis ntuples ready to be used in the next step

Physics analysis



- Differential cross section formula for bin i[th]

$$\left(\frac{d\sigma}{dX}\right)_i = \frac{1}{T\Phi}\frac{\sum_j U_{ij}(N_j^{data} - N_j^{bkg})}{\epsilon_i}$$

Common analysis tools

- MINERvA has developed common analysis tools to facilitate cross section extraction and systematic uncertainty handling
  - MnvH1D for handling histograms
  - MnvUnfold
  - MnvPlotter
- We have organized Minerva 101 in the last three years to get new people familiar with these analysis tools and the offline software framework

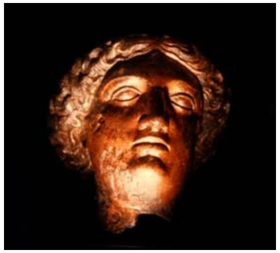
The MnvH1D class

- MINERvA customized 1D histogram class
- Based on ROOT TH1D class
  - The underlying TH1D is the central-value histogram
- Main additions are lists of “error bands” that are used to calculate systematic uncertainties
  - One error band exists for each source of systematic uncertainty
  - Each error band is independent of all others. When there are correlations, they are handled upstream by generating correlated weights

Error bands

- Each error band has a name and a list of universe histograms
- Vertical error bands
  - Fill each universe with the same value, but with different weight
  - Example: flux uncertainty
- Lateral error bands
  - Fill each universe with a different shifted value, but with the same weight
  - Example: Electromagnetic energy scale

“Many-universes” method



- Error propagation uses “many universes” method
  - For each parameter with known uncertainty, throw random, Gaussian distributed random numbers, then recalculate derived quantities
- Example
  - Imagine 100 universes that differ only in the NuMI flux
  - In each of these universes, calculate cross section result
  - The spread in the result is the uncertainty due to flux

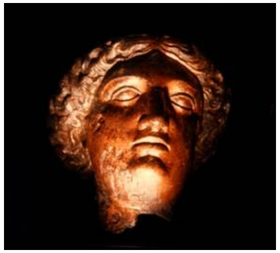
MnvH1D functions

- Add, Divide, Multiple, and Scale
- MnvH1D propagates these functions to all error bands
- An error band propagates these functions to all universe

MnvUnfold

- MnvUnfold: interface to external package RooUnfold (Tim Adye, PhysStat 2011) to support MnvH1D and MnvH2D (2D version of MnvH1D)

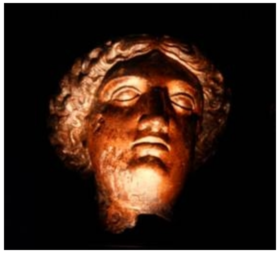
Physics analysis



- Differential cross section formula for bin i[th]

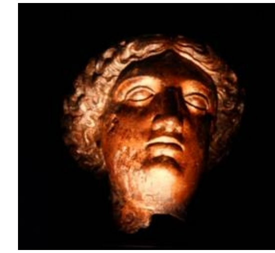
$$\left(\frac{d\sigma}{dX}\right)_i = \frac{1}{T\Phi}\frac{\sum_j U_{ij}(N_j^{data} - N_j^{bkg})}{\epsilon_i}$$

Event selection

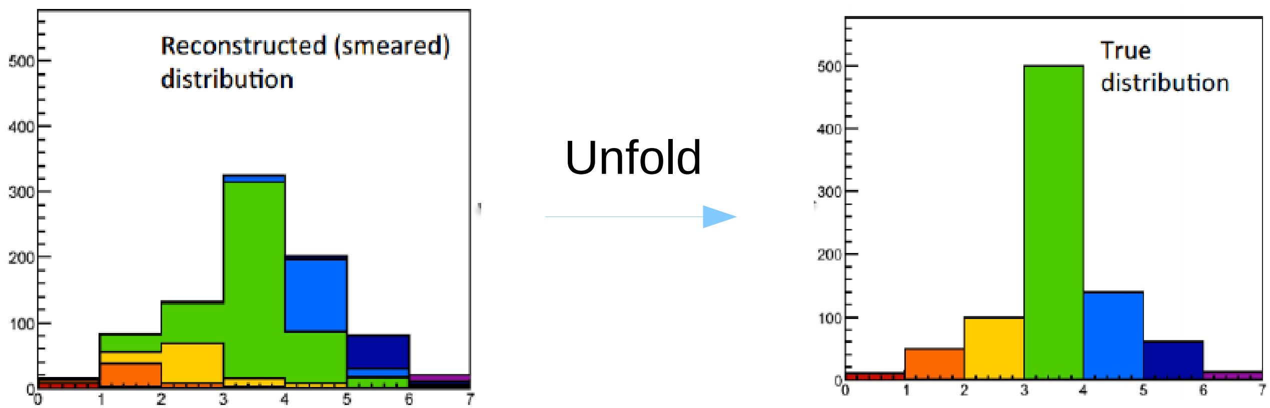


- Select signal events using their characteristic signatures in the detector using series of cuts
  - Look for two EM showers if it is a pi0
- Loop over MC events and make cuts
  - Fill histograms using selected events
  - Fill histograms in the universes
- Loop over DATA event and make cuts
  - Fill histograms using selected events

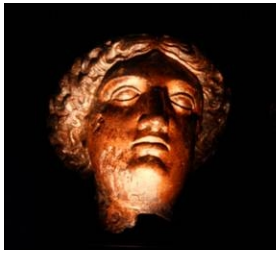
Unfolding



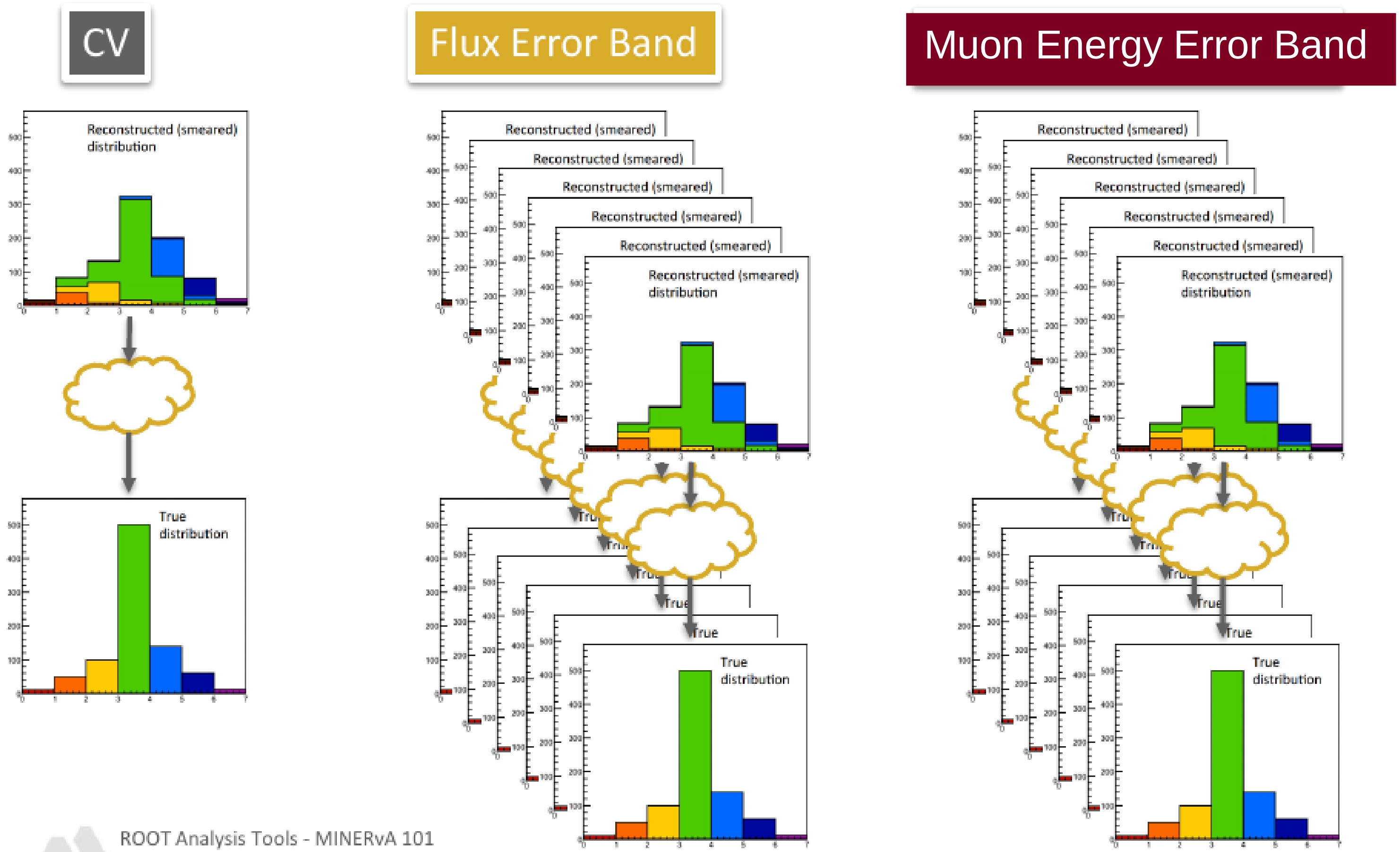
- Undo detector resolution effects
- Input: reconstructed 1D histogram and a migration matrix from MC
- Output: “true” 1D distribution



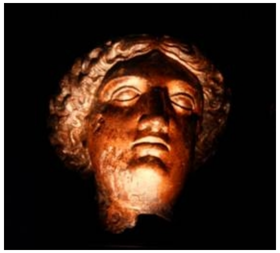
Unfolding



- Unfolding is done in each universe



Cross section



- Differential cross section formula for bin i[th]

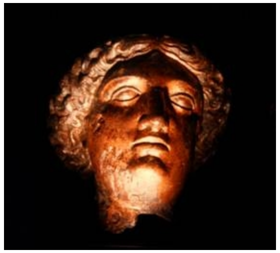
$$\left(\frac{d\sigma}{dX}\right)_i = \frac{1}{T\Phi}\frac{\sum_j U_{ij}(N_j^{data} - N_j^{bkg})}{\epsilon_i}$$

- Apply efficiency correction
- Divide by the number of nucleons and integrated flux to get the final cross section
- The cross section result is stored as an MnvH1D

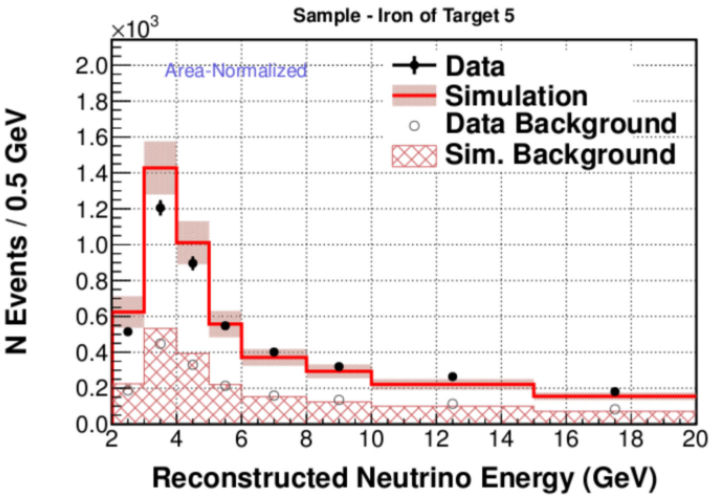
MnvPlotter –utility for making plots

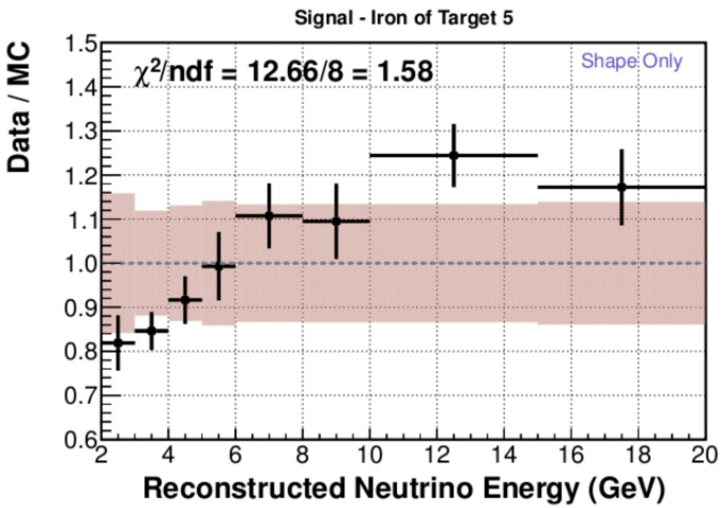
- Collection of functions that produce different types of plots

Different types of plots

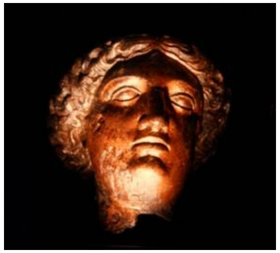


- Example

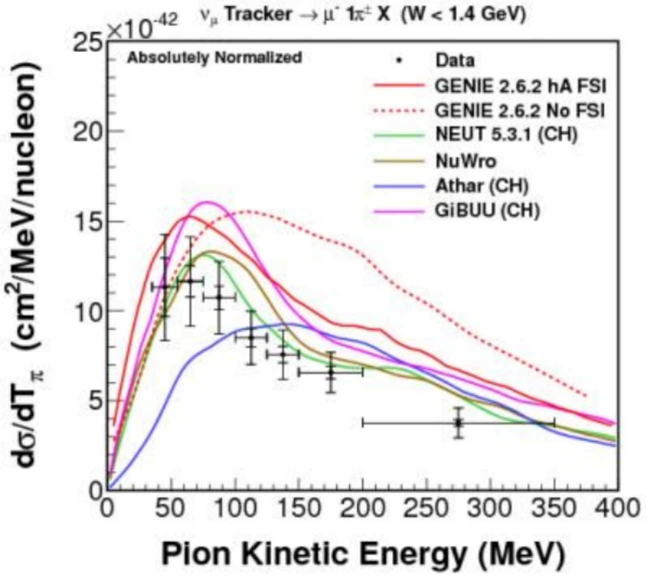


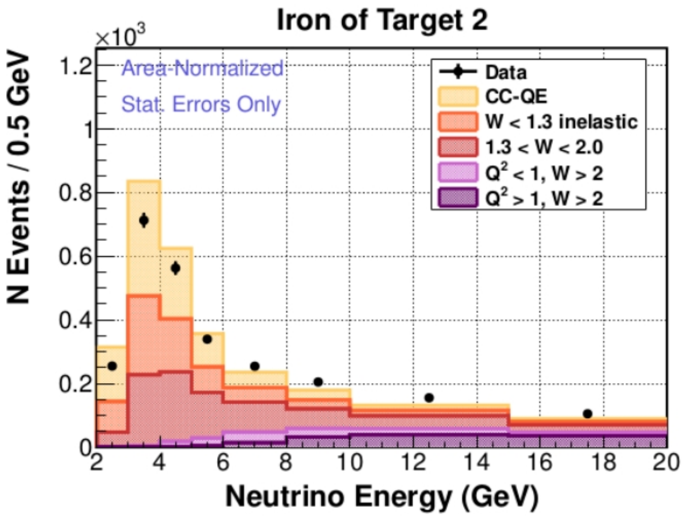


Different types of plots

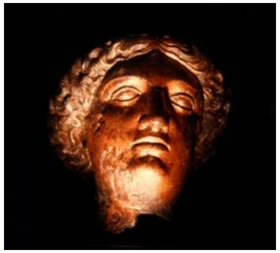


- Example

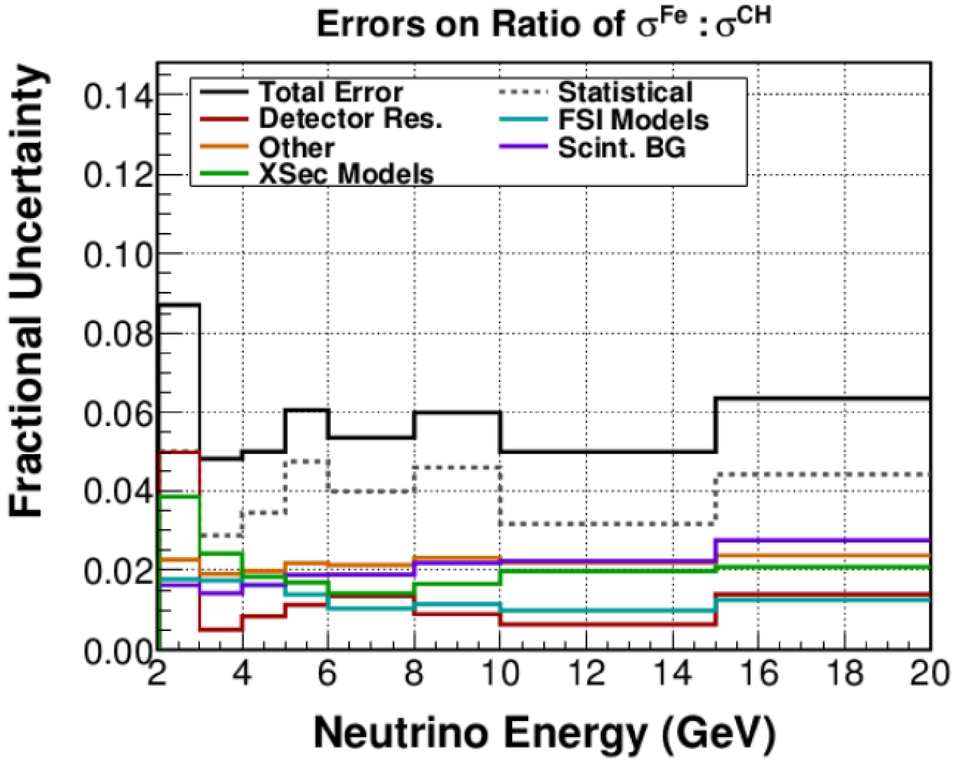




Different types of plots



- Example



- *MINERvA has done this many times sucessfully and we have developed common analysis tools that users can take advantage*