Tools for Estimating and Propagating Systematic Uncertainties

Daniel Cherdack
Colorado State University

LBNE LBPWG Systematics Session
CETUP* 2014
Monday July 14th, 2014

Introduction

- To calculate sensitivities of ELBNF to oscillation parameter measurements we need:
  - Simulations to predict event spectra
  - Oscillation analysis tools
  - Systematic uncertainty estimates
- The closer these are to reality, the better the sensitivity estimates

- What tools are available?
  - Up and running
  - In development
- Are these tools good enough?
  - Do they describe reality/data
  - Are we sensitive to improved modeling
- Where should we focus our efforts?
  - Will improvements effect calculations
  - Do the uncertainties give us sufficient coverage (are they detailed/conservative enough)

External Data

- Always the best option
- Tune models to data
- Well defined uncertainties

- Target hadronization / NA61-like experiments
- Previous neutrino beam (NuMI)
- Test beam experiments (LArIAT & CAPTAIN)
- R&D detectors (35kt)
- Previous/Running LAr experiments (ICARUS & MicroBooNE)
- Electron scattering experiments

Simulation Tools

- Beam simulations: G4LBNE
- Generators
  - GENIE:
    - Primary tool in LBNE
    - Tuned to data
    - Systematic uncertainty reweighting
  - NEUT: Primary generator for T2K
  - NuWRO: Cutting edge model implementations
  - GiBUU: Superior FSI treatment

- Detector Simulations
  - GEANT4
  - Full Simulations
    - LArSoft
    - ND simulations
  - Parameterizations
    - Fast MC
    - ND Fast MC
- Simulation chain
  - Protons on target → Reconstructed quantities
  - There is a lot going on in that "→"

Analysis Tools

- GLoBES
  - Used for LBNE sensitivity studies so far
  - Uses parameterized inputs
- My GLoBES Tools (MGT)
  - Built on GLoBES
  - Integrated with the Fast MC
  - Tools for propagation of realistic systematic uncertainties
  - Ability to do multitude of sensitivity studies
- VALOR
  - Software developed for T2K full 3-flavor oscillation analyses
  - Generalized and adapted for LBNE (and LBNO and T2HK) sensitivity studies
  - Constraints on flux + cross section from a multi-sample ND fits
    - Topologically based sample selections
    - Generates post-fit covariance matrix used in FD fits

GENIE

- Collection of neutrino cross section and related models
- Uncertainties on free parameters of the models
  - Tuned to data (somewhat involved process)
  - Set of reweighting functions to fluctuate free parameters without rerunning
- Areas of study and development crucial to ELBNF
  - Initial state of the nucleus
  - Final-state interactions
  - DIS hadronization model uncertainties
  - Single pion production rate and final-state kinematics
  - Cross section ratios ($\bar{\nu}/\nu$, ν_e/ν_μ, ν_τ/ν_μ)
  - Incorporation new models and data
  - **Updated/streamlined data tuning procedure**

Shamelessly stolen from Laura F.

G4LBNE

- We use a GEANT-4 based simulation of the LBNE beamline
  - Based on G4NuMI, a similar tool for simulating the NuMI beam line

The entire G4LBNE beamline, visualized:

S. Park

Target/ Horns/ Shielding

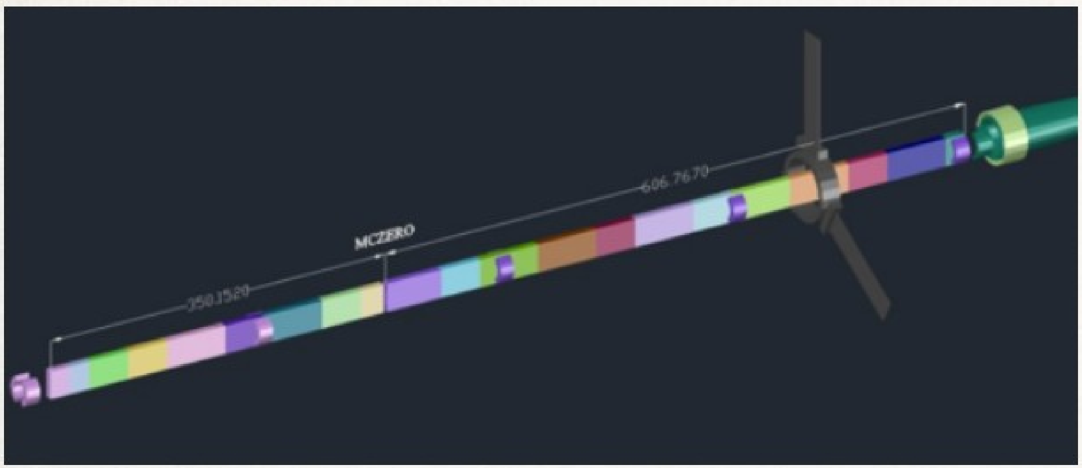
Decay Pipe

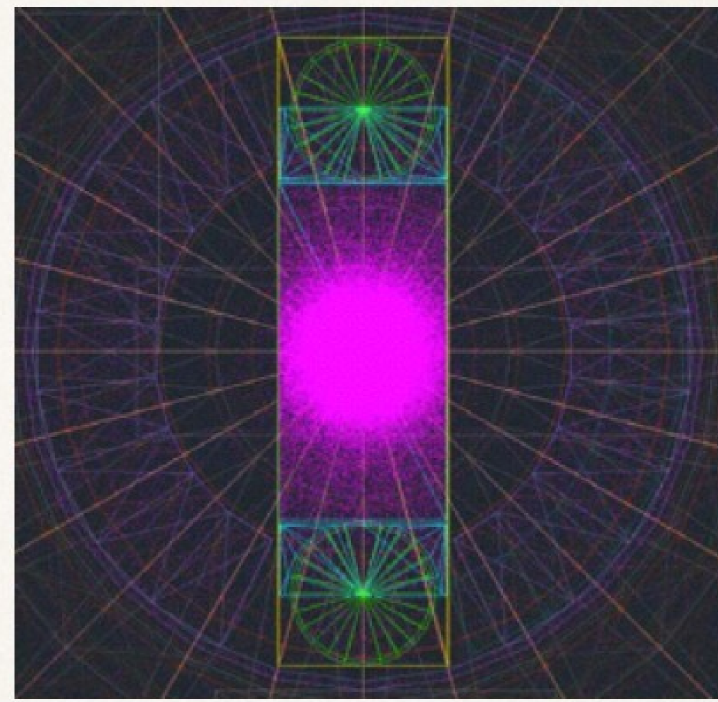
Hadron Absorber

Shamelessly stolen from Laura F.

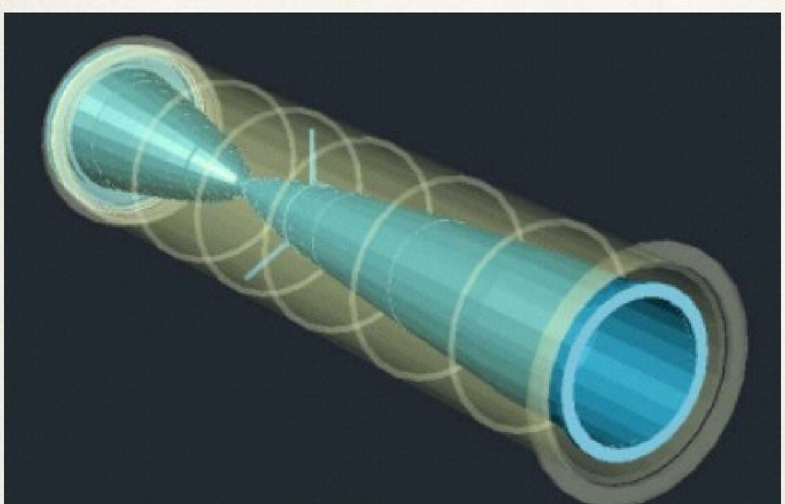
G4LBNE

- A closer look at the target and horns:





S. Park



Shamelessly stolen from Laura F.

G4LBNE

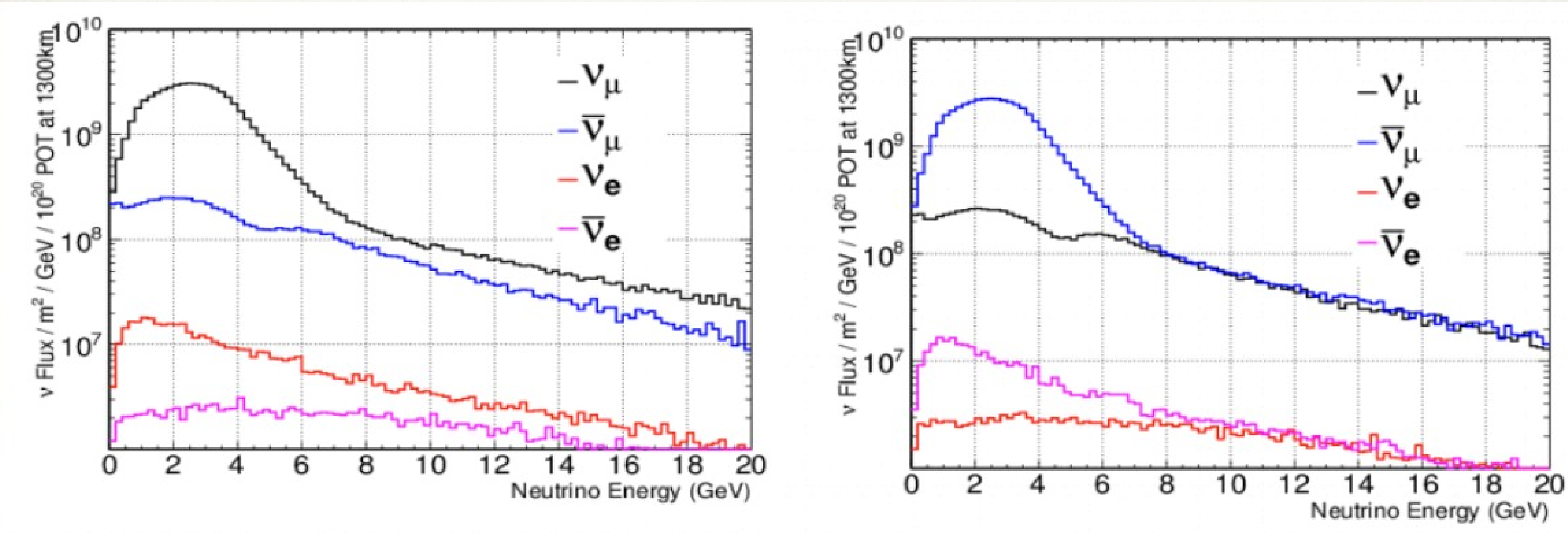
- 120 GeV protons are fired at the target and propagated through the entire beam line:

- Output is an ntuple containing information describing each neutrino produced along the beam line
    - Sufficient to produce neutrino flux distributions at any point in space (so the same simulation can be used to study flux at different locations — near and far detectors, off axis angles, etc)
- ~1 neutrino for every 2.5 protons
- Two types of weighting used to minimize CPU & disk resources — small samples (~1e8 protons) can be used to study far detector flux

Shamelessly stolen from Laura F.

G4LBNE

Nominal neutrino fluxes



Neutrino Mode
(Forward Horn Current)

Antineutrino Mode
(Reverse Horn Current)

Multiple alternate fluxes available with beam optics uncertainties
and alternate design choices

What is the Fast MC?

- A full simulation of LBNE from flux → oscillation parameter sensitivities
  - Flux (g4lbne)
  - Cross Sections and Nuclear models (GENIE)
  - Detector response (Fast MC)
  - Reconstruction (Fast MC)
  - Analysis Samples (Fast MC)
  - Systematics Uncertainties (g4lbne, GENIE reweighting, Fast MC, etc)
  - Sensitivity Studies (GLoBES)
- Allows the user to:
  - Simulate (almost) every aspect of the experiment
  - Accurately generate analysis samples
  - Propagate systemic uncertainties to physics sensitivities
    - Improve beam and detector design, and understand the ramifications of design tolerances
    - Understand leading sources of physics uncertainty, and work with theorist, current experiments, and ND designers to reduce them

How Does the Fast MC work

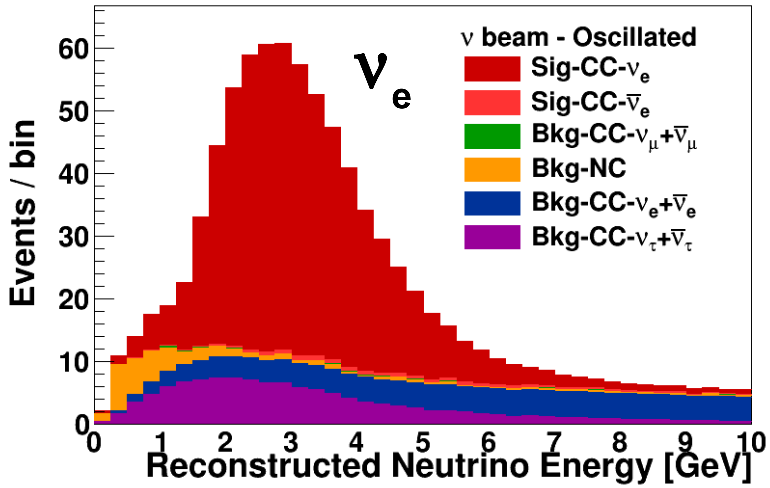
- Use flux files and GENIE to generate ν-nucleus interactions on LAr
  - List of final state particles (after FSI)
  - Truth level 4-vectors and kinematics
- Loop over events and:
  - Smear the energy/momentum/angle of each final state particle
  - Reconstruct event level kinematic quantities (E_ν, Q^2, x, y, etc)
  - Identify lepton candidate (CC-ν_μ: longest MIP track, CC-ν_e: largest EM shower, NC: neither)
  - Classify each event based on lepton candidate
  - Calculate weights for ±1,2,3 σ fluctuations in source of systematic uncertainty (cross section, nuclear model, flux, energy resolution, etc)
- Use output 'reconstructed' quantities and analysis variables to:
  - Plot 'reconstructed' energy spectra for the ν_e appearance and ν_μ disappearance event samples
  - Plot ratios of systematically fluctuated spectra to the nominal spectra
  - Generate inputs to a modified version of GLoBES
    - Energy spectra (true)
    - Smearing functions
    - **'Response functions' encoding systematic variations**

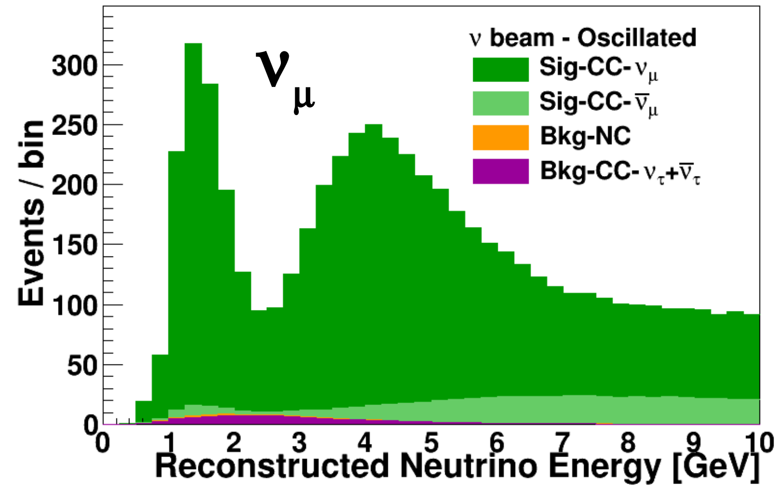
Detector Response and PID

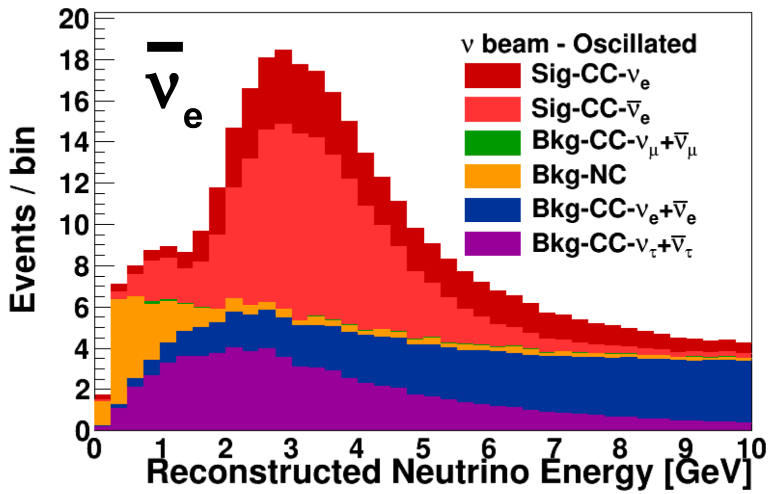
- Detector response based on:
  - GEANT4 simulations of particle trajectories in LAr
  - Resolutions (E/p/θ) determined from ICARUS papers and LArSoft
- Reconstruction
  - Straightforward
  - $E_\nu = E_{lep} + \Sigma E_{had}$
  - Missing energy from neutrons and particles below threshold
- Possible improvements:
  - Neutron response
  - Charged pion fates
  - Updated smearing and threshold numbers
  - Improved response with a photon detector
  - Updated detector and FV dimensions

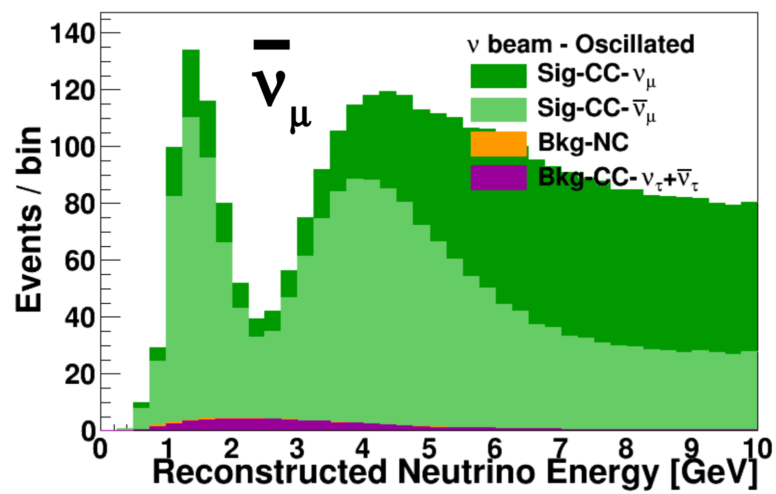
- Classification:
  - CC-ν_μ: MIP-like track > 2 m
  - CC-ν_e: e-like EM shower (no μ candidate)
  - NC: no μ or e candidate
- Low energy response
  - Efficiency of selection based on:
    - Energy of candidate lepton
    - Hadronic shower energy fraction (Y_{bj})
  - Selection probability = $[E_{lep}*(1-Y_{bj}+1) - E_{thr}] / [E_{lep}*(1-Y_b+1) - E_{thr}* \mathbf{m}]$
  - Scanning study results used to tune **m**
- E/γ separation
  - Based on very preliminary studies
  - Requires 95% signal efficiency
  - Applied to low multiplicity (<4 prongs) events
- kNN based ν_τ cut (also cuts NC)

Reconstructed Energy Spectra

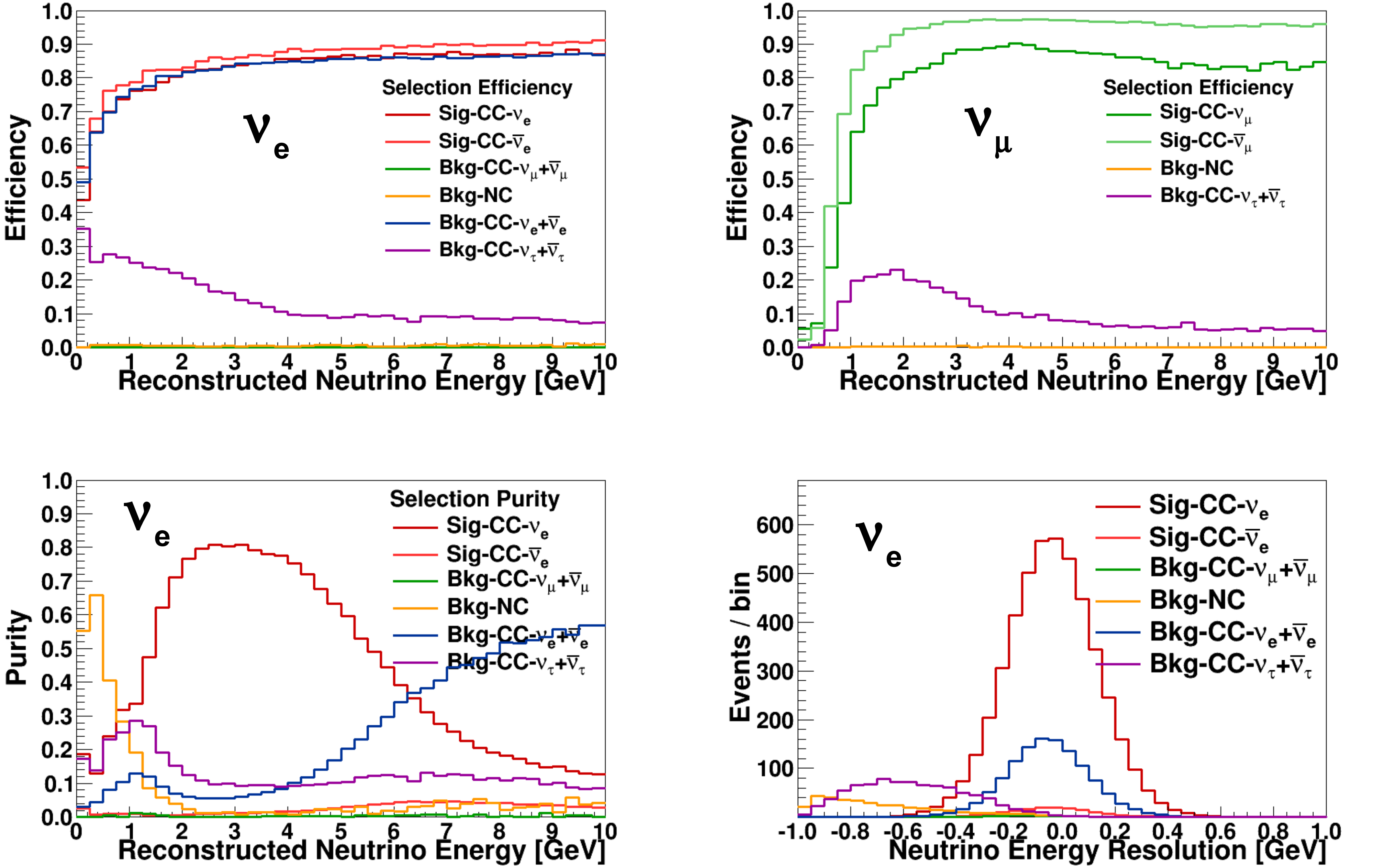




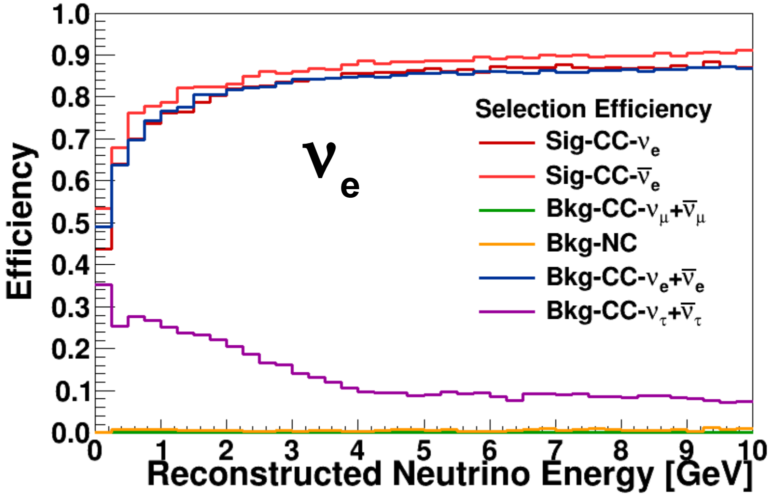




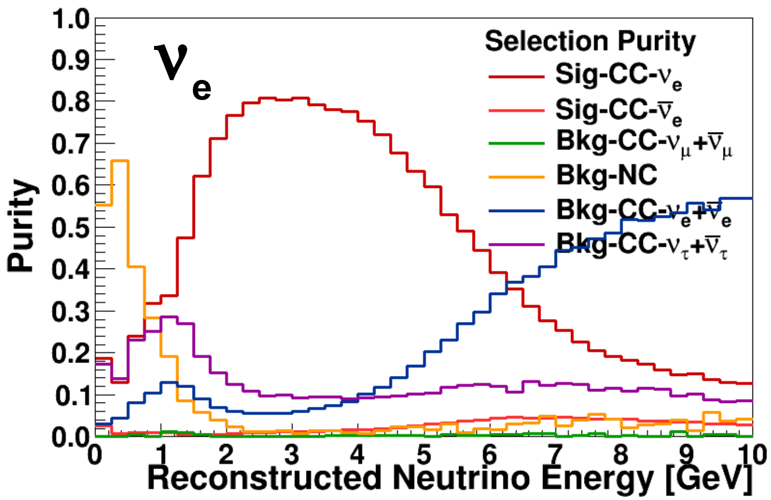
Purity, Efficiency, and Energy Resolution

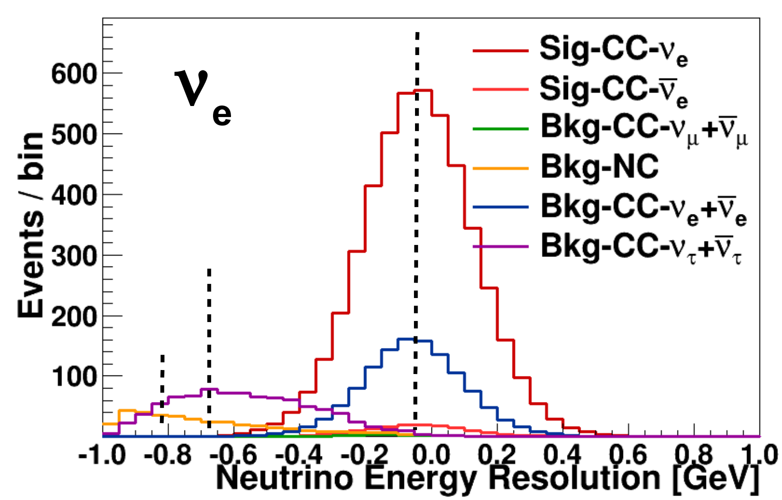


Purity, Efficiency, and Energy Resolution

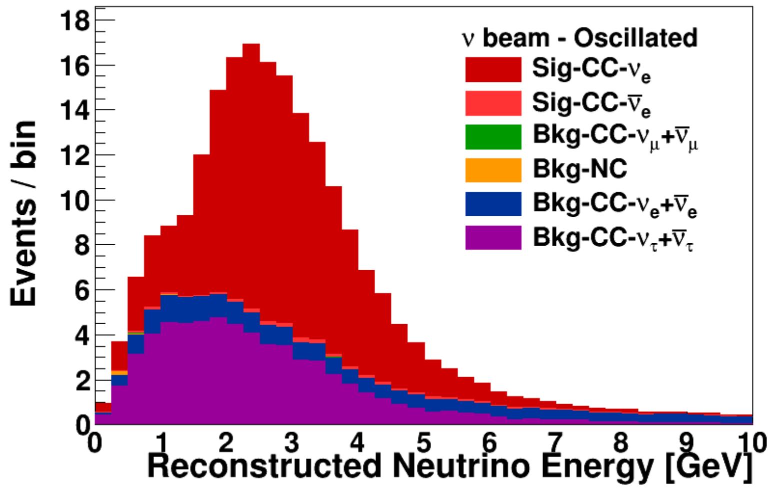


- Calorimetric energy response
- Bias in CC ν_μ and CC ν_e events mostly from missing energy from neutrons
- Bias in NC and CC ν_τ enhanced by final state neutrinos

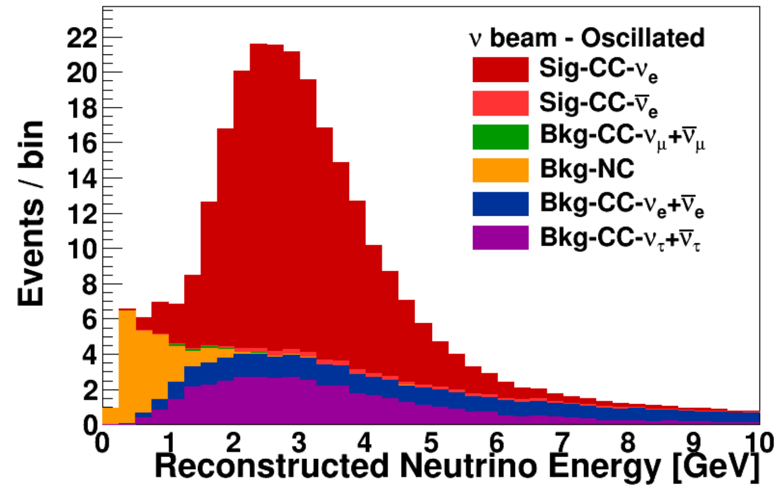




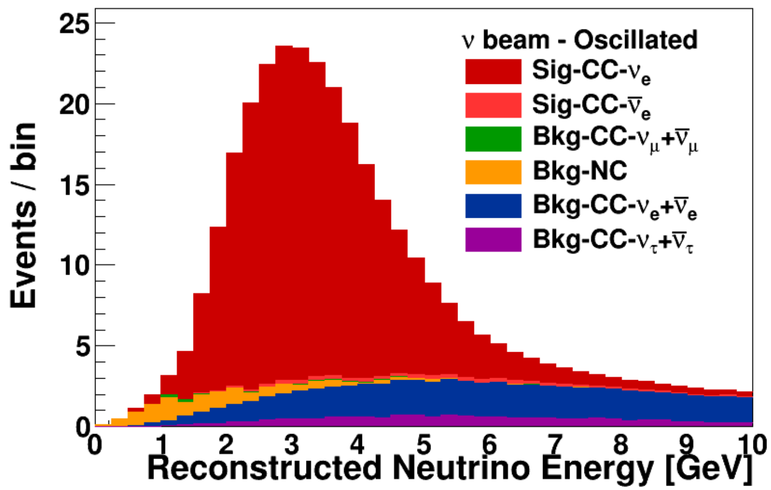
ν_e-Appearance by X-Sec Model



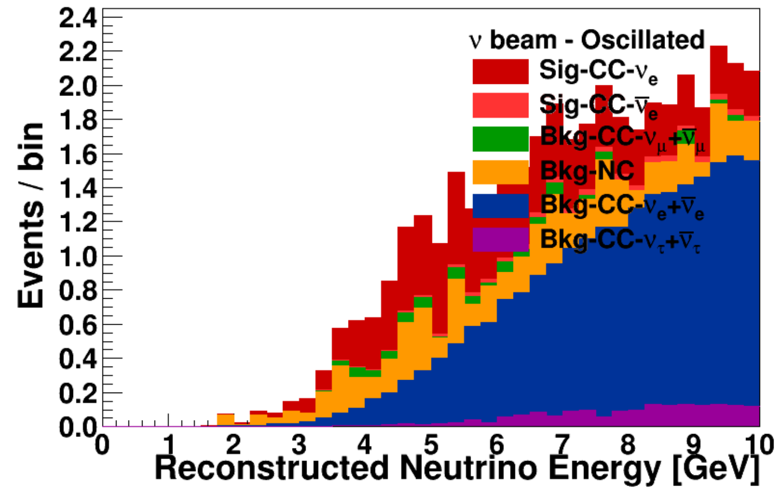
Quasi-elastic



Resonance Production



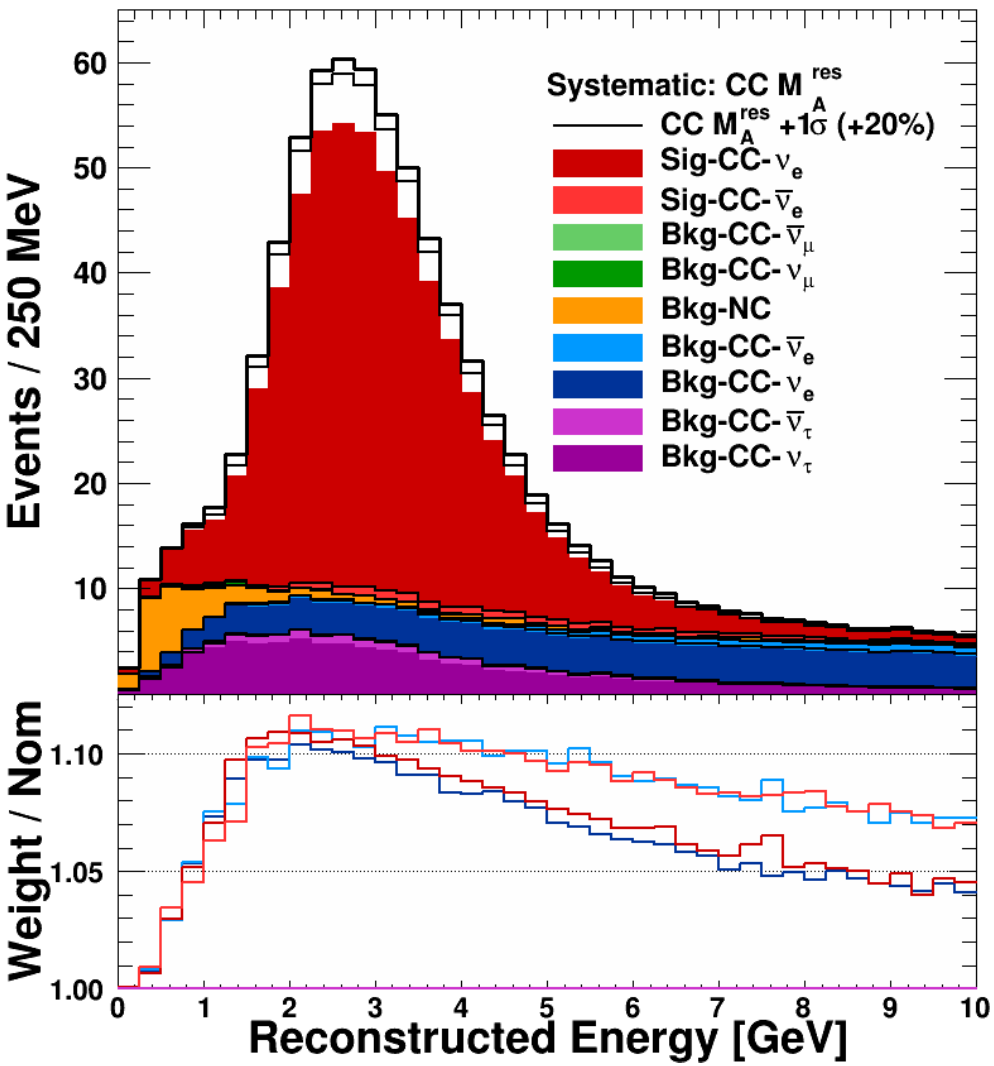
DIS W < 2.7



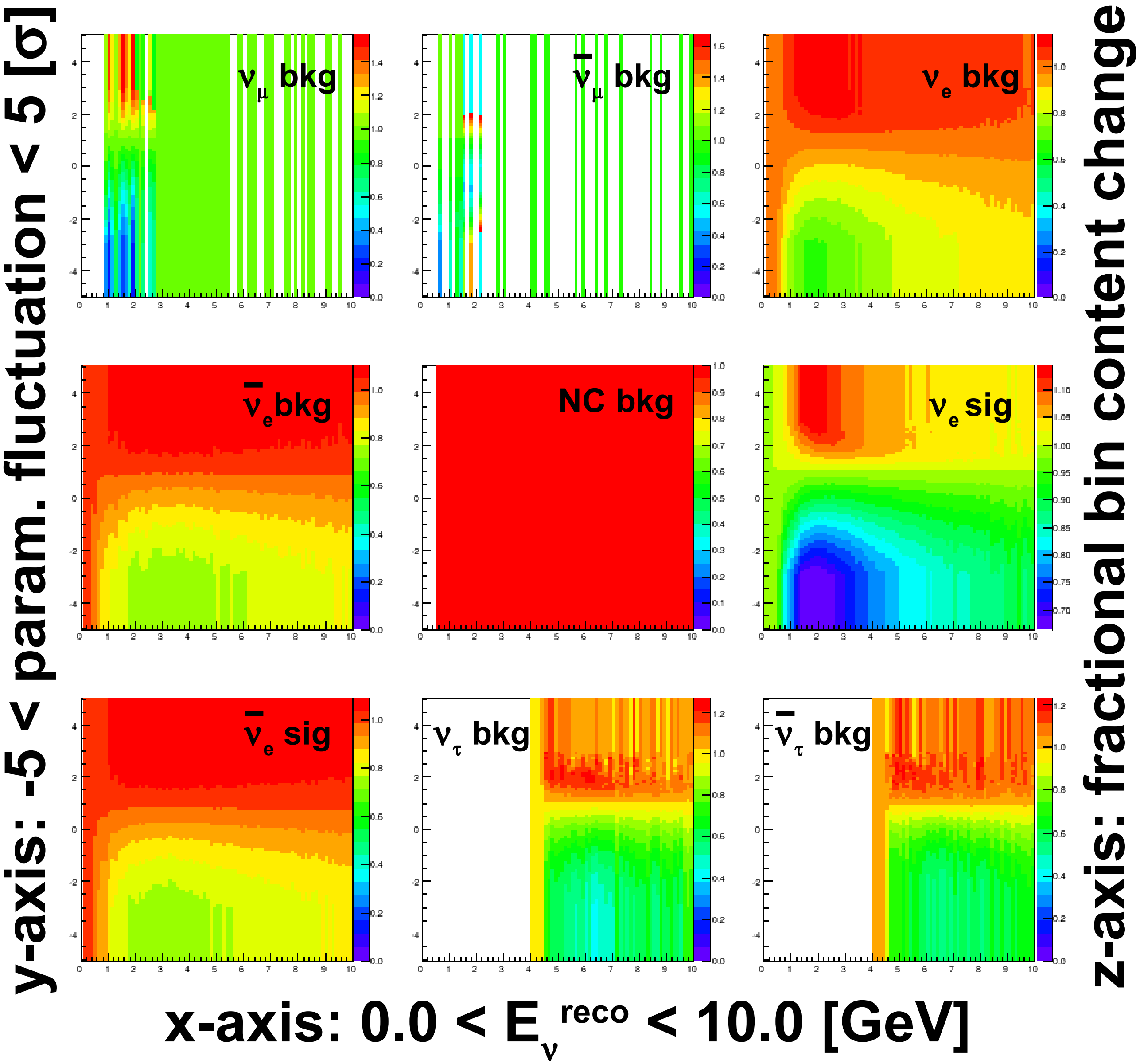
DIS W > 2.7

Systematic Weights

- Currently Considered
  - Flux: beam optics parameters, beam optimizations
  - Xsec: QE, RPA, res, res->DIS, Intranuke
- In development
  - Flux: hadronization model
  - Xsec: nuclear initial state, DIS and hadronization model
  - Detector response: reconstructed energy scale, detection and selection efficiencies

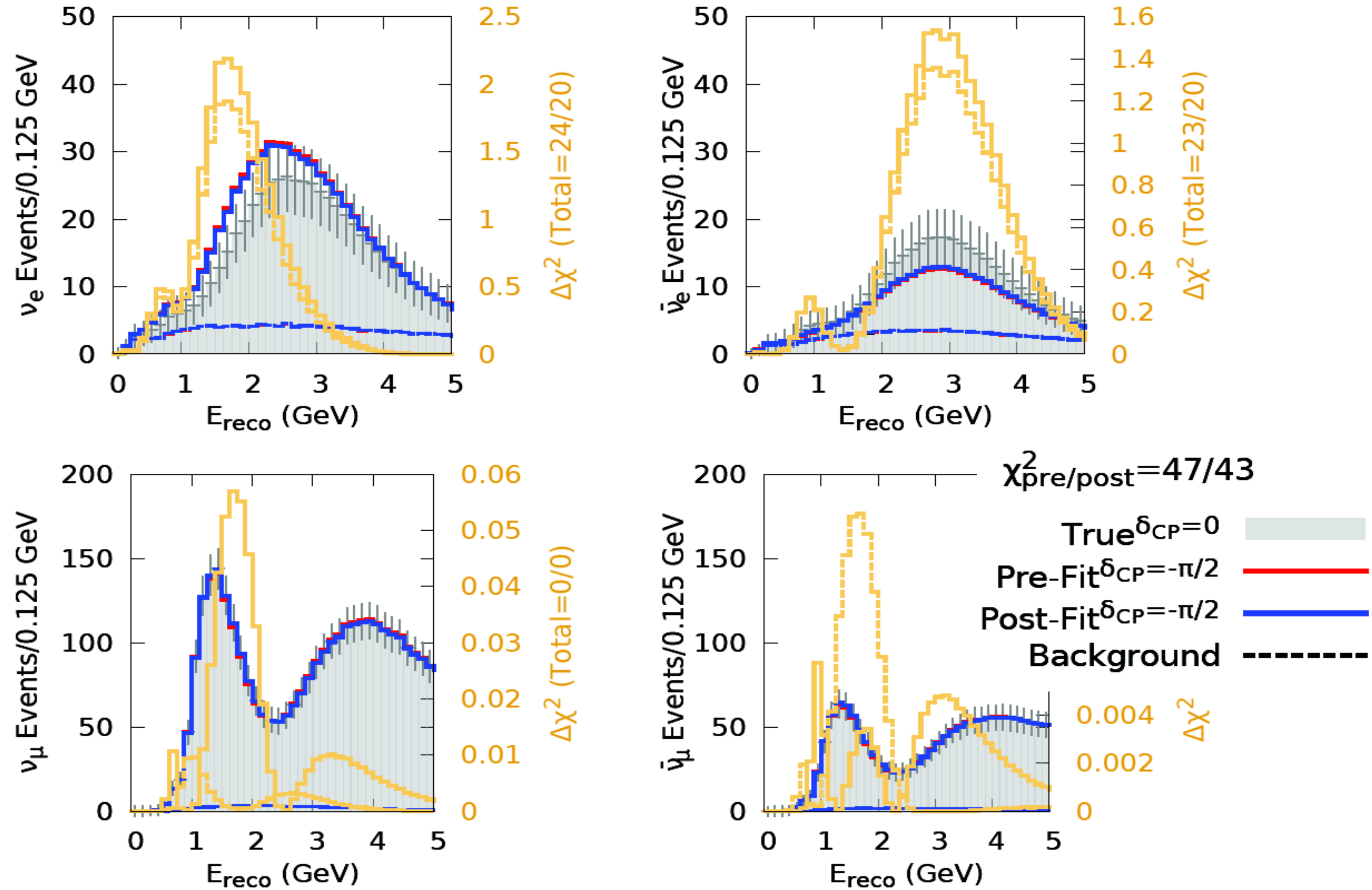


My GLoBES Tools (MGT)



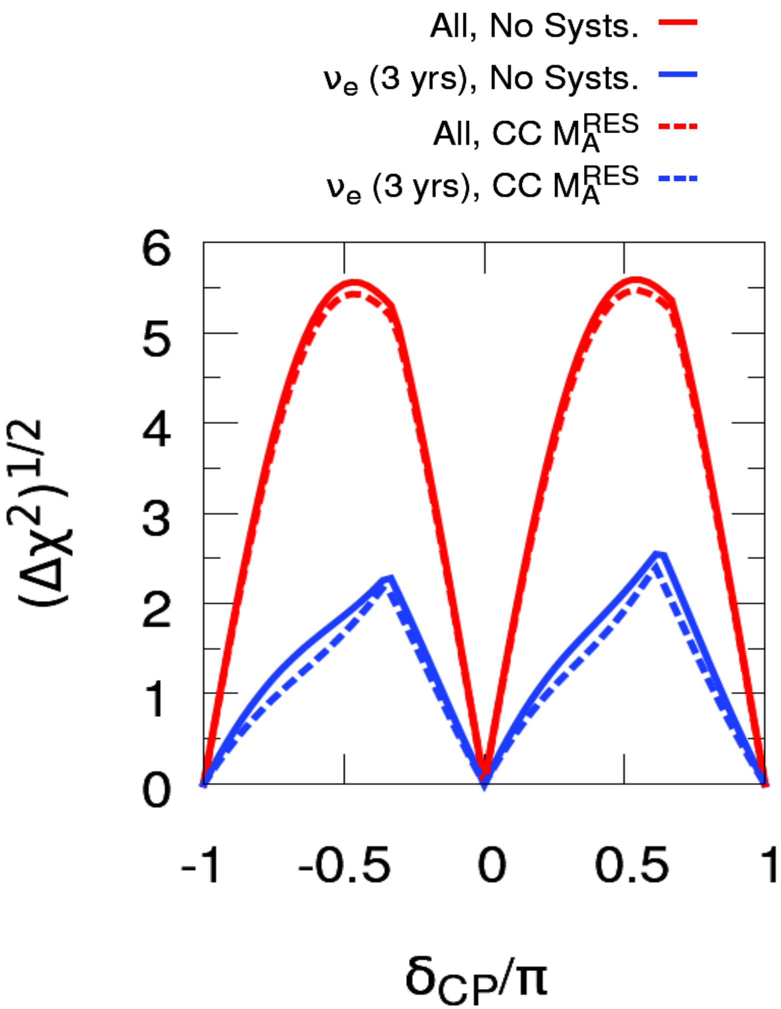
- Based on GLoBES fitter
- Takes inputs built event-by-event from Fast MC
  - Analysis sample true energy spectra
  - Smearing functions
  - Systematic error response functions (left)
- Determines sensitivity with detailed systematics

CPV Fit Spectra and χ^2 with Variations in M_A^{res} (w/ osc systs)



Sensitivity to CPV with Variations in M_A^{res}

- Fits to all 4 samples
- Exposure: 3yrs, 1.2MW, 34 kt
- No ND constraints
- **WITH oscillation systematics**
- Allow CC M_A^{res} to vary by ±20%
  - Current generator level uncertainty / no ND constraint
  - CC M_A^{res} is essentially a normalization on resonance production interaction in E_{reco}
- Degradation to the sensitivity is greatly decreased
  - Large constraint from $\bar{\nu}_e$ or ν_μ samples



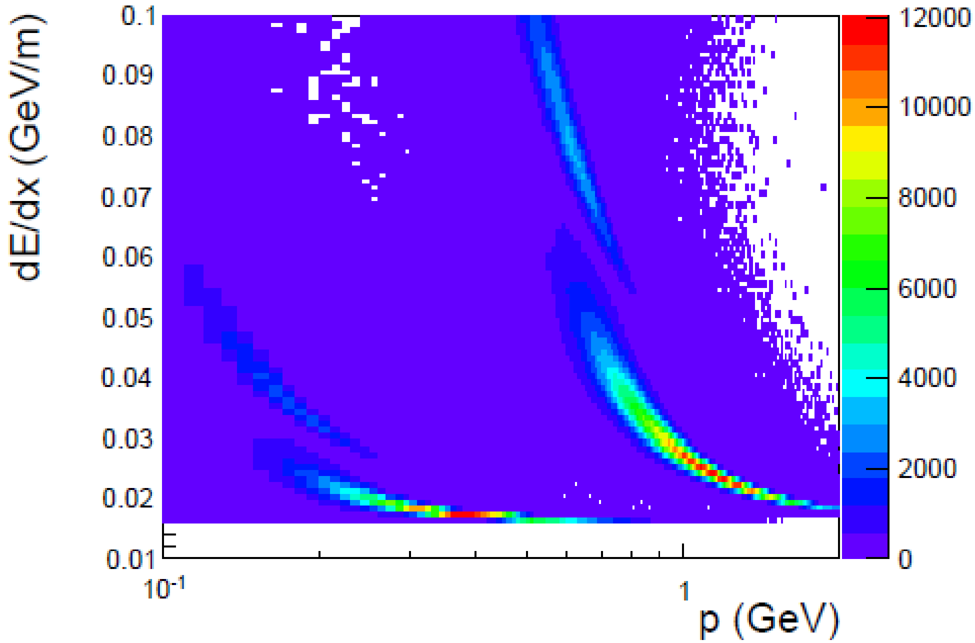
Shamelessly stolen from Xinchun T.

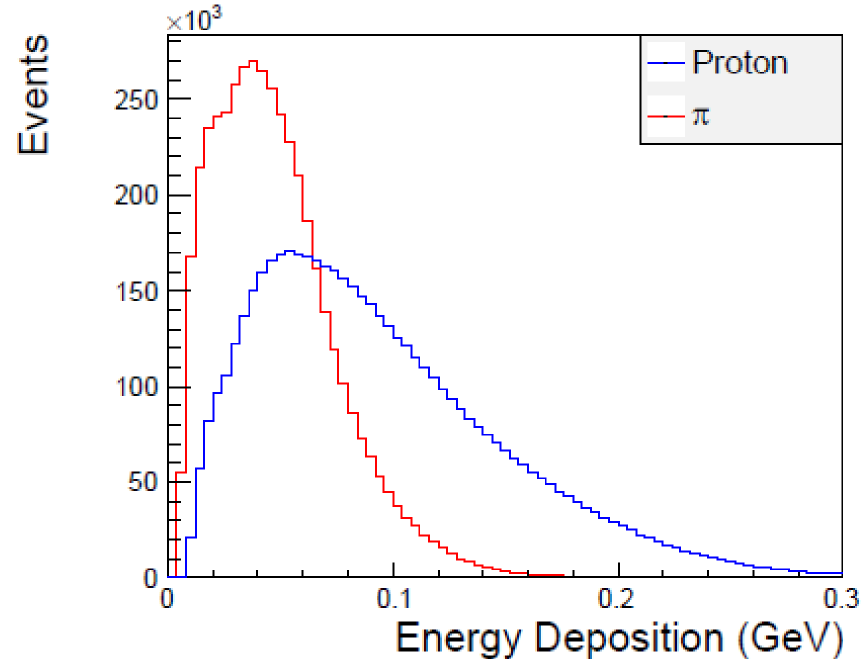
The FGT ND Fast MC

- Fast MC = Fast Detector Simulation + Fast Reconstruction
- The Chain: G4LBNE→GENIE→ND Fast MC→Analyzing the output ROOT files
    - G4LBNE produces the flux
    - GENIE produces the interactions with a homogeneous detector with approximately the same composition as the current design of HiResMν
    - ND Fast MC will mimic the detector simulation and recontruction to produce the "reconstructed" variables for downstream analysis
    - Analyzing the output "reconstructed" ROOT files for specific topics
- Use the exisiting NOMAD data to benchmark the whole chain
- Re-use as much as possible the existing Fast MC codes developed by Dan and Rik. It is also a good cross check of the existing code

Shamelessly stolen from Xinchun T.

The FGT ND Fast MC - Inputs



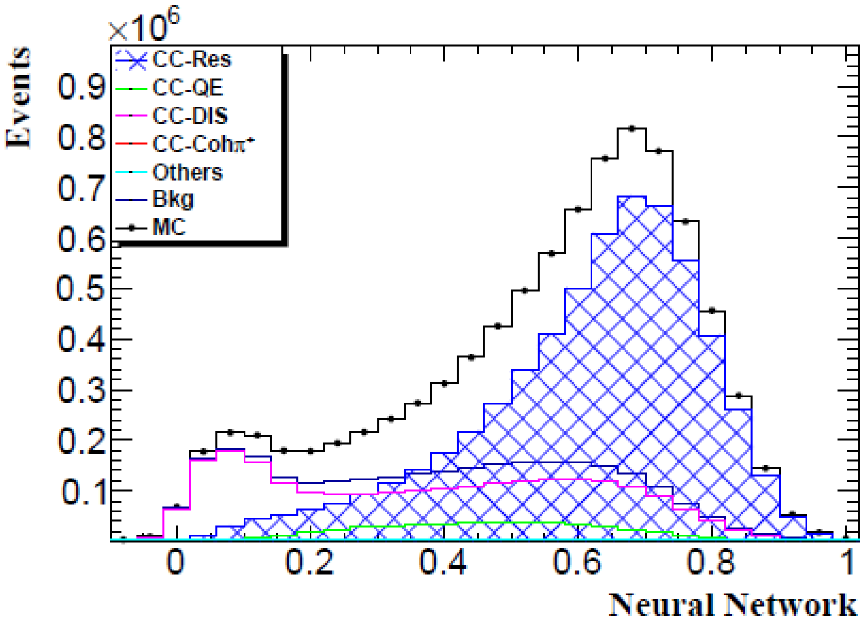


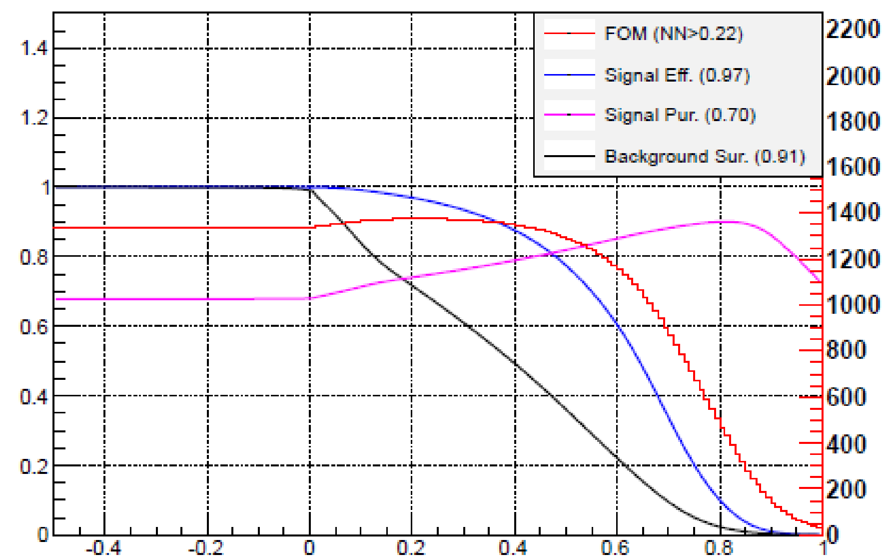
DE/dx inputs for PID tagging efficiencies

Shamelessly stolen from Xinchun T.

The FGT ND Fast MC - Analyses

- NN inputs: p_{μ}^{x}, p_{μ}^{y}, p_{μ}^{z}, p_{Proton}^{x}, p_{Proton}^{y}, p_{Proton}^{z}, p_{π}^{x}, p_{π}^{y}, p_{π}^{z}





Analyses use neural network based event selections using kinematic quantities

Shamelessly stolen from Costas A.

VALOR

Costas Andreopoulos[1,2], Fatih Bay[3], George Christodoulou[1], Thomas Dealtry[4], Steve Dennis[2,5], Debra Dewhurst[4], Lorena Escudero[6], Nick Grant[7], Silvestro Di Luise[3], Davide Sgalaberna[3] and Raj Shah[4].

[1]University of Liverpool, [2]STFC Rutherford Appleton Laboratory, [3]ETH Zurich, niversity of Oxford, [5]University of Warwick, [6]IFIC Valencia, [7]University of Lancas

VALOR is a well-established (EU) T2K oscillation fitting group (2010-present) with contribution to several published T2K oscillation results.

The VALOR code was adapted for HyperK at the end of last year.

(Recent contribution at the 5th Open HyperK Meeting, Vancouver: http://indico.ipmu.jp/indico/contributionDisplay.py?contribId=49&sessionId=24&confId=34)

The code is now adapted for LBNx ND (systematic constraint) and FD (3, 3+1 and 3+2 flavour oscillation fits

Objectives:

- **Physics-driven requirements for the LBNE, LBNO and T2HK designs.**
- Going beyond simple GloBES studies.
  - Using a framework deriving from a real analysis that produced the best constraints on θ_{23} and δ_{CP}, and one of the best constraints on Δm^2_{32}.
  - Using a common framework for all experiments.
  - Using a common framework for all proposed configurations within each experiment.

VALOR

Selections are based on LBNE FastMC
The LBNE FastMC (Dan Cherdack, Rik Gran) is a parameterized detector response package built on top of GENIE. The HiResMν version of FastMC was developed by Xinchun Tian. Several improvements were installed by George Christodoulou for this analysis.

- ν_μ CC inclusive
  - ν_μCC 1-track QE enhanced (FHC: μ^- only)
  - ν_μCC 2-track QE enhanced (FHC: μ^- + p)
  - ν_μCC $1\pi^\pm$ (FHC: $\mu^- + 1\pi^\pm$ + X)
  - ν_μCC $1\pi^0$ (FHC: $\mu^- + 1\pi^0$ + X)
  - ν_μCC $1\pi^\pm + 1\pi^0$ (FHC: $\mu^- + 1\pi^\pm + 1\pi^0$ + X)
  - ν_μCC other

    $\rightarrow$ in future, subdivide further (3-track Δ-enhanced, νe)
- Wrong-sign ν_μCC inclusive (FHC: μ^+ + X)

  $\rightarrow$ in future, subdivide further
- ν_eCC inclusive (FHC: e^- + X)

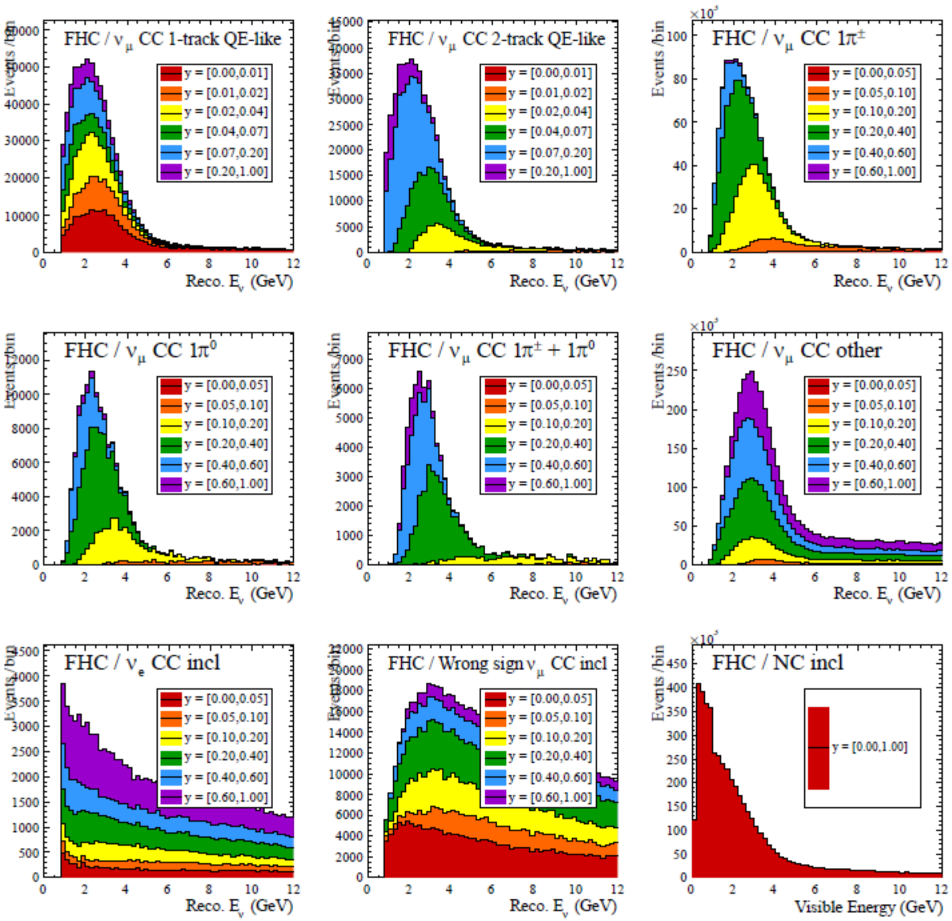
  $\rightarrow$ in future, subdivide further
- NC inclusive

  $\rightarrow$ in future, subdivide further (NCEL, NC $1\pi^\pm$, NC $1\pi^0$)

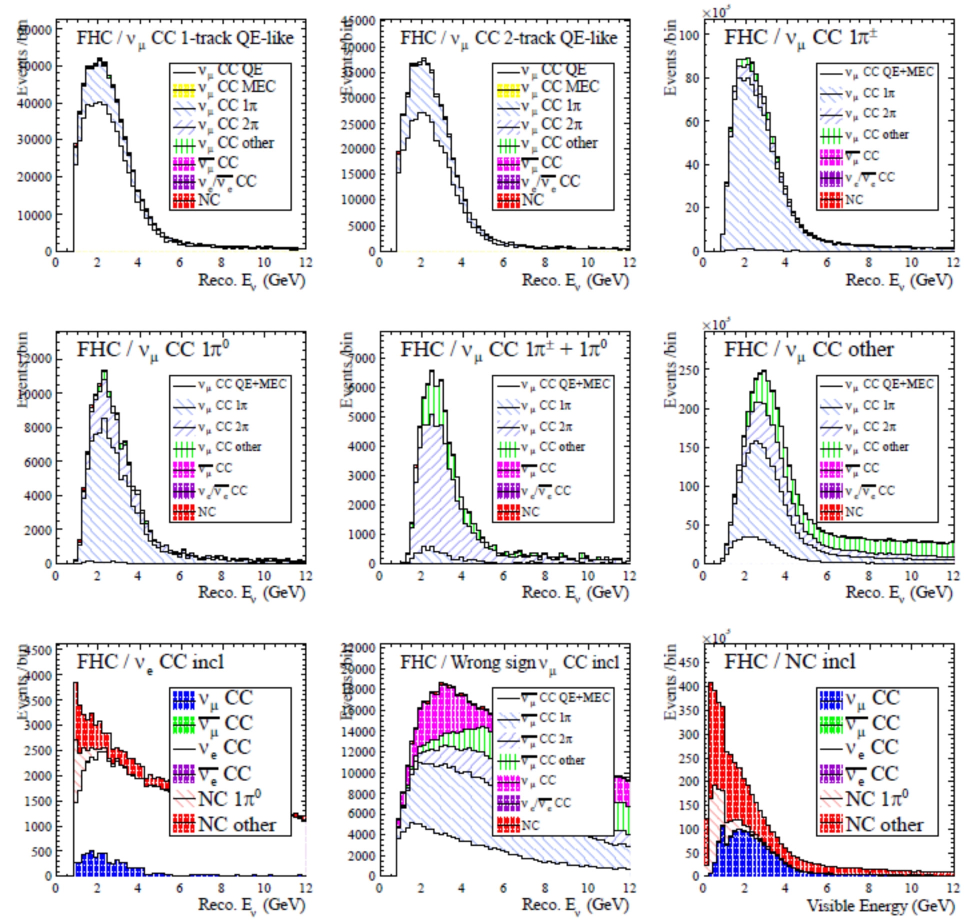
Samples in red are included in the *current (2014v1) version* of our ND systematics constraint fit.

Inclusion of other samples, and their utility in constraining systematic uncertainties, will be tested in future iterations of this work.

VALOR



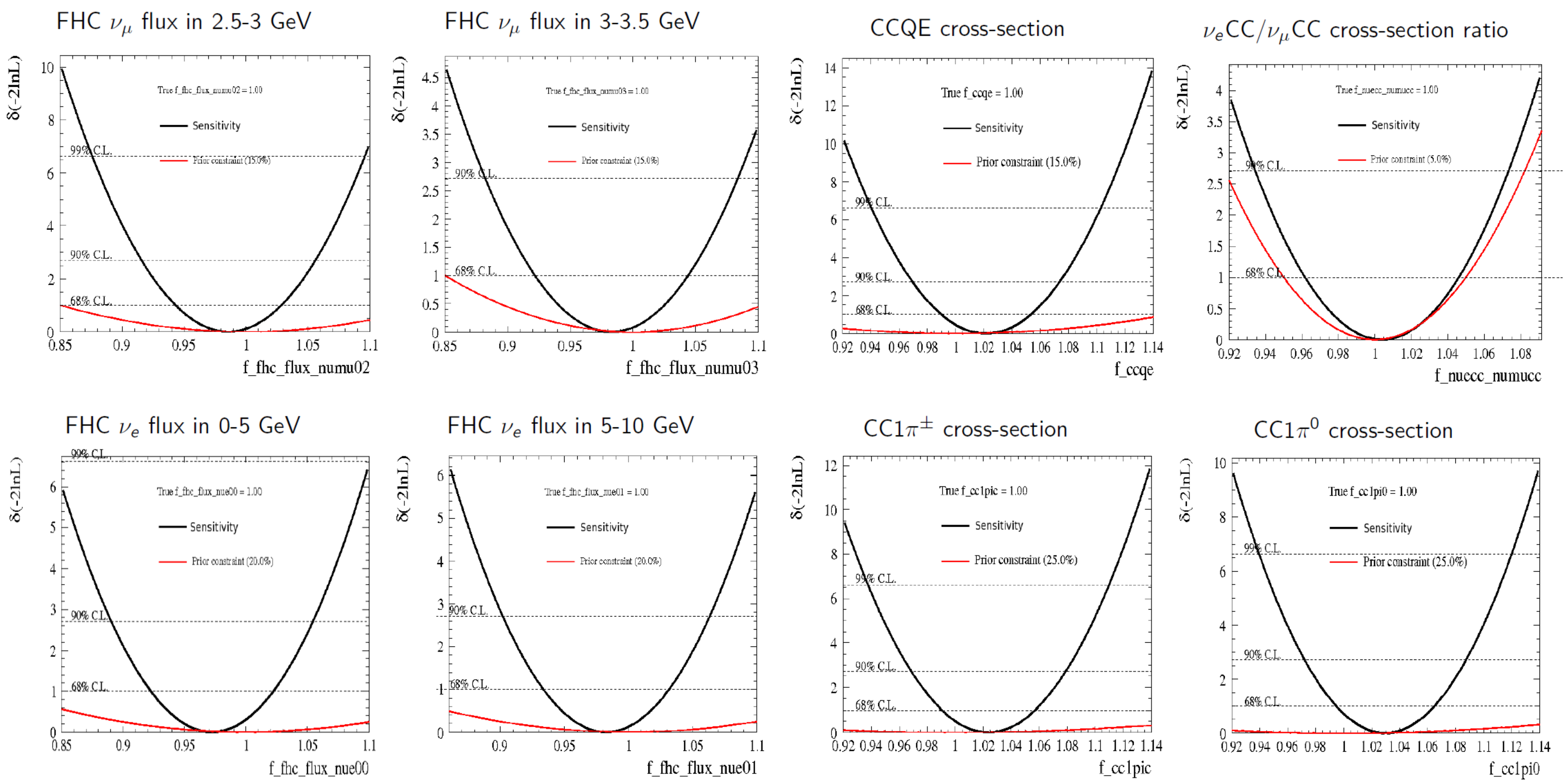
ND Event Samples by Bjorken y



ND Event Samples by Interaction channel

Shamelessly stolen from Costas A.

VALOR



Sample ND Fit Results

FNAL Redmine Project Links

- Systematics document: https://cdcvs.fnal.gov/redmine/projects/lbne-systematics/wiki/Status_of_Systematics
- Beam Simulations: https://cdcvs.fnal.gov/redmine/projects/lbne-beamsim
- Flux Utilities: https://cdcvs.fnal.gov/redmine/projects/nuutils
- GENIE: https://cdcvs.fnal.gov/redmine/projects/genie
- LArsoft general: https://cdcvs.fnal.gov/redmine/projects/larsoft/wiki
- LBNE sim/reco: https://cdcvs.fnal.gov/redmine/projects/lbne-fd-sim/wiki
- Fast MC: https://cdcvs.fnal.gov/redmine/projects/fast_mc/wiki/Fast_MC_Basics
- MGT: https://cdcvs.fnal.gov/redmine/projects/lbne-lblpwgtools/wiki