Solutions for Past Atmospherics Simulations

November 21st, 2022

DUNE HEP Working Group Meeting

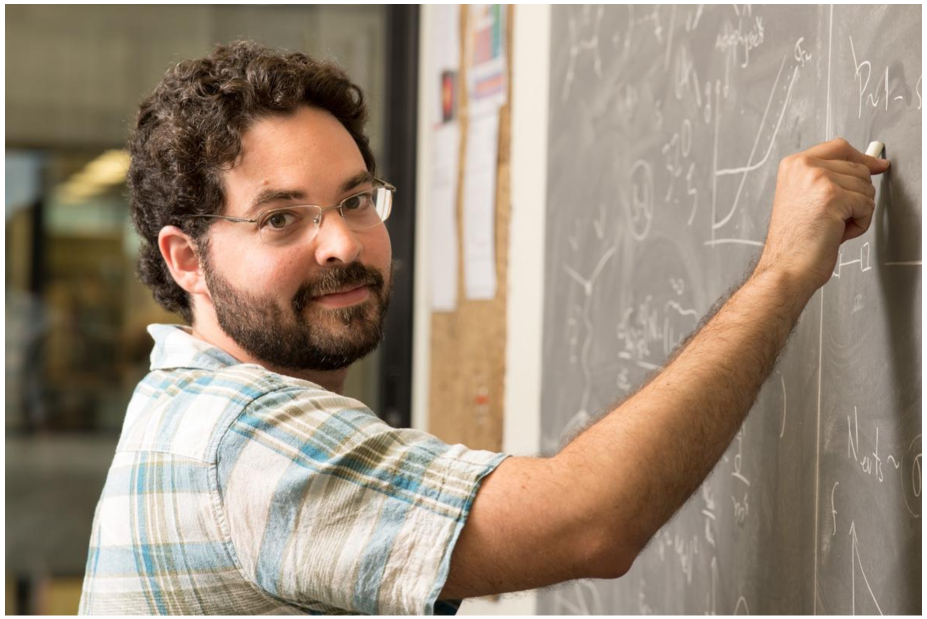
by J. L. Barrow

Zuckerman Postdoctoral Scholar

MIT & TAU

Collaborators

See Pedro and Ivan's atm. ν PRL



Pedro Machado

Ivan Martinez-Soler

Viktor Pěč

Worked together to calculate oscillated atmospheric neutrino fluxes in a DUNE 10 kt far detector module for all neutrino directions with $E\epsilon(0.1,100)\text{GeV}$ from 15km utilizing Honda fluxes

The Need to Understand Simulation Systematics

...in the context of automated reco. and analysis techniques (CNNs, BDTs)

- How do we understand the usefulness and stability of prospective, Monte Carlo-derived limits?
  - When we *simulate events*, how stable are signal efficiencies?
  - When we *simulate backgrounds*, how stable are rejection rates?
  - In some respect, these questions go beyond simple statistics…
- Do our final state signal and/or background topologies change depending on the underlying nuclear and intranuclear cascade models?
  - In GENIEv3.0.6, there are three main nuclear models *currently* available:

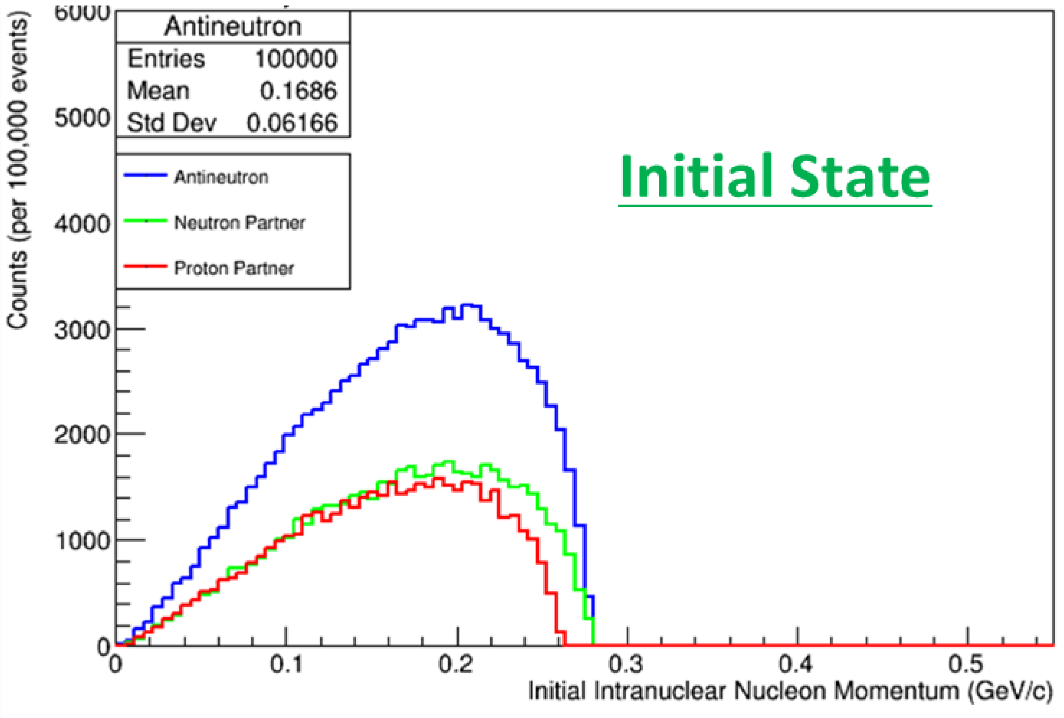
1. Bodek-Ritchie (relativistic) *nonlocal* Fermi gas
2. *Local* (nonrelativistic) Fermi Gas
3. Effective Spectral Function (*nonlocal*)

  - There are two main intranuclear cascades *currently* available:

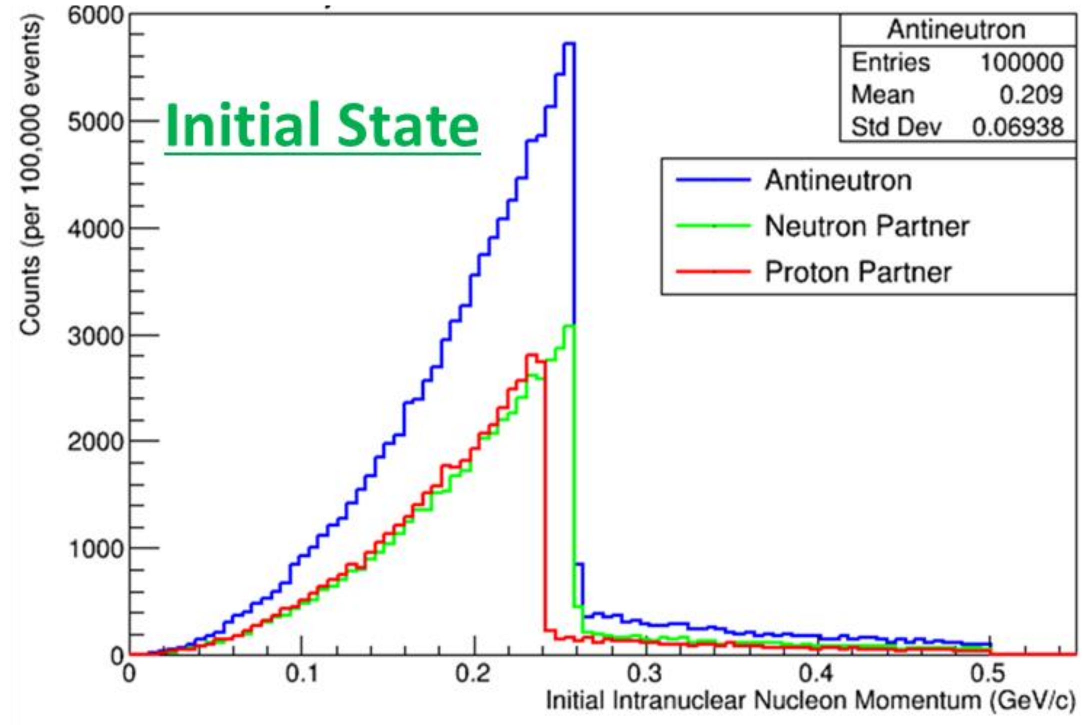
1. hA2018 (single effective interaction)
2. hN2018 (full intranuclear cascade model)

Some GENIEv3.0.6 Intranuclear $\bar{n}N$ *Initial* States

Single Nucleon Momentum



GENIEv3.0.6 hA/hN2018 and Local Fermi Gas

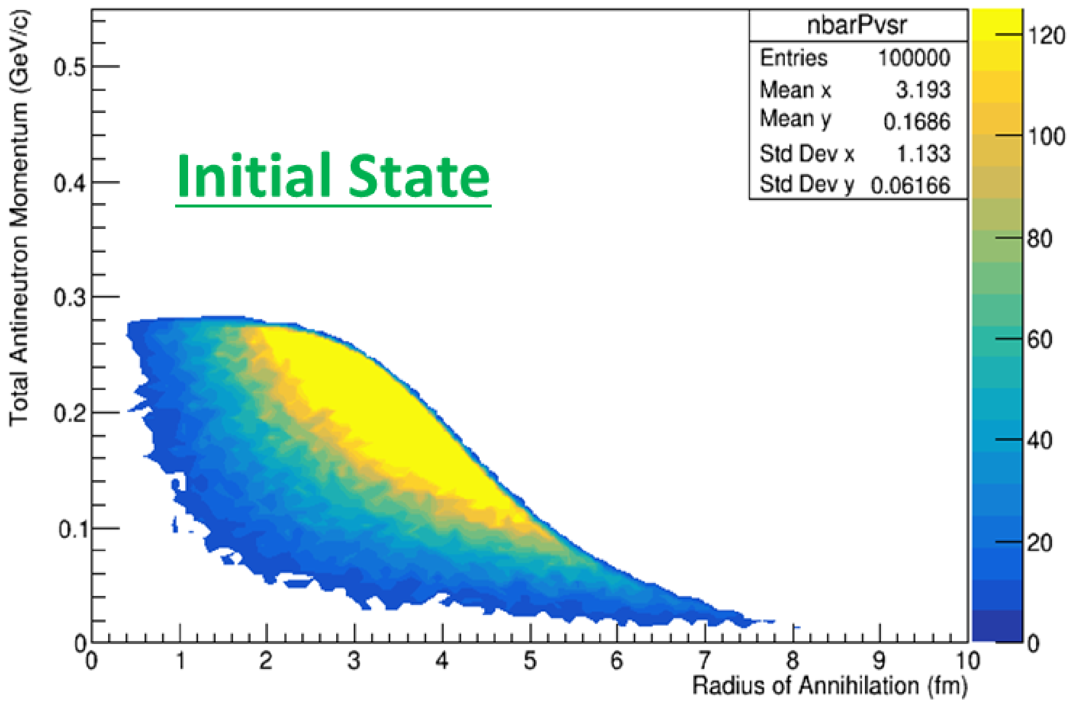


GENIEv3.0.6 hA/hN2018 and Bodek-Ritchie

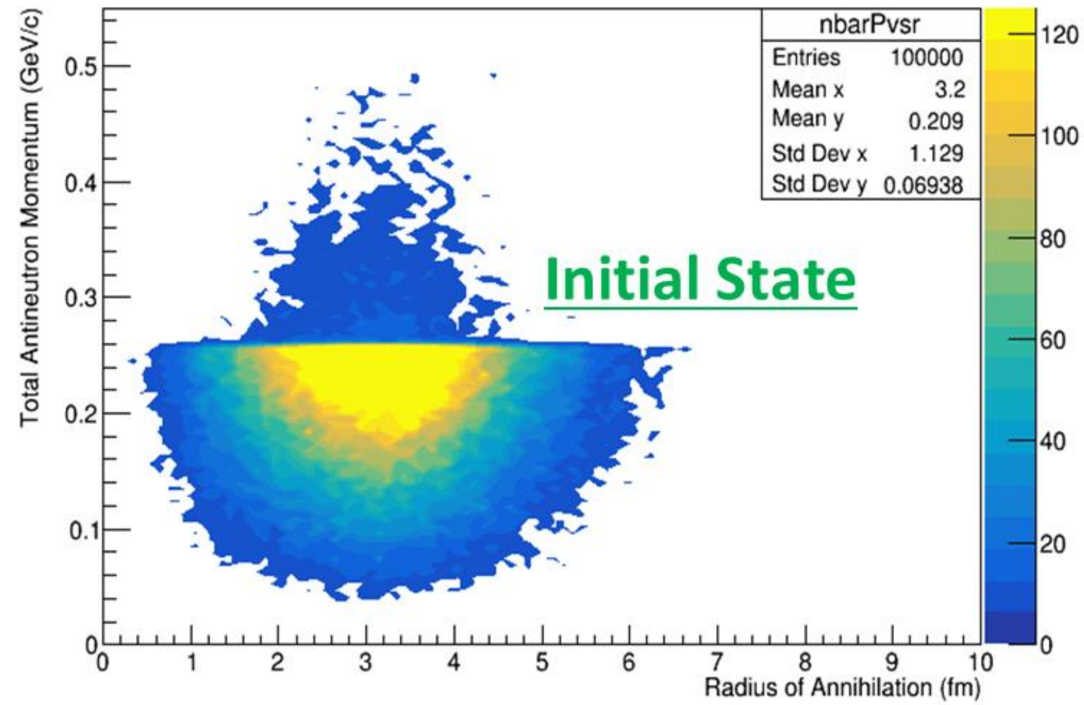
GENIE truth-level—no DUNE reconstruction yet

Some <u>GENIE</u>v3.0.6 Intranuclear $\bar{n}N$ *Initial* States

…Momentum vs. Radius…internal correlations or no?



GENIEv3.0.6 hA/hN2018 and Local Fermi Gas

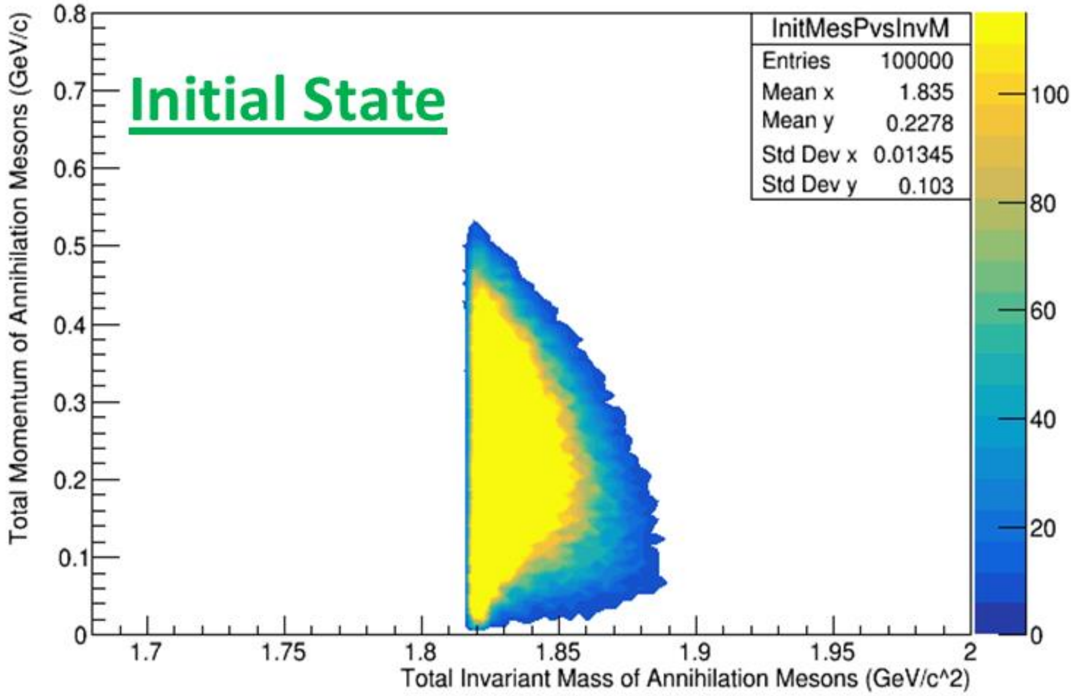


GENIEv3.0.6 hA/hN2018 and Bodek-Ritchie

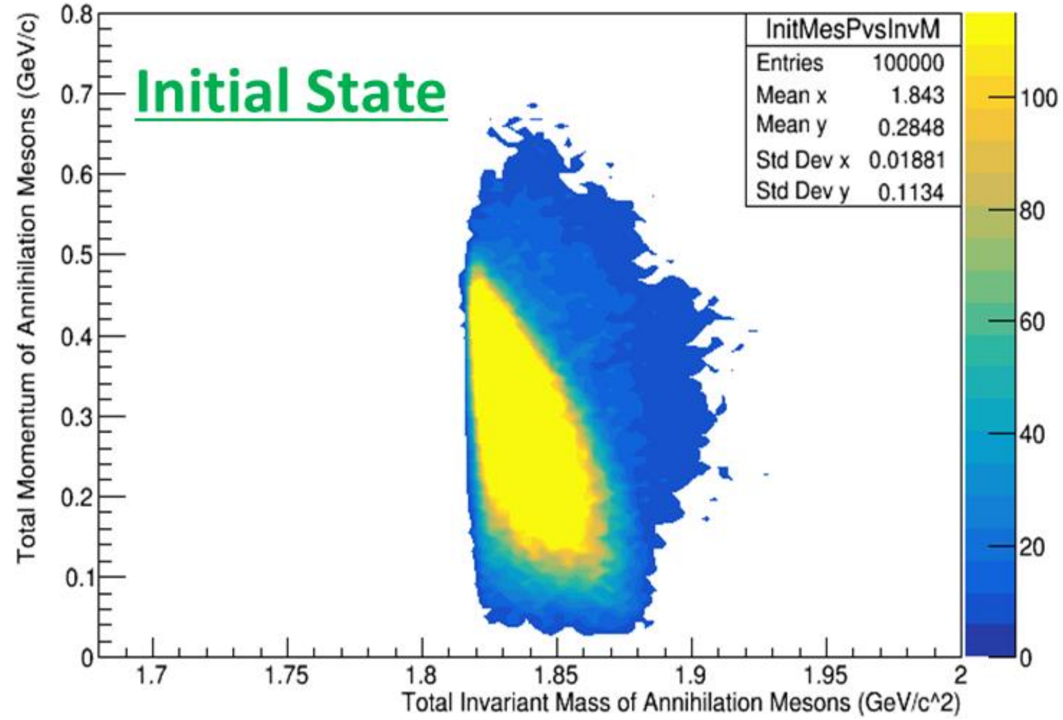
GENIE truth-level—no DUNE reconstruction yet

Some <u>GENIE</u>v3.0.6 Intranuclear $\bar{n}N$ *Initial* States

Mesonic Parameter Space



GENIEv3.0.6 hA/hN2018 and Local Fermi Gas

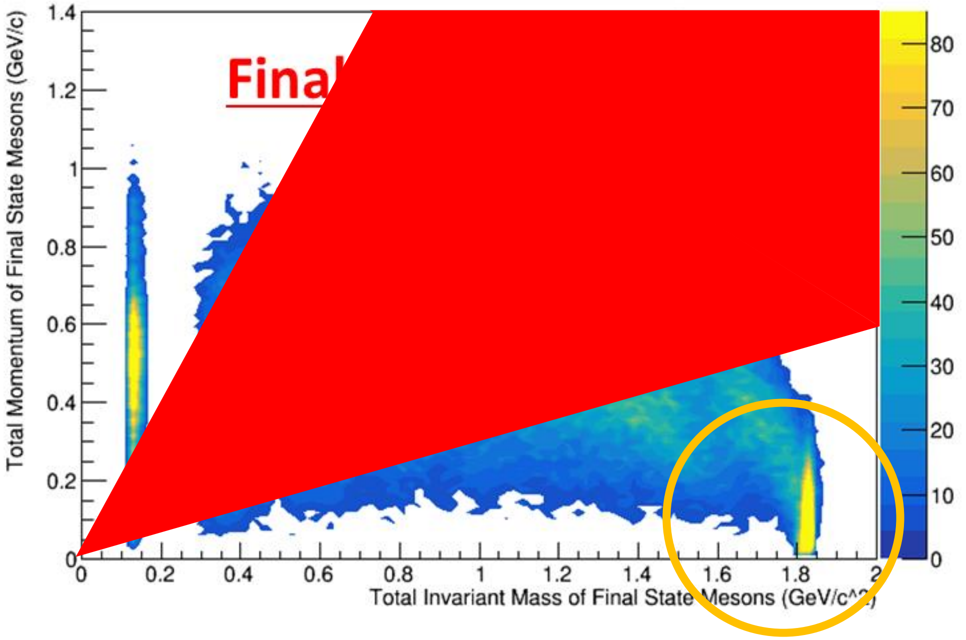


GENIEv3.0.6 hA/hN2018 and Bodek-Ritchie

GENIE truth-level—no DUNE reconstruction yet

Some GENIEv3.0.6 Intranuclear $\bar{n}N$ *Final* States

Pionic parameter space *only*



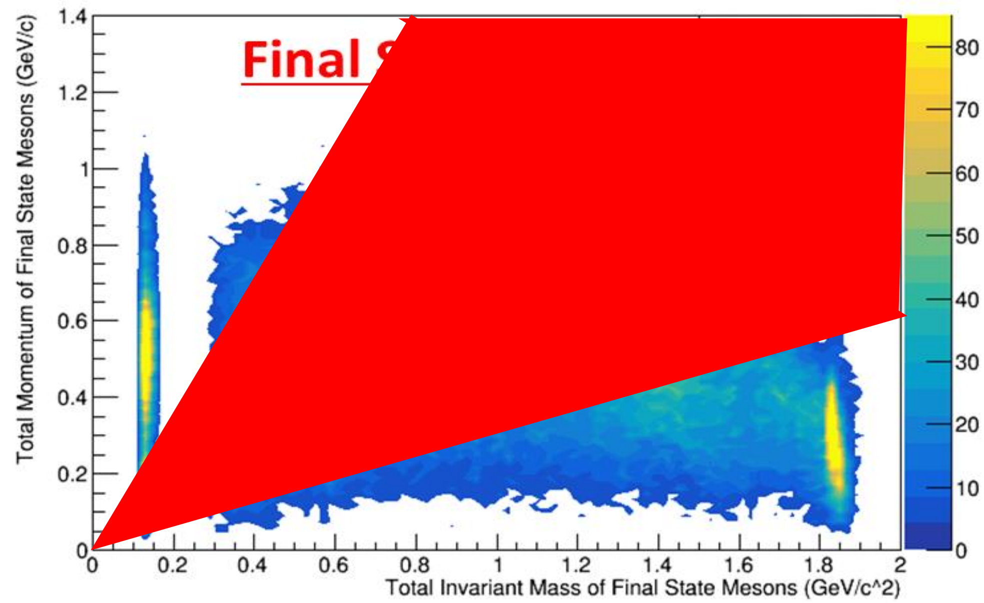
GENIEv3.0.6 hN2018 and Local Fermi Gas

The intranuclear cascade kills most of the original shape of the parameter space

***Approximate* atmospheric ν parameter space is overlayed**

Will greater localization matter?

Effects on FS topologies?



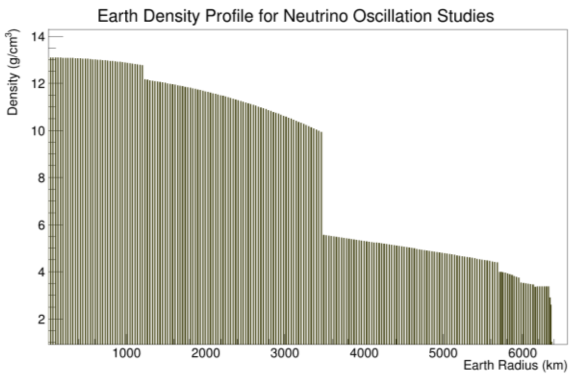
GENIEv3.0.6 hN2018 and Bodek-Ritchie

GENIE truth-level—no DUNE reconstruction yet

Goals for the atmospheric samples

- Generate oscillated Honda atmospheric samples
  - Do so across nuclear model configurations
  - Serve as background for rare event searches
    - Oscillations provide a portal for ν_τ appearance, can give rise to ***CC backgrounds*** to $n \rightarrow \bar{n}$ due to multipionic decays
  - Will use "nominal" oscillation parameters from world data best fits, but can be editable if you like
- $2 \times 6 \times 2{,}000{,}000+$ samples generated, reconstructed
  - Six nuclear model configurations: $\{\mathrm{hN2018}, \mathrm{hA2018}\} \otimes \{\mathrm{BR}, \mathrm{LFG}, \mathrm{ESF}\}$

Process for atmospheric oscillation calculations




```
/*/////////////////////////////////////////////////////////////////////
Originally written by Ivan Martinez-Soler (Northwestern, FNAL)
PLEASE GIVE CREDIT WHERE CREDIT IS DUE!!!

Provided to Joshua Barrow (UTK, FNAL) for use by the DUNE NDK HEP
Working Group for atmospheric background studies using Honda flux
files. This code produces a new atmospheric flux file containing
6 neutrino flavors/type (nu_tau/nu_taubar) of similar structure to
the nominal 4-flavor/type files; it is intended that these files
interface with a new GENIE Honda flux driver to create a standard
set of oscillated atmospheric neutrino background samples. In
general, this code could also be used for more intense atmospheric
studies, but this is left to other future users.

This file requires a build of GNU GSL to run.
/////////////////////////////////////////////////////////////////////*/
```

AtmFlxOsc.cpp
by Ivan Martinez-Soler

- After discussions, was given modified code by Ivan using GSL libraries
  - Calculates oscillation probabilities and total event numbers from given Honda flux file structures for any site in average ν_ℓ/yr from $0.1 - 10{,}000\ \mathrm{GeV}$
    - Assumes the normal hierarchy using Super-Kamiokande fits from NuFIT
    - $\theta_{12} = 33.82°, \theta_{13} = 8.6°, \theta_{23} = 48.6°, \Delta m^2_{21} = 7.39 \times 10^{-5}\ eV^2, \Delta m^2_{31} = 2.528 \times 10^{-3} eV^2, \delta_{CP} = 221°$
    - Includes density changes in the earth's geological makeup (as concentric shells in PREM)
    - Averages over individual angular bins via throws
    - Treats atmosphere as vacuum, ν production height set at $15\ \mathrm{km}$ (this is a parameter)
  - All parameters can be easily changed, and it is all scriptable

average flux in [cosZ = 0.90 -- 1.00, phi_Az = 0 -- 30]

Enu(GeV)	NuMu	NuMubar	NuE	NuEbar (m^2 sec sr GeV)^-1
1.0000E-01	9.8672E+03	1.0020E+04	4.8979E+03	4.5842E+03
1.1220E-01	8.8940E+03	8.9868E+03	4.4193E+03	4.0866E+03
1.2589E-01	7.9442E+03	8.0068E+03	3.9200E+03	3.5941E+03
1.4125E-01	7.0265E+03	7.0725E+03	3.4317E+03	3.1152E+03
1.5849E-01	6.1272E+03	6.1404E+03	2.9803E+03	2.6778E+03
1.7783E-01	5.2653E+03	5.2489E+03	2.5622E+03	2.2779E+03
1.9953E-01	4.4671E+03	4.4417E+03	2.1739E+03	1.9119E+03

hms-ally-20-12-solmax.d

See Honda group site and associated article

g++ -o flux AtmFlxOsc.cpp -lgsl –lgslcblas
./flux hms-ally-20-12-solmax.d > hms-ally-20-12-solmax_3FlavOsc.d

average flux in [cosZ = 0.90 -- 1.00, phi_Az = 0 -- 30]

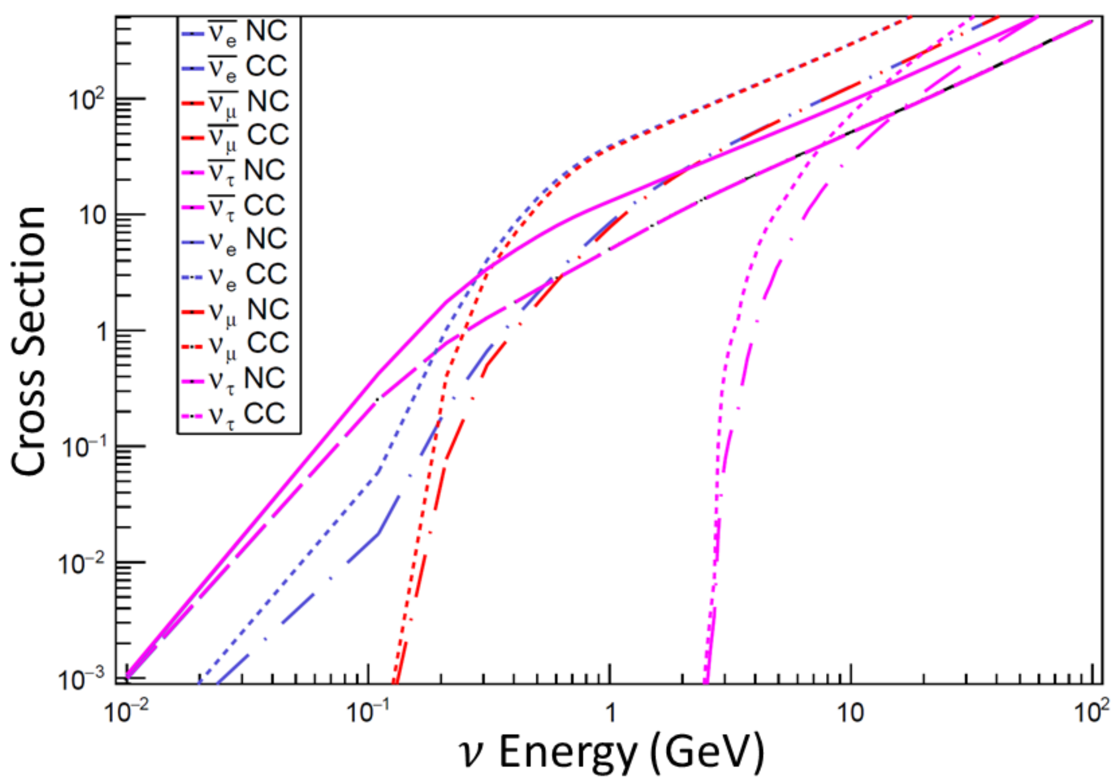
Enu(GeV)	NuMu	NuMubar	NuE	NuEbar	NuTau	NuTaubar (m^2 sec sr GeV)^-1
1.0000e-01	3.2243e+03	2.5527e+03	4.3507e+03	5.5255e+03	7.1901e+03	6.5260e+03
1.1220e-01	4.6411e+03	4.2804e+03	4.0232e+03	4.5436e+03	4.6490e+03	4.2493e+03
1.2589e-01	3.5965e+03	3.8676e+03	4.2707e+03	3.4236e+03	3.9970e+03	4.3097e+03
1.4125e-01	2.3155e+03	2.5628e+03	3.8968e+03	3.0621e+03	4.2459e+03	4.5628e+03
1.5849e-01	5.5881e+03	5.6154e+03	3.0665e+03	2.7240e+03	4.5285e+02	4.7885e+02
1.7783e-01	3.1063e+03	3.1578e+03	2.6391e+03	2.2205e+03	2.0821e+03	2.1485e+03
1.9953e-01	1.6847e+03	1.8388e+03	2.4418e+03	1.8112e+03	2.5145e+03	2.7036e+03

hms-ally-20-12-solmax_3FlavOsc.d

GENIE changes and sample production

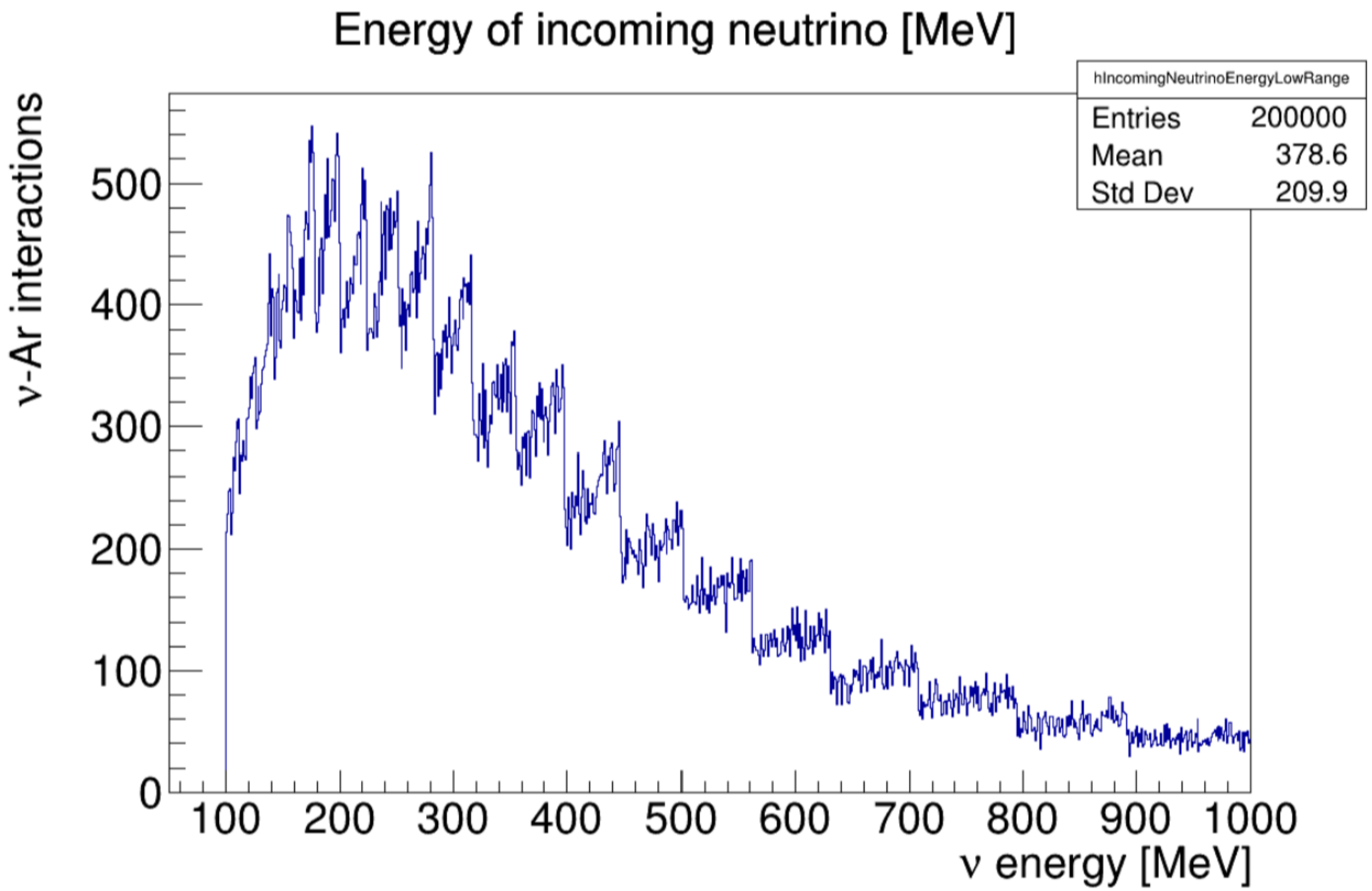
average flux in [cosZ = 0.90 -- 1.00, phi_Az = 0 -- 30]

Enu(GeV)	NuMu	NuMubar	NuE	NuEbar	NuTau	NuTaubar (m^2 sec sr GeV)^-1
1.0000e-01	3.2243e+03	2.5527e+03	4.3507e+03	5.5255e+03	7.1901e+03	6.5260e+03
1.1220e-01	4.6411e+03	4.2804e+03	4.0232e+03	4.5436e+03	4.6490e+03	4.2493e+03
1.2589e-01	3.5965e+03	3.8676e+03	4.2707e+03	3.4236e+03	3.9970e+03	4.3097e+03
1.4125e-01	2.3155e+03	2.5628e+03	3.8968e+03	3.0621e+03	4.2459e+03	4.5628e+03
1.5849e-01	5.5881e+03	5.6154e+03	3.0665e+03	2.7240e+03	4.5285e+02	4.7885e+02
1.7783e-01	3.1063e+03	3.1578e+03	2.6391e+03	2.2205e+03	2.0821e+03	2.1485e+03
1.9953e-01	1.6847e+03	1.8388e+03	2.4418e+03	1.8112e+03	2.5145e+03	2.7036e+03



- Quite simple changes were made to GENIE's src/Tools/Flux/GAtmoFlux.cxx and GHAKKMAtmoFlux.cxx flux drivers to allow for six ν types to be read in from newly calculated oscillated Honda flux files
  - These are available in a personal git
- Six new nuclear model configuration tunes were constructed
  - $\{hN2018, hA2018\} \otimes \{BR, LFG, ESF\}$
  - Tunes available, can be switched easily
  - Based on G18_10a/b
- Splines are generated across all ν types for all of these nuclear model configurations
  - Differences in cross sections are due to various momentum distributions
  - Cross sections available

Known Issue: Flux-Cross Section Convolution

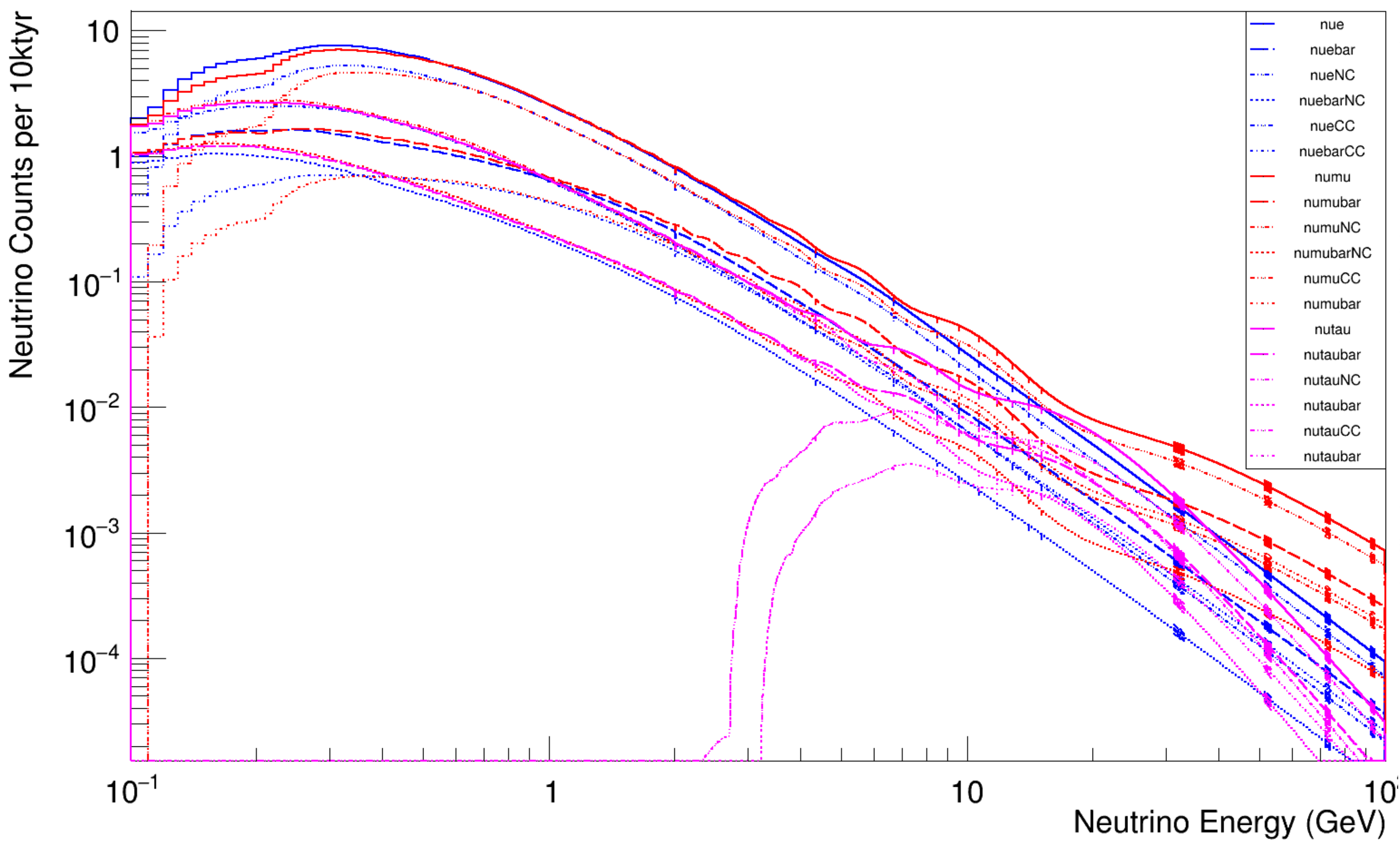


- Previously unknown behavior
  - C. Marshall and I discovered this almost simultaneously and independently
- Results from convolution of logarithmically spaced bins of energy and an approximately linearly increasing cross section
- What to do?
  - Interpolate, find actual spectrum, and then cull/reweight events appropriately

Spectral Interpolation

Expected count rates via integration

$$\log \Phi_\alpha(E) = \log \Phi_\alpha(E_i) + \frac{d \log \Phi_\alpha}{d \log E} \cdot (E - E_i) \longrightarrow N \sim \sum_{f=1}^{6} N_f \sim 40\,\text{kt of Argon} \times N_{\text{Avogadro}} \times \sum_{f=1}^{6} \sum_{E_i}^{E_m} \sigma\left(E_{i,f}\right) \cdot P_{\nu_f \to \nu_{f'}}(L, E) \cdot \Phi_{\text{Honda Interp.}}\left(E_{i,f}\right)$$



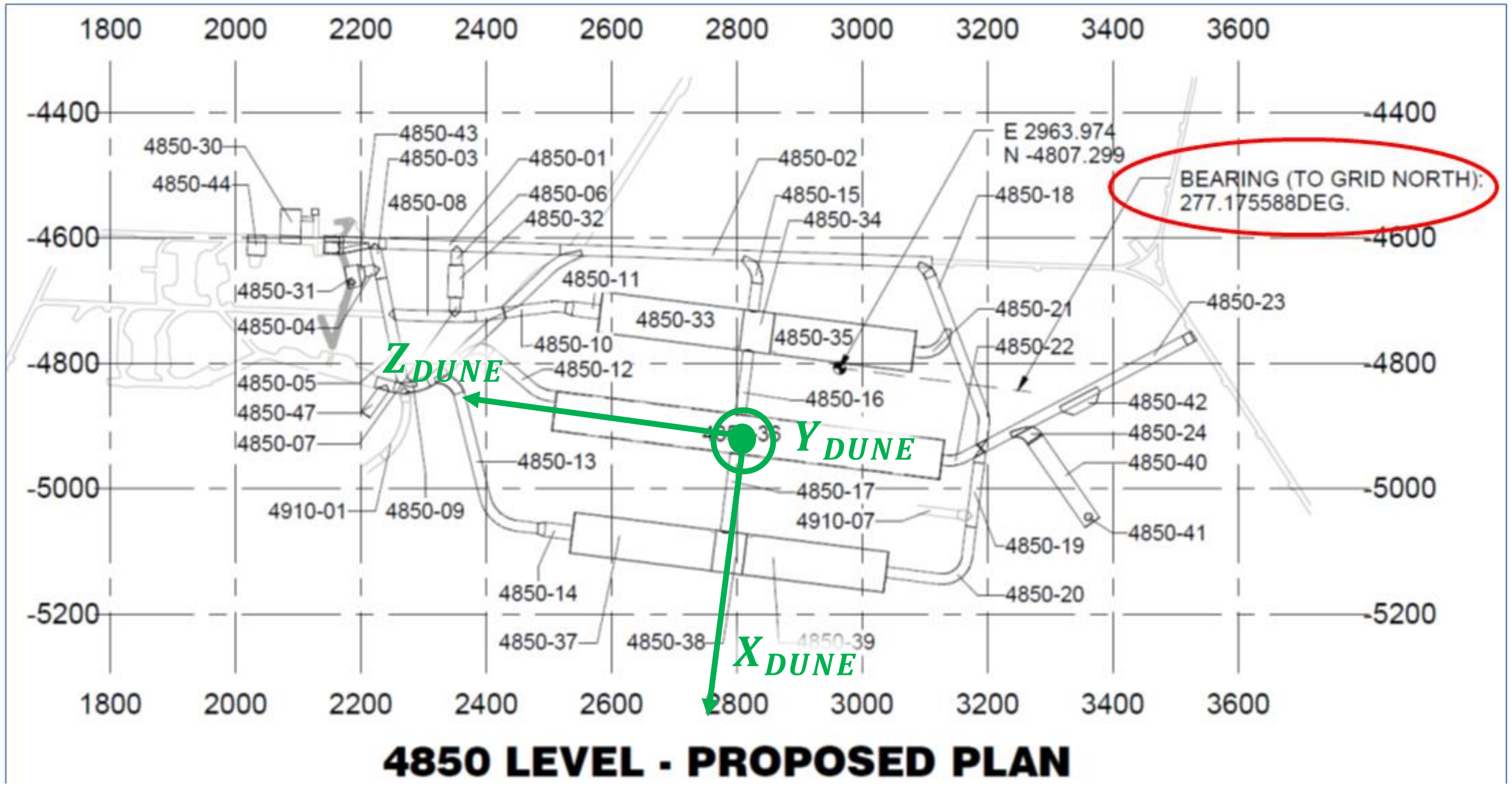
Bodek-Ritchie Nuclear Model	Total	CC	NC
ν_e	746.76	538.93	207.83
$\bar{\nu}_e$	188.833	113.893	74.9393
ν_μ	756.522	527.679	228.842
$\bar{\nu}_\mu$	216.493	126.104	90.3886
ν_τ	234.757	14.1756	220.581
$\bar{\nu}_\tau$	92.5186	5.14774	87.3709
Totals per $10\,\textbf{kt} \cdot \textbf{yr}$	2235.88	1325.93	909.953

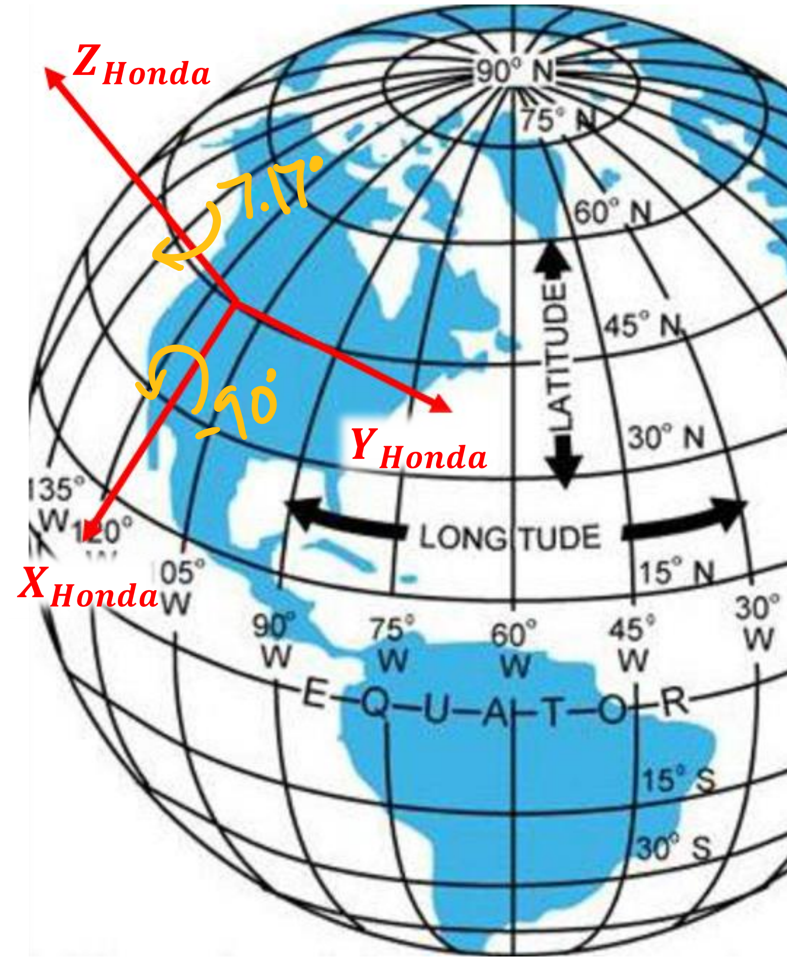
Local Fermi Gas Nuclear Model	Total	CC	NC
ν_e	782.185	578.756	203.43
$\bar{\nu}_e$	197.509	124.602	72.9071
ν_μ	788.383	564.348	224.036
$\bar{\nu}_\mu$	224.443	136.450	87.9928
ν_τ	230.162	14.2033	215.959
$\bar{\nu}_\tau$	90.2157	5.16265	85.053
Totals per $10\,\textbf{kt} \cdot \textbf{yr}$	2312.90	1423.52	889.377

Effective Spectral Function Nuclear Model	Total	CC	NC
ν_e	711.738	503.908	207.830
$\bar{\nu}_e$	182.029	107.089	74.9393
ν_μ	725.050	496.208	228.842
$\bar{\nu}_\mu$	209.751	119.362	90.3886
ν_τ	234.699	14.1175	220.581
$\bar{\nu}_\tau$	92.5227	5.15187	87.3709
Totals per $10\,\textbf{kt} \cdot \textbf{yr}$	2155.79	1245.84	909.953

Known Issue in GENIE 3.0.6: Detector Coordinates

ARUP drawings "Long Baseline Neutrino Facility – Far Site Conventional Facilities – Excavation Design", page 6 – Nov-2015

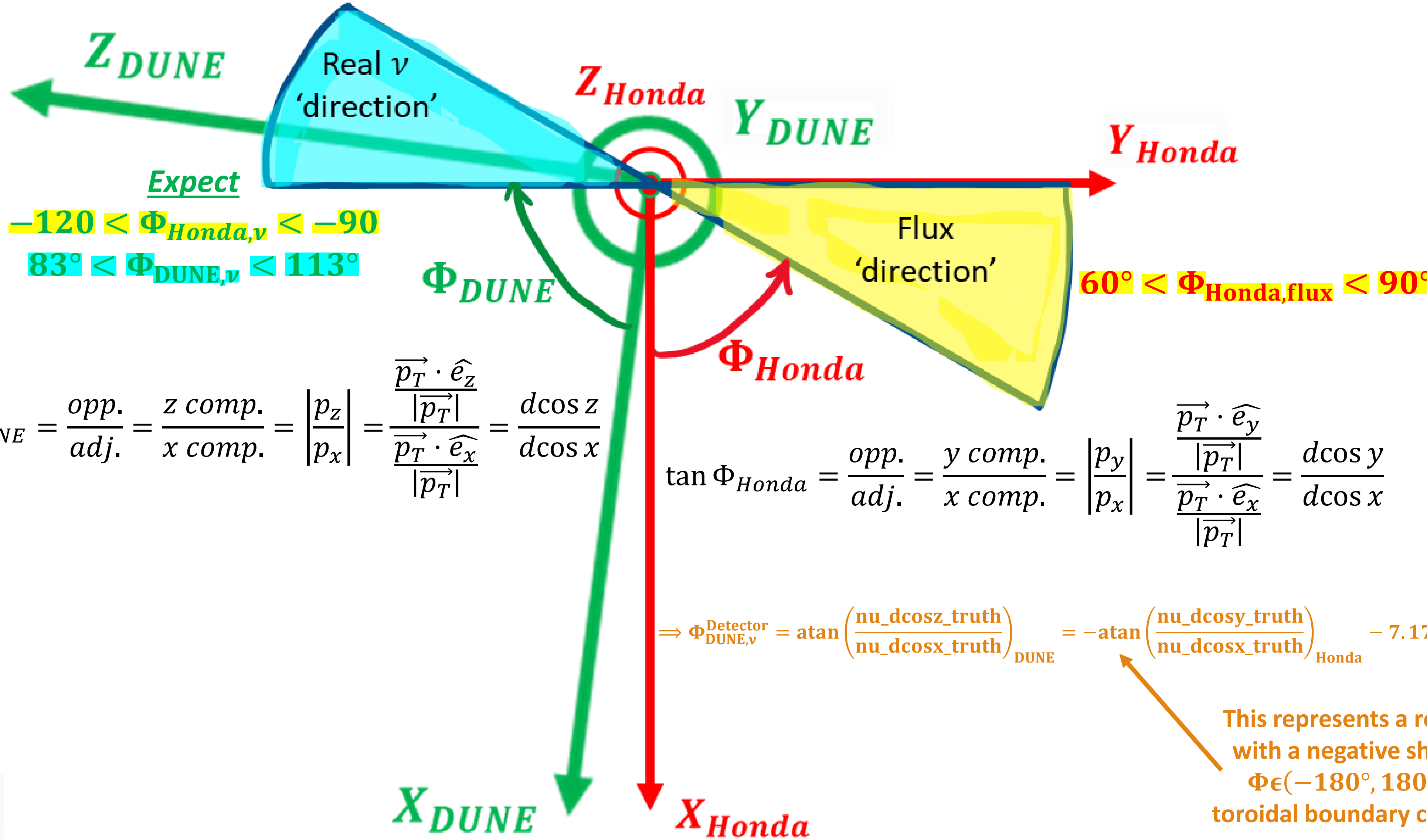




Topocentric Horizontal Coordinates vs. DUNE Coordinates

Rotations of right-handed coordinate systems must be performed within GENIE generation steps using two Euler angles: $-\mathrm{R}\ 0.125237636, -1.57079633, 0.0 \sim -\mathrm{R}\ 7.17°, -90°, 0°$

Known Issue in GENIE 3.0.6: Detector Coordinates



$$\tan\Phi_{DUNE} = \frac{opp.}{adj.} = \frac{z\ comp.}{x\ comp.} = \left|\frac{p_z}{p_x}\right| = \frac{\frac{\overrightarrow{p_T}\cdot\widehat{e_z}}{|\overrightarrow{p_T}|}}{\frac{\overrightarrow{p_T}\cdot\widehat{e_x}}{|\overrightarrow{p_T}|}} = \frac{\mathrm{dcos}\,z}{\mathrm{dcos}\,x}$$

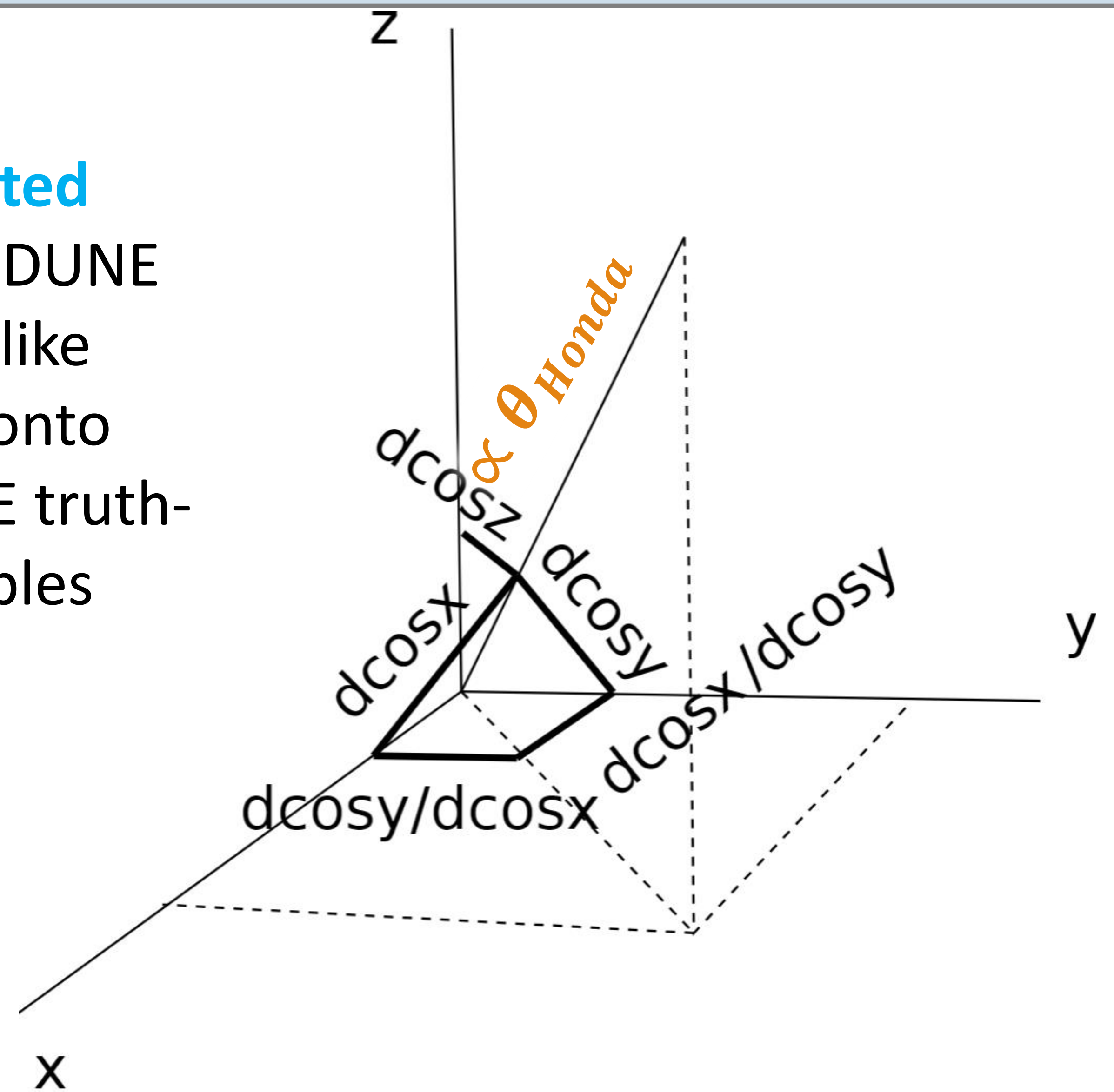
$$\tan\Phi_{Honda} = \frac{opp.}{adj.} = \frac{y\ comp.}{x\ comp.} = \left|\frac{p_y}{p_x}\right| = \frac{\frac{\overrightarrow{p_T}\cdot\widehat{e_y}}{|\overrightarrow{p_T}|}}{\frac{\overrightarrow{p_T}\cdot\widehat{e_x}}{|\overrightarrow{p_T}|}} = \frac{\mathrm{dcos}\,y}{\mathrm{dcos}\,x}$$

$$\Rightarrow \Phi^{\mathrm{Detector}}_{\mathrm{DUNE},\nu} = \mathrm{atan}\left(\frac{\mathrm{nu\_dcosz\_truth}}{\mathrm{nu\_dcosx\_truth}}\right)_{\mathrm{DUNE}} = -\mathrm{atan}\left(\frac{\mathrm{nu\_dcosy\_truth}}{\mathrm{nu\_dcosx\_truth}}\right)_{\mathrm{Honda}} - 7.17° = \Phi^{\mathrm{Detector}}_{\mathrm{Honda},\nu}$$

This represents a reflection with a negative shift with $\Phi\epsilon(-180°, 180°)$ and toroidal boundary conditions

Known Issue in GENIE 3.0.6: Detector Coordinates

Non-rotated mapping of DUNE detector-like variables onto Honda/GENIE truth-like variables



Further Refinements Required

- Allow for production height input (known to Honda)
- Allow reweighting of events to reshape spectra appropriately
  - Or fix GENIE directly (not sure how to do this given input flux's structure)
  - Need to add new weight tree branch/leaf to current ntuples
    - Normalize everything to NuFit 4.1 parameters
- Run same calculation for solar minimum (max. ν counts)
- Use new oscillation software?
- Or use current software method and make scriptable?
  - Each oscillation point takes ~3-4 minutes to run *and interpolate* using GSL
  - Modify for the grid, get thousands of points? Develop ntuple structure?

Previous Slides

Systematics for Rare Processes and Their Backgrounds

November 18th, 2022

DUNE HEP Working Group Meeting

by J. L. Barrow

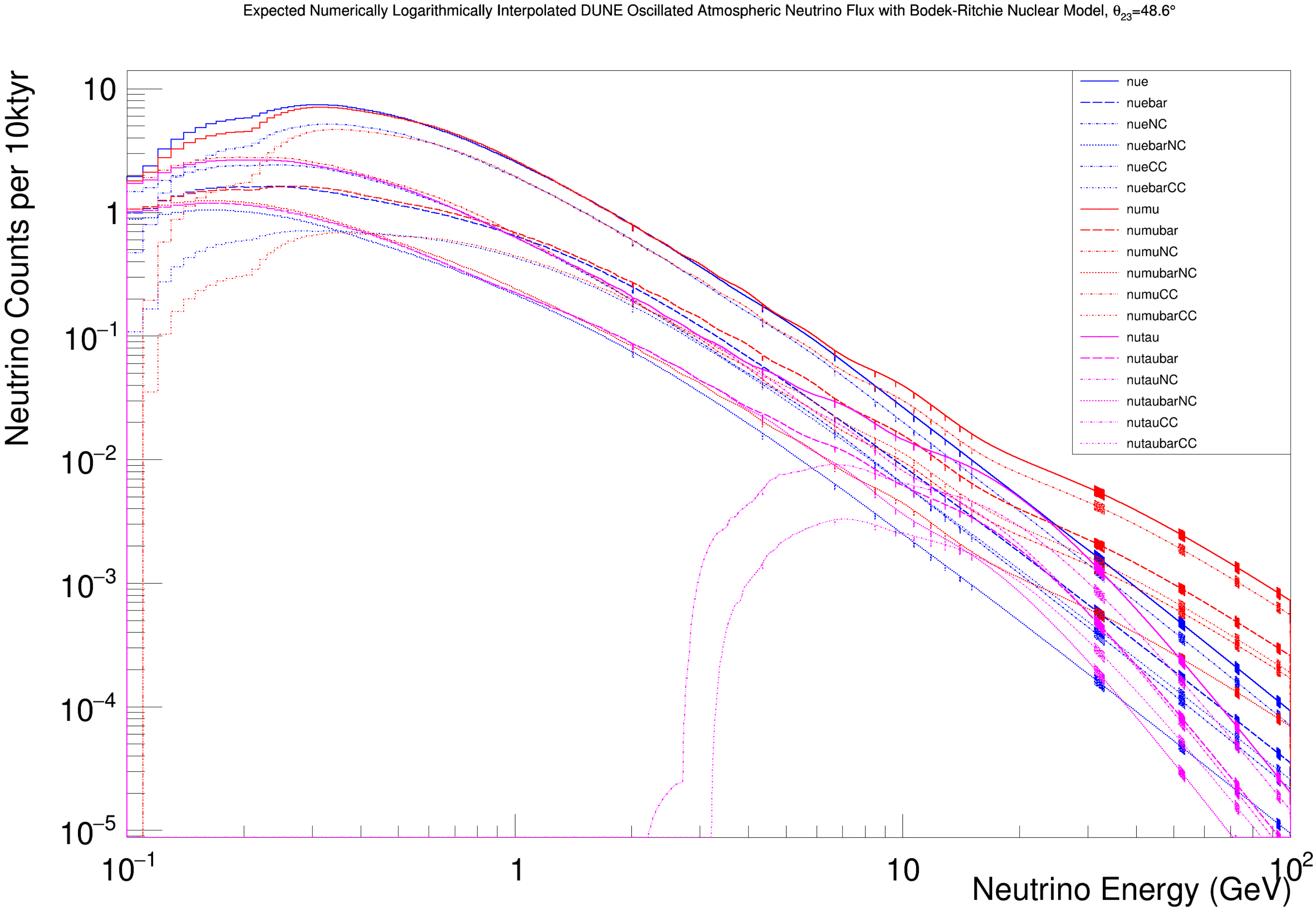
Zuckerman Postdoctoral Scholar

MIT & TAU

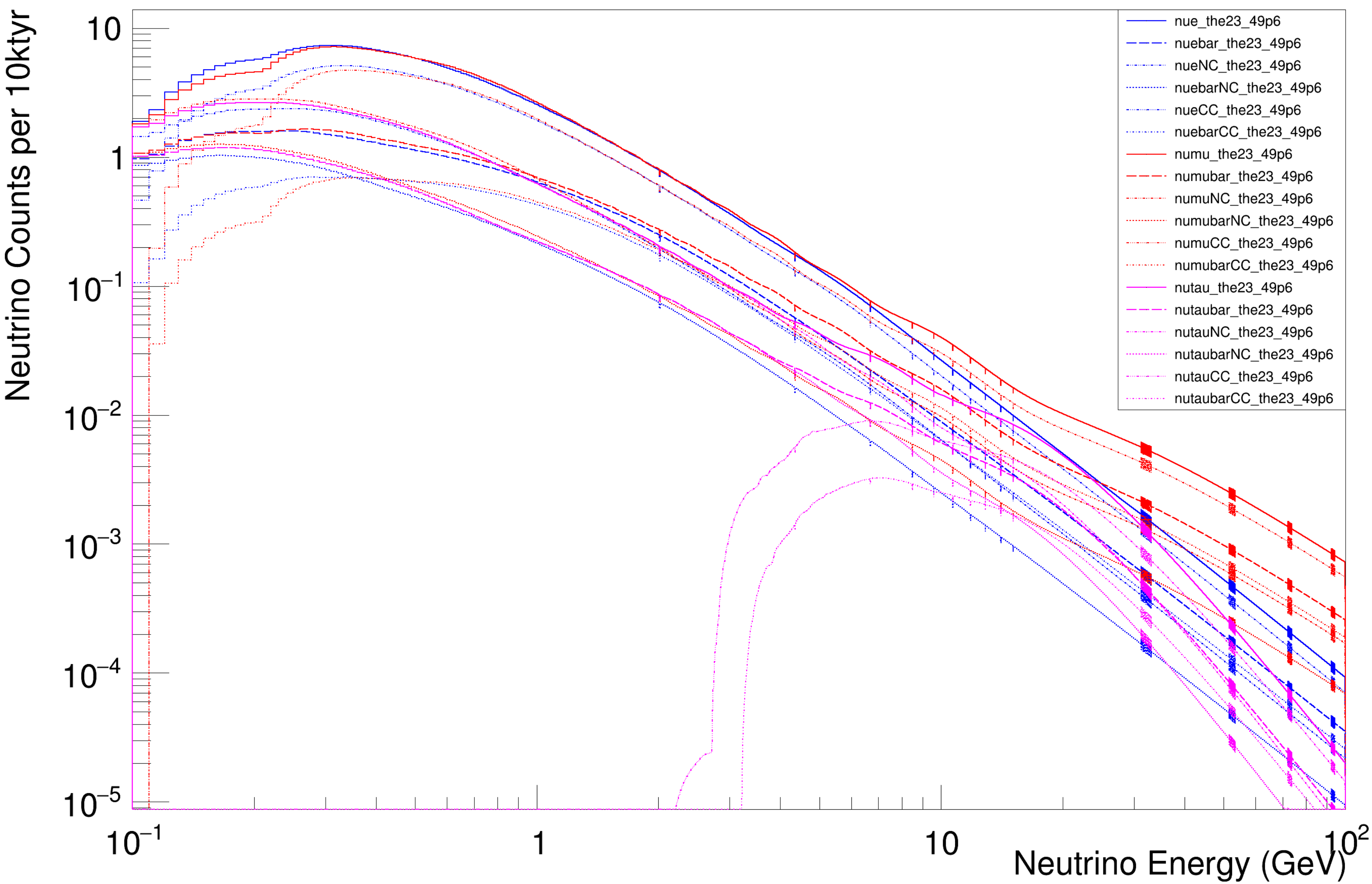
Atmospheric ν Background Systematics

Varying Oscillation Parameters

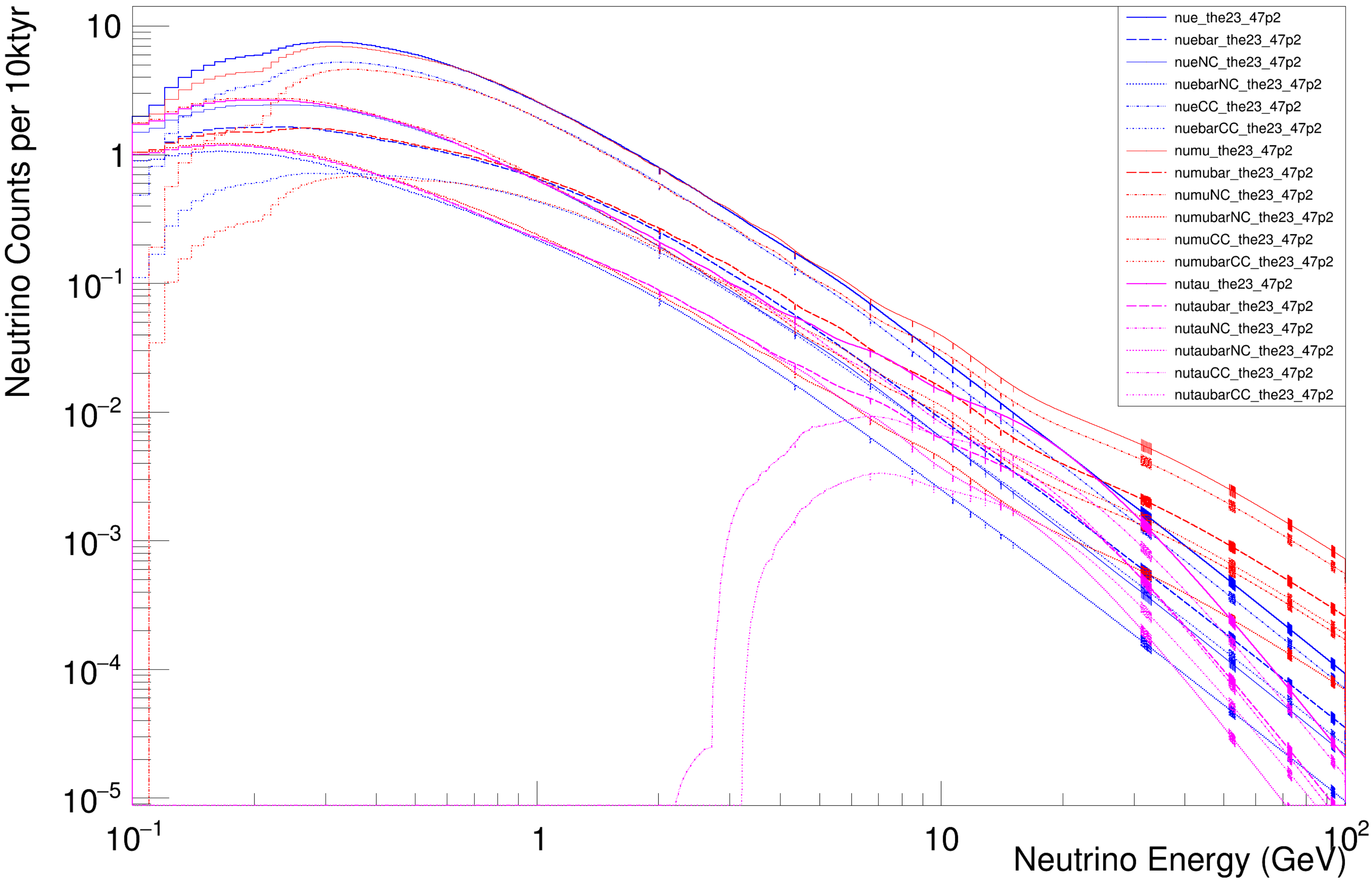
- Utilized NuFit4.1 parameters
- Largest uncertainty expected to be caused by θ_{23}
  - Ran these ranges in an uncorrelated way to see approximate behavior
  - Does not account for particular degeneracies that are possible


Expected Numerically Logarithmically Interpolated DUNE Oscillated Atmospheric Neutrino Flux with Bodek-Ritchie Nuclear Model, θ23=48.6°
Neutrino Counts per 10ktyr
10
1
10^-1
10^-2
10^-3
10^-4
10^-5
10^-1
1
10
10^2
Neutrino Energy (GeV)
nue
nuebar
nueNC
nuebarNC
nueCC
nuebarCC
numu
numubar
numuNC
numubarNC
numuCC
numubarCC
nutau
nutaubar
nutauNC
nutaubarNC
nutauCC
nutaubarCC

Expected Numerically Logarithmically Interpolated DUNE Oscillated Atmospheric Neutrino Flux with Bodek-Ritchie Nuclear Model, θ_{23}=49.6°



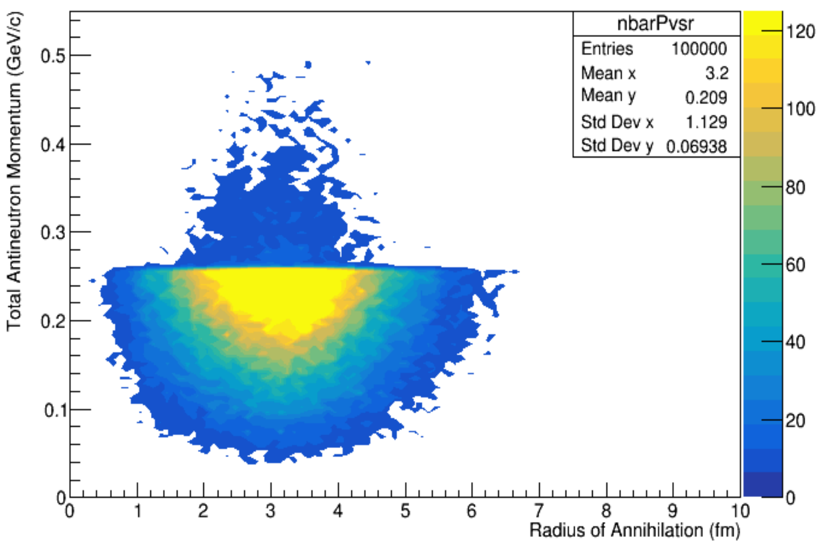
Expected Numerically Logarithmically Interpolated DUNE Oscillated Atmospheric Neutrino Flux with Bodek-Ritchie Nuclear Model, θ_{23}=47.2°

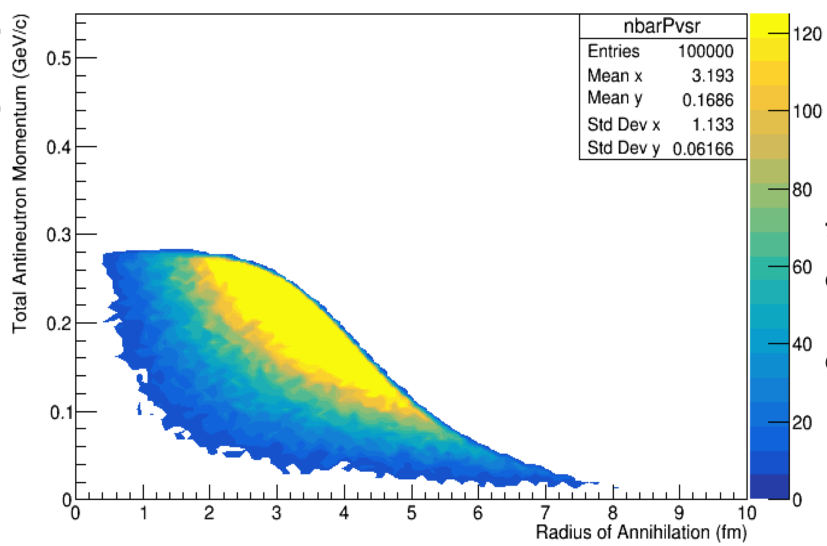


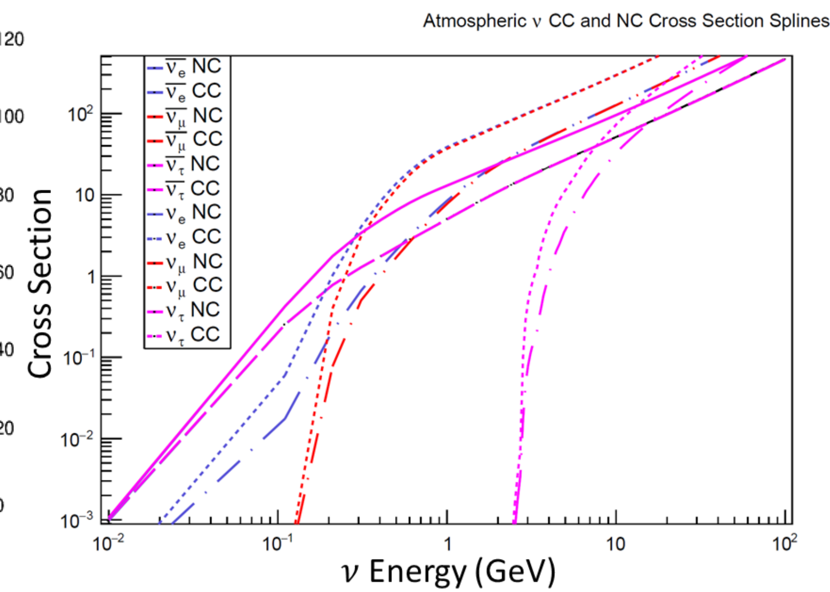
Varying Nuclear Model (Cross Section)

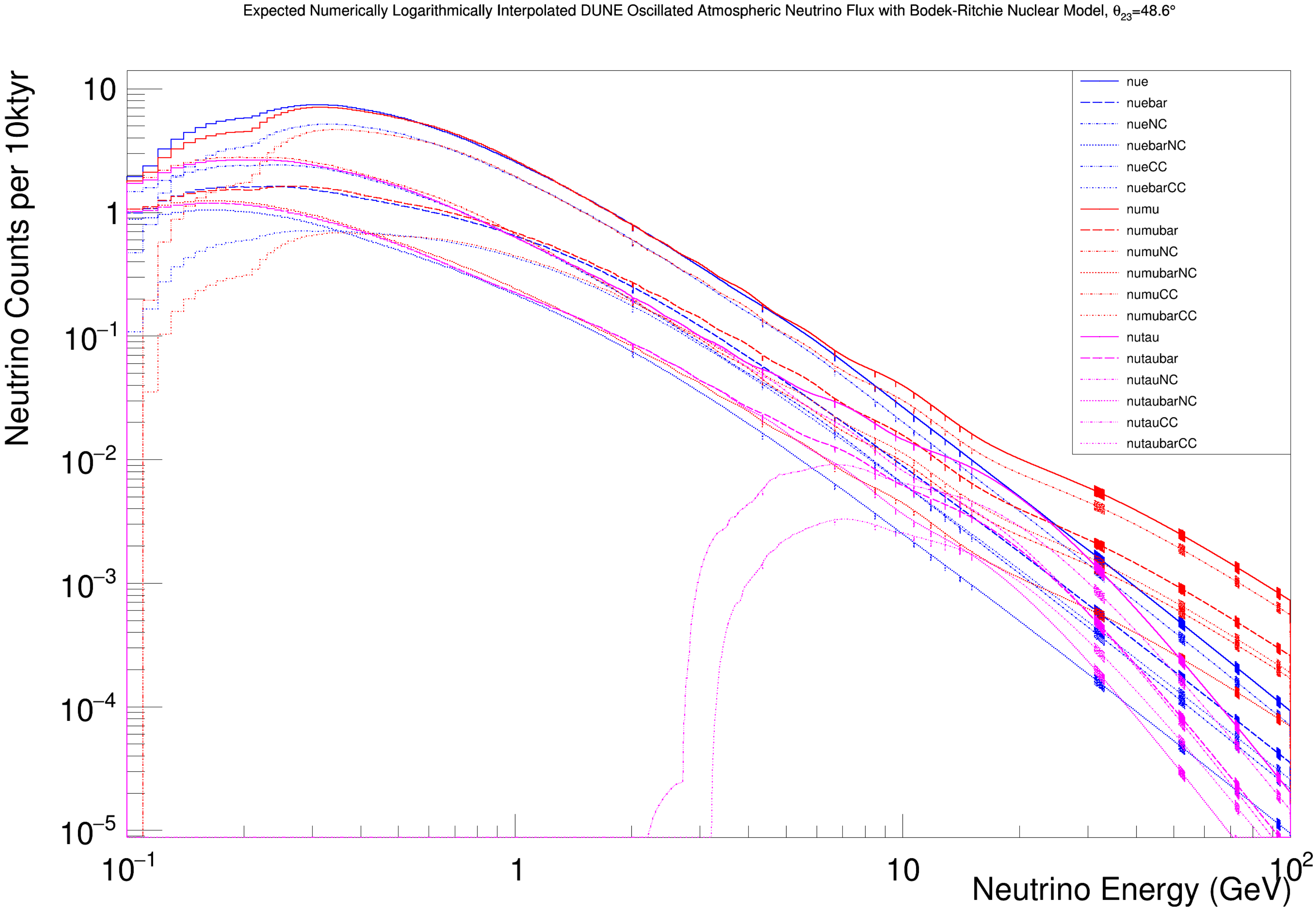
- Nuclear model of Fermi motion enters cross section calculation
  - Factorized generator model doesn't change cross section despite FSI effects

- Three were available at the time (more now)
  - **Bodek-Ritchie (BR)** nonlocal relativistic Fermi gas with pheno. SRC tail
  - **Local Fermi gas (LFG)** with nonrelativistic physics
  - Effective spectral function (ESF) with nonlocal, nonrelativistic physics

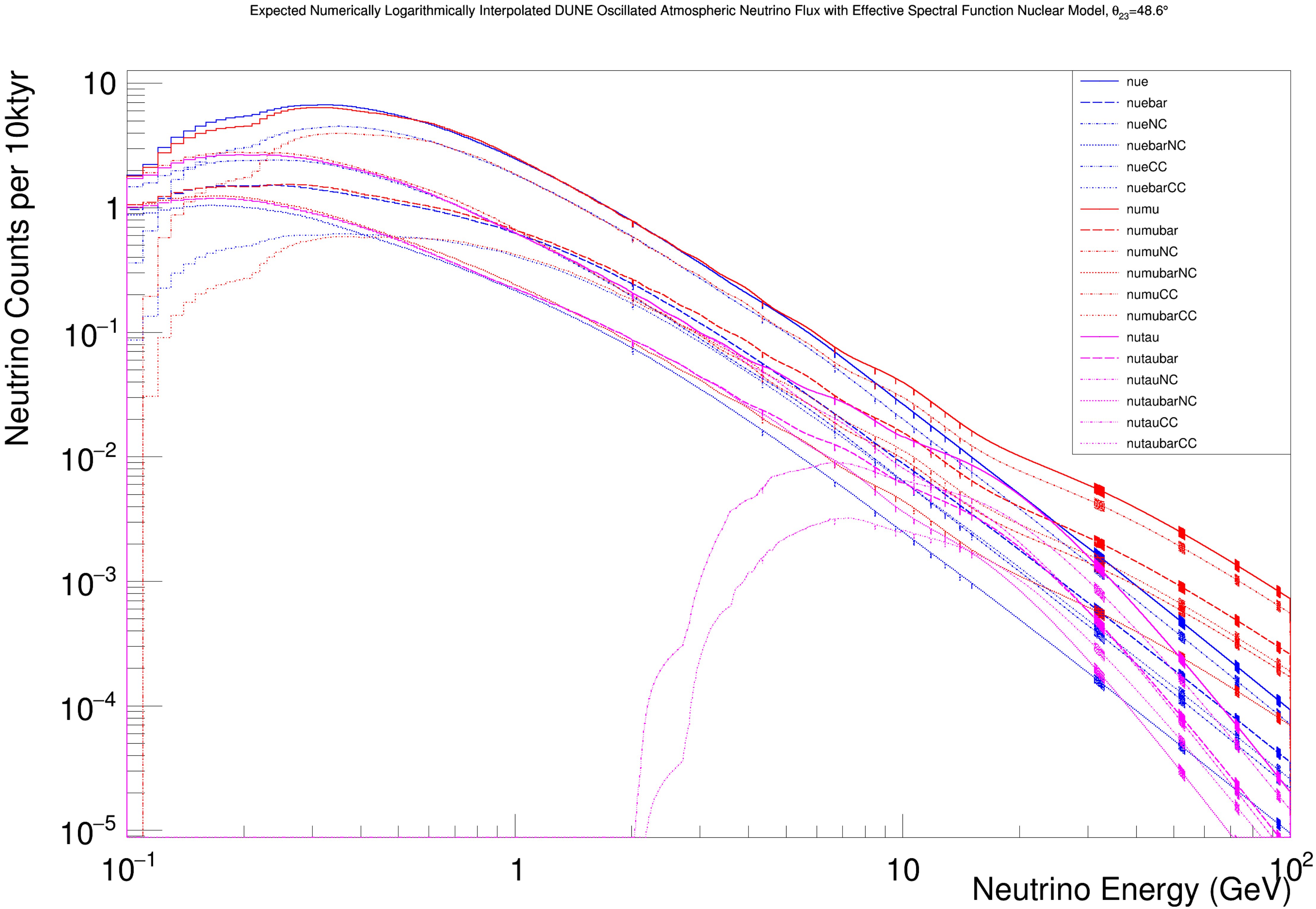




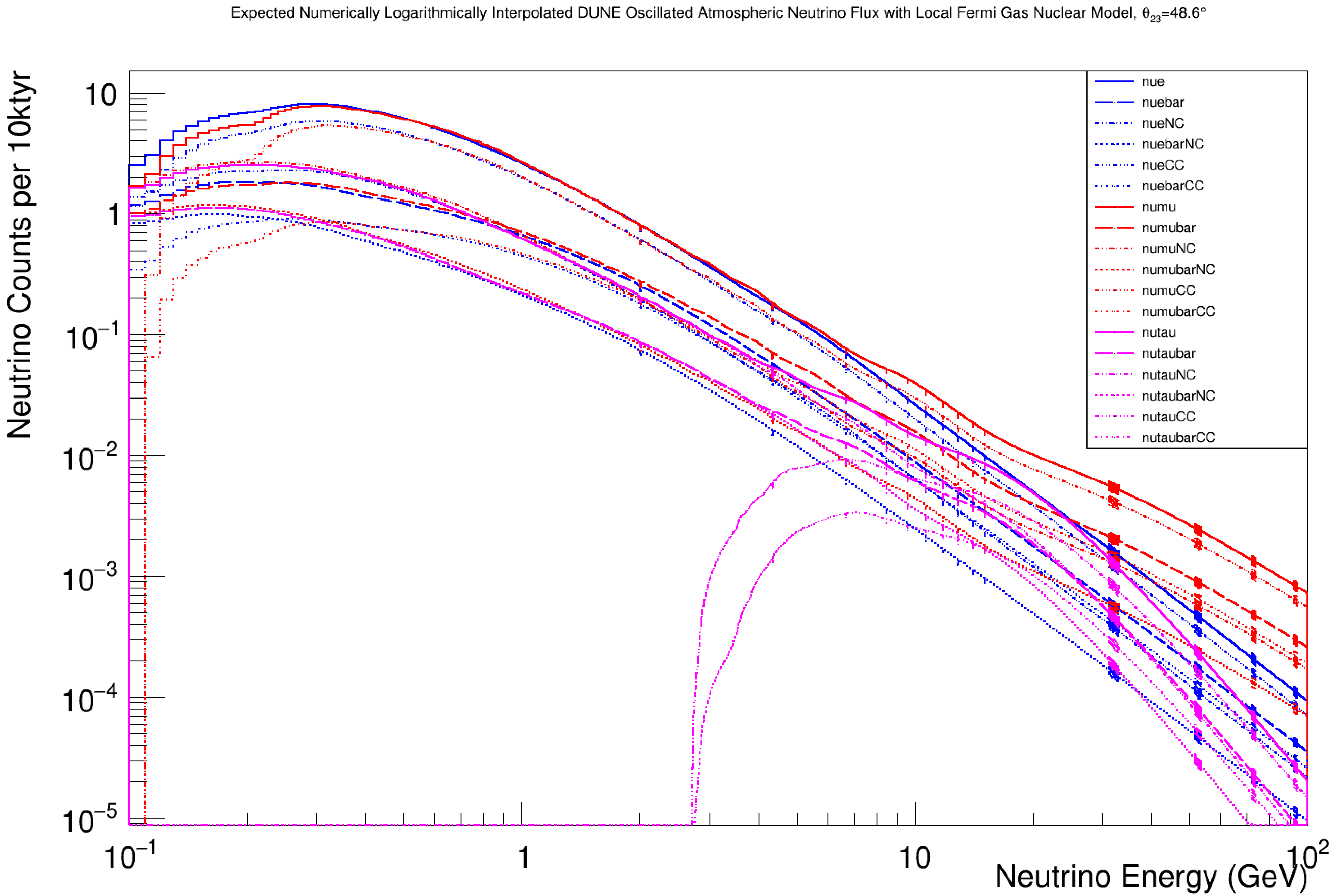




Expected Numerically Logarithmically Interpolated DUNE Oscillated Atmospheric Neutrino Flux with Bodek-Ritchie Nuclear Model, θ23=48.6°
Neutrino Counts per 10ktyr
Neutrino Energy (GeV)
10
1
10^-1
10^-2
10^-3
10^-4
10^-5
10^-1
1
10
10^2
nue
nuebar
nueNC
nuebarNC
nueCC
nuebarCC
numu
numubar
numuNC
numubarNC
numuCC
numubarCC
nutau
nutaubar
nutauNC
nutaubarNC
nutauCC
nutaubarCC



Expected Numerically Logarithmically Interpolated DUNE Oscillated Atmospheric Neutrino Flux with Effective Spectral Function Nuclear Model, θ_{23}=48.6°
Neutrino Counts per 10ktyr
Neutrino Energy (GeV)
10
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-1}
1
10
10^{2}
nue
nuebar
nueNC
nuebarNC
nueCC
nuebarCC
numu
numubar
numuNC
numubarNC
numuCC
numubarCC
nutau
nutaubar
nutauNC
nutaubarNC
nutauCC
nutaubarCC


Expected Numerically Logarithmically Interpolated DUNE Oscillated Atmospheric Neutrino Flux with Local Fermi Gas Nuclear Model, θ23=48.6°
Neutrino Counts per 10ktyr
Neutrino Energy (GeV)
nue
nuebar
nueNC
nuebarNC
nueCC
nuebarCC
numu
numubar
numuNC
numubarNC
numuCC
numubarCC
nutau
nutaubar
nutauNC
nutaubarNC
nutauCC
nutaubarCC

How to handle model variations?

- Not all nuclear model configurations are reweightable
  - Theoretically motivated FSI's are stochastic (hN Intranuke)
- Only play with 3/6 versions?
  - Even here, LFG is likely "best"

Bodek-Ritchie Nuclear Model	Total	CC	NC
ν_e	746.76	538.93	207.83
$\bar{\nu}_e$	188.833	113.893	74.9393
ν_μ	756.522	527.679	228.842
$\bar{\nu}_\mu$	216.493	126.104	90.3886
ν_τ	234.757	14.1756	220.581
$\bar{\nu}_\tau$	92.5186	5.14774	87.3709
Totals per $10\,\mathbf{kt}\cdot\mathbf{yr}$	2235.88	1325.93	909.953

Local Fermi Gas Nuclear Model	Total	CC	NC
ν_e	782.185	578.756	203.43
$\bar{\nu}_e$	197.509	124.602	72.9071
ν_μ	788.383	564.348	224.036
$\bar{\nu}_\mu$	224.443	136.450	87.9928
ν_τ	230.162	14.2033	215.959
$\bar{\nu}_\tau$	90.2157	5.16265	85.053
Totals per $10\,\mathbf{kt}\cdot\mathbf{yr}$	2312.90	1423.52	889.377

Effective Spectral Function Nuclear Model	Total	CC	NC
ν_e	711.738	503.908	207.830
$\bar{\nu}_e$	182.029	107.089	74.9393
ν_μ	725.050	496.208	228.842
$\bar{\nu}_\mu$	209.751	119.362	90.3886
ν_τ	234.699	14.1175	220.581
$\bar{\nu}_\tau$	92.5227	5.15187	87.3709
Totals per $10\,\mathbf{kt}\cdot\mathbf{yr}$	2155.79	1245.84	909.953

Missing Systematics?

- Beyond θ_{23} and other smaller oscillation effects, we have…

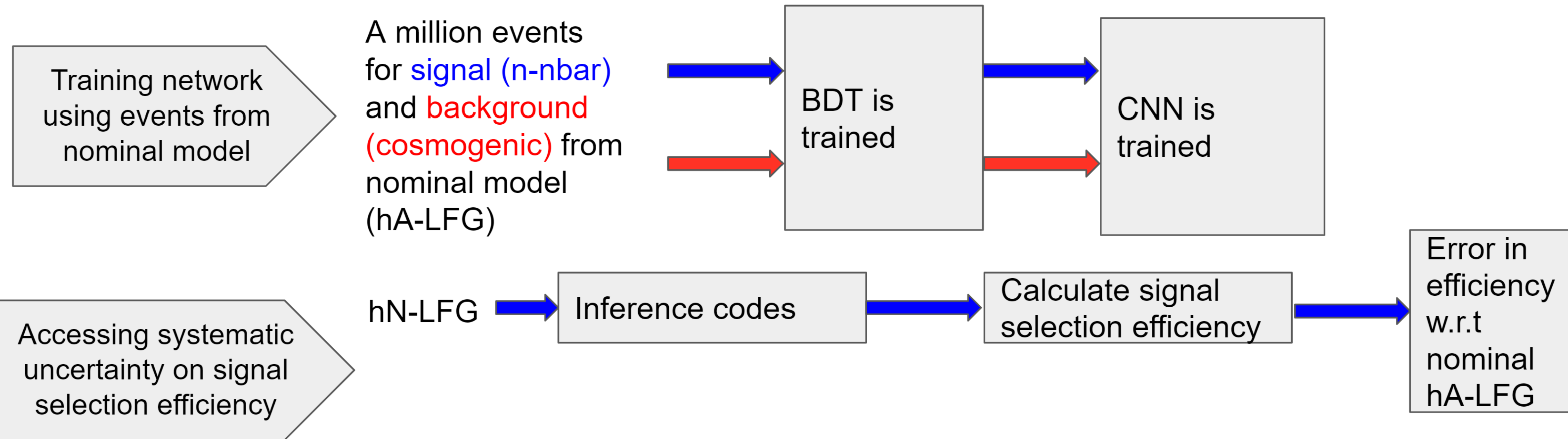
- Must study $E^{-\gamma}$ spectral dependence $\gamma \neq 3$ on expected counts
  - Should probably source directly from Honda?
- Confidence in logarithmic interpolation scheme?
- Must vary between solar minimum and maximum (over a cycle?)
- Small changes to PREM model densities are allowed
- Iterations on FSI parameters? Make things softer/harder?
- Effects of reconstruction via slight misalignments of the detector?

- How to study all these in a correlated fashion?
- How does ML factor into all of this???

$n \rightarrow \bar{n}$ Systematics

What is being done now in MicroBooNE?

❖ Uncertainty due to GENIE modeling is accessed by producing different samples from different models (hN-LFG, hA-BR and hN-BR) and passed them through the analysis framework (BDT and CNN) to find the uncertainty w.r.t nominal model (hA-LFG)



Same procedure for other models (hN-BR, hA-BR)

What is being done now in MicroBooNE?

Systematic uncertainties on signal selection efficiency	
GENIE - hN-LFG	**1.14%**
GENIE - hA-BR	**1.17%**
GENIE - hN-BR	**4.57%**
Total GENIE (adding in quadrature)	**4.85%**
Detector	**6.73%**
GEANT4	**2.32%**
Total systematic uncertainty on signal	**8.61%**

Similar procedure as followed for GENIE (→ Detector)

Used re-weighting scheme (→ GEANT4)

Can we ignore the ML-aspects of this selection on our systematic effects?

A Grand Scheme?

or

A Terrifying Prospect?

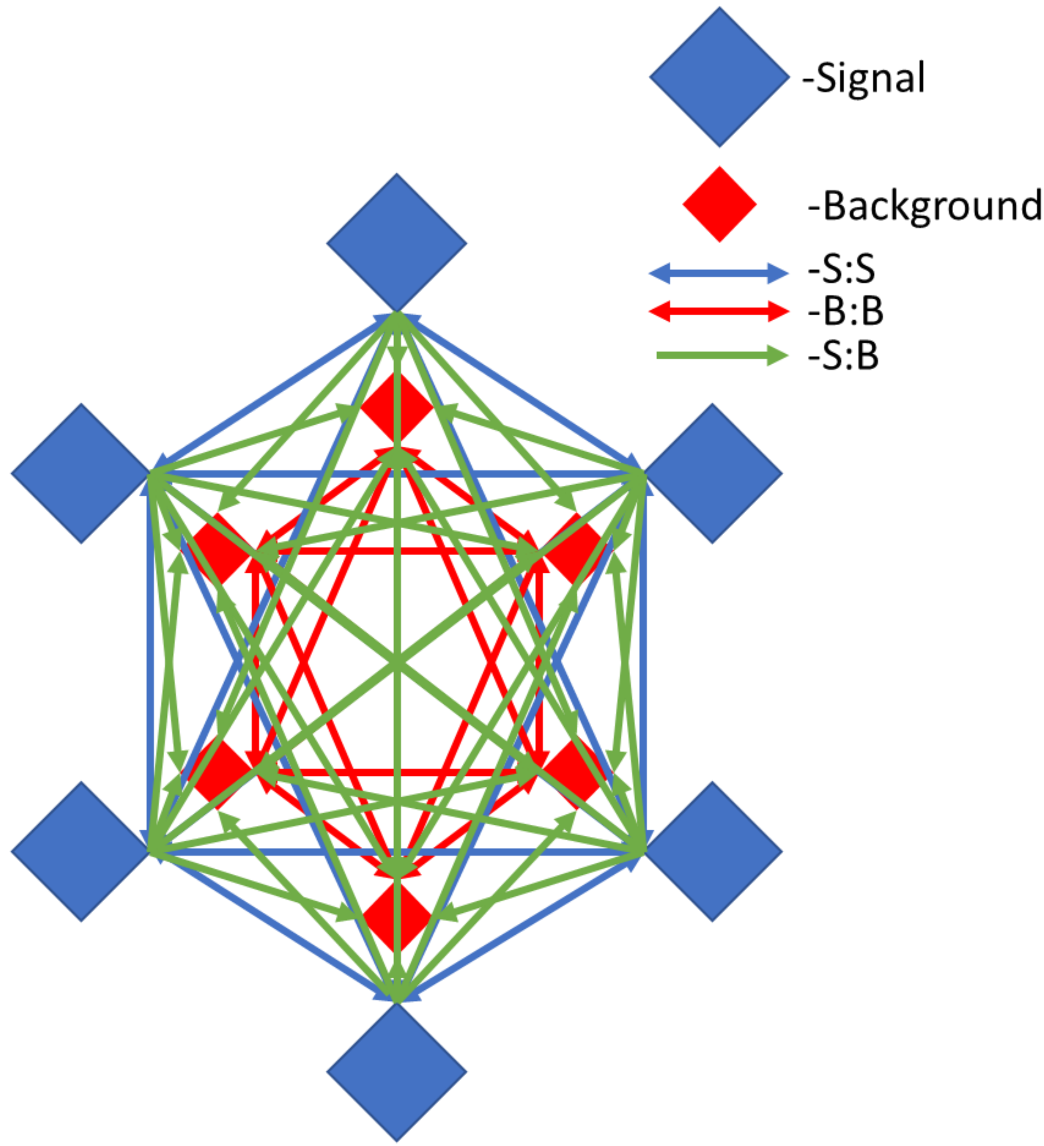
Model Configuration Flows for Signal and Background Sample Comparisons to Better Determine Model Systematics

The best way to understand modeling systematics of an unknown process is to comparatively iterate

S:S	hA_BR	hA_LFG	hA_E SF	hN_BR	hN_LFG	hN_ESF
hA_BR		Kinematic Distributions (BDT inputs)	⋯			
hA_LFG	Kinematic Distributions (BDT inputs)		⋱			
hA_ESF	⋮	⋱				
hN_BR						
hN_LFG						
hN_ESF						

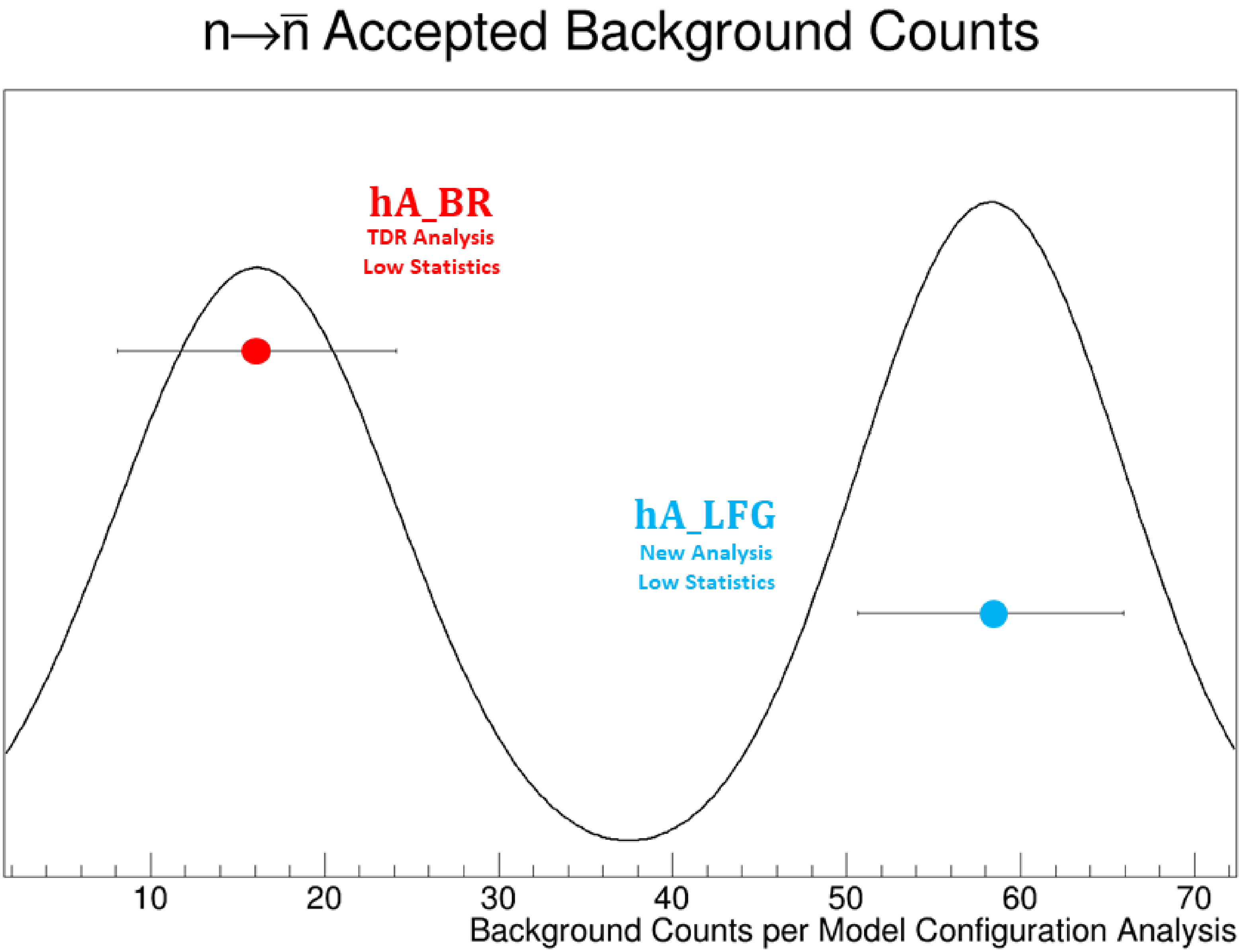
S:B	hA_BR	hA_LFG	hA_ESF	hN_BR	hN_LFG	hN_ESF
hA_BR	$\tau_{n\bar{n}}$	⋯				
hA_LFG	⋮	⋱				
hA_ESF						
hN_BR						
hN_LFG						
hN_ESF						

B:B	hA_BR	hA_LFG	hA_E SF	hN_BR	hN_LFG	hN_ESF
hA_BR		Kinematic Distributions (BDT inputs)	⋯			
hA_LFG	Kinematic Distributions (BDT inputs)		⋱			
hA_ESF	⋮	⋱				
hN_BR						
hN_LFG						
hN_ESF						



Effectively a "universe" approach

Background Count Ideogram



- We can compare counts model-by-model
  - Ideogram supposes gaussian distributed errors for these samples
  - Error is estimated as

$$\frac{\sqrt{\text{Res. Bkgr. Cnts. per 400kt} \cdot \text{yr}}}{\left(\frac{\text{MC Exposure in kt} \cdot \text{yr}}{400\ \text{kt} \cdot \text{yr}}\right)}$$

Model Analyzed	MC Exposure (kt·yr)	Counts/10 kt·yr	Tot. ν_{atm} Counts	Bkgr. Acc. Rate	Res. Bkgr. Counts/400 kt·yr	Est. Bkgr. Count Error
hA_BR	200	2014	40280	0.02%	16.11	8.03
hA_LFG	400	2312.9	92516	0.06%	58.29	7.63