Long-baseline physics analysis: Overview, status, future

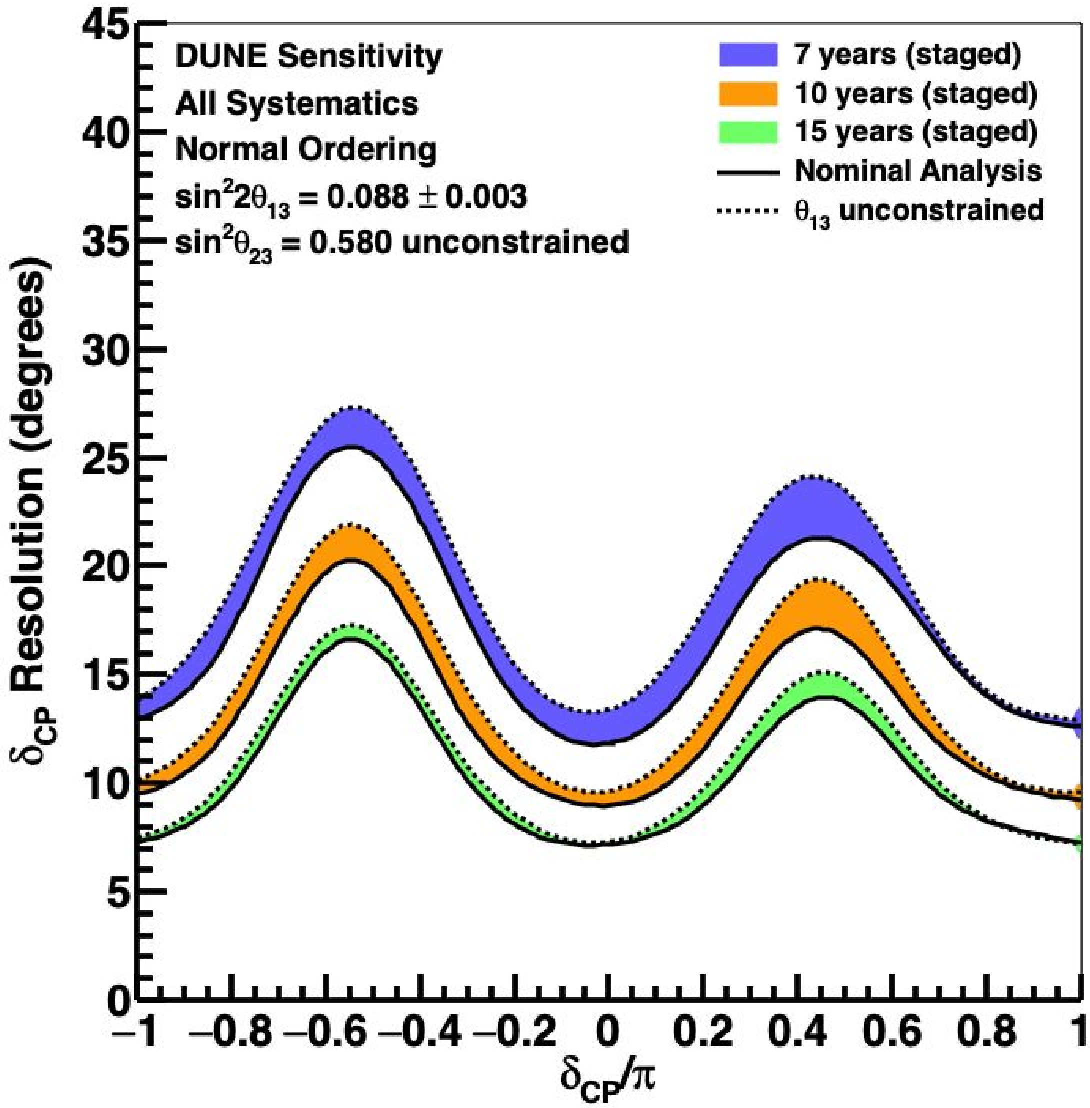
Chris Marshall
Lawrence Berkeley National Laboratory

TIFR workshop
28 February, 2020

Outline

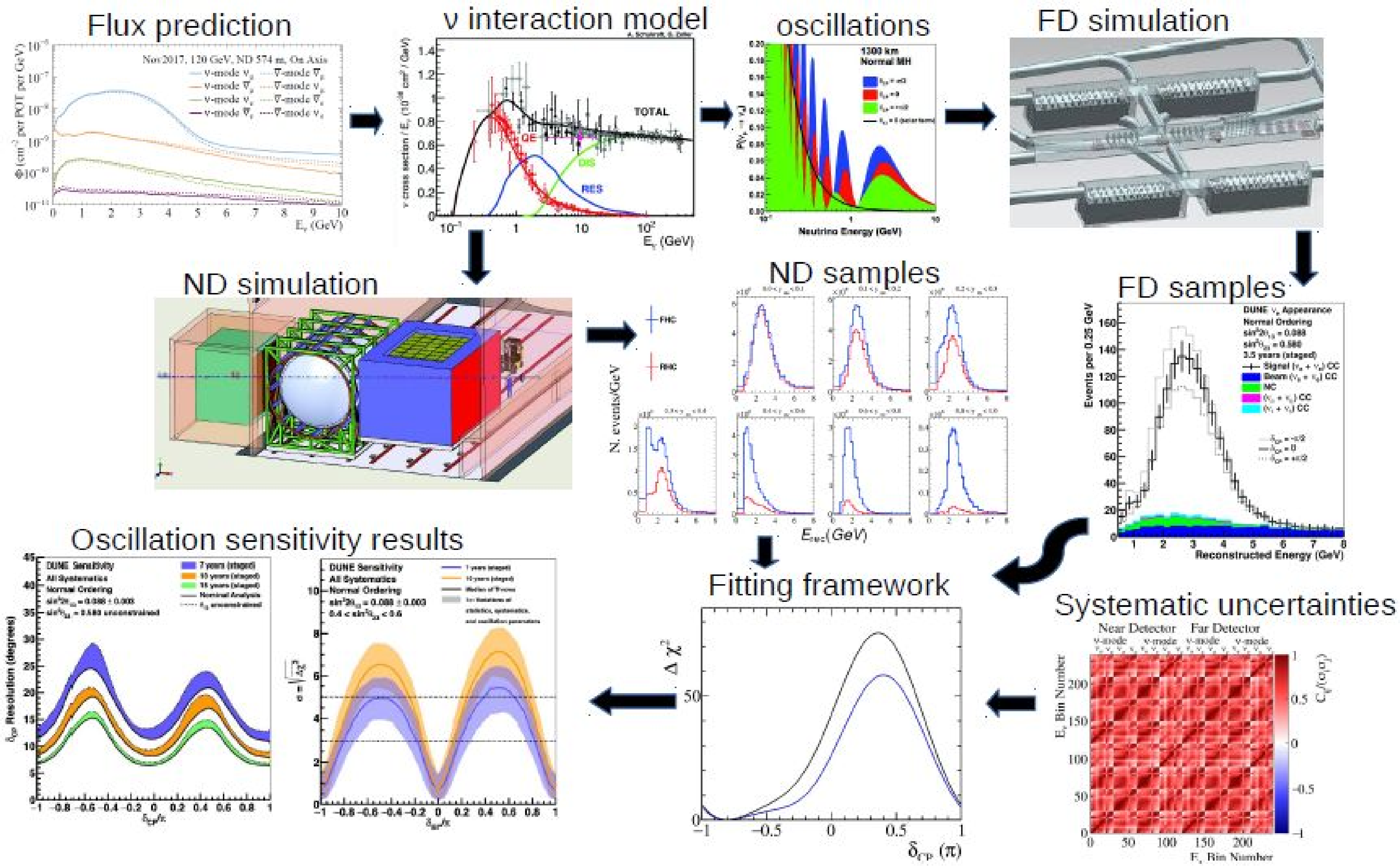
- Long-baseline oscillation fitting in DUNE
- Overview of analysis as implemented in FD TDR
- Challenges & limitations
- Looking to the future: new developments
- Opportunities for new groups
- Comments regarding computing

The end result: sensitivities & parameter resolutions

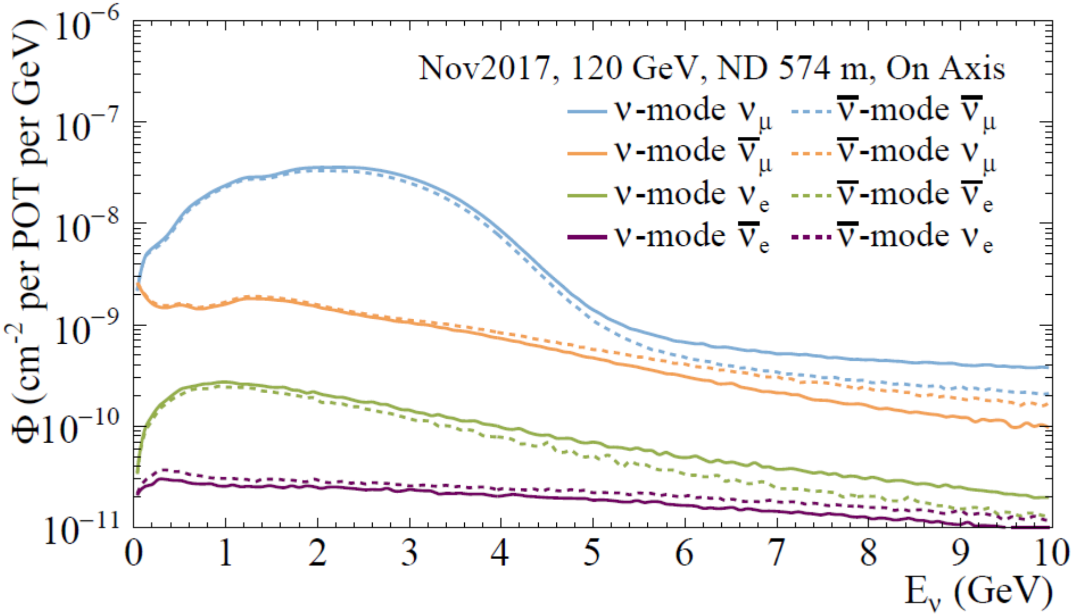


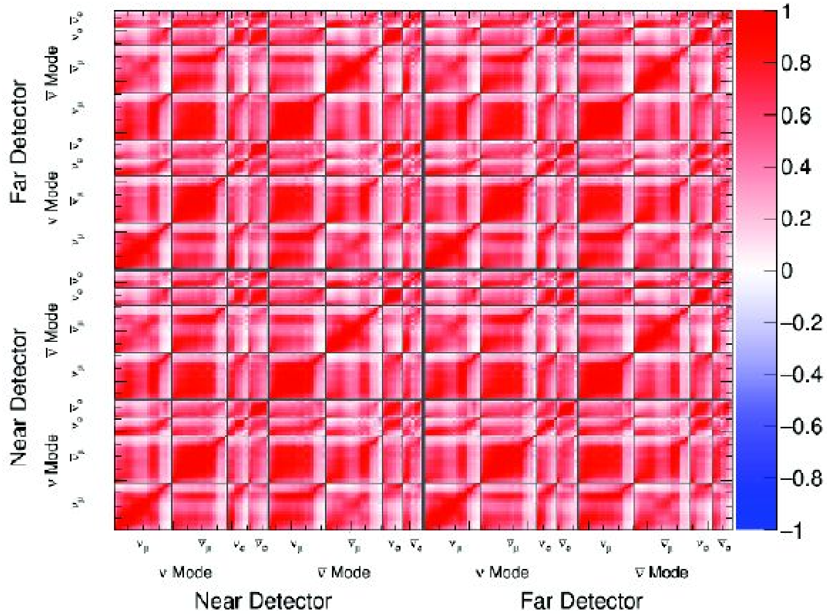
- For three different exposures, the resolution on δ_{CP} as a function of its true value
- Band represents the impact of using the reactor θ_{13} constraint as a prior, which improves our δ_{CP} resolution, especially for shorter exposures

How we get there



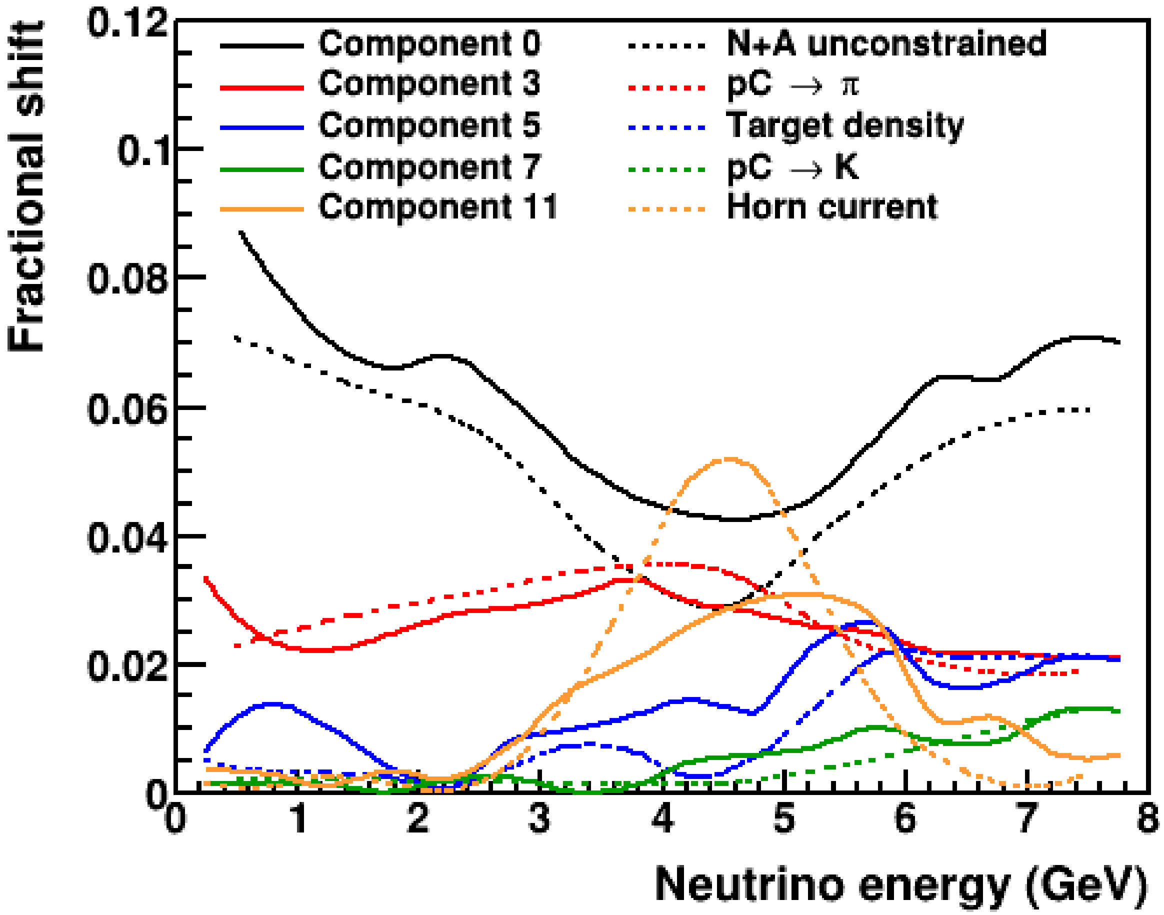
Geant4-based flux prediction and full systematic uncertainties





- Simulation of meson production in proton-carbon interactions and full focusing system
- Meson production is tuned to external proton-carbon data, focusing uncertainties come from varying many systematic parameters in the model
- Full covariance matrix between energy bins of 4 neutrino species ($\nu_\mu/\bar{\nu}_\mu/\nu_e/\bar{\nu}_e$), 2 beam modes (FHC/RHC), 2 detector locations (near/far)

Principal components of covariance matrix are used in analysis



- 208x208 matrix with only ~20 significant eigenvalues → use principal components
- Largest components match up with some of the largest hadron production and focusing uncertainties

Cross section uncertainties

MaCCQE
VecFFCCQEshape
CCQEPauliSupViaKF
MaNCEL
MaCCRES
MvCCRES
MaNCRES
MvNCRES
Theta_Delta2Npi
AhtBY
BhtBY
CV1uBY
CV2uBY
FrCEx_pi
FrElas_pi
FrInel_pi
FrAbs_pi
FrPiProd_pi
FrCEx_N
FrElas_N
FrInel_N
FrAbs_N
FrPiProd_N

Mnv2p2hGaussEnhancement
MKSPP_ReWeight
E2p2h_A_nu
E2p2h_B_nu
E2p2h_A_nubar
E2p2h_B_nubar
BeRPA_A
BeRPA_B
BeRPA_D
C12ToAr40_2p2hScaling_nu
C12ToAr40_2p2hScaling_nubar
nuenuebar_xsec_ratio
nuenumu_xsec_ratio
SPPLowQ2Suppression

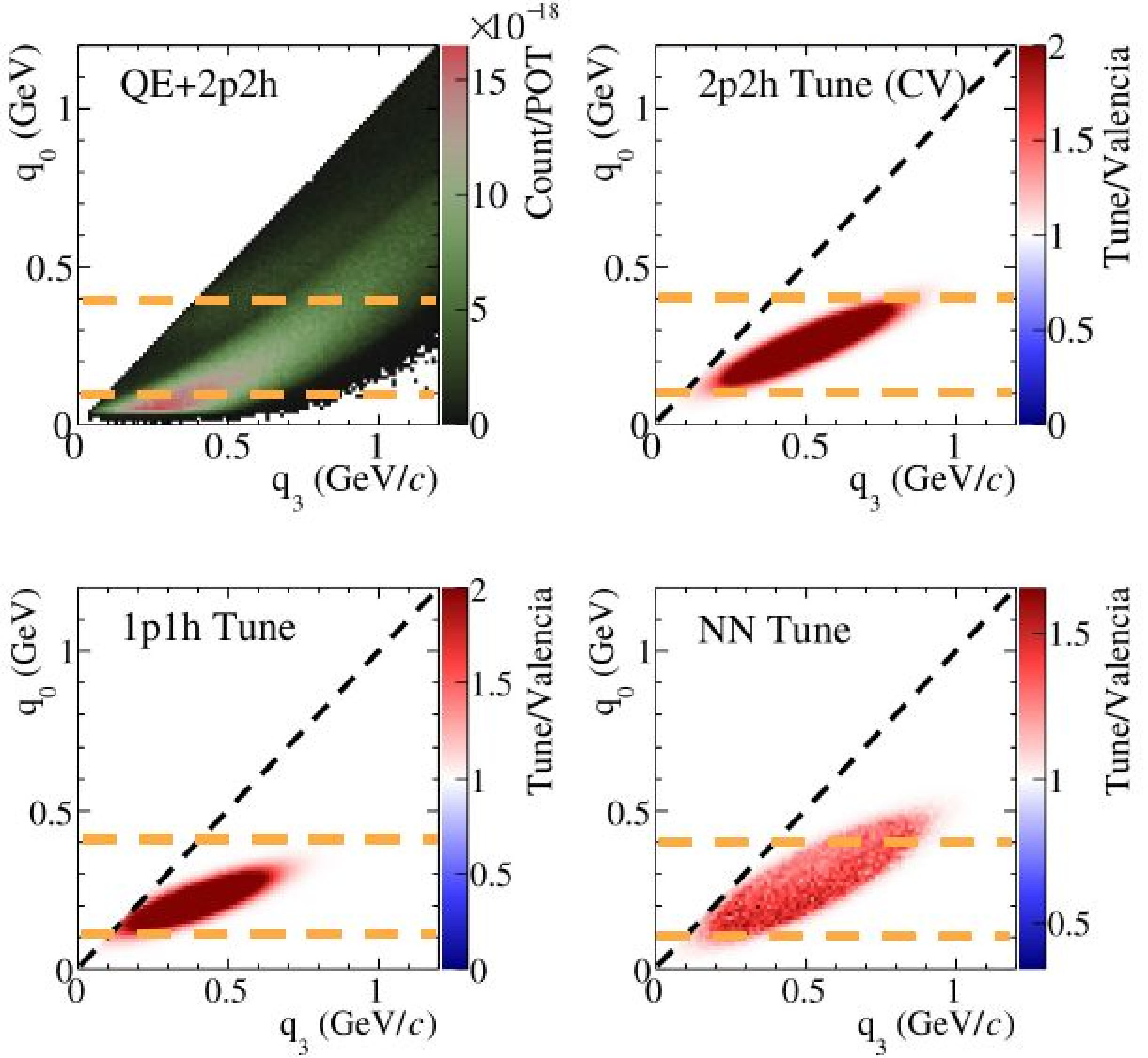
NR_nu_n_CC_2Pi
NR_nu_n_CC_3Pi
NR_nu_p_CC_2Pi
NR_nu_p_CC_3Pi
NR_nu_np_CC_1Pi
NR_nu_n_NC_1Pi
NR_nu_n_NC_2Pi
NR_nu_n_NC_3Pi
NR_nu_p_NC_1Pi
NR_nu_p_NC_2Pi
NR_nu_p_NC_3Pi
NR_nubar_n_CC_1Pi
NR_nubar_n_CC_2Pi
NR_nubar_n_CC_3Pi
NR_nubar_p_CC_1Pi
NR_nubar_p_CC_2Pi
NR_nubar_p_CC_3Pi
NR_nubar_n_NC_1Pi
NR_nubar_n_NC_2Pi
NR_nubar_n_NC_3Pi
NR_nubar_p_NC_1Pi
NR_nubar_p_NC_2Pi
NR_nubar_p_NC_3Pi

Cross section uncertainties strategy

- Many reweightable uncertain parameters are implemented in GENIE, but these uncertainties are insufficient
- Add additional “knobs” based on a combination of data/generator comparisons, alternate theory models, etc.
- Critical to get this right, and lots of work to do → need additional effort in this area

Example: uncertainty on "2-particle 2-hole" interactions

- MINERvA and NOvA see an enhancement in cross section that is consistent with multi-nucleon 2p2h scattering, i.e. $\nu_\mu(np) \rightarrow \mu^- nn$
- MINERvA can fit in 4 different ways: as 1p1h, nn only, pp only, 2p2h
- Implemented parameter moves events between nn $\rightarrow$ 2p2h $\rightarrow$ 1p1h

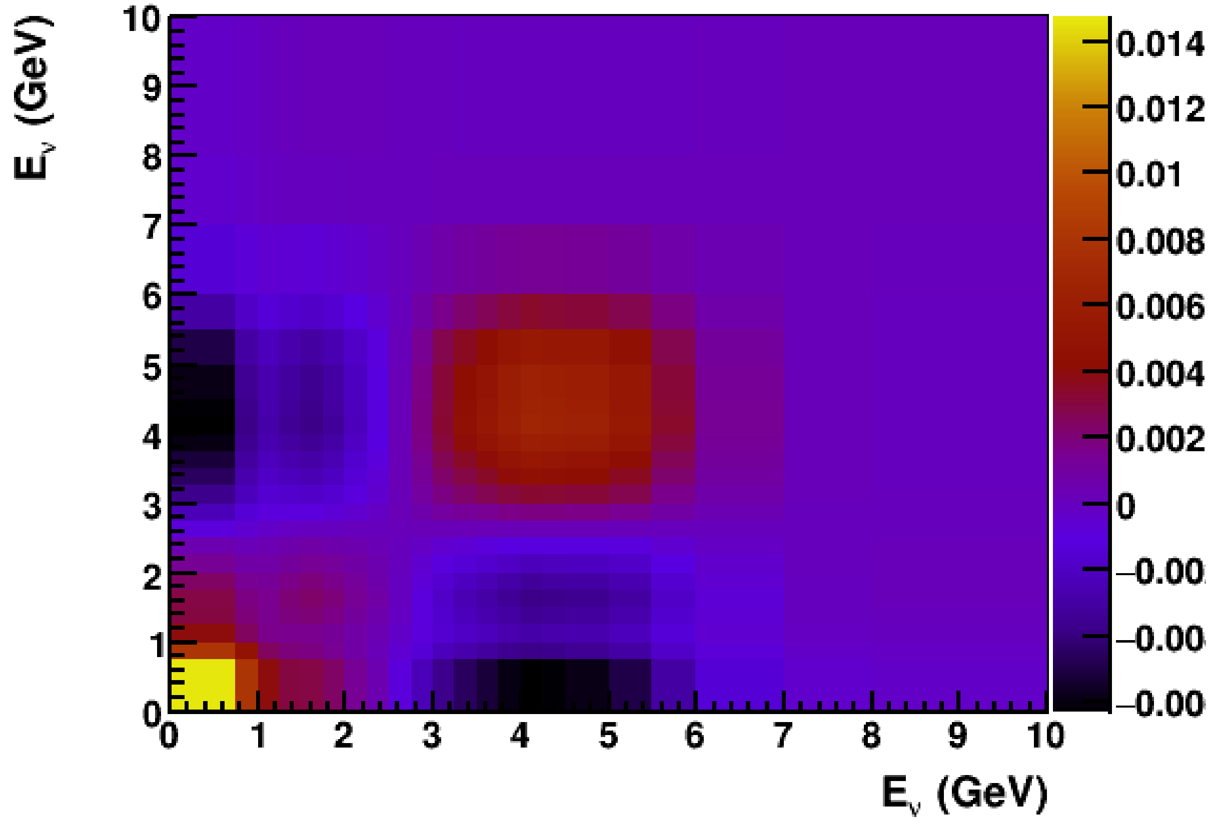


BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY

Detector uncertainties

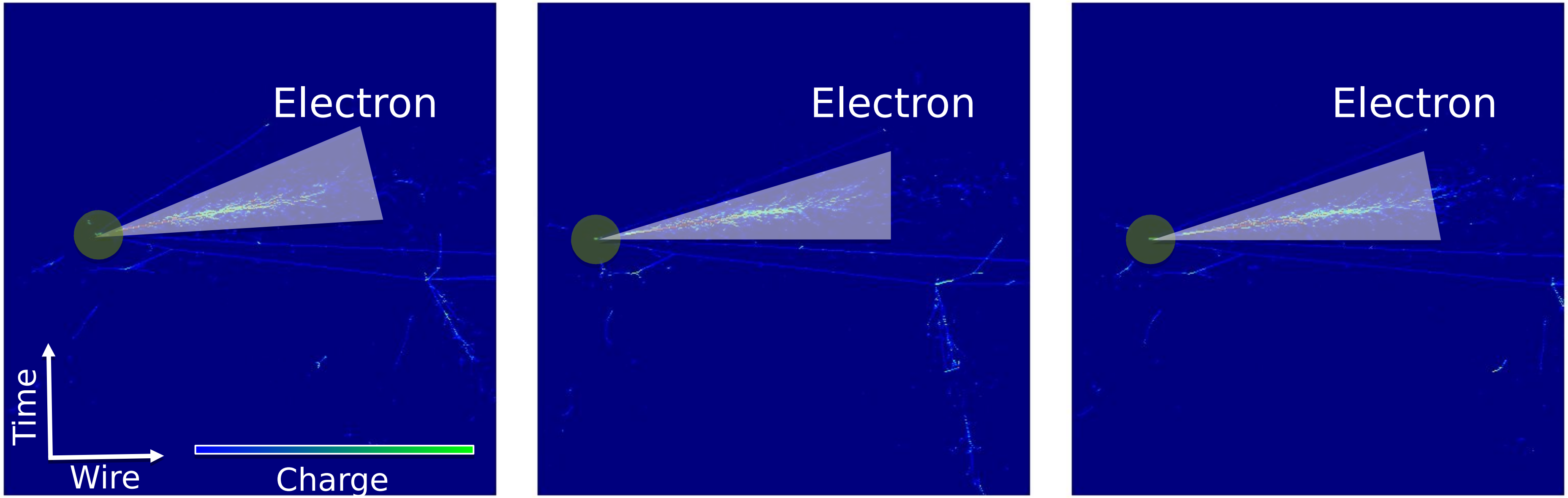
- We implement uncertainties on
  - Reconstructed energy scale
  - Reconstructed energy resolution
  - Detector acceptance corrections
  - NC background rejection
- Currently included only for FD and LAr ND – need to develop model for uncertainties in HPgTPC, SAND, correlations between detectors

Near detector uncertainties are described by covariance matrix



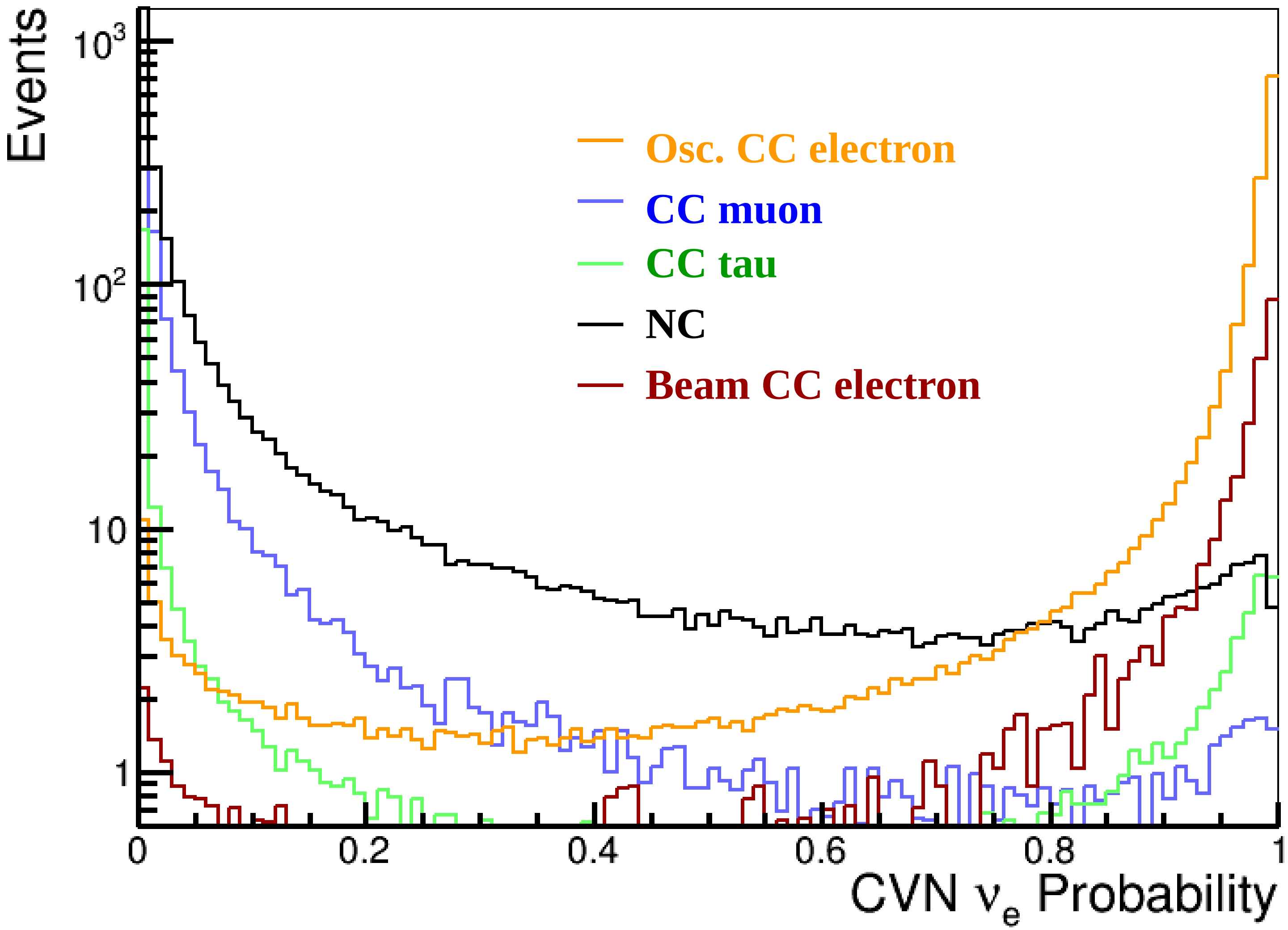
- FD uncertainties are implemented as nuisance parameters constrained in the fit
- This approach is difficult for ND because high statistics, lack of realism leads to overfitting
- Lots of work to be done in making ND model more realistic

FD event selection with convolutional neural network



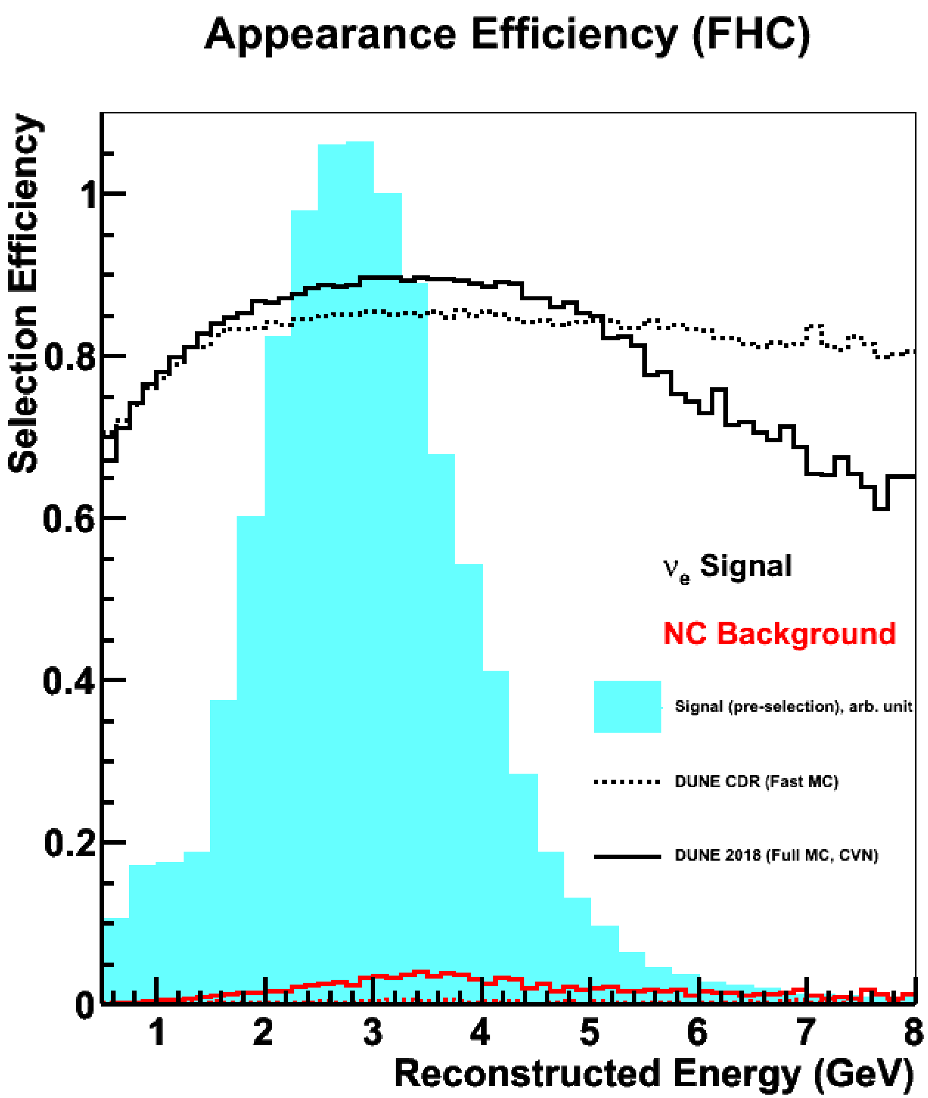
- CVN is trained on event images with known flavor
- Three wire readout planes in far detector → three 2-dimensional “images” of each interaction
- ν_e CC event shown, electron-induced shower highlighted

Far detector event selection: FHC ν_e CVN probability



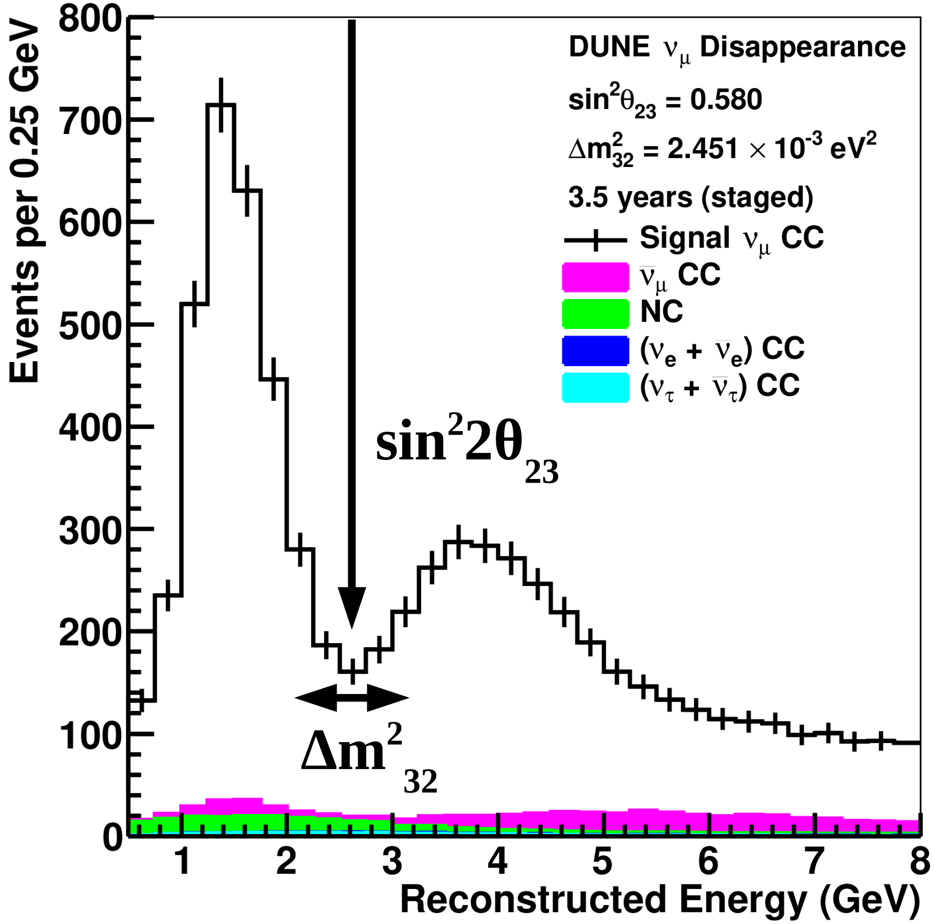
- FHC event probabilities from CVN
- Cut at 0.85 for this analysis
- Selects oscillated and intrinsic electrons

Far detector ν_e selection efficiency

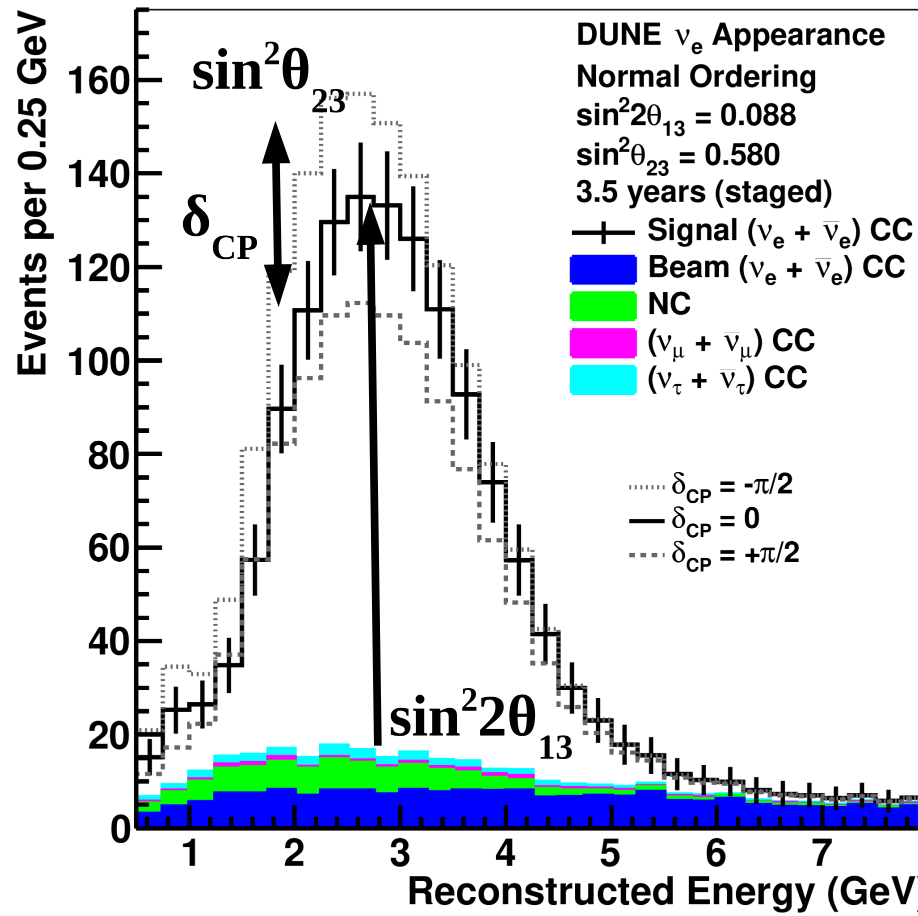


- Full MC with CVN event selection (solid curve) is comparable to fast MC from CDR (dashed curve)
- 85-90% efficient in the region where most events are expected

FHC selected event samples

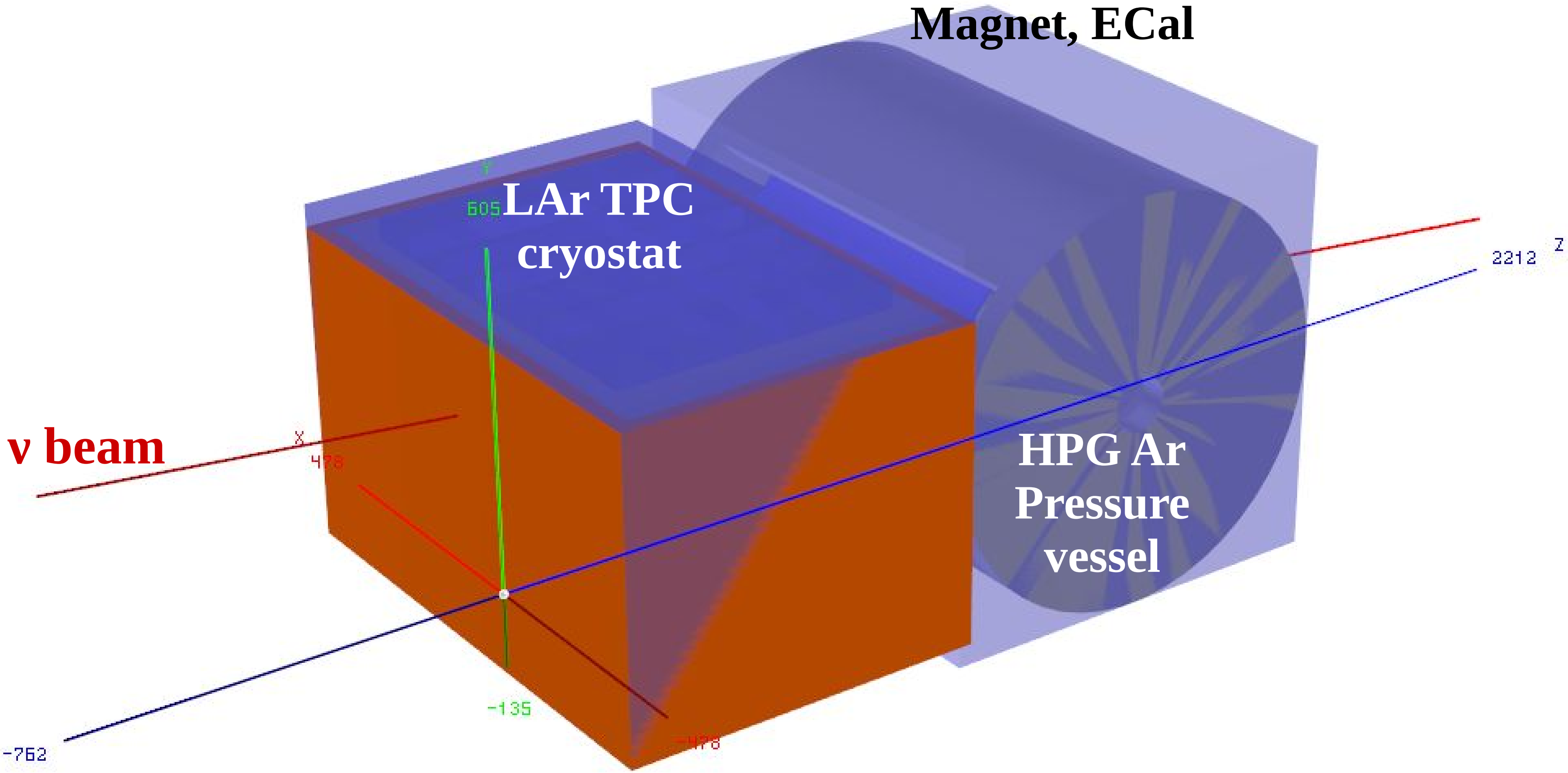


ν_μ disappearance

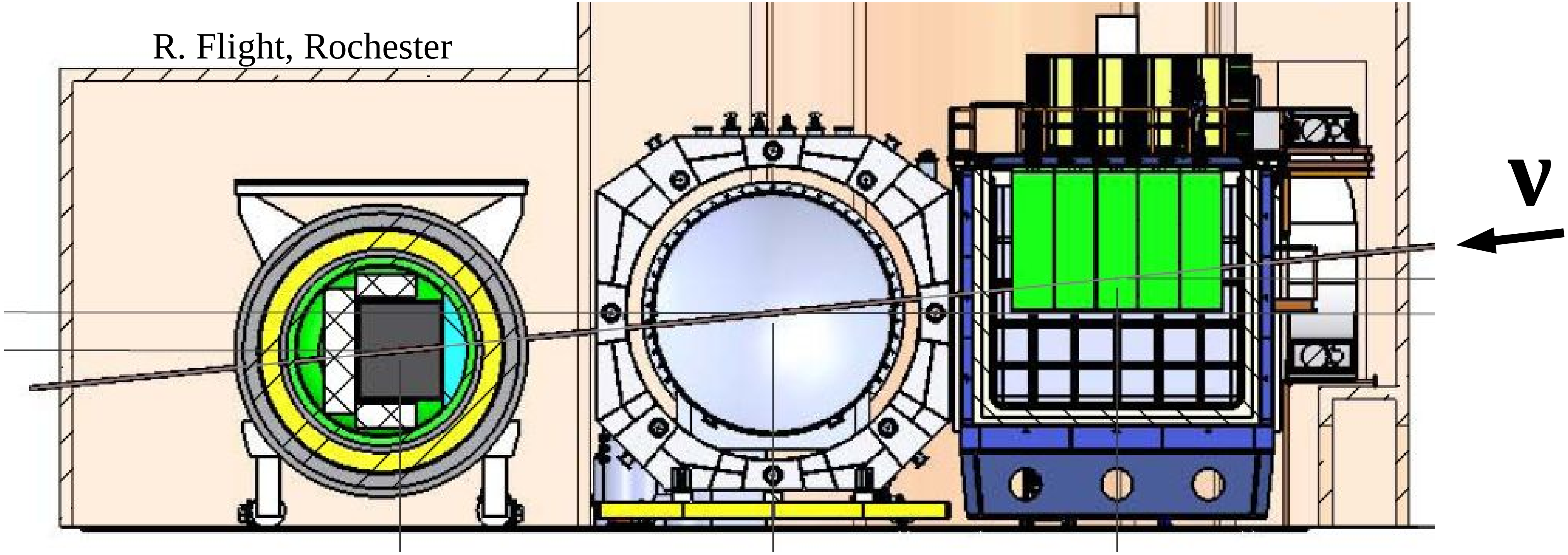


ν_e disappearance

ND samples for TDR LBL analysis use an outdated geometry

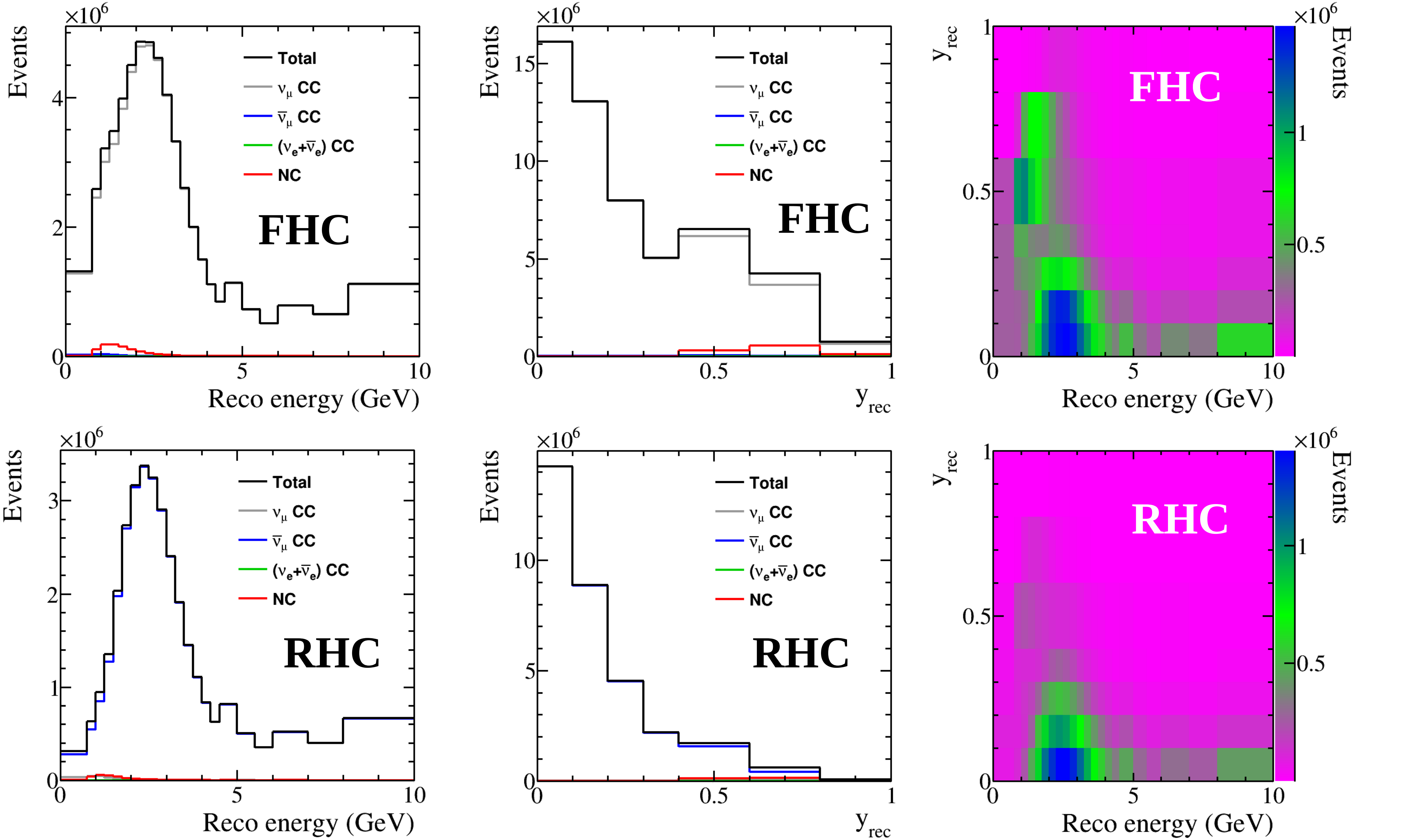


Current simulations use the updated geometry



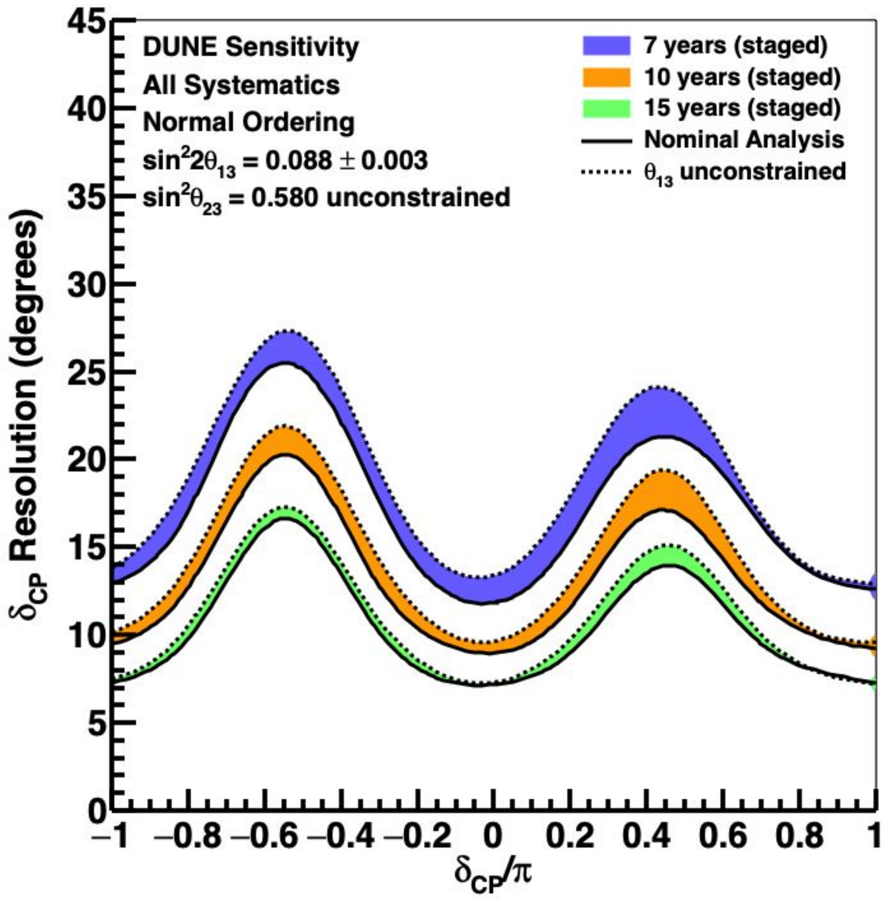
- TDR: parameterized reconstruction of LAr samples using LAr + HPgTPC detectors
- Moving forward: full simulation+reconstruction, directly incorporate HPgTPC(+ECAL+μID) + SAND samples

TDR analysis ND samples: CC inclusive binned in 2D



Chris Marshall

TDR analysis was successful!

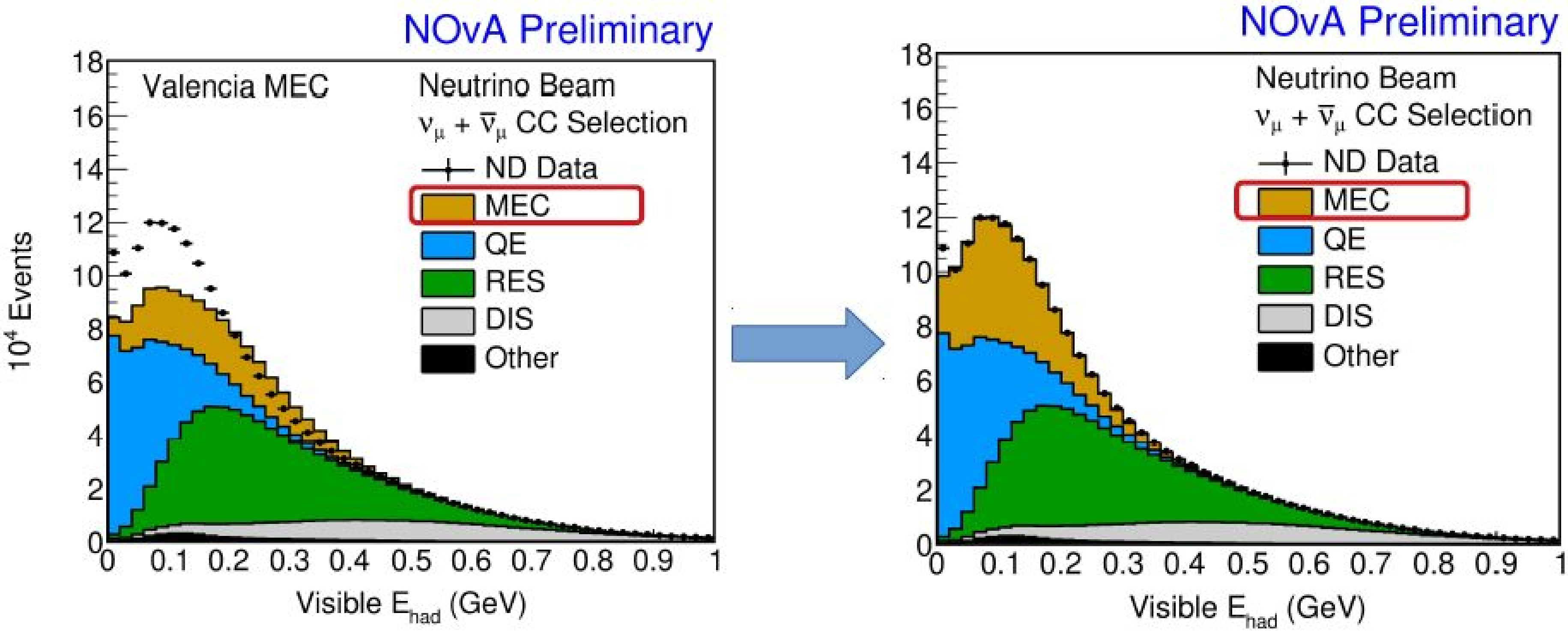


- Produced full suite of oscillation sensitivity results from an end-to-end analysis with full reconstructed FD samples, explicit ND constraints, and realistic systematics
- This was a ton of work and a huge accomplishment

Limitations

- Uses a single ND sample – not practical to directly implement dozens of possible selected samples in LAr, GAr, 3DST
- Implicitly assumes that interaction and detector models are correct and describe the data, up to the included uncertainties – not the experience of every experiment ever
- Very difficult to describe shape uncertainties – most “knobs” have a very particular shape in some kinematic space, and with enough statistics the ND can “measure” the correct value

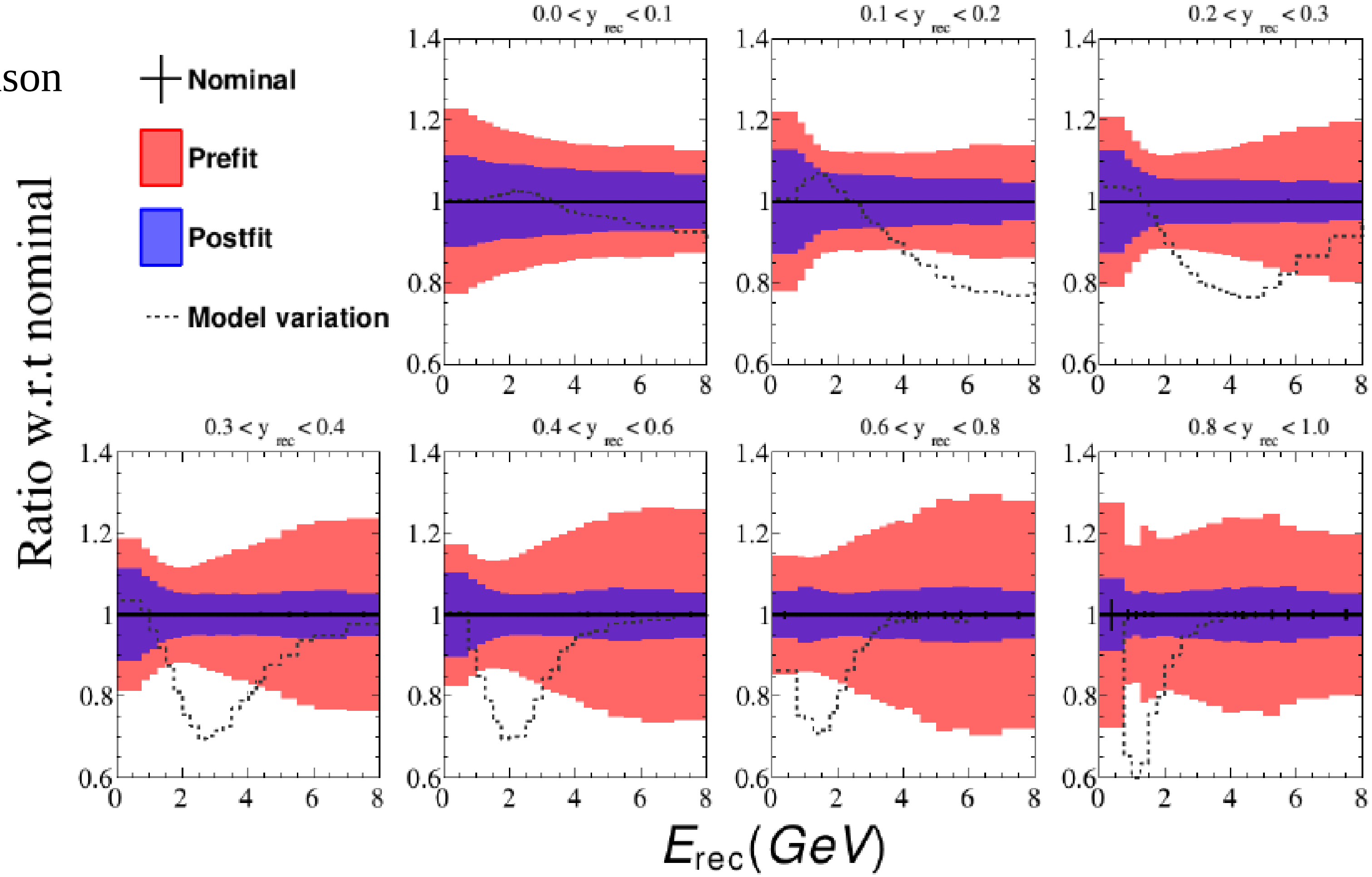
How it works in experiments



- ND data will **not** be described by our model
- We will modify our model to describe the ND data in many different projections, and add systematic uncertainties for the many different ways this can be done

Example: MK single pion

C. Wilkinson



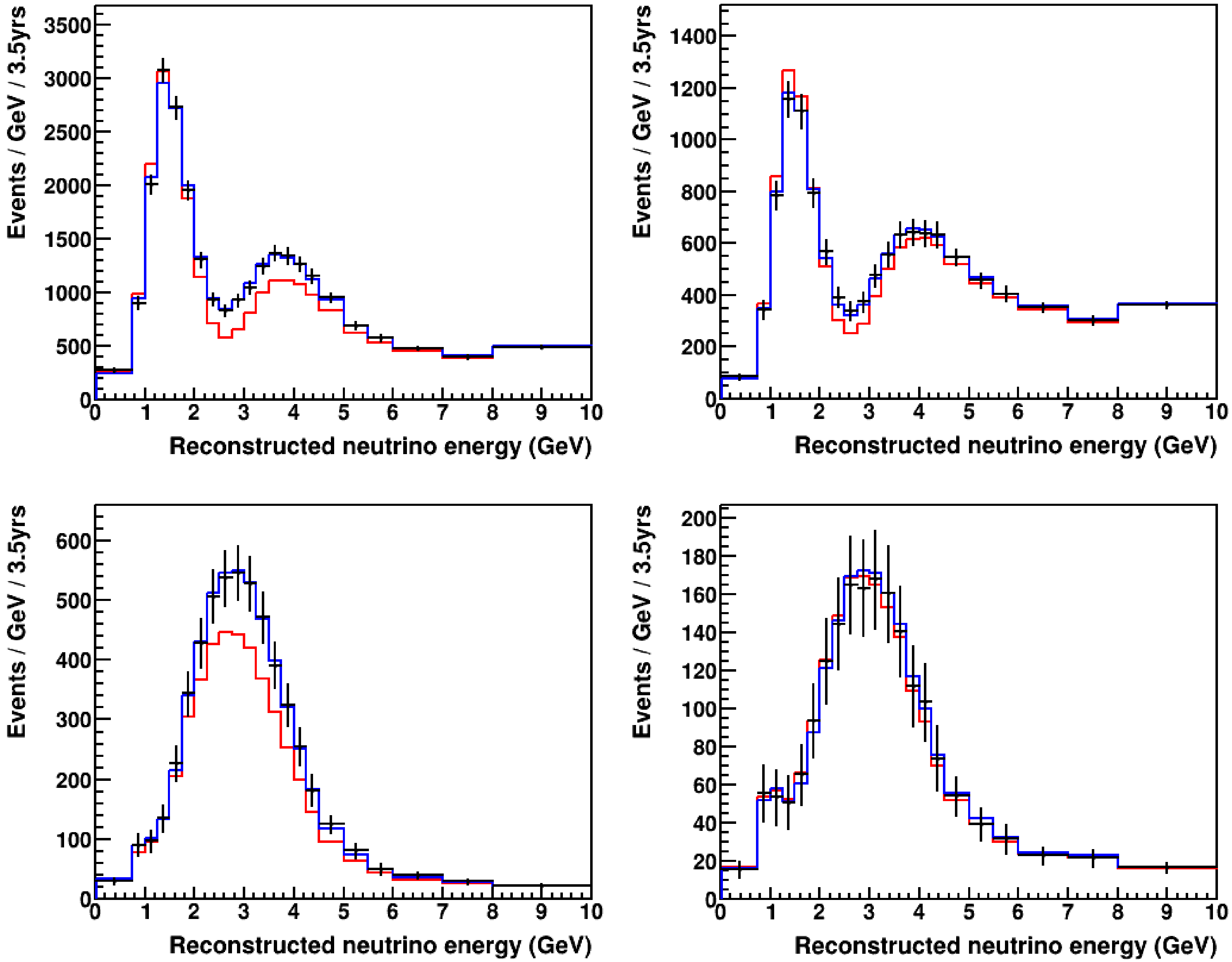
- Easy to see why this on/off dial (MK SPP reweight) is simply resolved by the ND… it simply knows whether it's on or off.

Bias studies with mock data

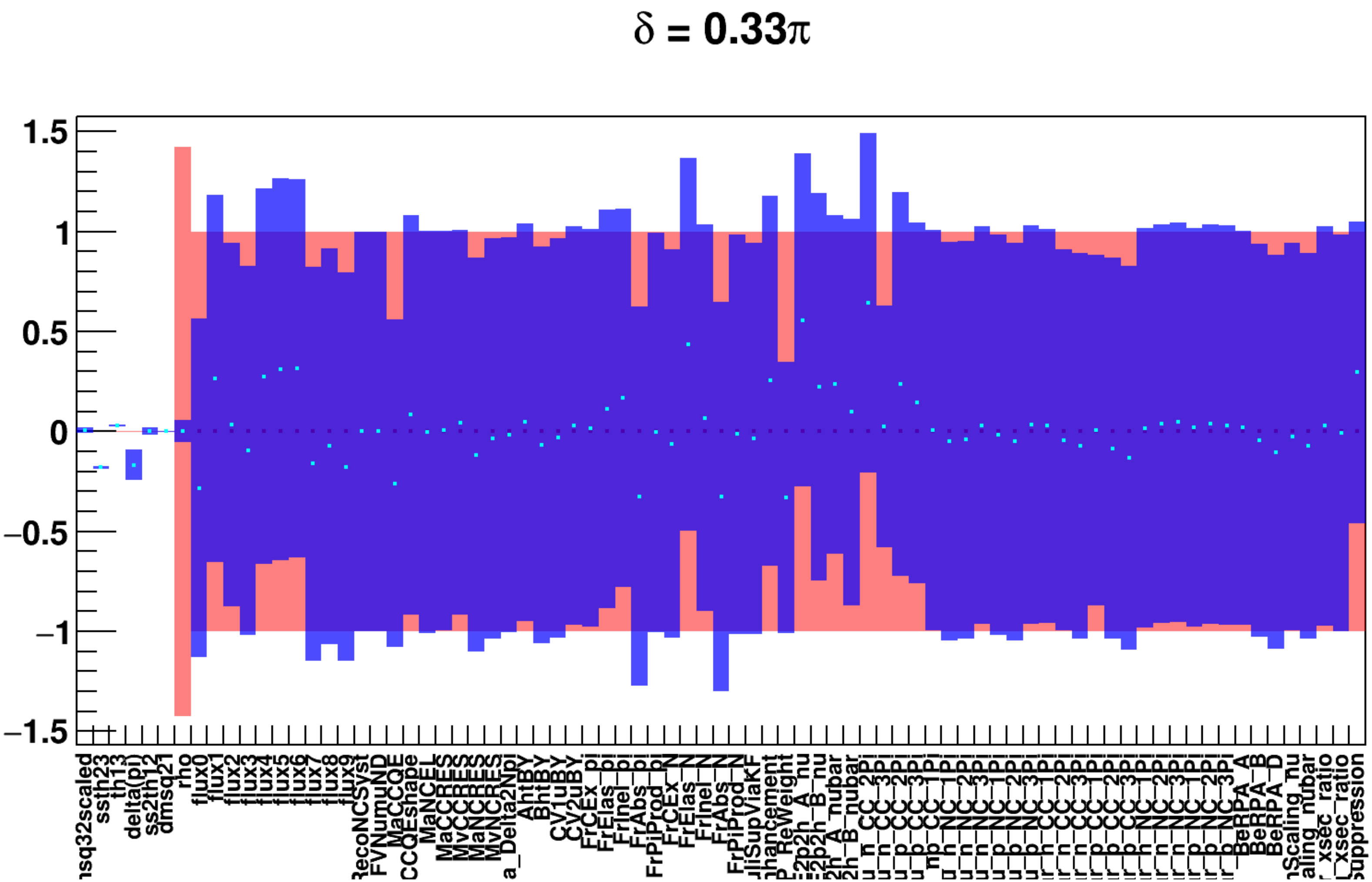
- Consider alternative MC “mock data” samples, and evaluate potential bias on analysis
  - “NuWro mock data”, where a BDT is trained to generate event weights to make GENIE reproduce NuWro prediction in 18 kinematic quantities
  - “Missing proton energy”, where 20% of proton energy is removed (i.e. converted to unobserved neutrons), and cross sections are adjusted so that on-axis hadronic energy spectrum is unchanged

FD-only fits

- FD-only we get very good fit, with $\chi^2 \sim 10$
- No evidence of any problems with model

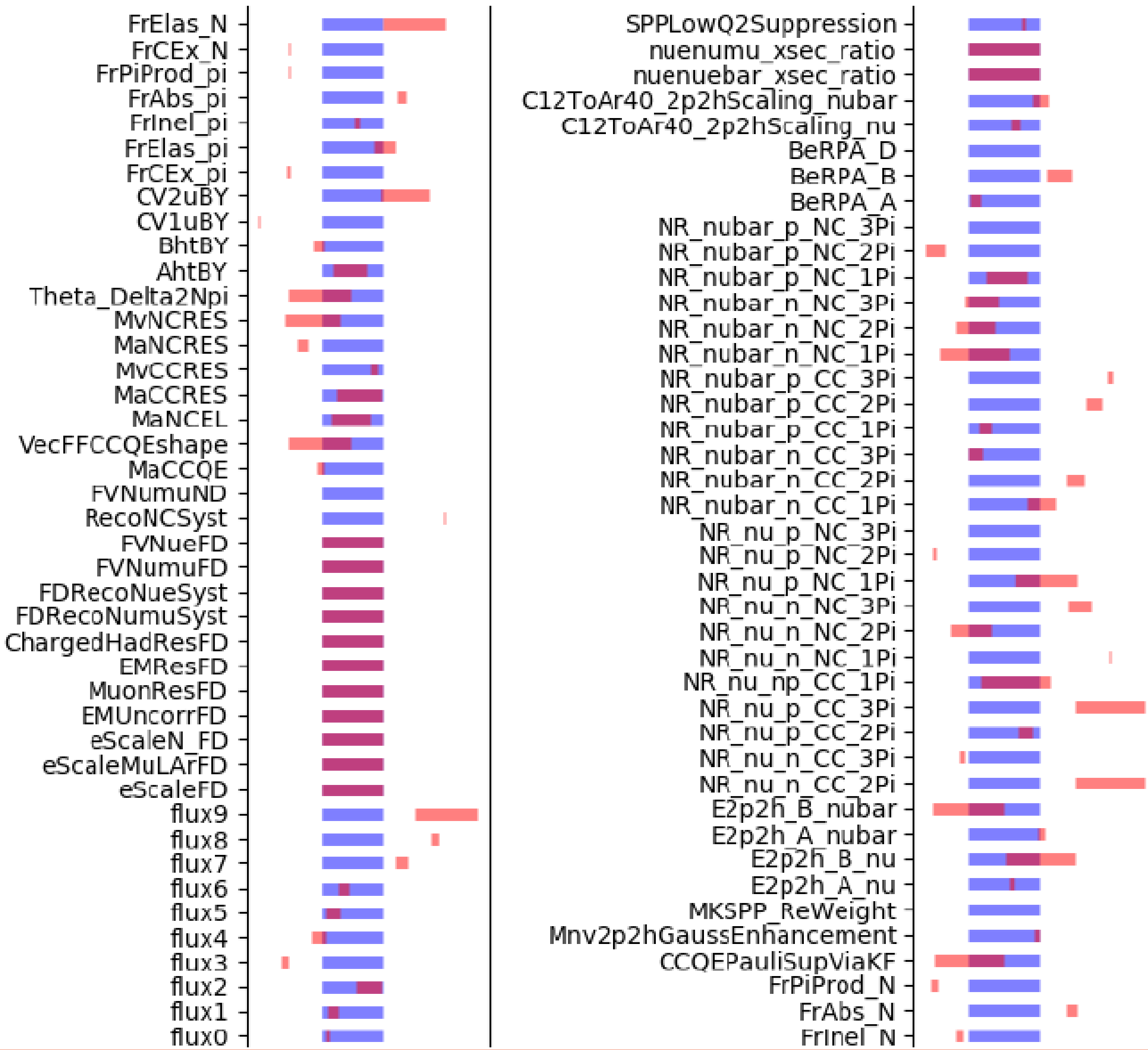


FD-only nuisance parameter post-fits are < 0.5σ of pre-fit values

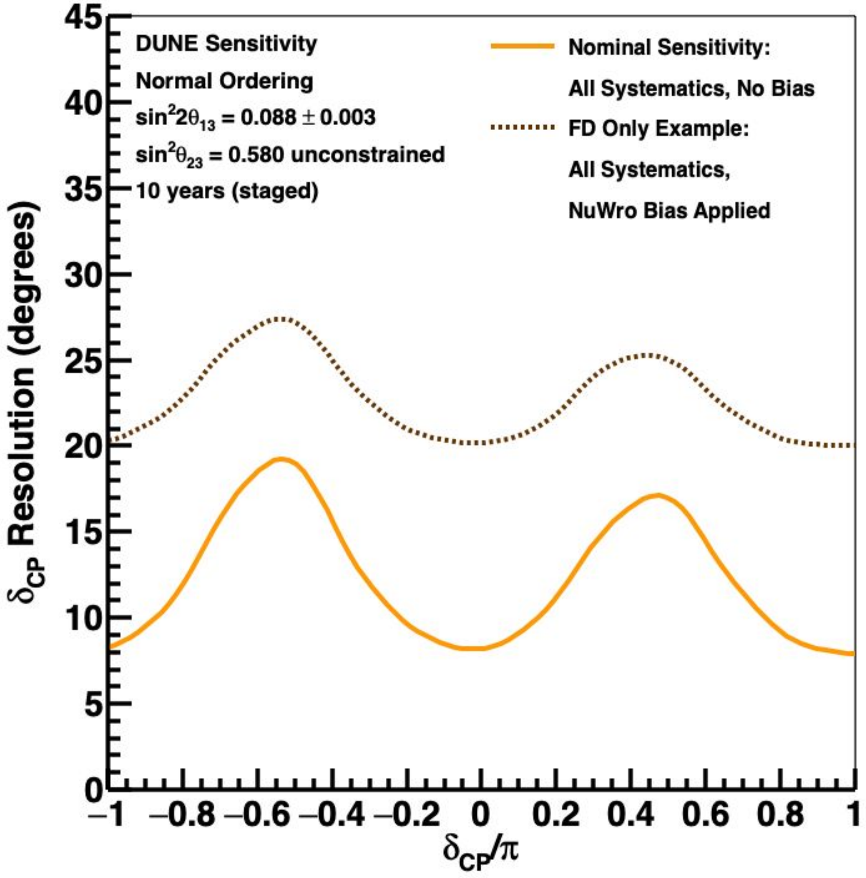


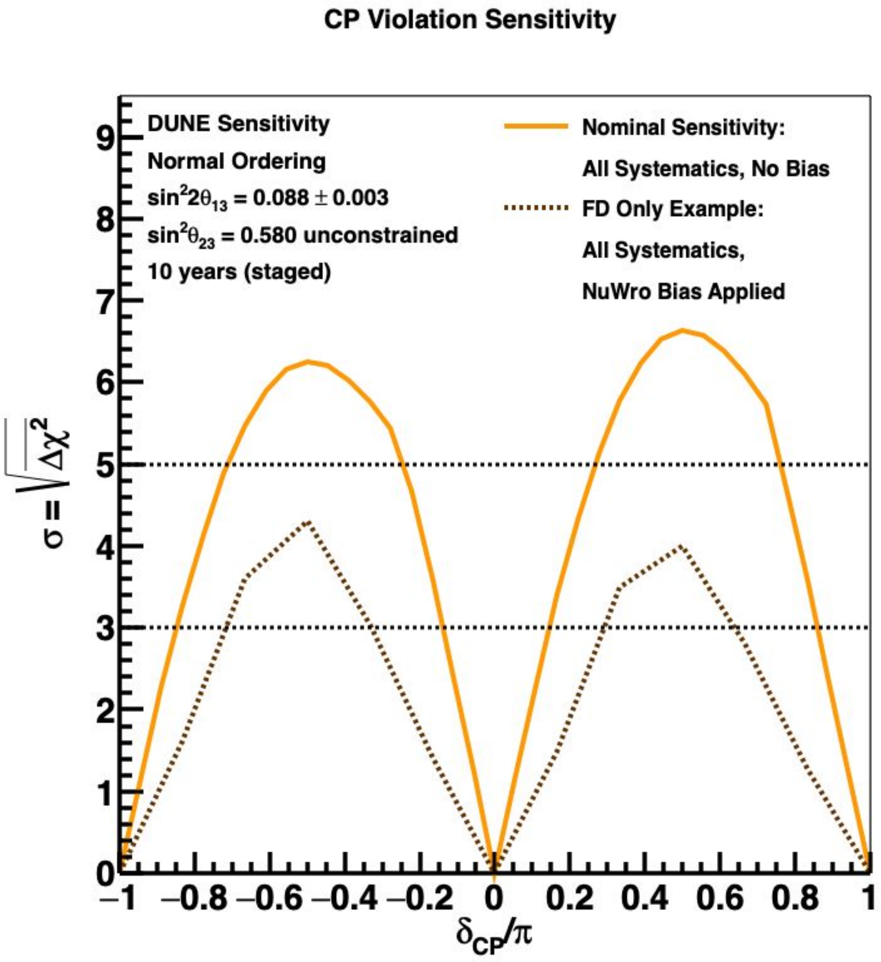
ND+FD fit χ^2 = 10879.2

- Post-fit parameter uncertainties are shown as red bands
- Parameters get pulled way outside their pre-fit ranges, with tiny constraints
- Fit to ND data is terrible – we would definitely know there is a problem, although we do not yet show how we would fix it

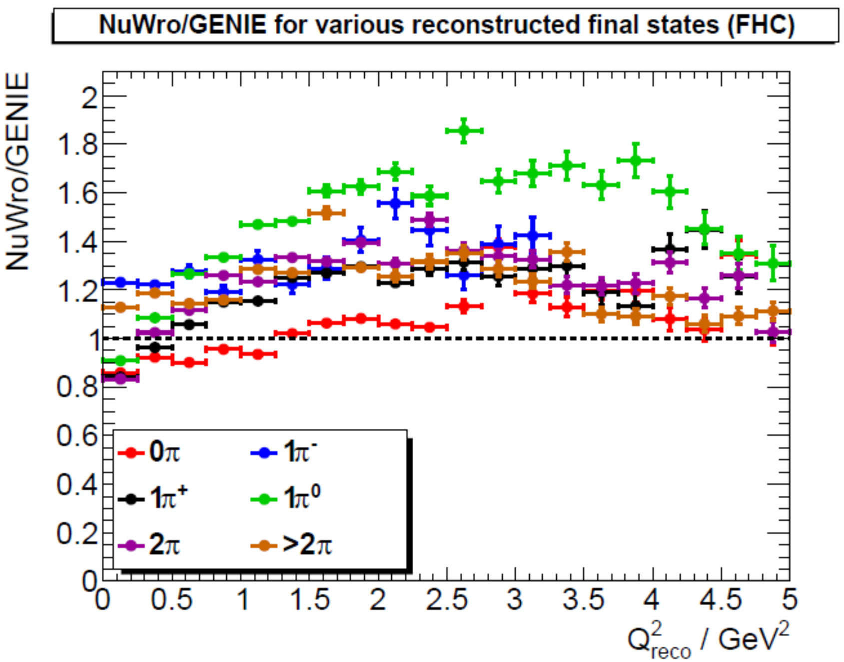


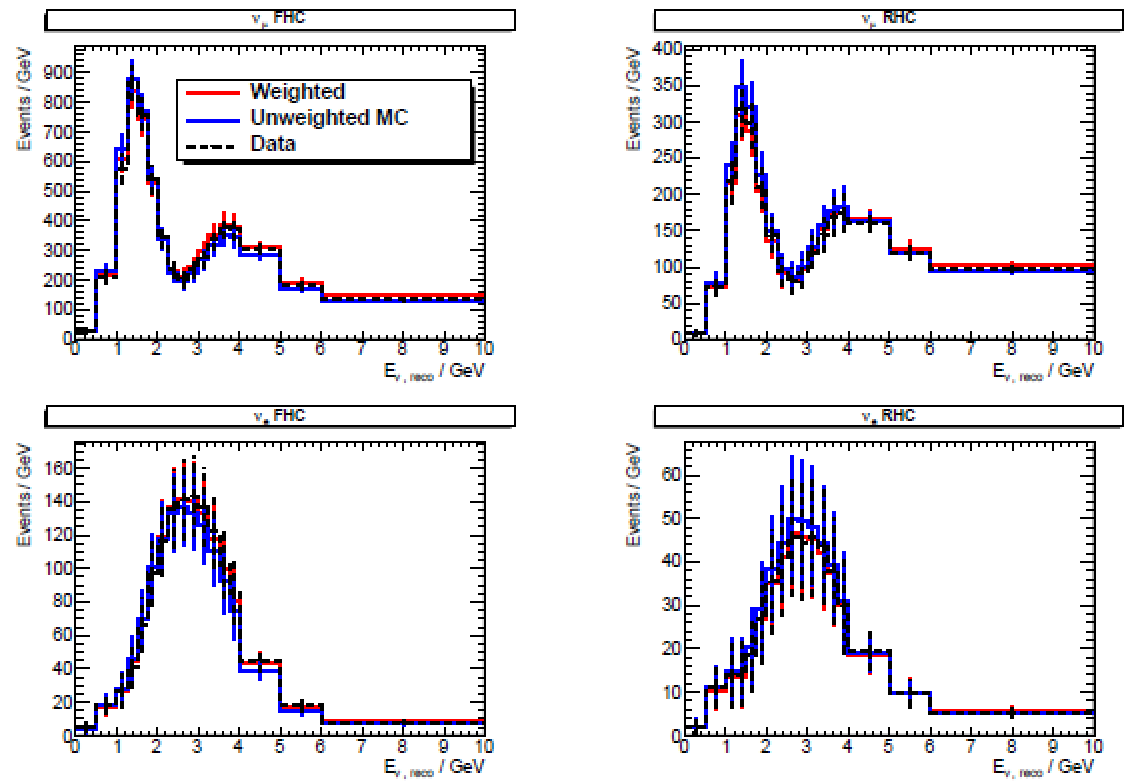
Sensitivities with bias applied





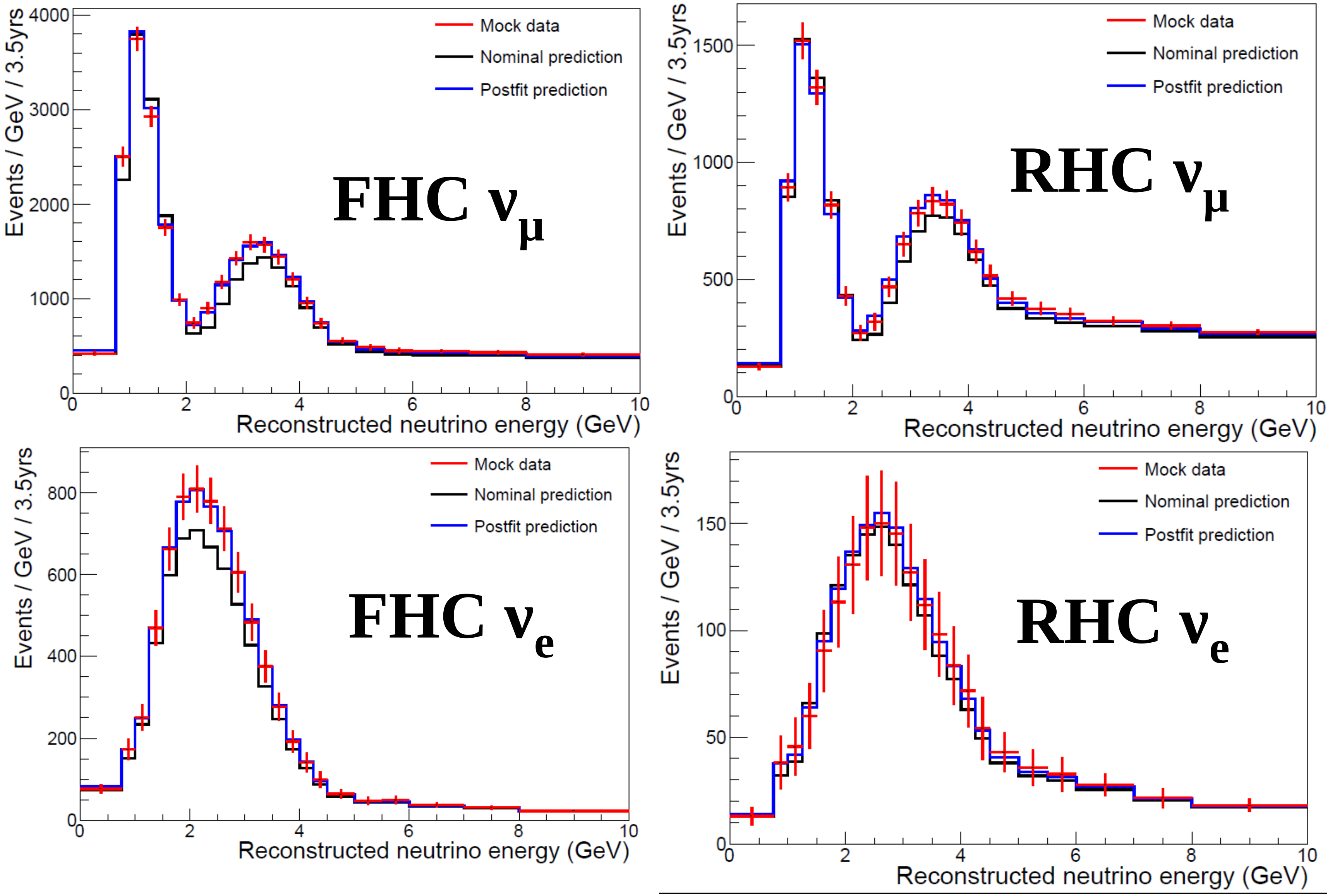
Reweighting the CV prediction with HPgTPC data





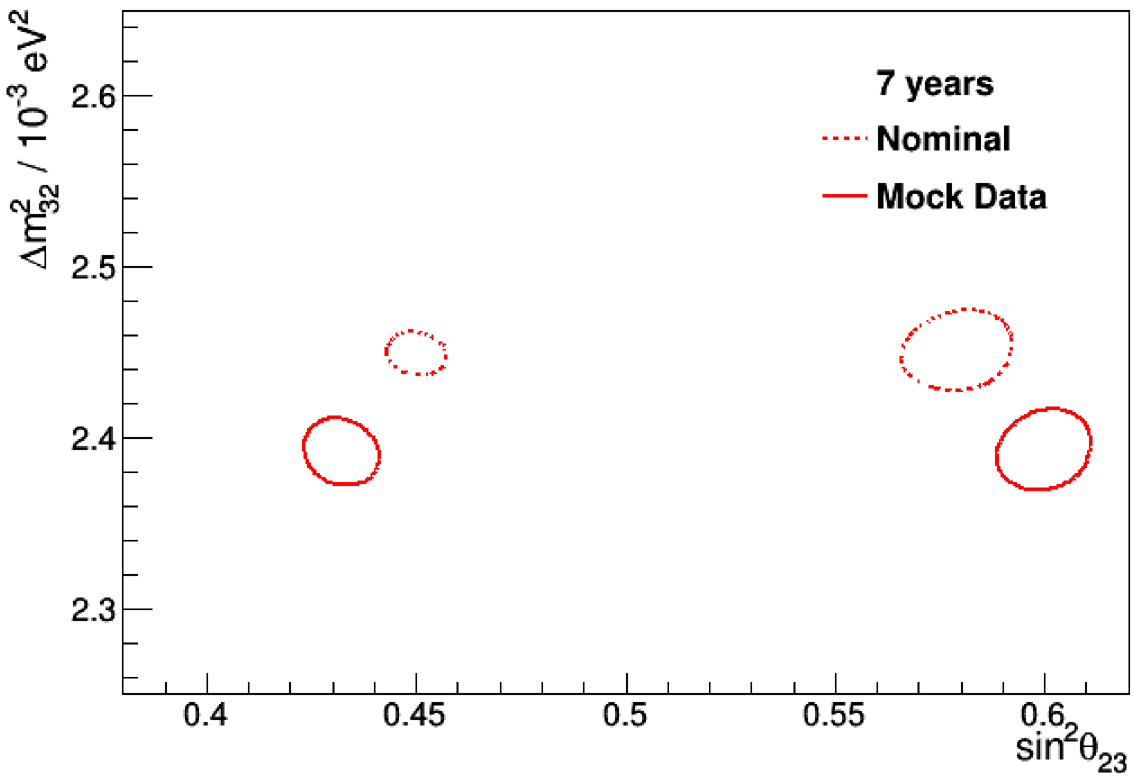
- Derive Data/GENIE ratio for different reconstructed samples by number of pions
- Applying this to the *a priori* MC prediction improves the result

Missing proton mock data



Missing proton bias

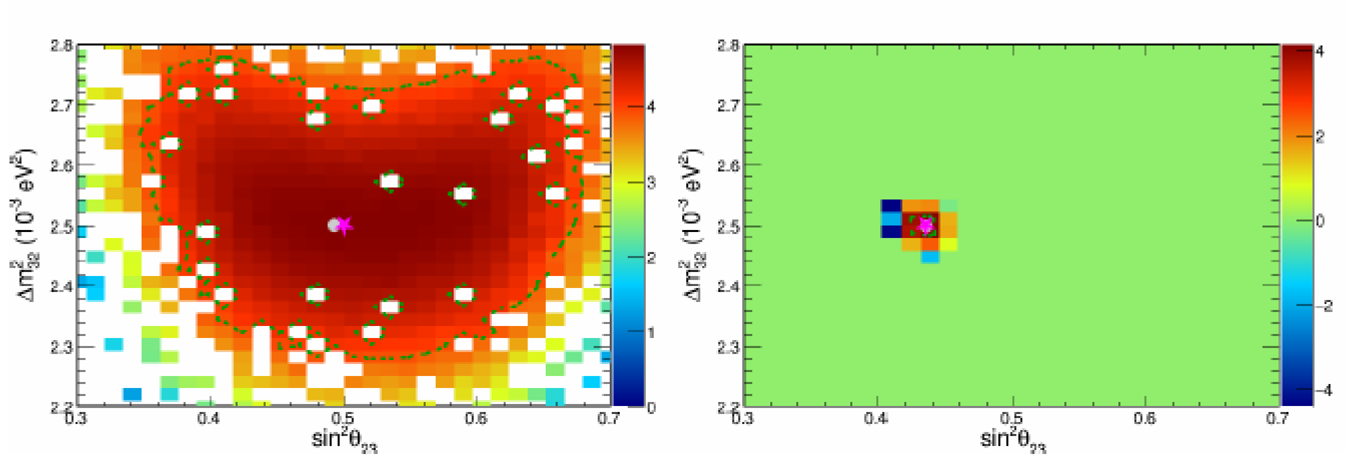
- Best-fit gives significant bias to Δm^2 and θ_{23}, several sigma outside uncertainties



- Additional uncertainty would be required to cover the bias
- Effect is easily detected with off-axis ND data

New thread: Markov Chain MC

- Ongoing work on Bayesian analysis using Stan, with initial results looking very promising
- Another similar effort on using Mach3 package developed in T2K
- This approach scales much better as fits become more complicated



Near-term & Longer-term plans

- Near-term: Near detector IDR & TDR (now-2021)
  - Continue to explore ND constraints with mock data
- Longer-term:
  - Incorporate fully-reconstructed ND samples
  - Include additional LAr, HPgTPC, SAND constraints
  - Pursue alternate analysis approaches, such as MCMC
  - Other ideas?

Backups

GENIE ReWeight

MaCCQE
VecFFCCQEshape
CCQEPauliSupViaKF
MaNCEL
MaCCRES
MvCCRES
MaNCRES
MvNCRES
Theta_Delta2Npi
AhtBY
BhtBY
CV1uBY
CV2uBY
FrCEx_pi
FrElas_pi
FrInel_pi
FrAbs_pi
FrPiProd_pi
FrCEx_N
FrElas_N
FrInel_N
FrAbs_N
FrPiProd_N

GENIE reweight parameters affecting
CC quasi-elastic
CC resonance production
CC deep inelastic scattering
Final-state interactions
Neutral currents

DUNEint not covered in GENIE

Additional parameters:

CC QE

CC Resonance

2p2h

Scaling C → Ar

ν_e/ν_μ or $\nu_e/\bar{\nu}_e$

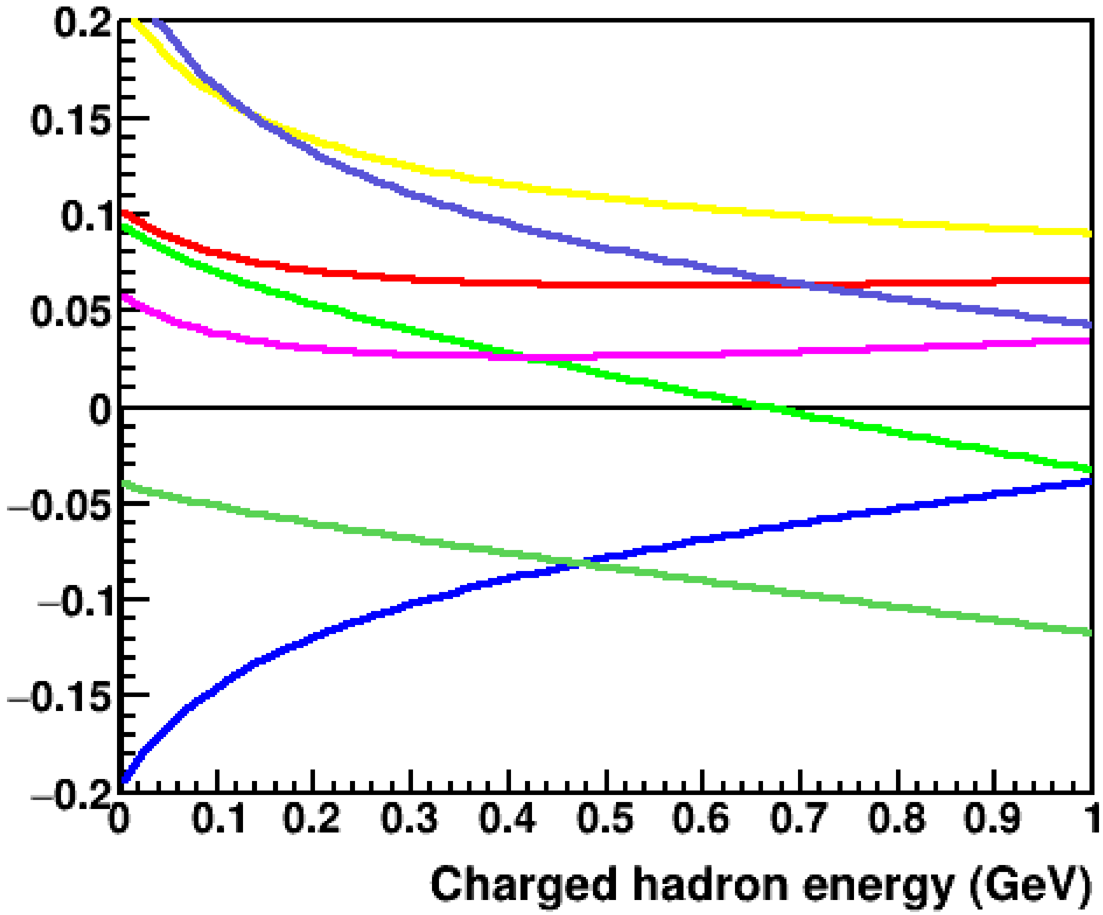
Mnv2p2hGaussEnhancement
MKSPP_ReWeight
E2p2h_A_nu
E2p2h_B_nu
E2p2h_A_nubar
E2p2h_B_nubar
BeRPA_A
BeRPA_B
BeRPA_D
C12ToAr40_2p2hScaling_nu
C12ToAr40_2p2hScaling_nubar
nuenuebar_xsec_ratio
nuenumu_xsec_ratio
SPPLowQ2Suppression

DUNEint not covered in GENIE

Additional parameters affecting non-resonant pion production

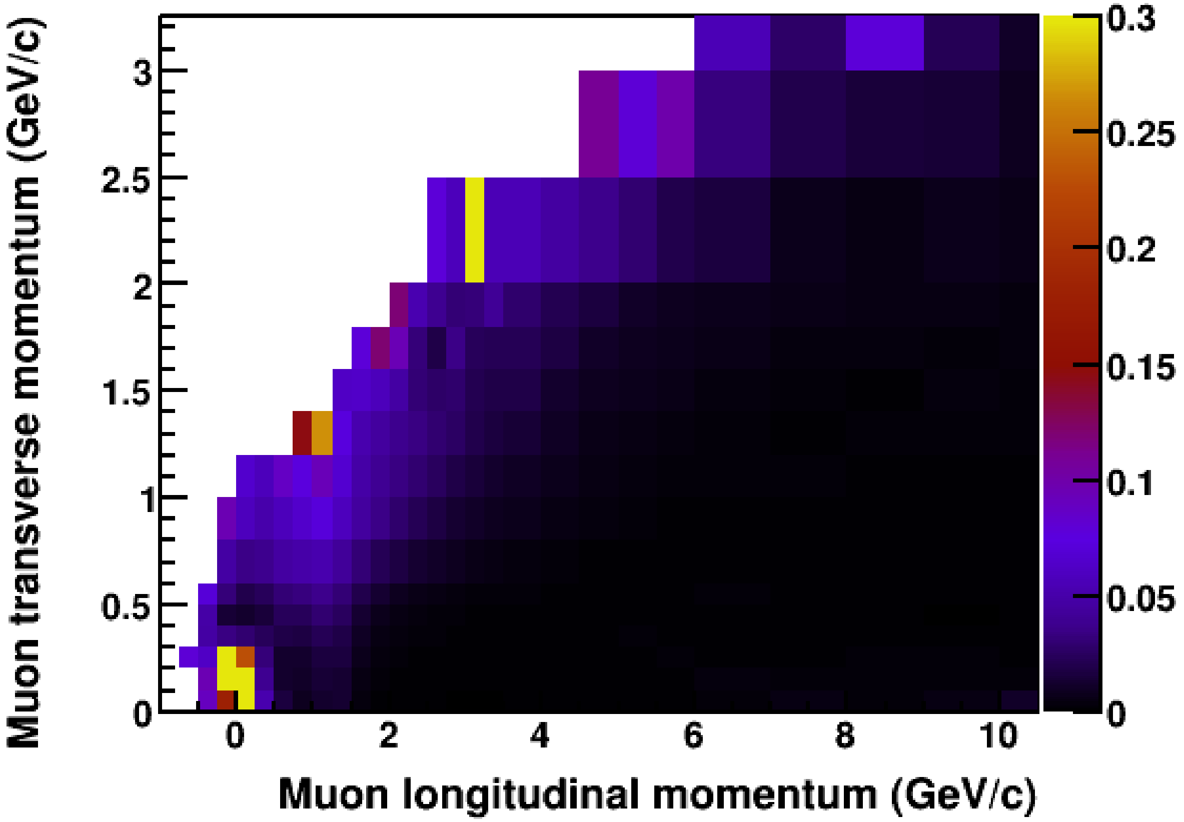
NR_nu_n_CC_2Pi
NR_nu_n_CC_3Pi
NR_nu_p_CC_2Pi
NR_nu_p_CC_3Pi
NR_nu_np_CC_1Pi
NR_nu_n_NC_1Pi
NR_nu_n_NC_2Pi
NR_nu_n_NC_3Pi
NR_nu_p_NC_1Pi
NR_nu_p_NC_2Pi
NR_nu_p_NC_3Pi
NR_nubar_n_CC_1Pi
NR_nubar_n_CC_2Pi
NR_nubar_n_CC_3Pi
NR_nubar_p_CC_1Pi
NR_nubar_p_CC_2Pi
NR_nubar_p_CC_3Pi
NR_nubar_n_NC_1Pi
NR_nubar_n_NC_2Pi
NR_nubar_n_NC_3Pi
NR_nubar_p_NC_1Pi
NR_nubar_p_NC_2Pi
NR_nubar_p_NC_3Pi

Example energy scale uncertainty: charged hadron response



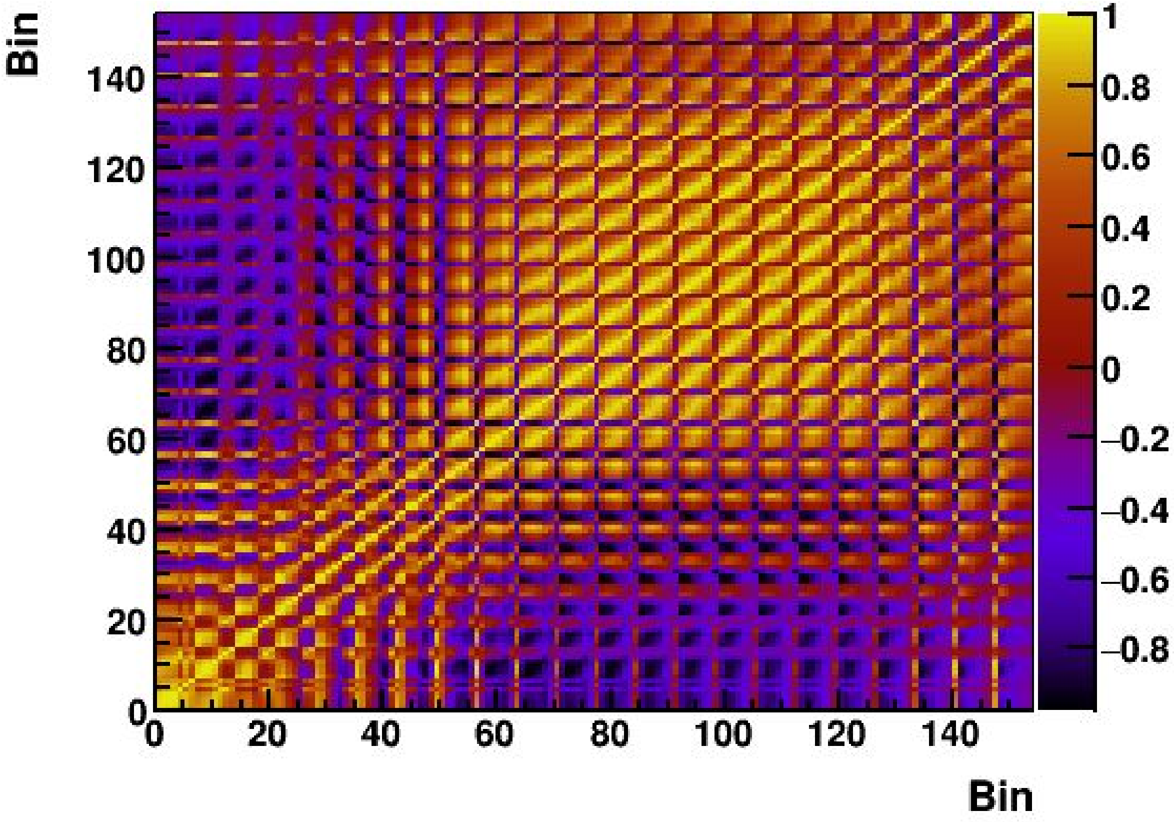
- Each curve represents the energy response bias in a particular universe, where the parameters have been chosen randomly consistent with the energy-dependent uncertainty

ND CC ν_μ acceptance fractional uncertainty



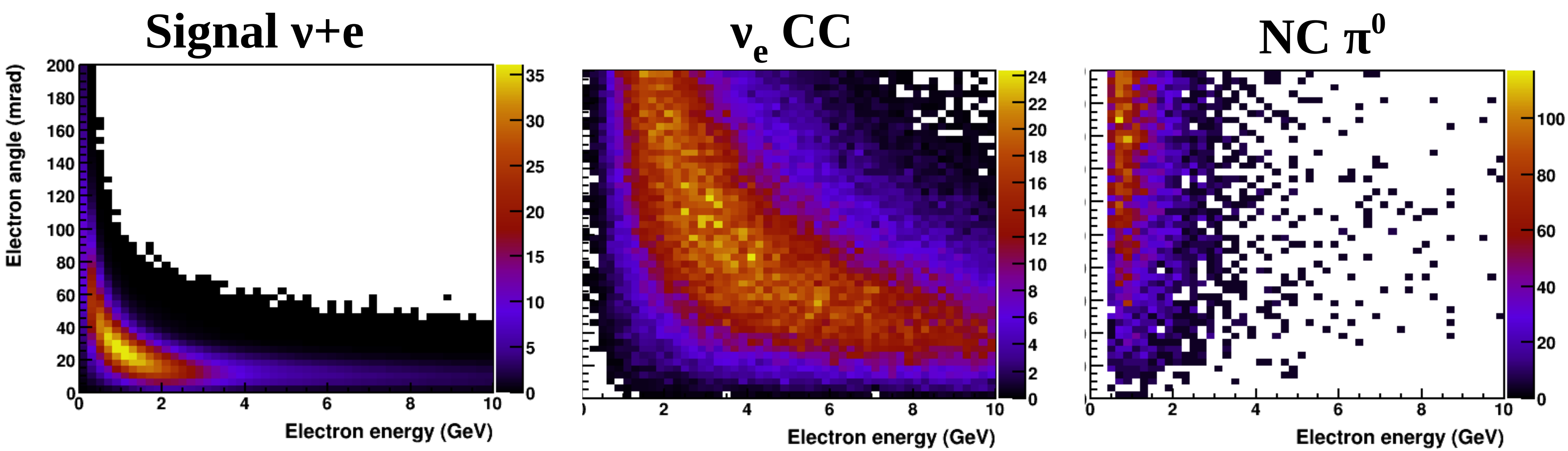
- CC events are rejected when
  - Muon is reconstructed as $\pi^\pm$ (low energy)
  - Muon exits sides
  - Muon exits downstream but does not enter gas TPC
- Acceptance is sensitive to detector modeling in phase space where muon acceptance is rapidly changing
- Uncertainty is evaluated as a function of muon momentum in transverse and neutrino direction (equivalently, energy and angle)

The actual matrix, in the analysis 2D binning



- The ND binning in the fit is two-dimensional in E_ν and y, so the full covariance matrix includes this full binning

Additional LAr sample: ν+e scattering



- Pure EW process with known cross section → sensitive to flux only
- Signal is subject to kinematic constraint $E_e\theta_e^2 < 2m_e$
- Dominant background is ν_e CC at low Q^2
- Signal and background samples are ready, but have yet to be included in fit