DUNE analysis methods & systematic uncertainties

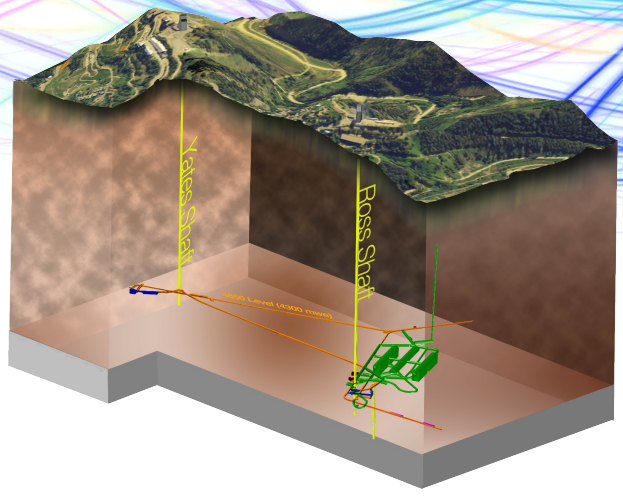
NNN18 – Vancouver
November 2, 2018

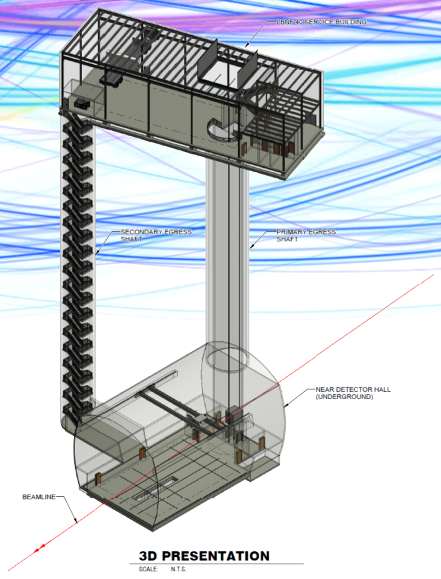
Chris Backhouse
University College London
for the DUNE collaboration

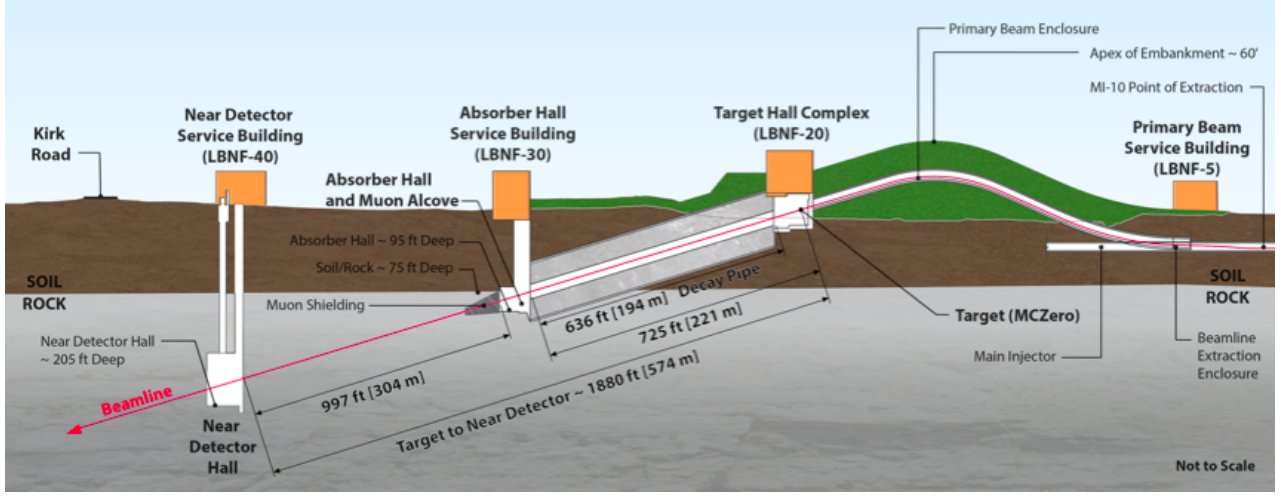
DEEP UNDERGROUND
NEUTRINO EXPERIMENT

Overview

- Analysis methods
  - Reconstruction
  - CVN classifier
- Systematics
  - Flux
  - Cross-section
  - Det. simulation
- Analysis methods
  - Fitter
  - Near Detector
  - DUNEPrism







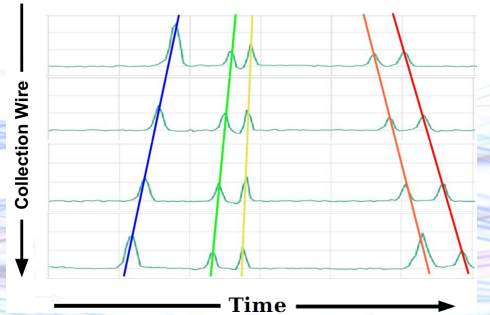
Scope

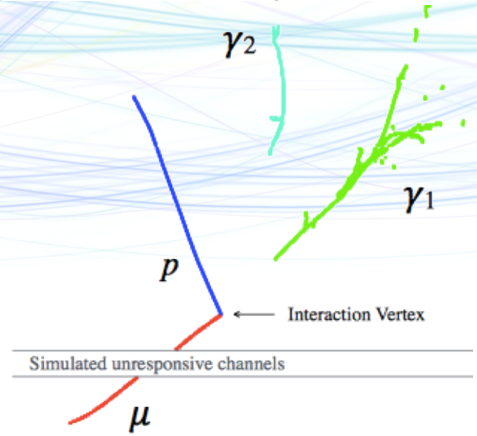
- Aiming for a full end-to-end simulation, reconstruction, systematics treatment for the Technical Design Report

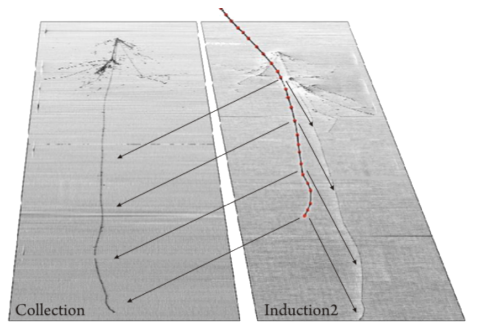
- Full LArSoft Monte Carlo
  - GEANT4 beam simulation
  - GENIE event generator
  - GEANT4 particle propagation
  - Readout simulation with realistic waveforms and noise – MicroBooNE/ProtoDUNE experience
- Automated signal processing and hit finding
- Automated reconstruction and event classification

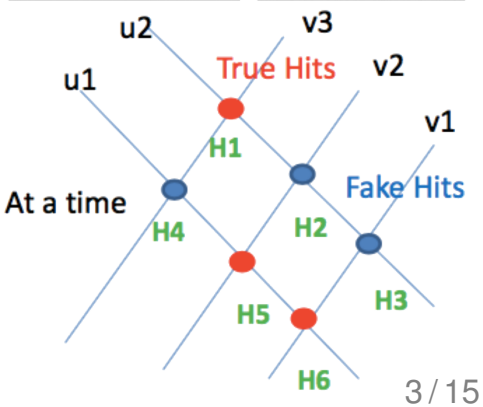
Reconstruction

- Hit finding is a whole talk to itself – **Hannah**
- Pandora – multi-algorithm approach
- Project matching
- 2D→3D unfolding
- . . .
- Being given a workout on ProtoDUNE

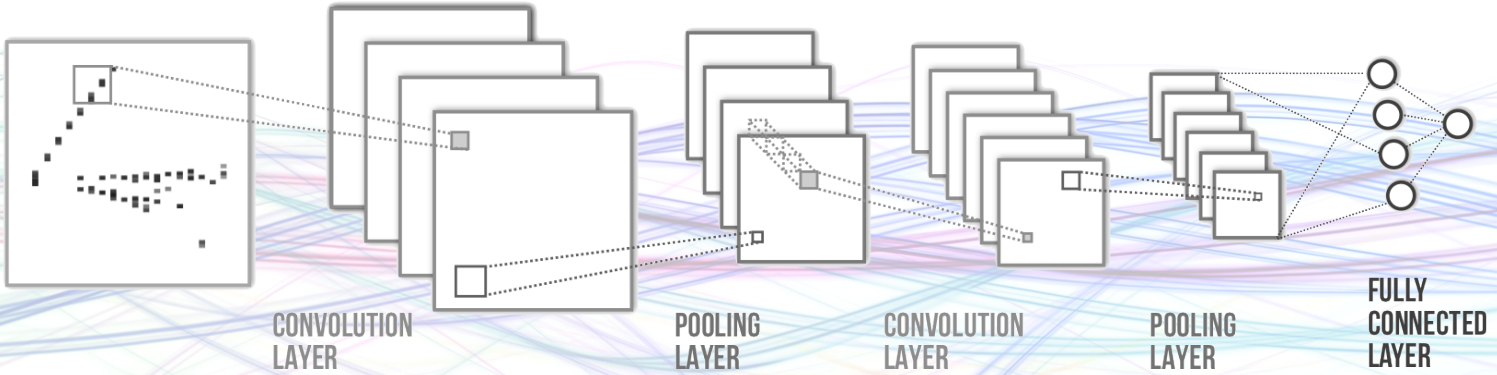


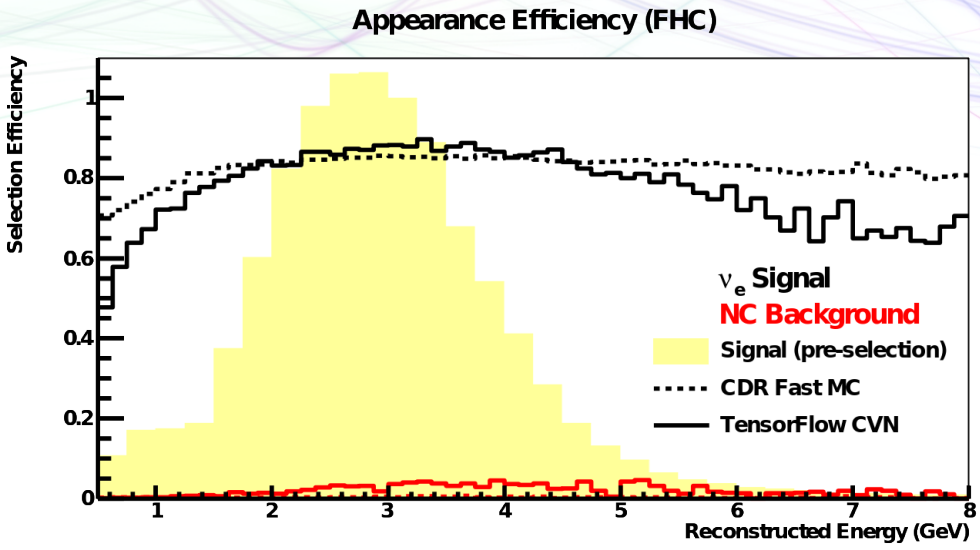


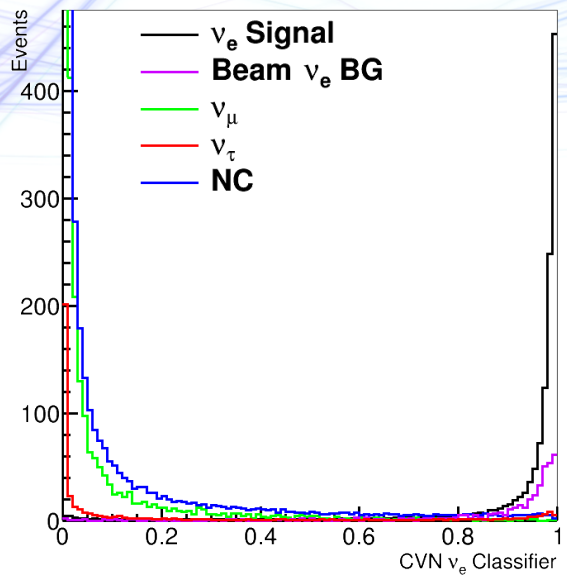




CVN



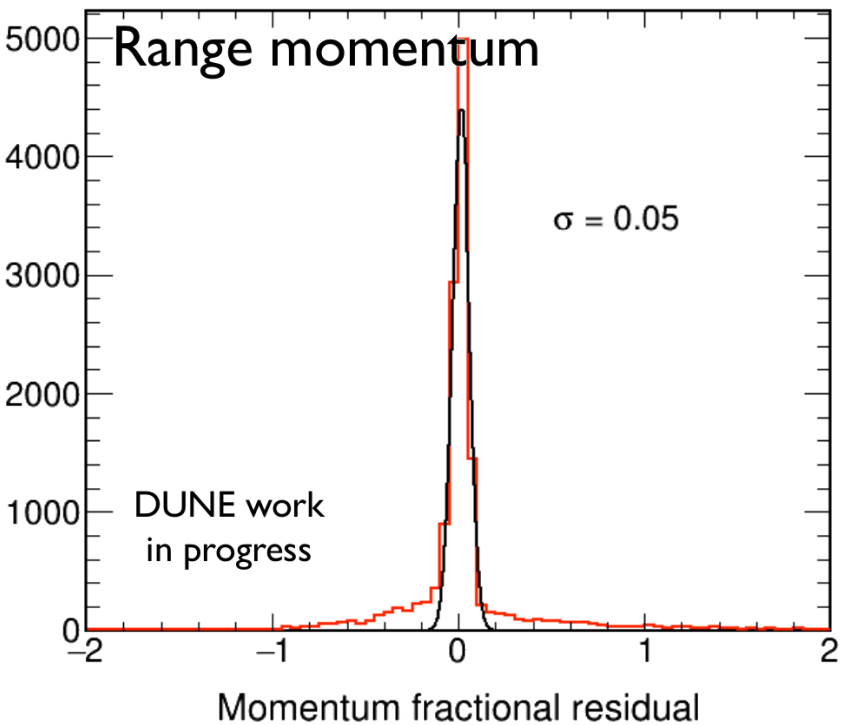


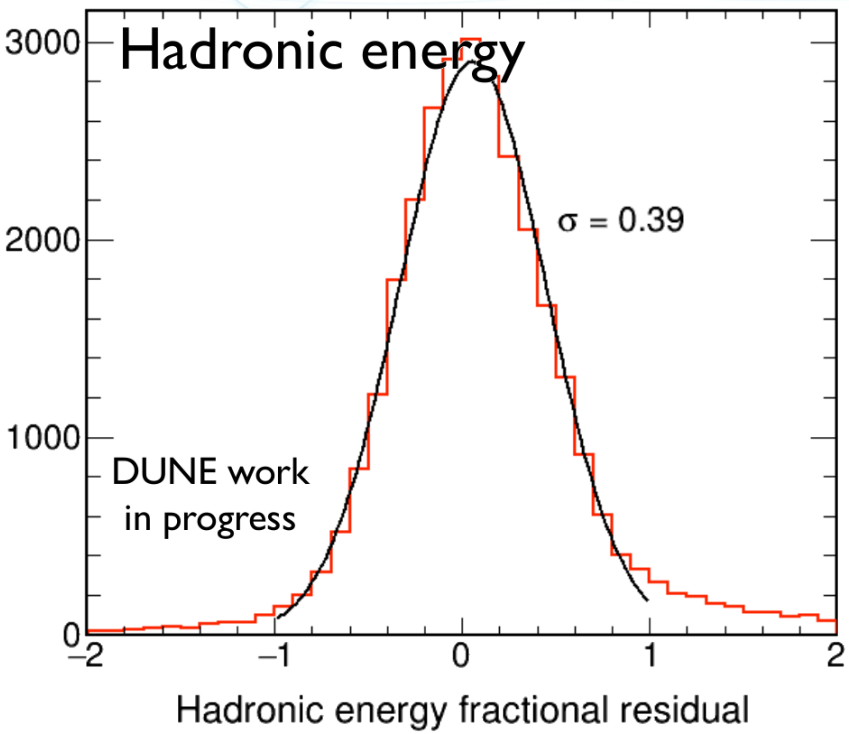


- See **Kazu**'s overview from Thursday
- ResNet architecture in TensorFlow
- Train with 500 $\times$ 500 MC images
- Currently best-performing classifier
  - Conventional techniques not fully exhausted yet
- Investigate systematic dependence just like other classifiers
- Also hit-by-hit CNN as an input to traditional reconstructions

Energy reconstruction

- $E_{\text{reco}} = E_{\text{lep}} + E_{\text{had}}$
- Muon energy from range
  - Use multiple coulomb scattering if uncontained
- Electron and hadronic energies estimated calorimetrically
- Corrections for recombination, electron lifetime, invisible energy





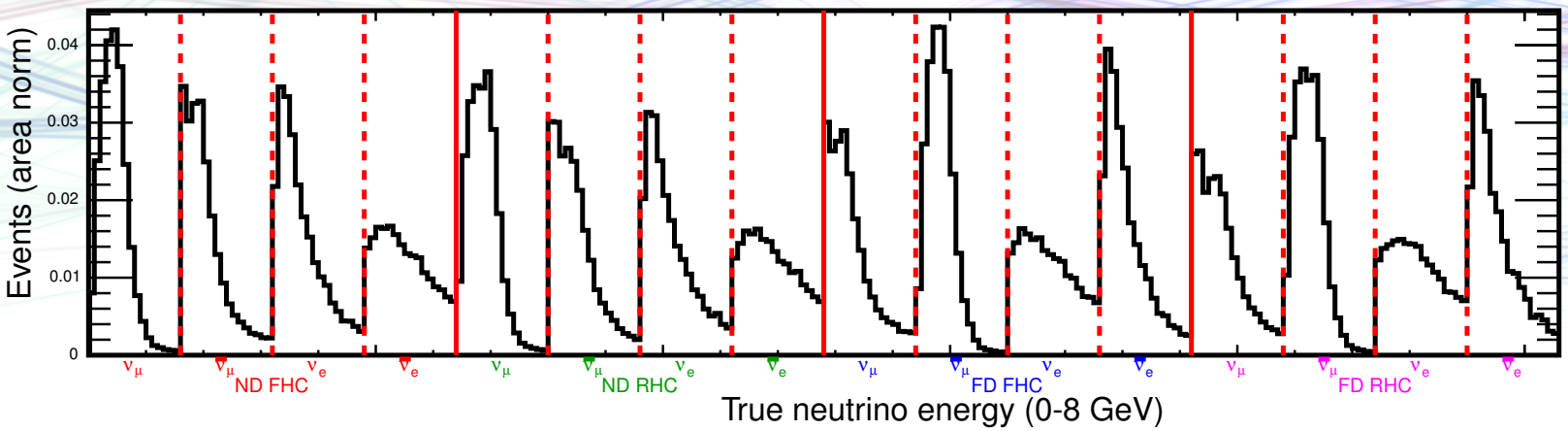
- Crude, but realistic
- Can be elaborated in future (track individual hadronic particles?)

Systematics

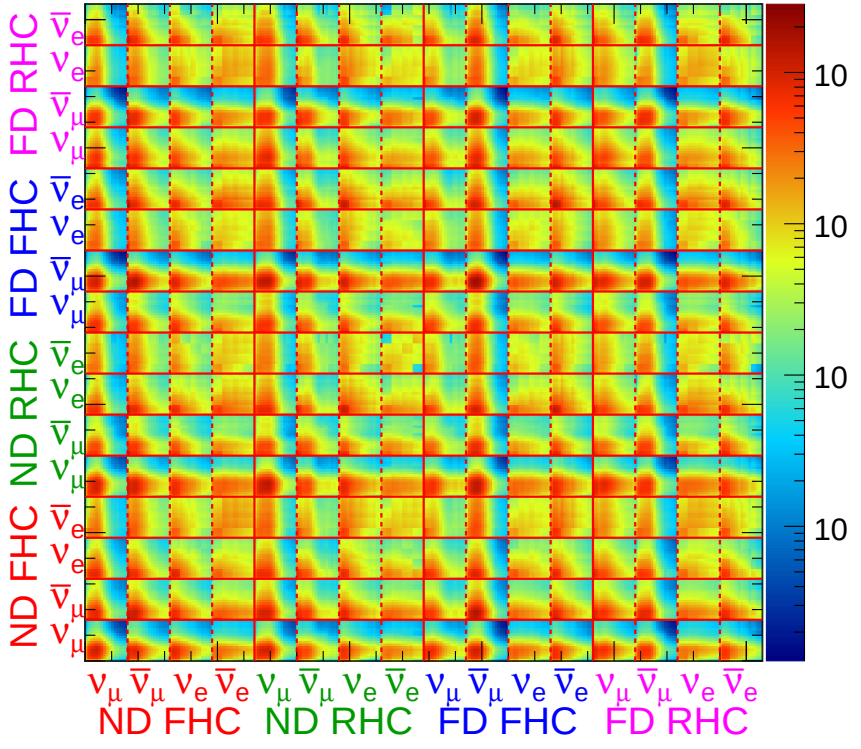
- Oscillation parameters from comparing data with FD Monte Carlo
- How different can the FD prediction reasonably be?
- Need ability to generate variant predictions
- Many systs are reweightable, a few require separate samples

- Constraints on systematic parameters from
  - External inputs
  - ND measurements
- Try to be robust, *i.e.* insensitive to small weaknesses in the systematic model

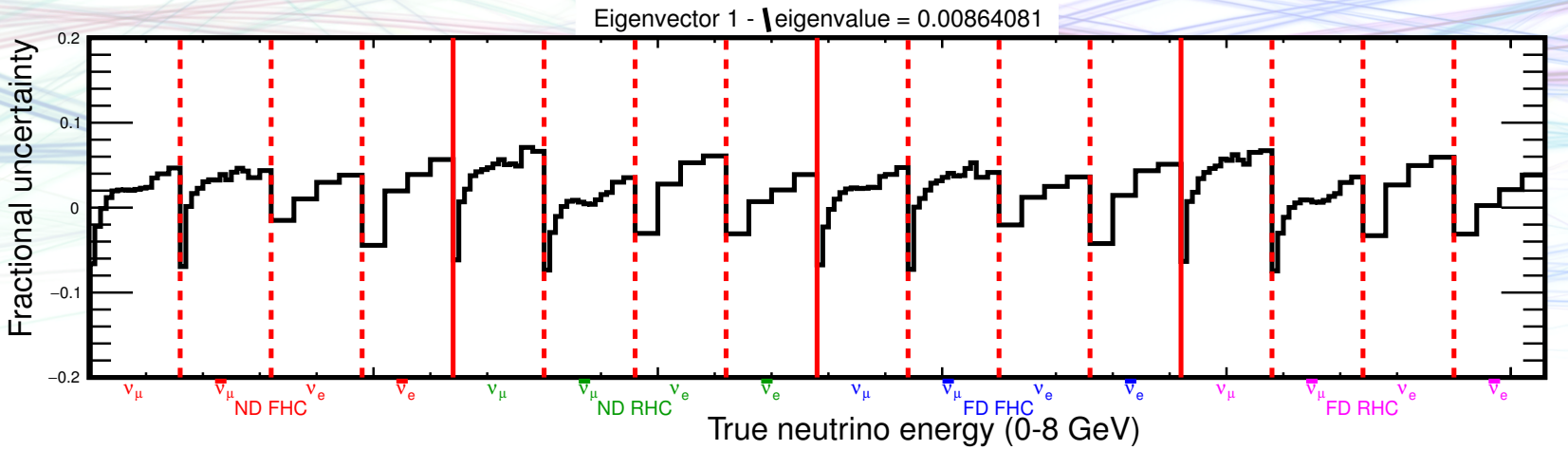
Flux



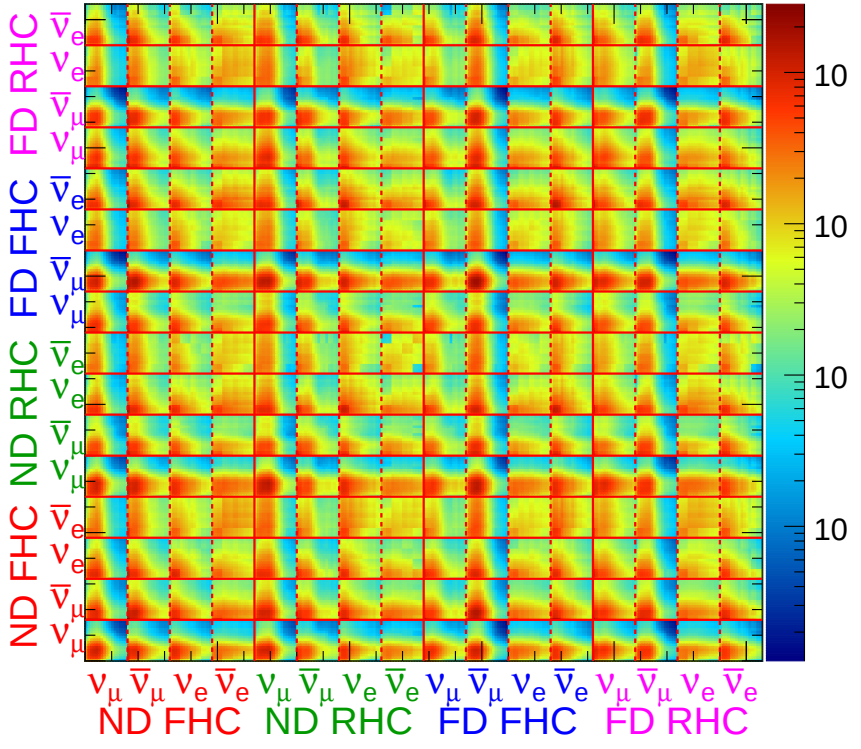
- Uncertainties from hadron production and focusing
- Encoded in a covariance matrix connecting flavours, detectors, beam modes
- Use principal component analysis to find smaller basis that covers most of the effect



Flux



- Uncertainties from hadron production and focusing
- Encoded in a covariance matrix connecting flavours, detectors, beam modes
- Use principal component analysis to find smaller basis that covers most of the effect



Cross sections

- Informed by NOvA, T2K, MINERvA groups
- Aim to provide full set of orthogonal knobs

GENIE dials (v2.12.10c)

- Default priors where they don't double count
- Plus. . .

QE-like

- Z-expansion axial FF
- MINERvA's 2p2h enhancement (low recoil data)[1]
- 2p2h energy dependence – MINERvA/DUNE energies not equal

Low-W

- Swap MK model for Rein-Sehgal – interference of RES+non-RES
- Empirical fit to low Q^2 suppression for RES needed by NuMI expts

[1] https://arxiv.org/abs/1511.05944

Cross sections

High-W

- Uncorrelated normalization uncertainties for non-resonant pion production for 1,2,3+ pions, up to W=5 GeV

FSI-like

- Inflation of smearing of E_{avail} reflecting C$\rightarrow$Ar

Other

- Potential $\nu_e/\bar{\nu}_e$ xsec differences
- ν_μ/ν_e differences from lepton mass differences[2]

- Combination of smaller effects can be treated with PCA
- Ideally also swap in entirely different models

[2] Phys.Rev. D86 (2012) 053003

Detector effects

- Have a lot of handles for calibration to characterize the detector
- But there will always be some residual uncertainty
- Least likely to cancel between detectors

- Most of these require independent samples to be simulated

Actively pursuing

- *E*-field distortions
- Alignment

Others

- Calibrations – absolute scale, channel-to-channel variations
- Dead channels
- Neutron-Ar cross-section
- . . .

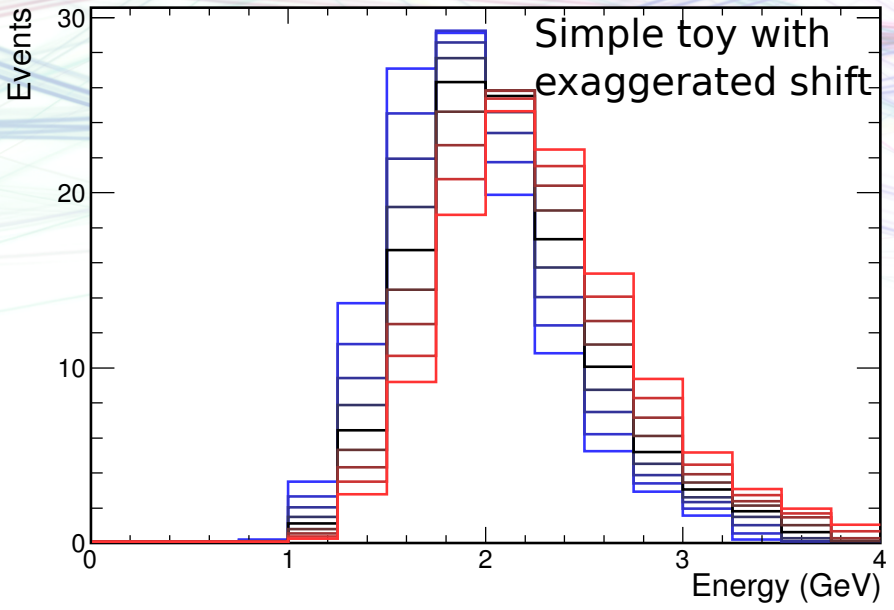
Fitter

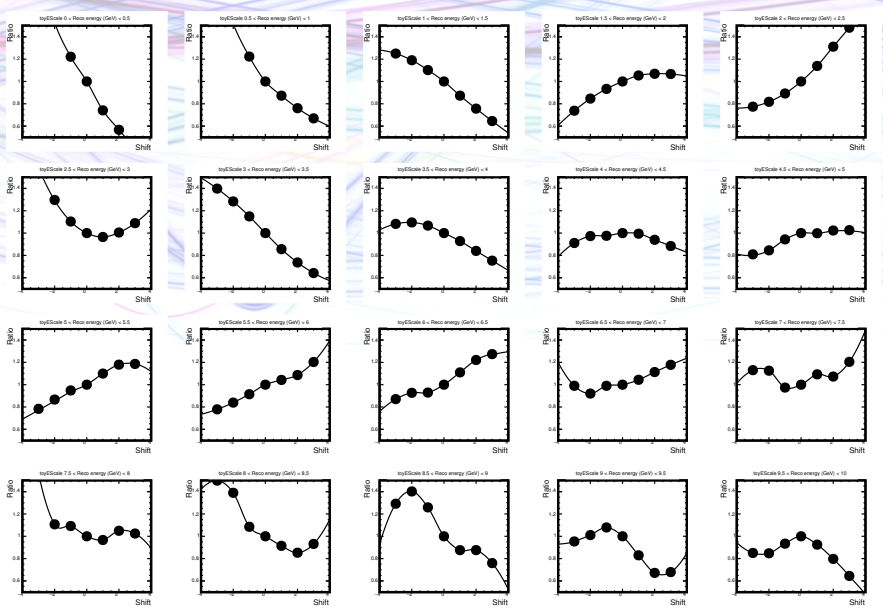
- For speed, *CAFAna* fit is based on reco-vs-true templates for each oscillation channel
- To oscillate, reweight with $P_{\nu_\alpha \to \nu_\beta}(E_{\text{true}})$ and sum

- Systematically-shifted matrices required for systematics fit
- Reweights, rewrite event record, or specially-generated samples

- Profile systematic pulls and subdominant oscillation parameters

Template interpolation



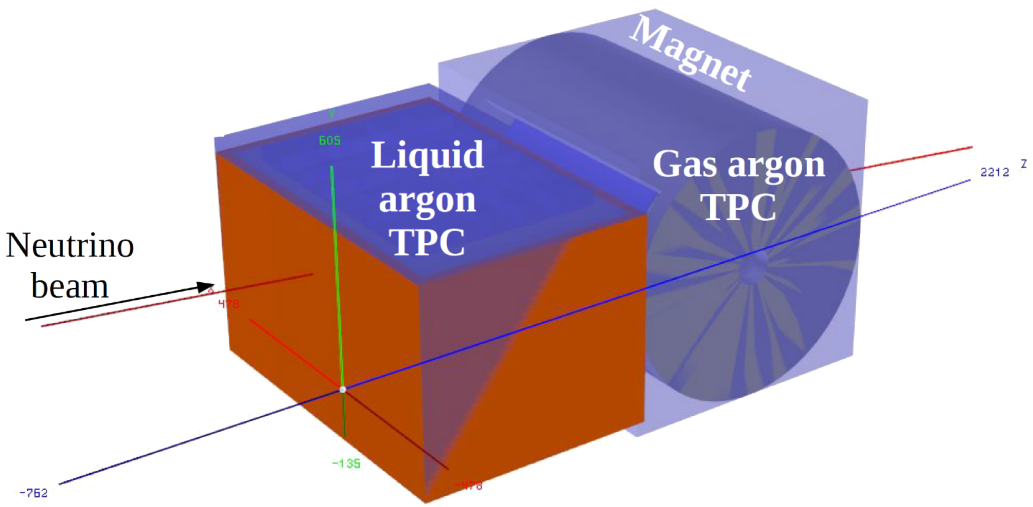


- Covariance matrix equivalent to pull terms for linear bin changes
- Direct pull term approach can deal with non-linearity
- Cubic interpolation (differentiable) between templates

- Alternate approaches available:
- VALOR – covariance matrices, more explicit about correlations
- GLoBES – parameterized, accesible to outside community

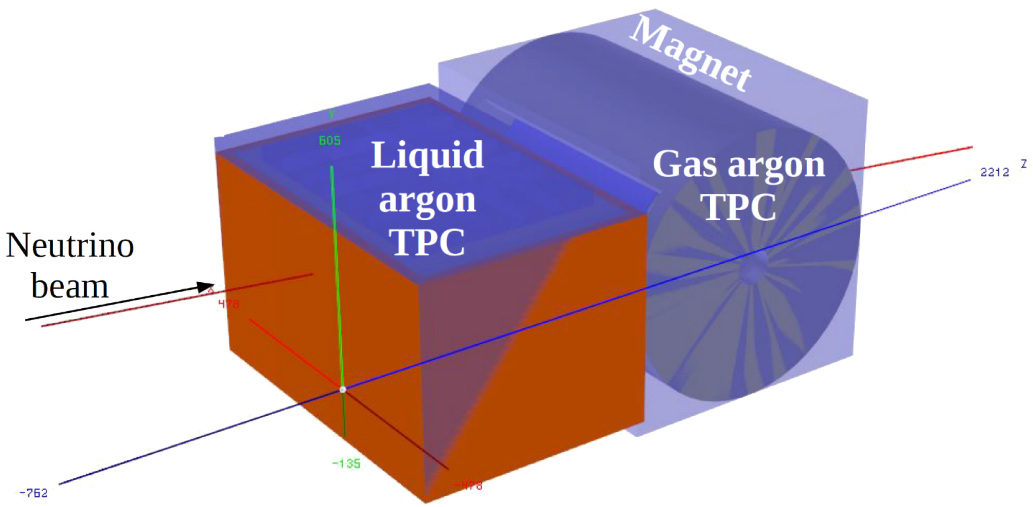
Near Detector

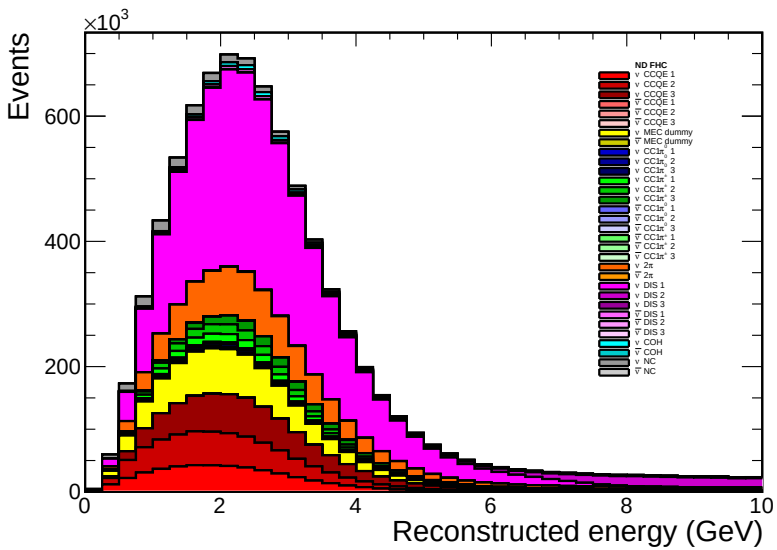
- Systematics not just constrained with external priors
- Use Near Detector to measure exactly what we want
- Similar flux $\times$ xsec $\times$ eff $\rightarrow$ correlations $\sim$ cancel in FD
- Rival philosophies – match FD or detailed study of nuclear effects
- Do both! Unmagnetized LAr plus magnetized low-density tracker
- See **Chris**' talk tomorrow for much more



Near Detector

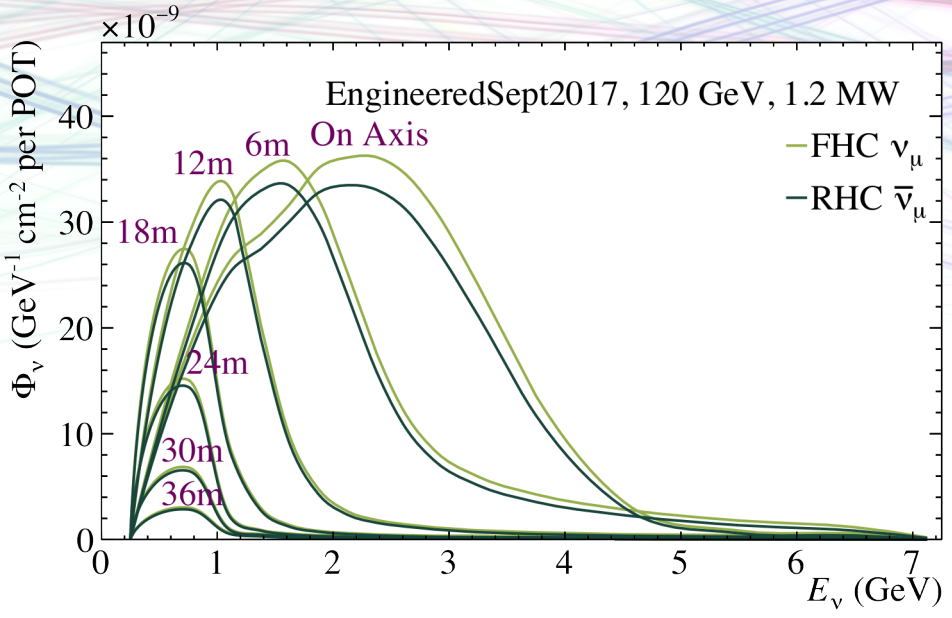
- Systematics not just constrained with external priors
- Use Near Detector to measure exactly what we want
- Similar flux $\times$ xsec $\times$ eff $\rightarrow$ correlations $\sim$ cancel in FD
- Rival philosophies – match FD or detailed study of nuclear effects
- Do both! Unmagnetized LAr plus magnetized low-density tracker
- See **Chris**' talk tomorrow for much more

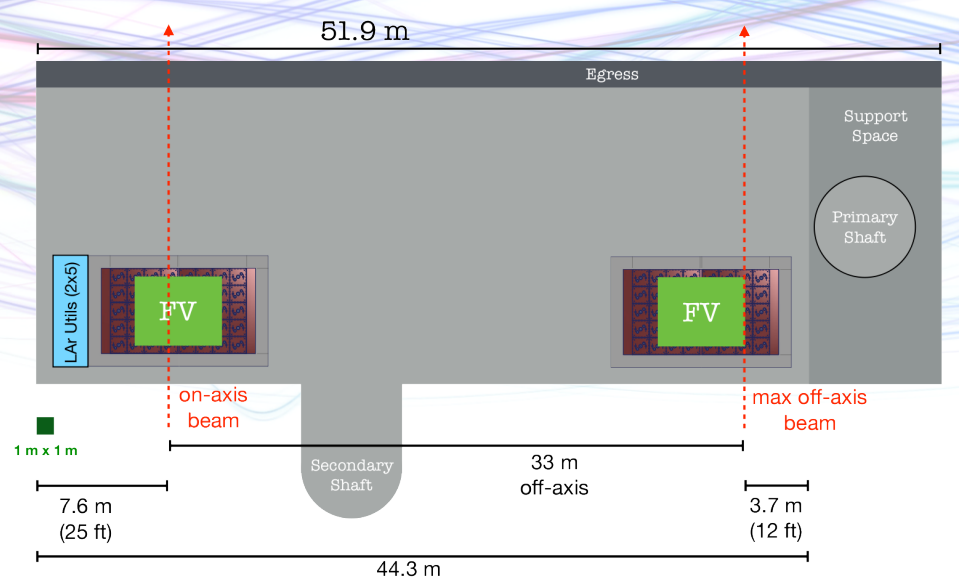




- Careful not to overconstrain model with $\sim$ infinite ND stats
- Start by including simple inclusive energy spectrum in fit
- Will add additional samples carefully as need becomes apparent

DUNEPrism

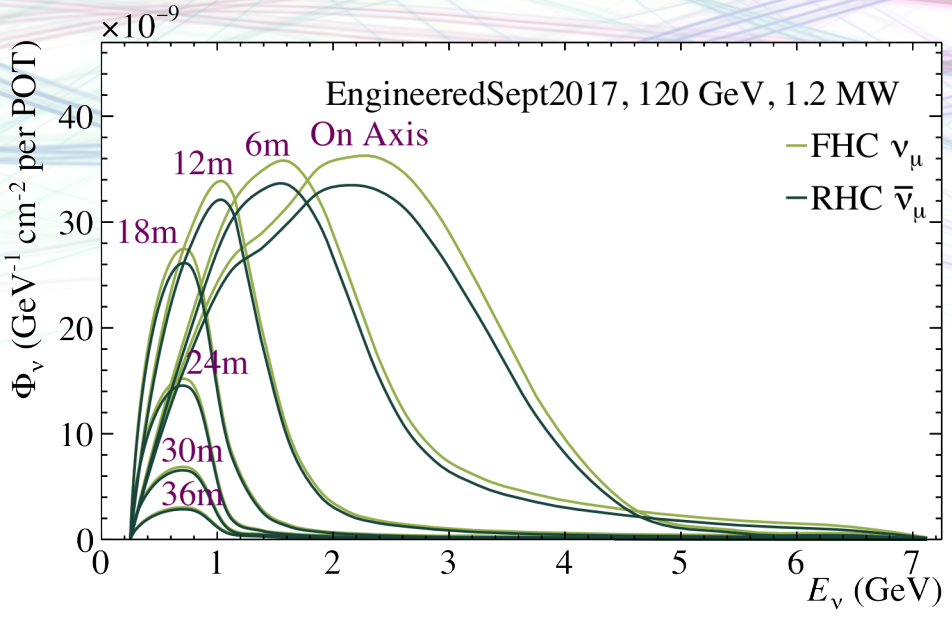


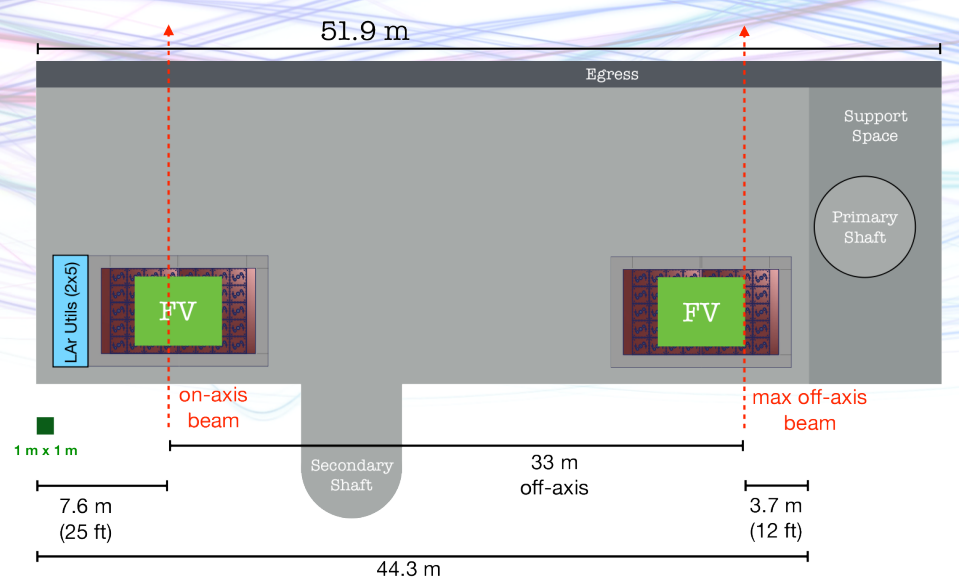


Possible layout

- Viewing flux from off-axis locations provides additional information
- Benefits convincing enough that it's now in the baseline plan

DUNEPrism

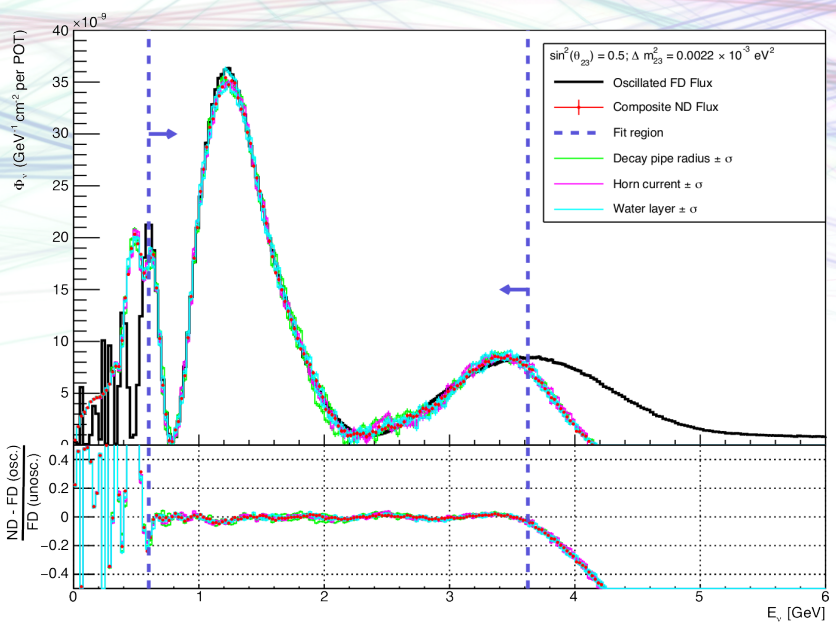


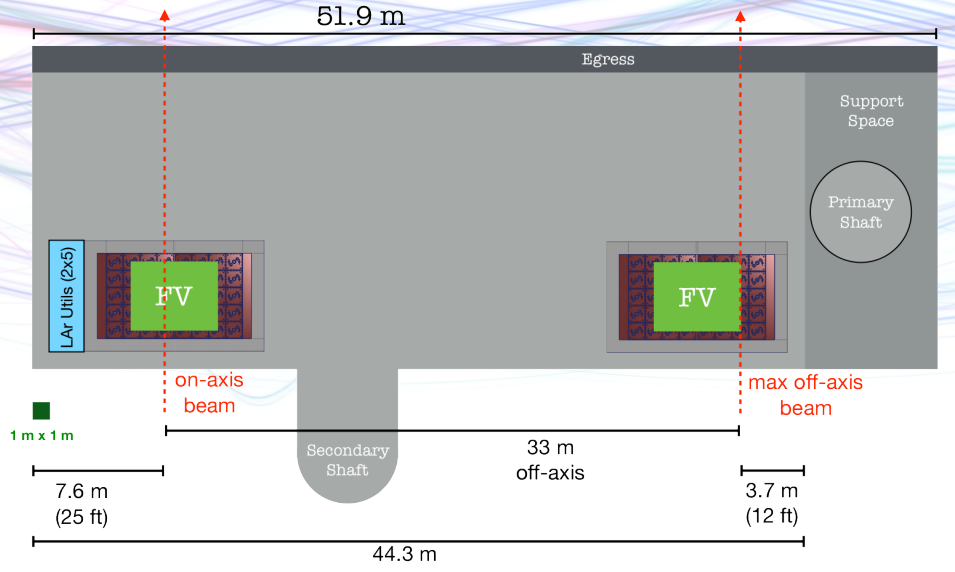


Possible layout

- Viewing flux from off-axis locations provides additional information
- Benefits convincing enough that it's now in the baseline plan
- In principle also able to deal with unknown unknowns
- Can imagine a conspiracy that alters reco vs true but leaves all on-axis observables unchanged

DUNEPrism





Possible layout

- Viewing flux from off-axis locations provides additional information
- Benefits convincing enough that it's now in the baseline plan
- In principle also able to deal with unknown unknowns
- Can imagine a conspiracy that alters reco vs true but leaves all on-axis observables unchanged
- Direct "extrapolation" summing ND spectra, no fit parameters

Conclusion

- Working towards a full end-to-end analysis
  - Full simulation
  - Real reconstruction and PID
  - Sophisticated flux, xsec, detector systematics
  - Full fit
- Focus on robustness of our conclusions
- Can always add complexity later

- Near Detector design taking shape
- Sophisticated systematic studies required for guidance

- Things are looking good!


Thank you!
DUNE

Backup

Interpolation

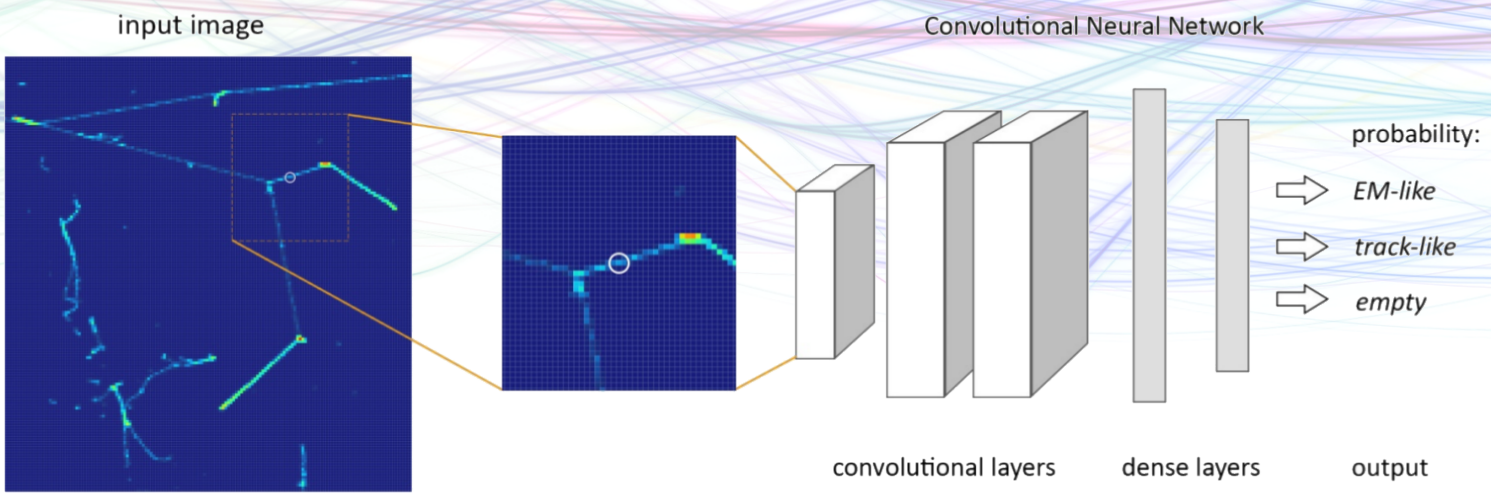
- Assumptions: combination of systs is product of their effects, effects on bins independent of osc pars

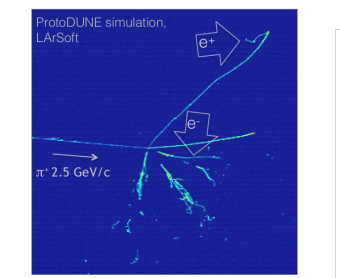
$$p_i(\vec{\theta}, \vec{s}) = p_i(\vec{\theta}, \vec{0}) \prod_j f_{ij}(s_j)$$

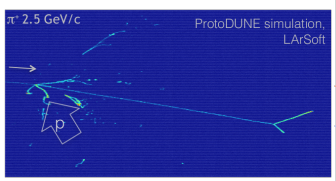
where

$$f_{ij}(N) = \frac{p_i(\vec{\theta}_0, [0, \ldots, N, \ldots, 0])}{p_i(\vec{\theta}_0, \vec{0})}$$

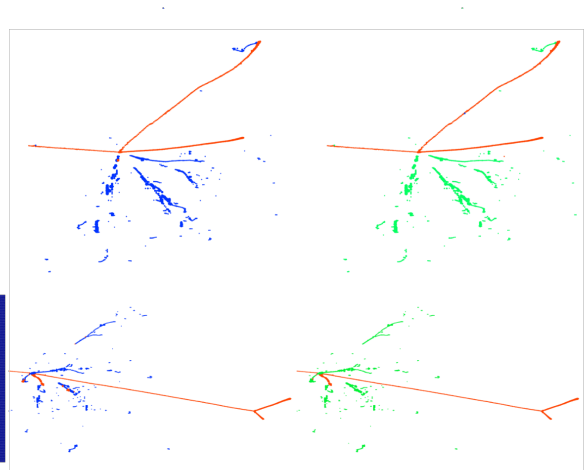
CNN for reconstruction







input: 2D ADC



CNN output:
EM-like (blue) / track-like (red)

MC truth:
EM-like (green) / track-like (red)

- Can apply related techniques to other parts of the analysis
- Here classify hits as track / shower as input to std. reco