Detector Variation Studies

Elizabeth Worcester (BNL)

Calibration Task Force Meeting

October 16, 2018

Introduction

- Goal is to study impact of variations in detector performance/ specifications on CPV analysis
- List of variations chosen based on importance and ease of simulation:
  - **Dead channels** } This talk
  - **Drift field magnitude**
  - Drift field non-uniformities   Simulation planned/in progress
  - Noise   To do
- Method is to re-run simulation changing some parameter and repeat nominal CPV analysis using full MC simulation, nominal energy reconstruction and CVN event selection
  - Energy reconstruction not re-tuned, crude energy scaling is applied (more later)
  - CVN is not re-trained
  - CVN selection cut is re-optimized
  - Second order effects (eg: impact on systematics) not considered

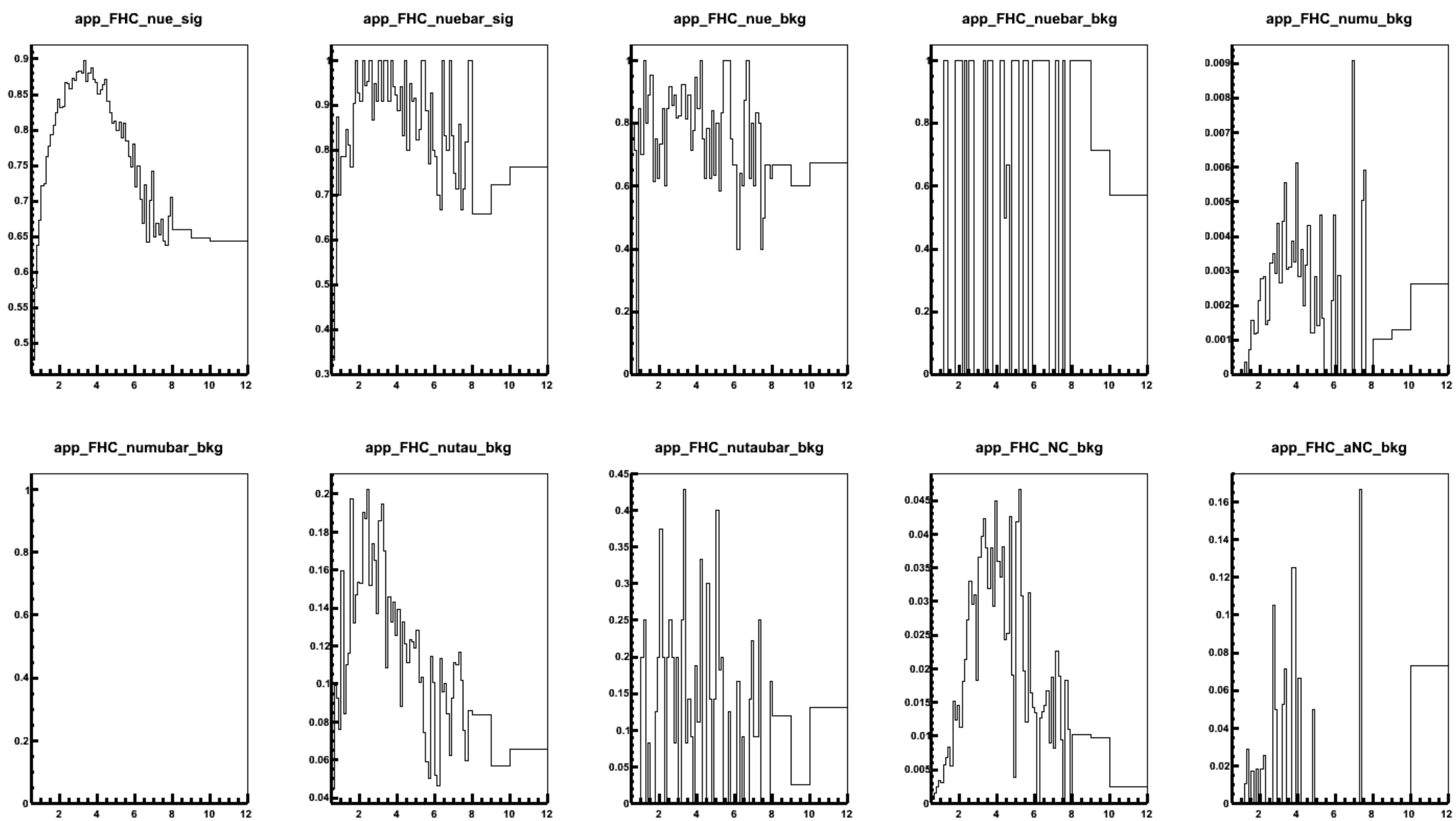
Warning

- Do not over-interpret! These studies can give a sense of what specifications are important for the CPV analysis, but, without fully repeating the analysis including reconstruction tuning and network training and without including the impact on our ability to calibrate and constrain systematics, they do not tell anything like a full story!

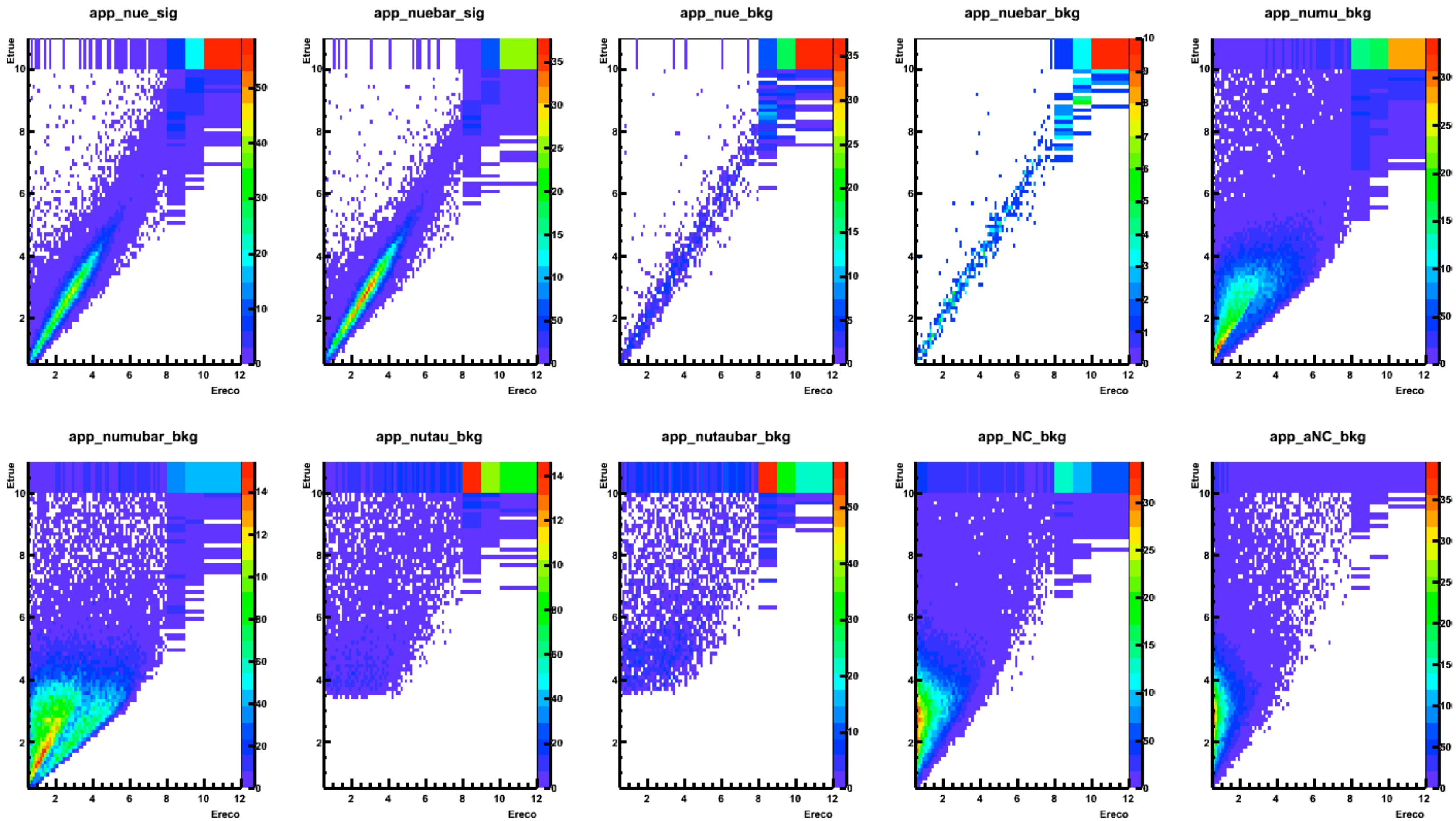
Analysis Details

- Full LArSoft Monte Carlo simulation
  - GENIE event generator
  - GEANT4 particle propagation
  - Detector readout simulation including "realistic" waveforms and white noise
- Automated signal processing and hit finding
- Automated energy reconstruction
  - Muon momentum from range (contained) or multiple Coulomb scattering (exiting)
  - Electron and hadron energy from calorimetry
- Event selection using convolutional visual network (CVN)
- MC truth used to produce efficiency histograms and true-reco smearing matrices that are used as input to GLoBES analysis
- GLoBES analysis identical to CDR (with the efficiency and smearing coming from previous step) performed to obtain CPV sensitivity
  - Includes normalization systematics chosen to represent residual uncertainties after ND constraints and sample-sample cancellations applied

Example Efficiency



Example Smearing



Analysis Input Details

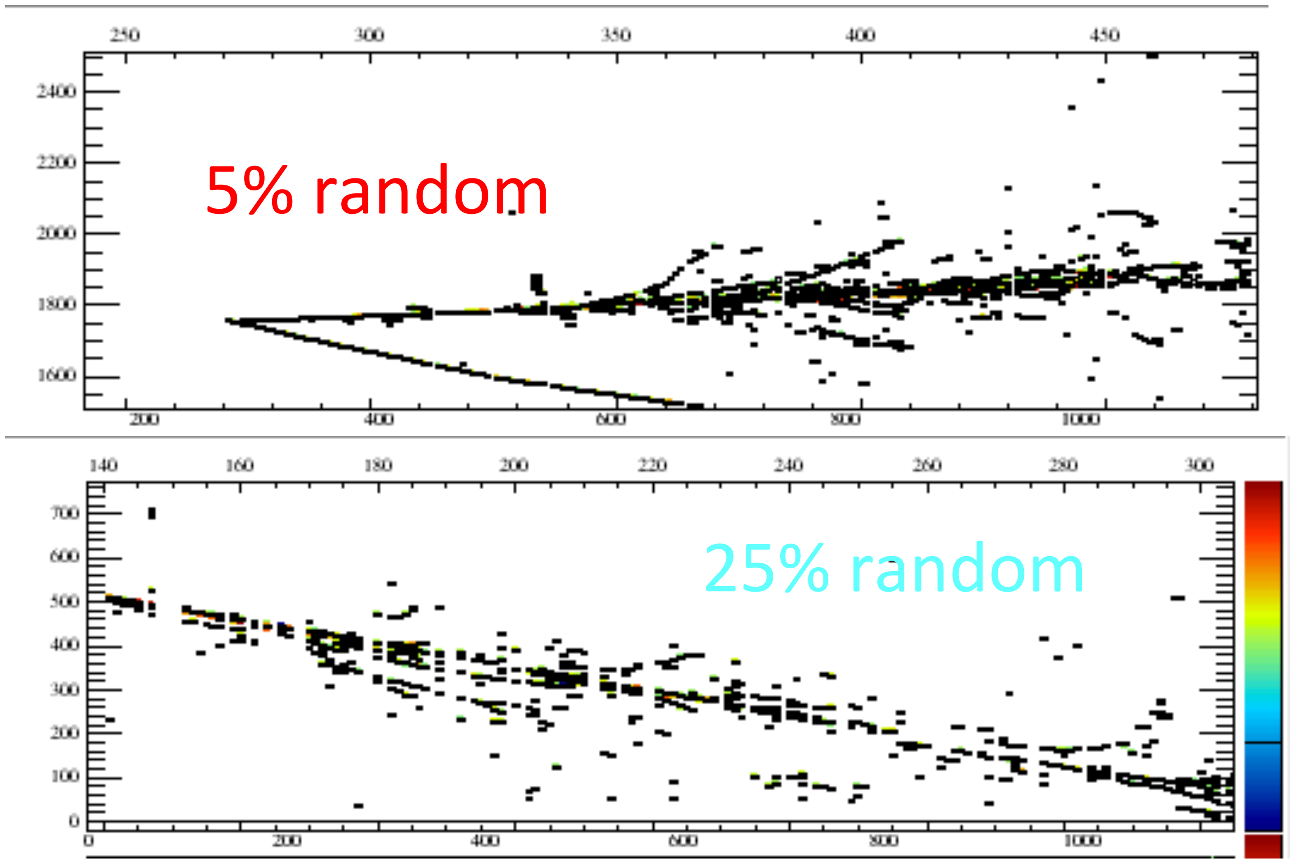
- Bad channel study:
  - MCC10.1
  - Remove 5% or 25% of all channels (Backhouse)
    - At random
    - Grouped by chip (groups of 16)
    - Grouped by board (groups of 128)
    - For groupings, ~follow protoDUNE mapping scheme
- Drift field study:
  - May 2018 CVN update (improved efficiency relative to MCC10.1)
  - Reduce the drift field in sim/reco fcl files
  - Generate new field response histograms given reduced field (Garfield, thanks Yichen!)

```
services.DetectorPropertiesService.Efield: [0.25, 0.333, 0.4]
services.SignalShapingServiceDUNE.FieldResponseFname: "dune_response_250vpercm.root"
```

  - Compare to nominal settings:

```
dunefd_detproperties.Efield:          [0.5,0.666,0.8]  #(predicted for microBooNE)
```

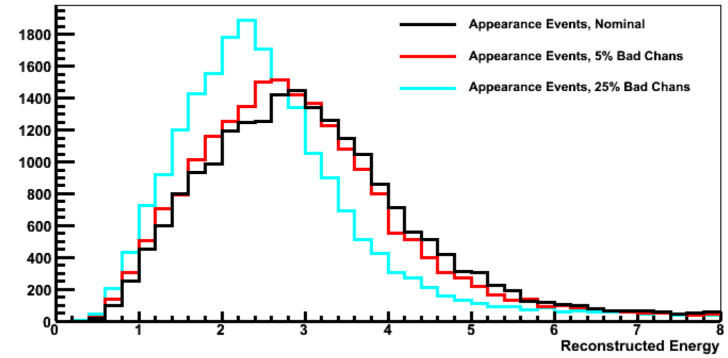
Example Event Displays: Bad Channels

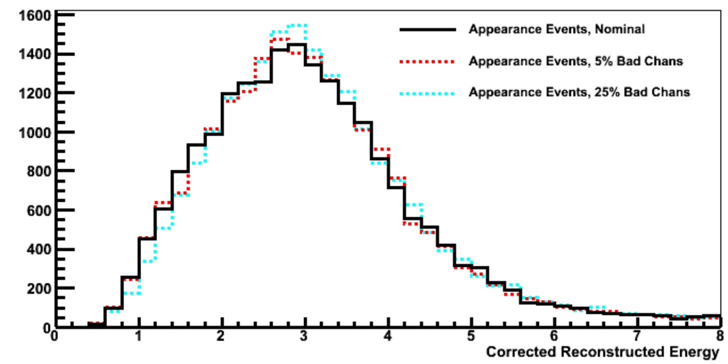


Energy Correction Details

- Run only nominal energy reconstruction, which is calorimetric for EM showers
  - Expect to be missing ~5% or 25% of energy for random bad channels
  - Scaling reconstructed energy by 1.05 or 1.25 approximately corrects back to nominal reconstructed energy as expected
- For studies where expectation is less obvious (ie: all the others) I just scale to force a Gaussian fit to the peak to match nominal
- Note: scaling performed at plot level only, does not impact event selection

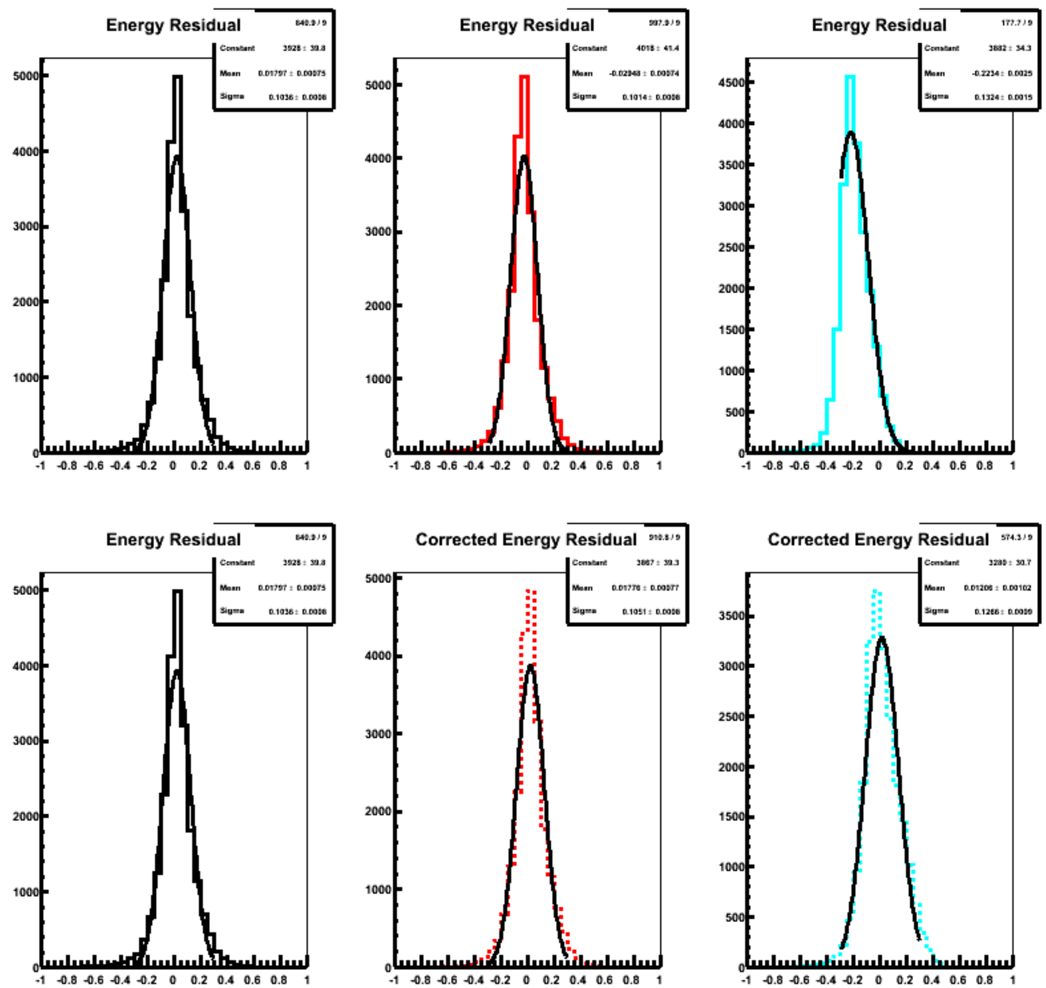
Example for randomly distributed bad channels:





Energy Resolution Crosscheck

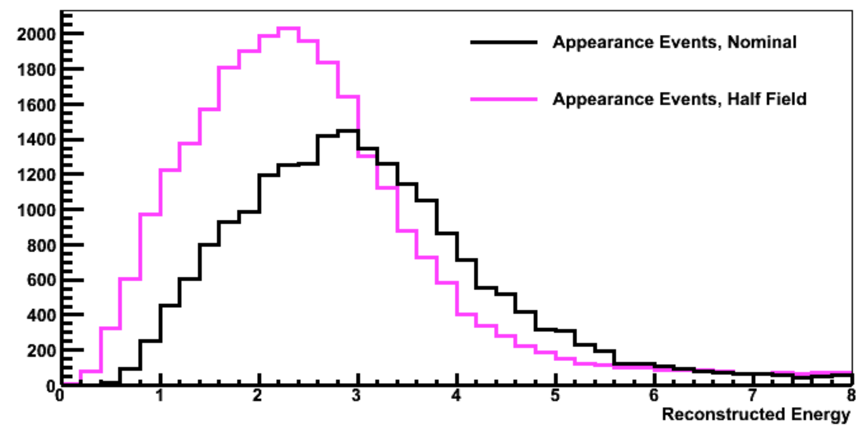
- Check fractional energy residuals for each case
- See bias as expected in uncorrected residuals
- Nominal resolution similar to that reported by N. Grant
- See some increase in resolution for 25% missing channels case

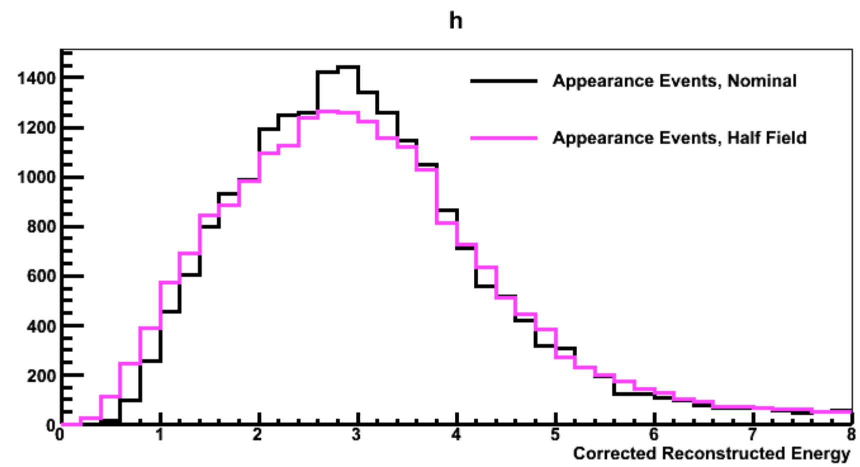


Nominal | 5% Random | 25% Random

Effect of Drift Field on E_{reco}

- Effects simulated in MC:
  - Larger diffusion (drifting for longer)
  - Larger loss due to attachment (drifting for longer)
    - Of course this depends on purity – we simulate 3 ms electron lifetime
  - Larger loss due to recombination (Birks' Law)
- Using https://lar.bnl.gov/properties/) for dE/dx = 4 MeV/cm, 3.6 m drift distance, nominal (3 ms) electron lifetime:
  - At 500 V/cm: $R_a = 0.48$, $R_c = 0.63$
  - At 250 V/cm: $R_a = 0.33$, $R_c = 0.51$
- Calorimetric energy reconstruction uses average $R_c = 0.63$
  - Expect to reconstruct around 25% less energy at lower field
- Lifetime correction is handled correctly
- Impact of diffusion on E_{reco} should be small?





PID Cut Tuning

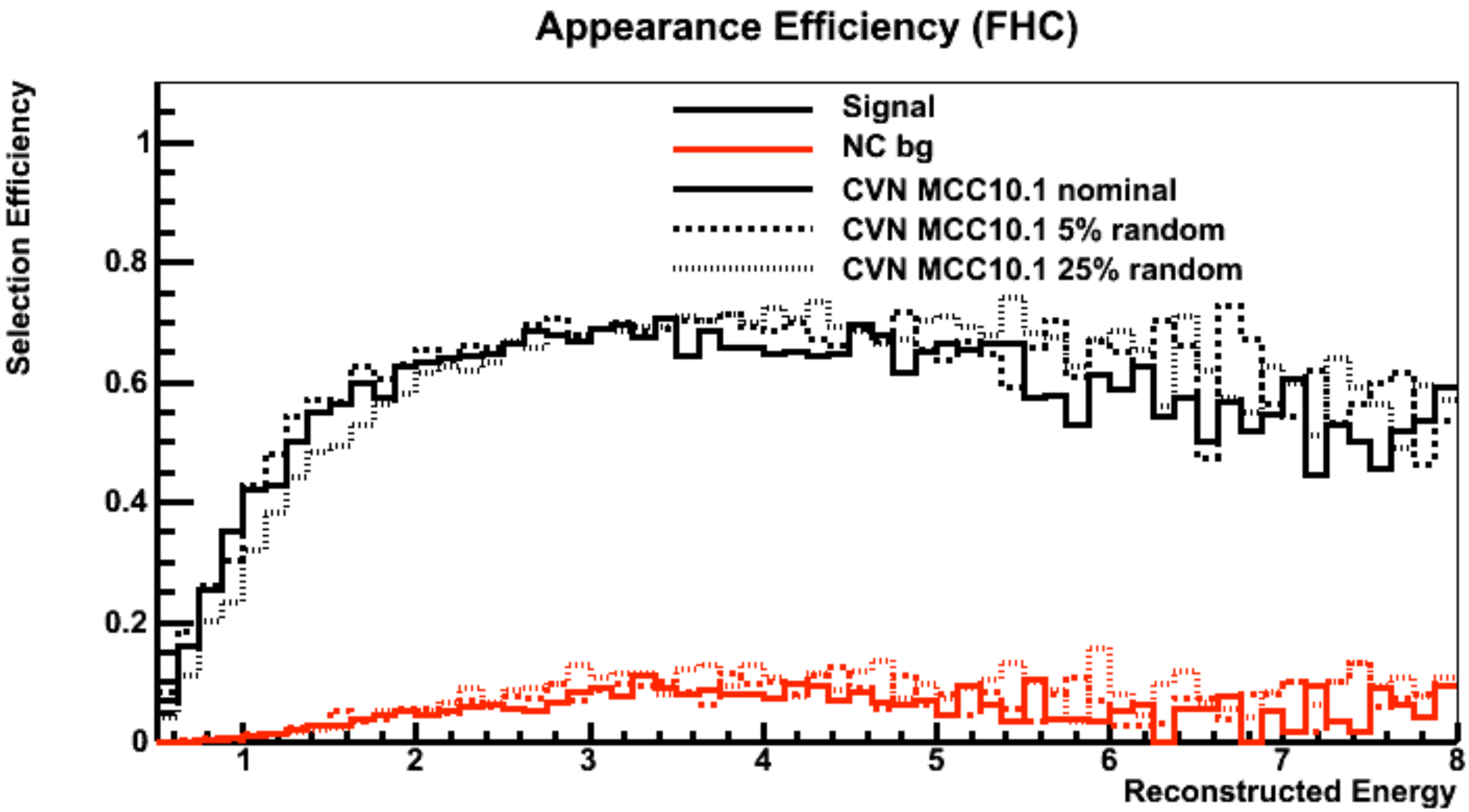
- Method:
  - Set numu PID cut at 0.5 (no tuning)
  - Require numu PID < 0.5 and nue PID > x for nue event to be selected (and inverse for numu events)
  - Scan through nue PID cut values in steps of 0.05
  - Evaluate CPV sensitivity at $\delta_{CP} = -\pi/2$
  - Define optimum cut value as the one before sensitivity starts to decrease (so might miss 2nd optimum value if sensitivity is not monotonic above and below optimum)
- Resulting optimized nue PID cut (CVN)
  - Nominal: 0.7
  - 5% bad channels: 0.65
  - 25% bad channels: 0.6
- With more bad channels, prefer cut with slightly higher acceptance such that signal efficiencies are more similar than with un-tuned cuts
  - Balanced by increasing background with looser cuts

Summary of Corrections/Optimization

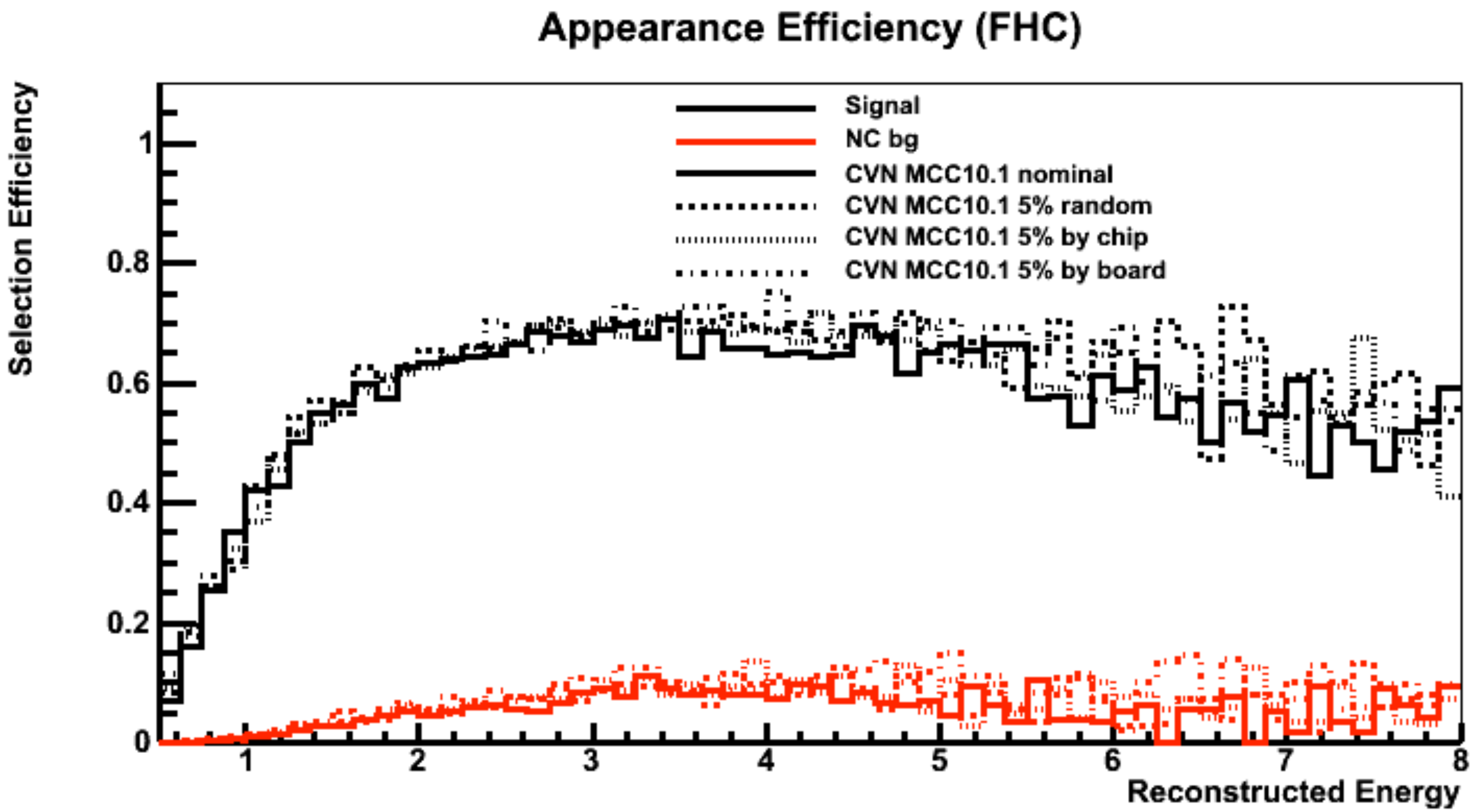
Sample	Energy Scale Factor	Optimized ν_e Cut
MCC10.1 Nominal	1.0	0.7
5% Bad – Random	1.05	0.65
5% Bad – By Chip	1.03	0.65
5% Bad – By Board	1.02	0.6
25% Bad – Random	1.25	0.6
25% Bad – By Chip	1.21	0.55
25% Bad – By Board	1.13	0.45

Sample	Energy Scale Factor	Optimized ν_e Cut
May 2018 Nominal	1.0	0.85
250 V/cm	1.264	0.8

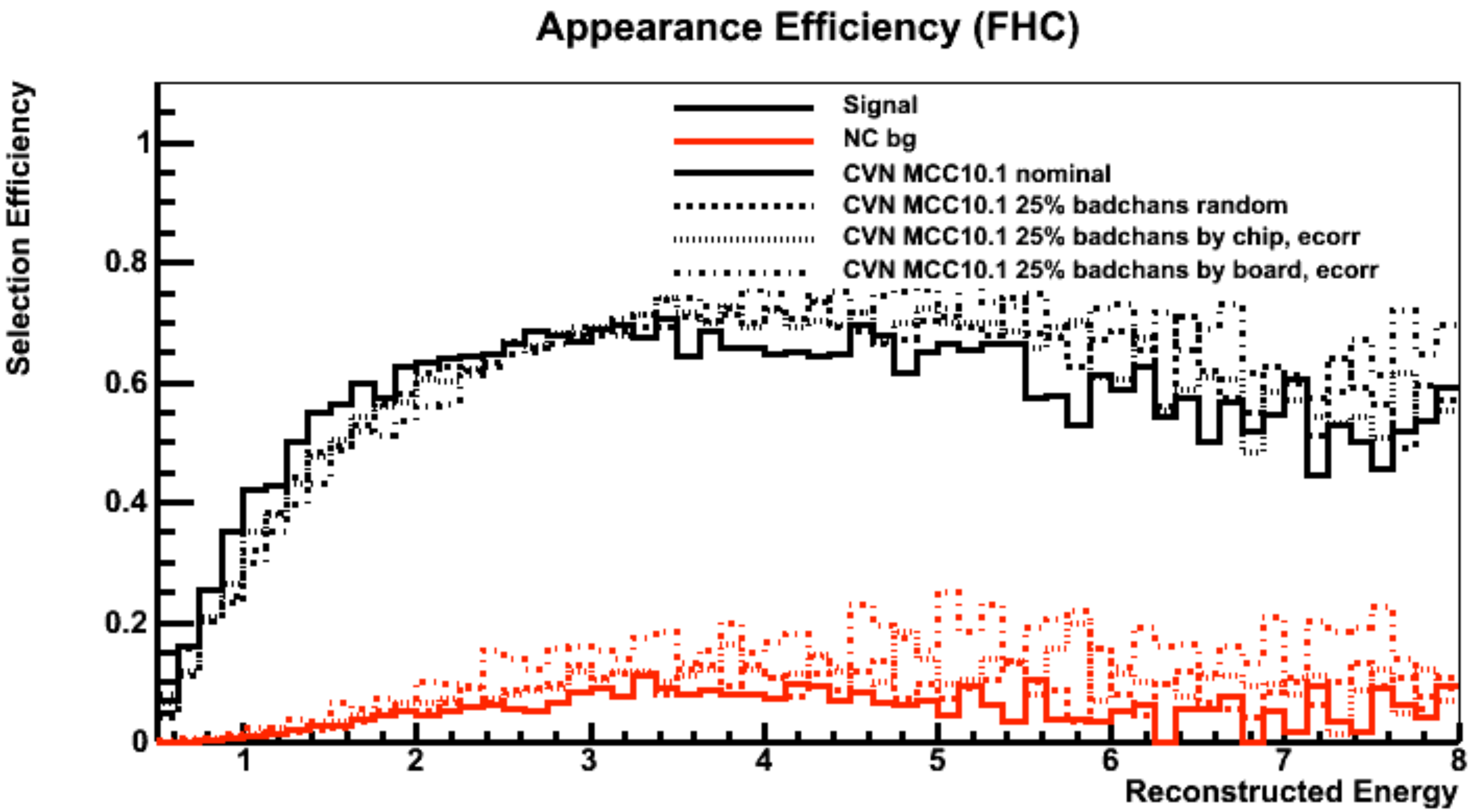
Efficiency: Random Bad Channels



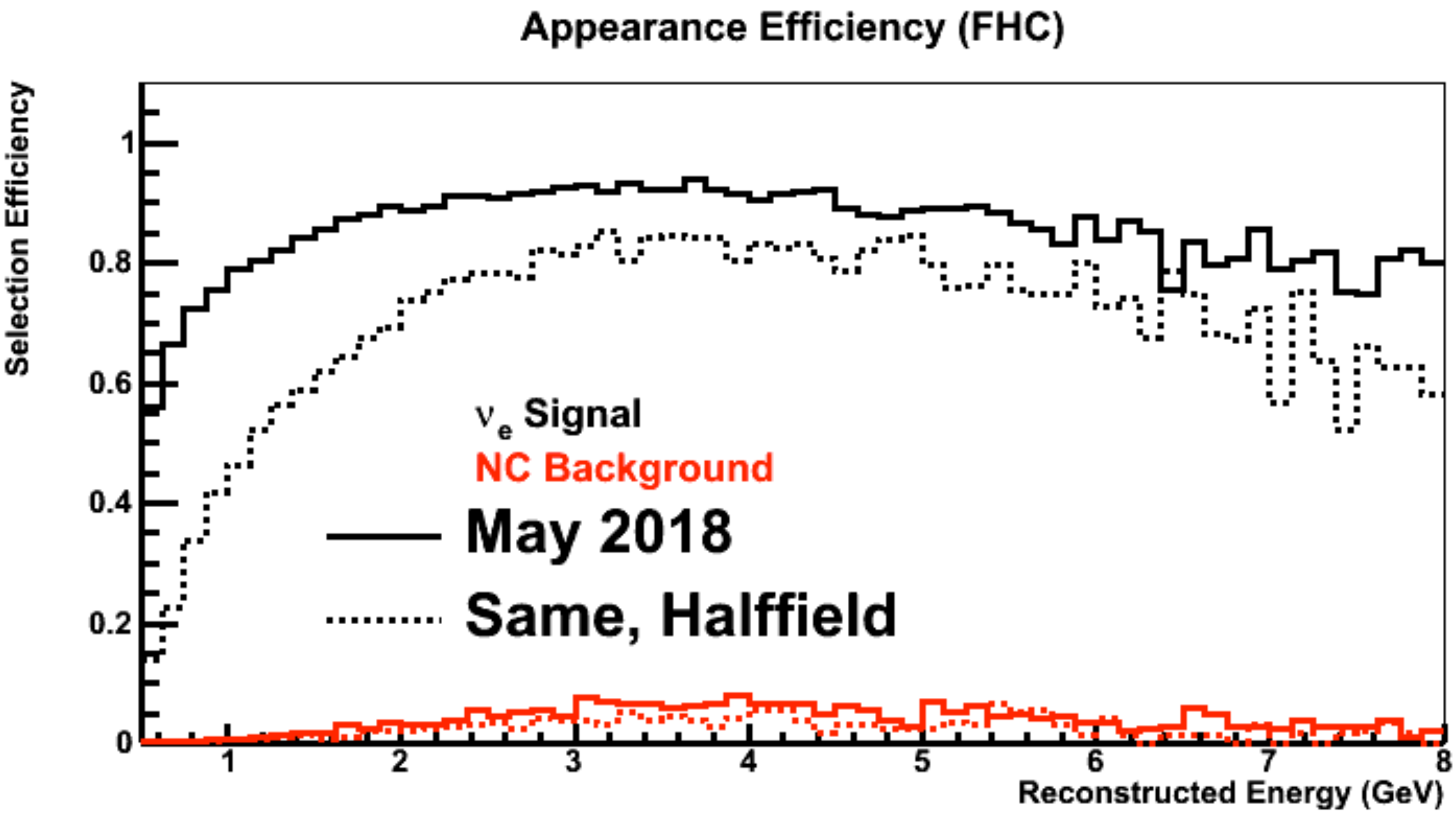
Efficiency: 5% Bad Channels



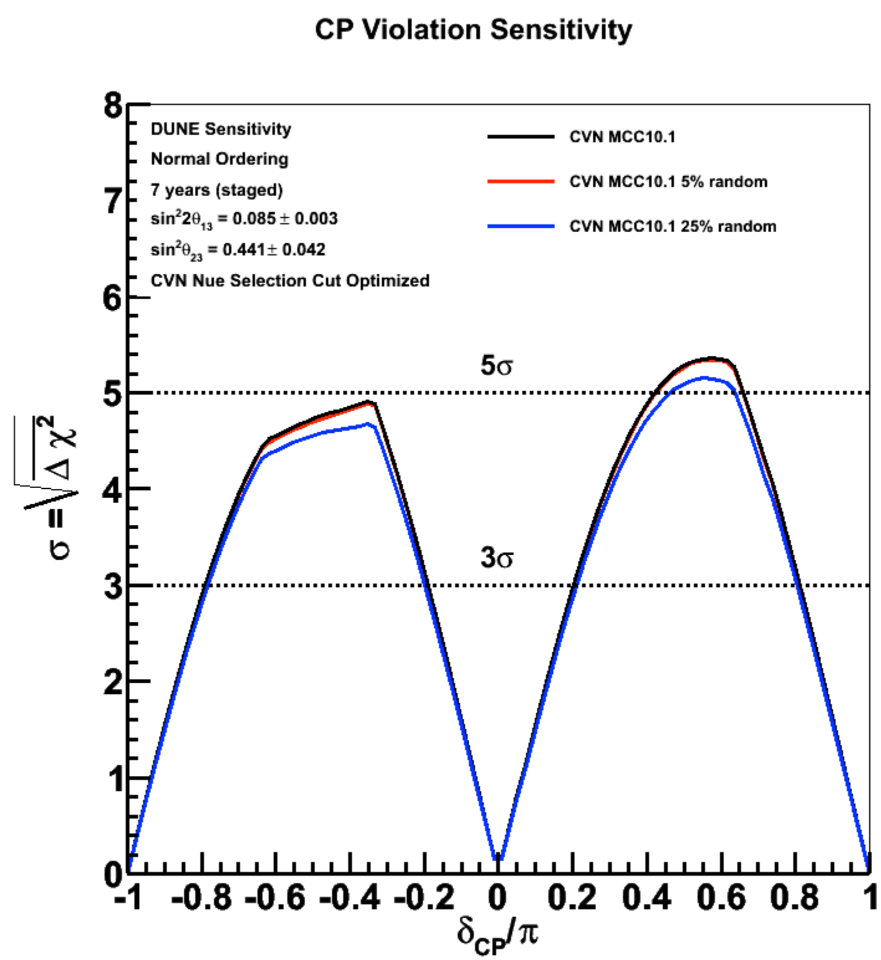
Efficiency: 25% Bad Channels



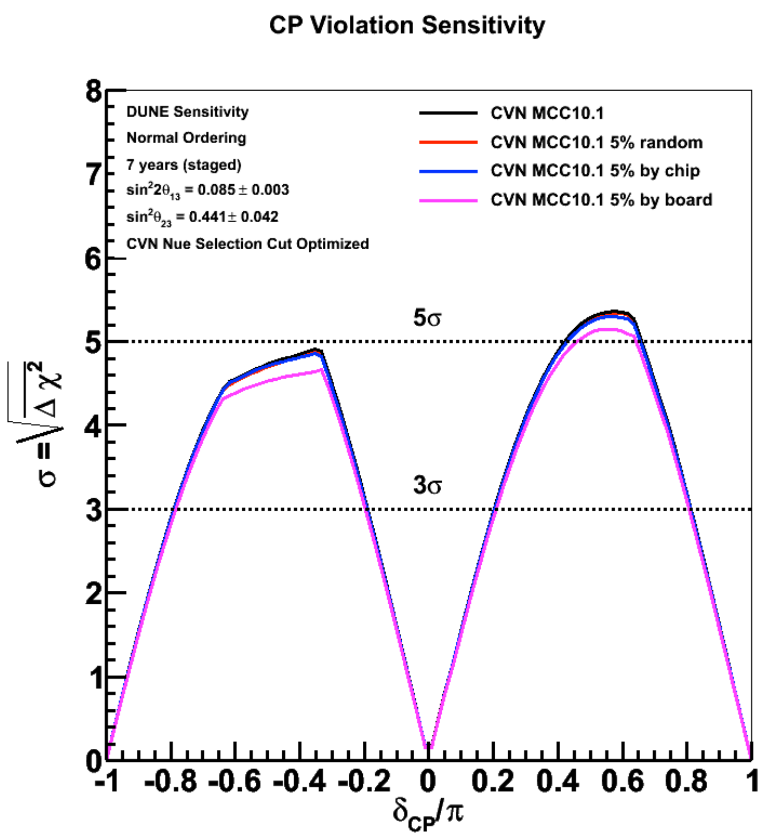
Efficiency: Drift Field



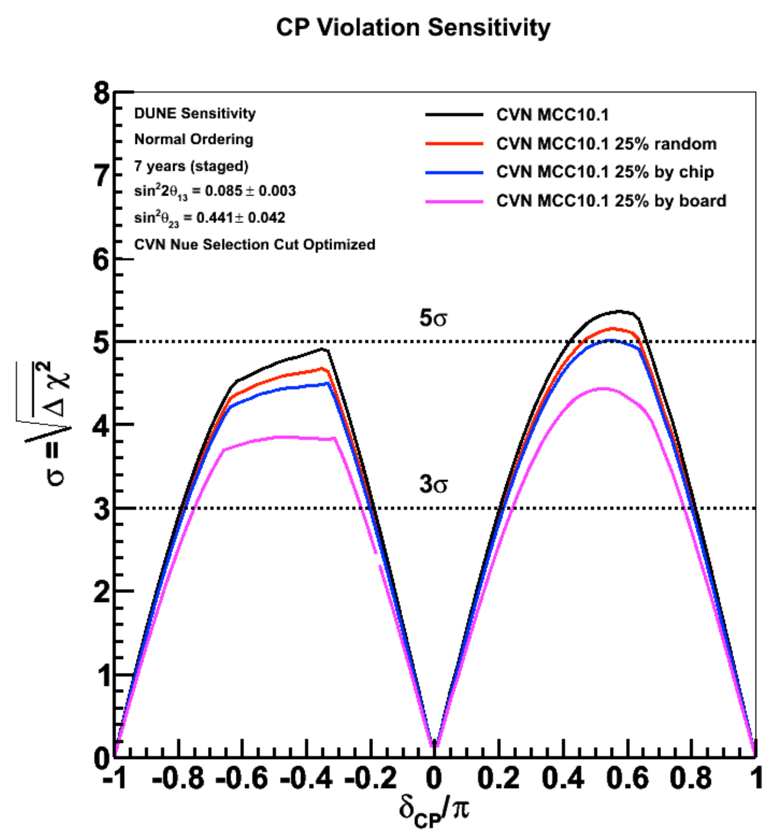
Sensitivity: Random Bad Channels



Sensitivity: Bad Channel Groups

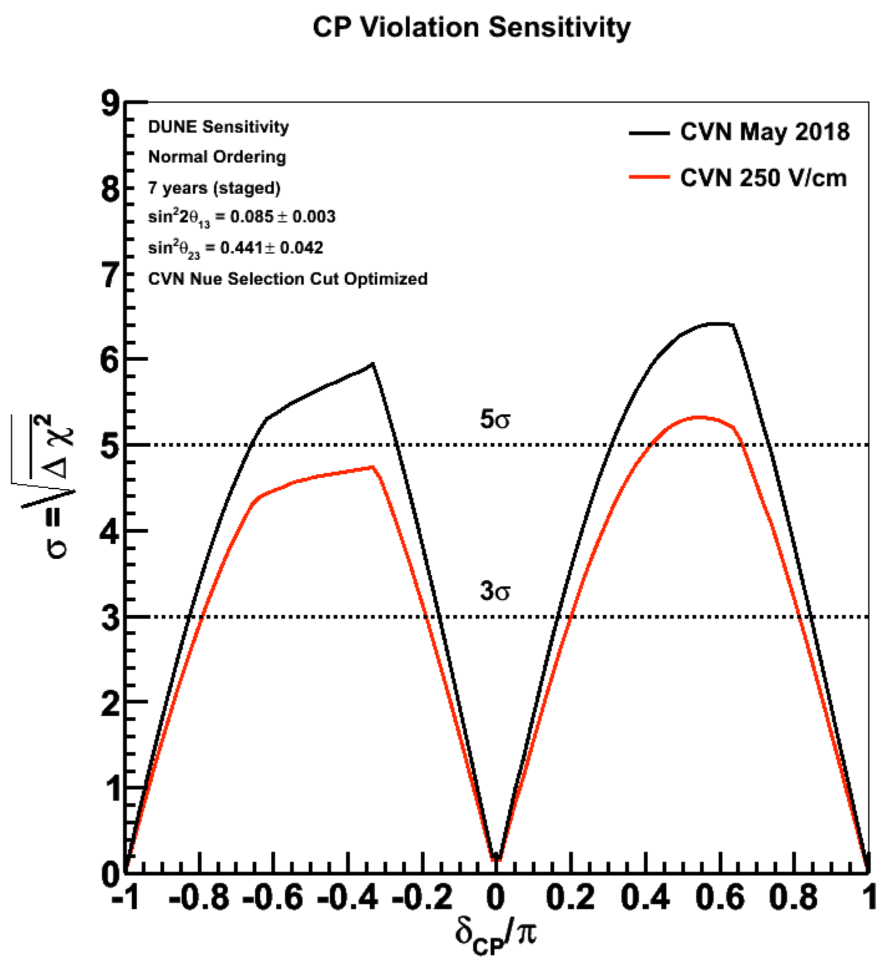


5% Bad Channels



25% Bad Channels

Sensitivity: Drift Field



Conclusion from Existing Studies

- CPV analysis pretty insensitive to bad channels, particularly when grouped randomly
  - Seems reasonable: by eye, differentiation between tracks and showers is quite easy even with many missing hits
- Larger groups of bad channels (eg: boards) cause more trouble than random bad channels
  - Again seems reasonable: when missing a whole board, significant portions of showers could be missing
  - Energy correction method also less robust in this case – overcorrecting showers that don't happen to have bad boards and undercorrecting those for which a large fraction of hits are missing
  - Some loss of sensitivity could be regained by detailed energy corrections and/or retraining CVN
- 250 V/cm drift field produces large degradation in sensitivity relative to full 500 V/cm
  - Would network retraining recover some of this effect?

Going Forward

- I claim bad channel studies are complete: conclusion is CPV analysis reasonably robust to any realistic set of bad channels even under worst case scenario of no retraining or detailed energy corrections
- Retraining network not feasible for many variations – if very important we could pick a few to study more carefully (250 V/cm?)
  - First step would be to look at some event displays – do we see major visual differences between showers in the different samples such that we suspect retraining will help?
- Field non-uniformity studies
  - Implement in MC using space charge effect simulation machinery (just dummy up "space charge" maps)
  - In progress – Bo is making the initial field distortion maps for us
- Noise studies
  - Study several levels of increased white noise and a data-driven noise simulation
  - Tools exist within MC, but work not yet actively underway to generate/analyze these samples