CAFAna for DUNE

DUNE LBL/ND meeting

Christopher Backhouse

California Institute of Technology

April 20, 2017

What is CAFAna?

- Used by ν_e+ν_μ+NC analysis groups in NOvA + several xsec analyses → battle-tested, plenty of experienced users, code well-commented
- CAF=NOvA's **C**ommon **A**nalysis **F**iles ≈ DSTs, Ntuples, AnaTree. . .

- Structured as a bag of tools to plug together, flexible, not monolithic
- Fast turnaround for plot-making as well as oscillation fits
- Fast enough to use interactively. $\mathcal{O}$(minutes)

- Oscillations are histogram-based
- Include systematics by profiling over pull terms

- This talk is focused on being an extremely short orientation/tutorial

Port

- Basic port accomplished at LBL hackdays
- Copy-pasted subset of NOvA files with adjustments for file loading
    - If something seems missing, ask! Maybe we have it but I didn't port it
- Retained SRT build system for convenience

```
git clone https://github.com/DUNE/lblpwgtools.git
cd lblpwgtools/code/CAFAna/
./checkout.sh
source setup.sh # on each login
cafe test/test_dune.C
```

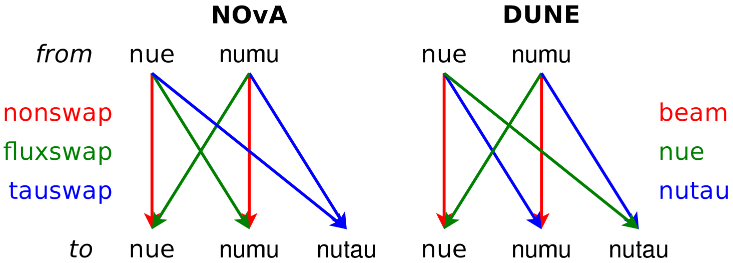
Input files

- Port reads FD MVASelect files
  `/pnfs/dune/persistent/TaskForce_AnaTree/far/train/v2.1/`
- Variables `Ev_reco`, `mvaresult`, `Ev`, `ccnc`, `beamPdg`, `neu`

- Chris Marshall has made ND TF files available in a compatible format
  `/dune/data/users/marshalc/NDTF_FGT*.root`
- Found there's no place for reconstructed charge
- Not storing anything about topology as yet

- MVASelect format has a lot of warts
- Stop-gap until we have a proper set of common files

MVASelect stumbling blocks



NOvA

from nue numu

nonswap
fluxswap
tauswap

to nue numu nutau

DUNE

nue numu

beam
nue
nutau

nue numu nutau

- MVASelect puts unswapped and swapped events all in the same tree
- Identified by magic run numbers
- Impedance mismatch with CAFAna – most of the effort in the port

- Separate ν_μ and ν_e files where `mvaResult` means different things
- Whereas in Chris M's ND files we use ±1 for that

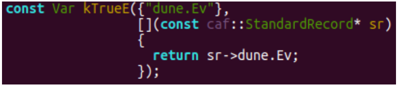
- FD POTs are scaled for 1.13kt

Resources

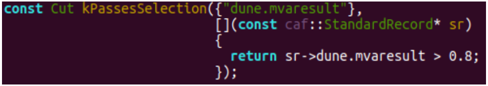
- NOvA-centric documentation at `http://cdcvs.fnal.gov/redmine/projects/novaart/wiki/CAFAna_resources`
- Most of this is applicable. Ignore references to specific variables and to "decomposition" or "extrapolation"
- Includes:
  - 2× introductory tutorial videos and slides (currently behind NOvA passwords, working on it...)
  - A prose high-level overview of the whole thing
  - A link to the low-level doxygen documentation of every class
  - An essay on how the systematics work

- Obviously the best way to learn is by doing
- Start with `test/test_dune.C` – NB this is *not* a very correct analysis
- Feel free to ask me and Kirk – I realize this talk isn't sufficient
- Definitely before re-inventing any wheels

Cuts and Vars

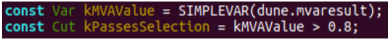


```
const Var kTrueE({"dune.Ev"},
                 [](const caf::StandardRecord* sr)
                 {
                   return sr->dune.Ev;
                 });
```



```
const Cut kPassesSelection({"dune.mvaresult"},
                           [](const caf::StandardRecord* sr)
                           {
                             return sr->dune.mvaresult > 0.8;
                           });
```

OR



```
const Var kMVAValue = SIMPLEVAR(dune.mvaresult);
const Cut kPassesSelection = kMVAValue > 0.8;
```

- `Var` is a recipe to extract a value from the file
- Can include arbitrary code if necessary
- Must list variables in use so we know which branches to activate

- `Cut` is the same thing, but returning a bool
- `SIMPLEVAR` macro for the most trivial Vars
- Can combine Vars with arithmetic and Cuts with logic operations

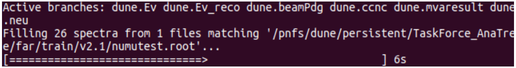
- See some pre-existing cuts in `Cuts/TruthCuts.h`

Spectrum and SpectrumLoader

- `Spectrum` is a wrapper around `TH1`. It also knows its POT
- Arithmetic ops between spectra handle POT scaling automatically
- 2D spectra also exist – flattened to 1D internally
- Construct with a `Cut`, `Var`, `Binning` and `SpectrumLoader`

- `SpectrumLoader` is responsible for looping over input files, filling all spectra and their POT
- Complicated in FD due to magic run numbers. See `test_dune.C`

- Create `SpectrumLoader`
- Create all `Spectrum` objects, passing loader
- Call `SpectrumLoader::Go()`



```
Active branches: dune.Ev dune.Ev_reco dune.beamPdg dune.ccnc dune.mvaresult dune
.neu
Filling 26 spectra from 1 files matching '/pnfs/dune/persistent/TaskForce_AnaTre
e/far/train/v2.1/numutest.root'...
[===============================>                           ] 6s
```

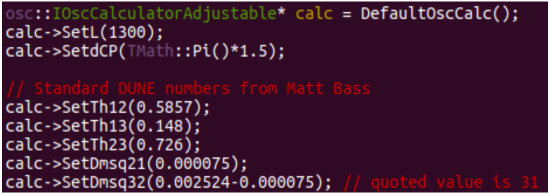
- Work with filled histograms

OscillatableSpectrum

- OscillatableSpectrum wraps a TH2 and POT
- *y*-axis is true energy
- Construct the same as a Spectrum

- Functions to obtain true energy spectrum or oscillated reco spectrum

- Oscillation calculators in OscLib/func/
- OscCalculatorPMNSOpt from João Coelho is generally the best



```
osc::IOscCalculatorAdjustable* calc = DefaultOscCalc();
calc->SetL(1300);
calc->SetdCP(TMath::Pi()*1.5);

// Standard DUNE numbers from Matt Bass
calc->SetTh12(0.5857);
calc->SetTh13(0.148);
calc->SetTh23(0.726);
calc->SetDmsq21(0.000075);
calc->SetDmsq32(0.002524-0.000075); // quoted value is 31
```

Prediction and Experiment

- A `Prediction` is a way to convert some oscillation parameters into a `Spectrum`
- You mostly want `PredictionNoExtrap` which has an `OscillatableSpectrum` for each channel and will oscillate them and sum them up

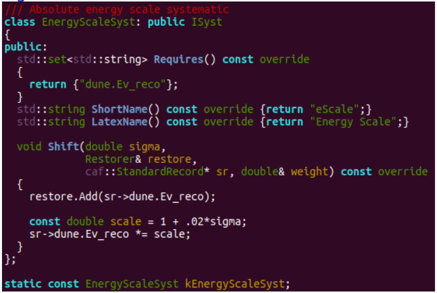
- An `Experiment` turns oscillation parameters into a χ^2
- `SingleSampleExperiment` takes a `Prediction` and a data spectrum
- `MultiExperiment` sums several χ^2s for a joint fit
- Include external constraints with `ReactorExperiment` and `SolarConstraints`, or add your own (or ask what NOvA have)

Fitter and Surface

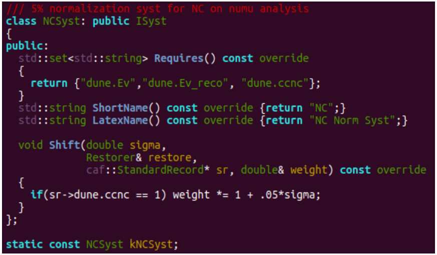
- A `Fitter` takes an `Experiment` and an oscillation calculator and finds the best fit parameters using MINUIT
- A `Surface` takes an `Experiment`, two oscillation parameters (`IFitVar`) to plot, and parameters to profile over, and makes a $\Delta\chi^2$ surface

- FitTheta13
- FitSinSq2Theta13
- FitDeltaInPiUnits
- FitSinSqTheta23
- FitSinSq2Theta23
- FitDmSq32
- FitDmSq32Scaled
- FitTanSqTheta12
- FitSinSq2Theta12
- FitDmSq21

Systematics



```
/// Absolute energy scale systematic
class EnergyScaleSyst: public ISyst
{
public:
  std::set<std::string> Requires() const override
  {
    return {"dune.Ev_reco"};
  }
  std::string ShortName() const override {return "eScale";}
  std::string LatexName() const override {return "Energy Scale";}

  void Shift(double sigma,
             Restorer& restore,
             caf::StandardRecord* sr, double& weight) const override
  {
    restore.Add(sr->dune.Ev_reco);

    const double scale = 1 + .02*sigma;
    sr->dune.Ev_reco *= scale;
  }
};

static const EnergyScaleSyst kEnergyScaleSyst;
```

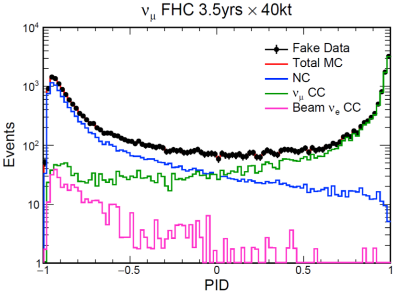


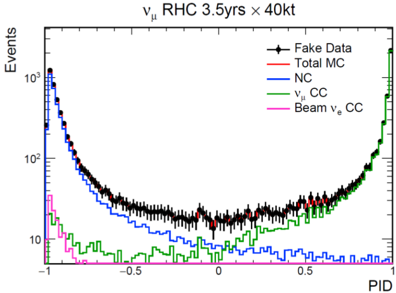
```
/// 5% normalization syst for NC on numu analysis
class NCSyst: public ISyst
{
public:
  std::set<std::string> Requires() const override
  {
    return {"dune.Ev","dune.Ev_reco", "dune.ccnc"};
  }
  std::string ShortName() const override {return "NC";}
  std::string LatexName() const override {return "NC Norm Syst";}

  void Shift(double sigma,
             Restorer& restore,
             caf::StandardRecord* sr, double& weight) const override
  {
    if(sr->dune.ccnc == 1) weight *= 1 + .05*sigma;
  }
};

static const NCSyst kNCSyst;
```

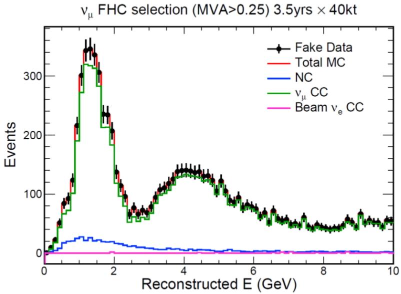
- An `ISyst` modifies or weights an event record as it's being loaded in
- Optional argument to `Spectrum` constructor taking a `SystShifts`

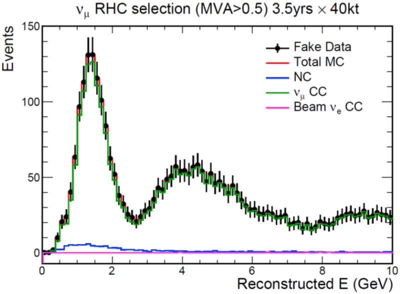
- `PredictionInterp` takes Predictions with various systematics applied and uses cubic interpolation between them
- If you only need scale systematics try `PredictionScaleComp`
- NOvA heritage means this machinery is a bit FD-centric (though ND sterile analyses have worked out), focus of upcoming development

Gallery 1/2

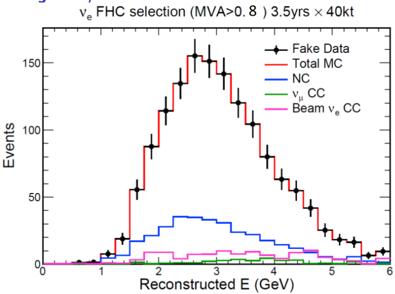

ν_μ FHC 3.5yrs × 40kt

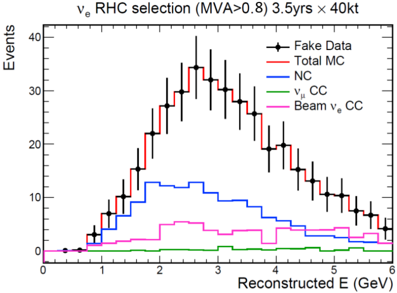

ν_μ RHC 3.5yrs × 40kt

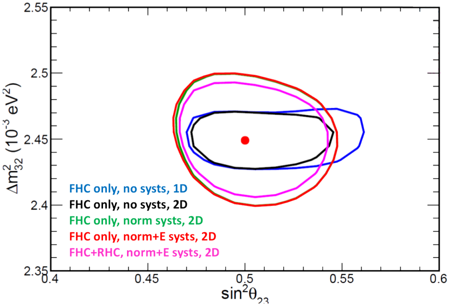

ν_μ FHC selection (MVA>0.25) 3.5yrs × 40kt

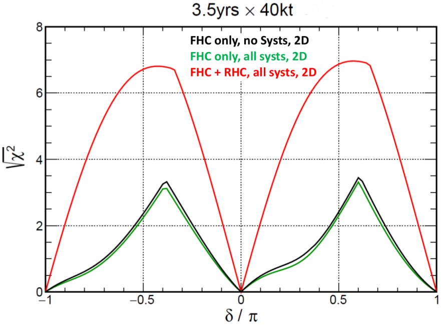

ν_μ RHC selection (MVA>0.5) 3.5yrs × 40kt

Gallery 2/2


ν_e FHC selection (MVA>0. 8) 3.5yrs × 40kt
Events
150
100
50
0
0
1
2
3
4
5
6
Reconstructed E (GeV)
Fake Data
Total MC
NC
ν_μ CC
Beam ν_e CC


ν_e RHC selection (MVA>0.8) 3.5yrs × 40kt
Events
40
30
20
10
0
0
1
2
3
4
5
6
Reconstructed E (GeV)
Fake Data
Total MC
NC
ν_μ CC
Beam ν_e CC


2.55
2.5
2.45
2.4
2.35
Δm^2_{32} (10^{-3} eV^2)
0.4
0.45
0.5
0.55
0.6
$\sin^2\theta_{23}$
FHC only, no systs, 1D
FHC only, no systs, 2D
FHC only, norm systs, 2D
FHC only, norm+E systs, 2D
FHC+RHC, norm+E systs, 2D


3.5yrs × 40kt
8
6
4
2
0
$\sqrt{\chi^2}$
−1
−0.5
0
0.5
1
δ / π
FHC only, no Systs, 2D
FHC only, all systs, 2D
FHC + RHC, all systs, 2D

Future developments

- Provide input to design of new unified file format

- Better story for ND/FD joint fits with systematics
- Flux systs, GENIE reweights

- Comparisons with other frameworks

- Serious studies to generate suite of tools (`Vars`, `Systs`, common macros)

- Collaborate with DUNE learners to produce DUNE-specific teaching materials and demo macros