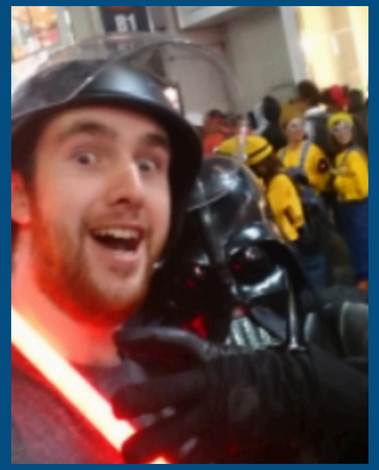
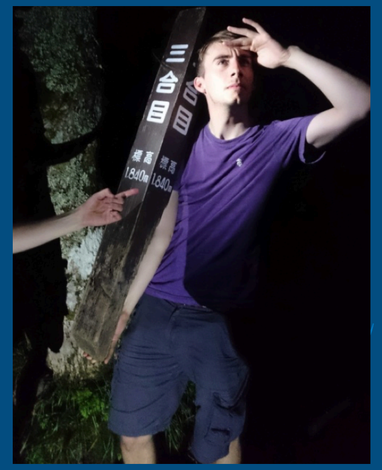
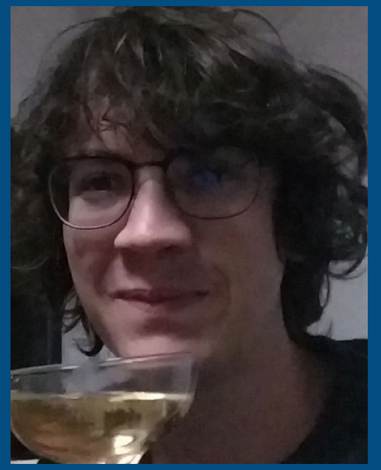
NUISANCE ♡ HEPDATA

WLCG/HSF Workshop 2024

p.stowell@sheffield.ac.uk

DOI: 10.1088/1748-0221/12/01/P01016



P. Stowell L. Pickering C. Wret C. Wilkinson S. Dolan

University of Sheffield

Science and Technology Facilities Council

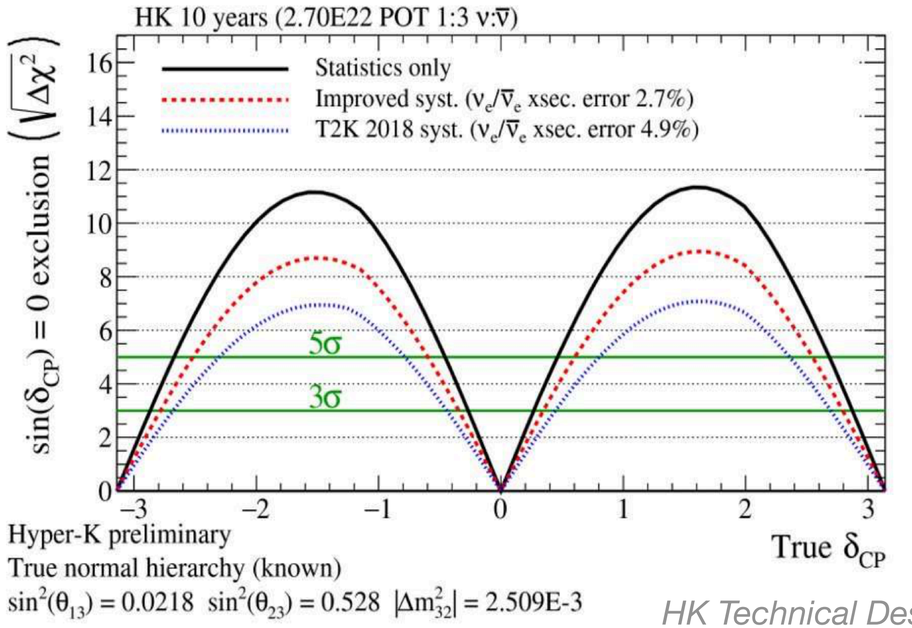
UNIVERSITY OF OXFORD

BERKELEY LAB Lawrence Berkeley National Laboratory

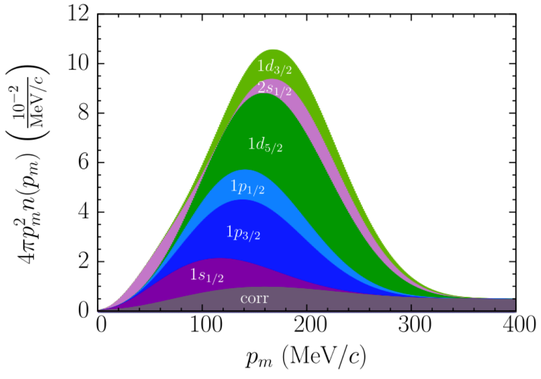
CERN

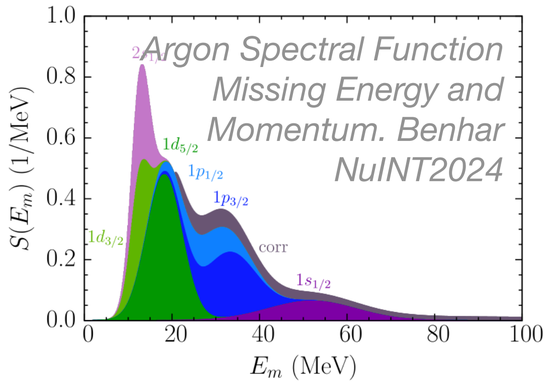
NEUTRINO CROSS-SECTION REQUIREMENTS

- Interaction cross-section modelling expected to be one of the major systematic uncertainties for future CP violation searches in the neutrino sector.
- Future Hyper-K and DUNE oscillation experiments are at en energy where nuclear effects are significant.
- Complex many body nuclear problem which scales exponentially with number of nucleons.



HK Technical Design Report

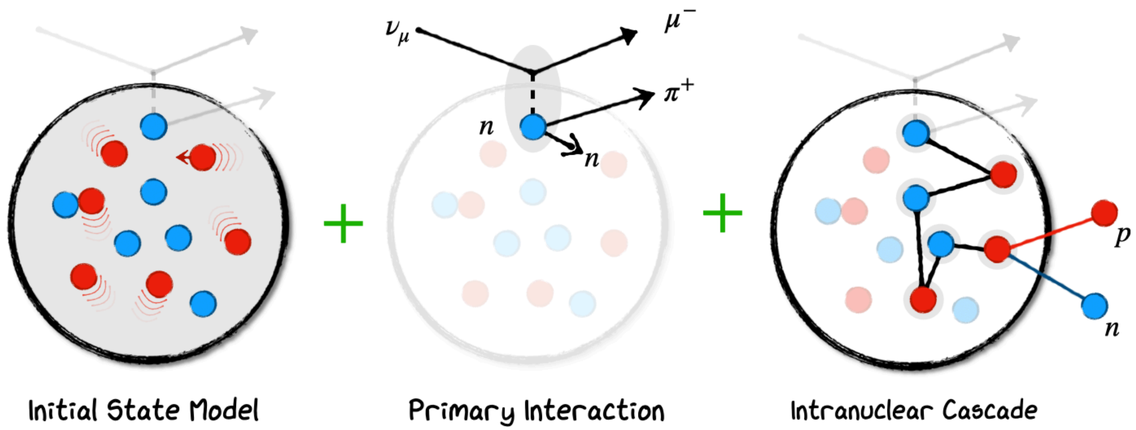




Argon Spectral Function Missing Energy and Momentum. Benhar NuINT2024

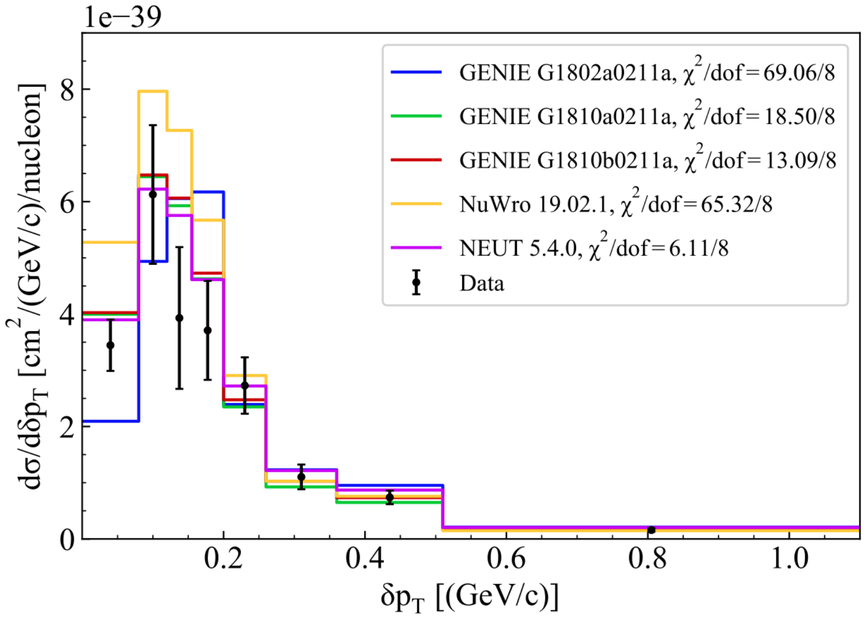
GENERATORS

- Neutrino community relies on interaction generators to bridge the gap between inclusive theoretical models and exclusive final state topology predictions.
- Five possible generators used in the community : GENIE, NEUT, NuWro, GiBUU, ACHILLES
- **The challenge:** Major tensions in the global dataset, no neutrino interaction generator model gets good agreement.



Generator Factorisation

Huge model differences in some regions of phase space



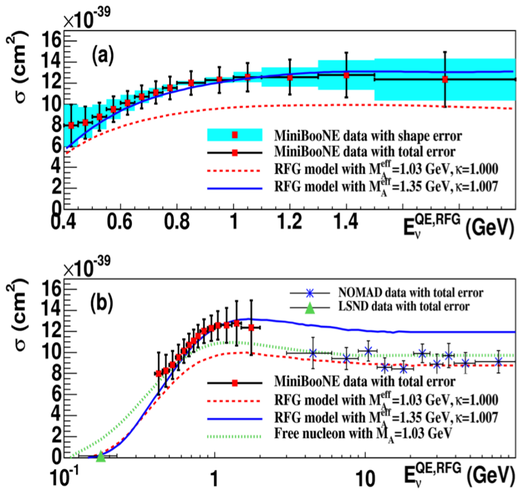
T2K Transverse Mom. Comparisons TENSIONS2019

CROSS-SECTION MEASUREMENTS

- Transformation in the field over the last decade from the realisation that flux unfolded measurements for specific interaction channels are highly model dependent.

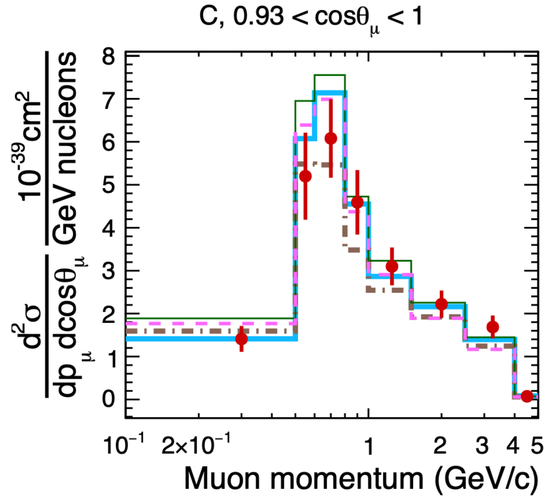
- Push to experiments releasing only exclusive final state topologies and avoiding unfolding into regions of phase space with poorly understood efficiency.
- Reduction in model dependence at a cost : harder for theory community to utilise new data, need initial state model, model of final state interactions, background models.

MiniBooNE True CCQE Neutrino Energy



Increasing Comparison Complexity

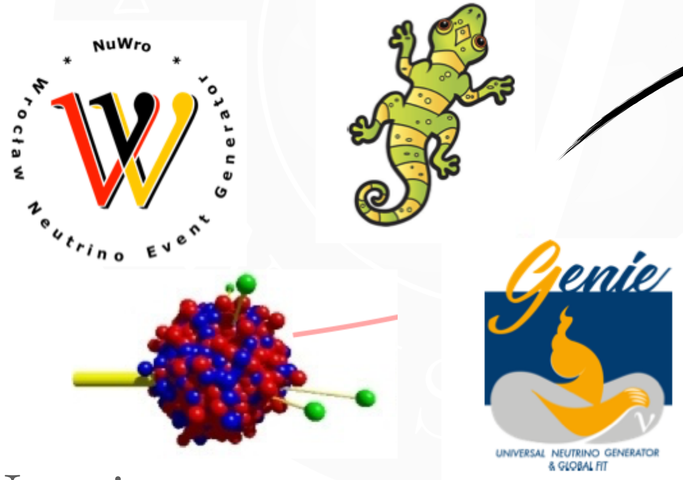
Decreasing Model Dependence



T2K FS CC0π Muon Energy

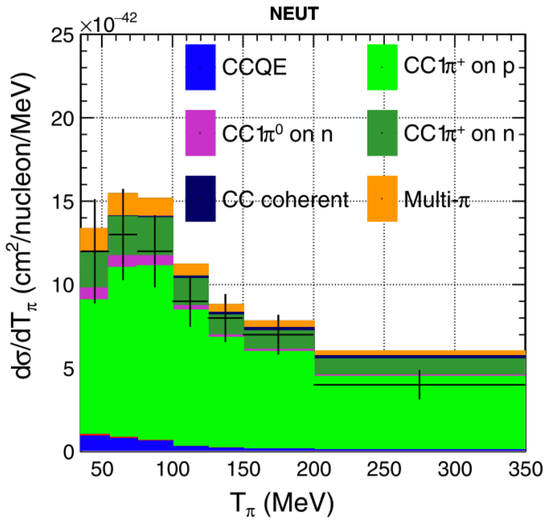
NUISANCE

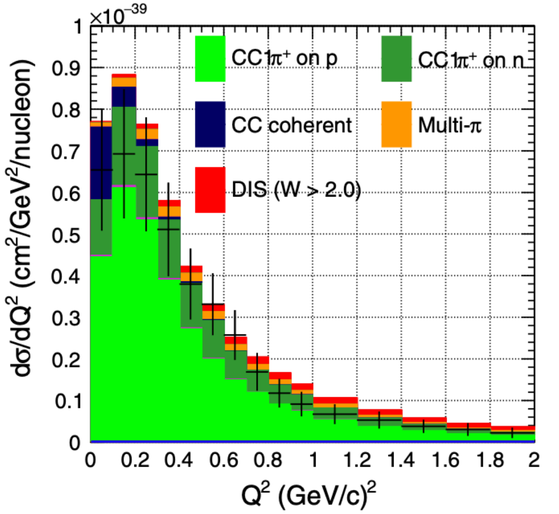
- Developed the NUISANCE framework as a neutrino focussed tuning tool to standardise comparing generators to data (neutrino equivalent of RIVET in collider community).
- V1 T2K external data fitter (2014). V2 community release (2017).
- Only open source tool for comparing/fitting neutrino generators.



Neutrino Generator Events

> 300 ν - A Standardised Comparisons



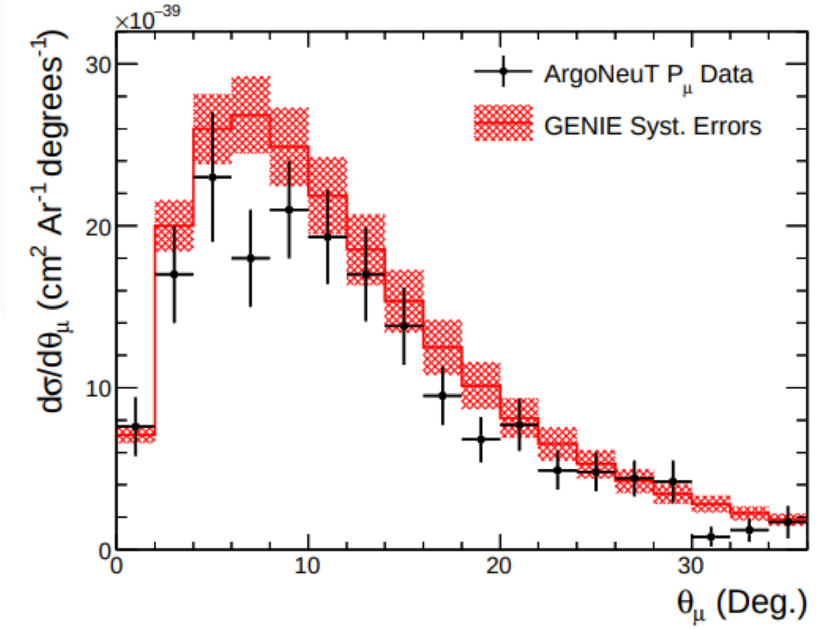


Pion Modelling in NEUT : Comparisons and Challenges of Modern Neutrino Scattering Experiments.

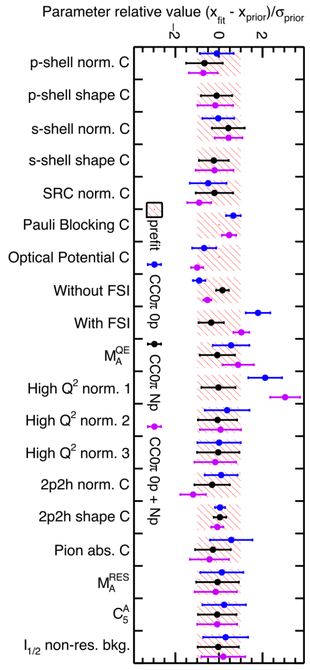
DOI: 10.1088/1748-0221/12/01/P01016

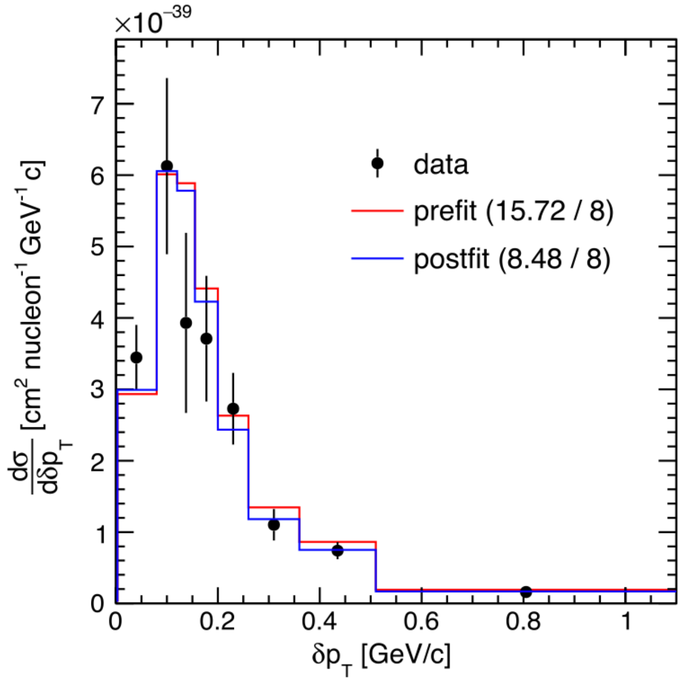
NUISANCE Analyses

- Original interface written aimed at flux averaged total cross-section neutrino experiments.
- Have required extensions to the analysis methodology in NUISANCE as new experiments move to novel unfolding/detector smearing approaches.
- Recently many measurements implemented into NUISANCE by experiments themselves, similar user model to RIVET analyses.



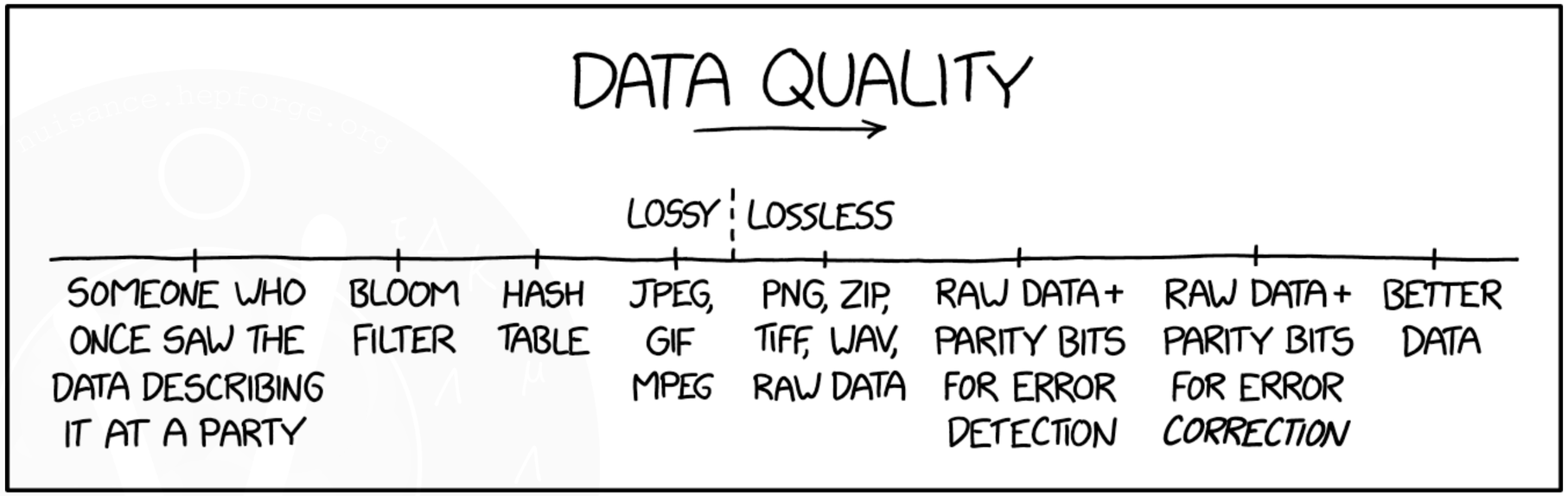
DOI: 10.1088/1748-0221/12/01/P01016

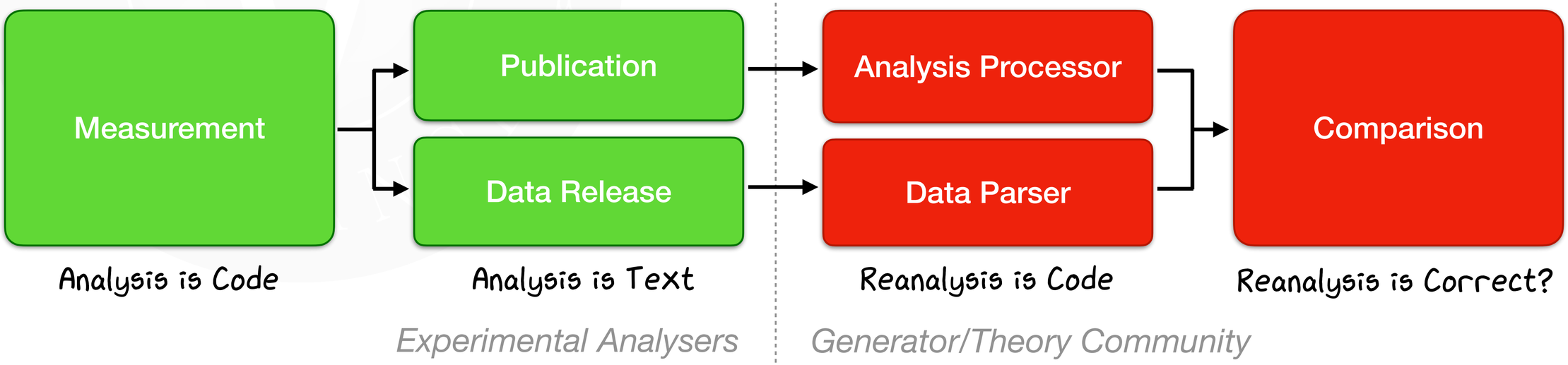




DOI: 10.1103/PhysRevD.109.072006

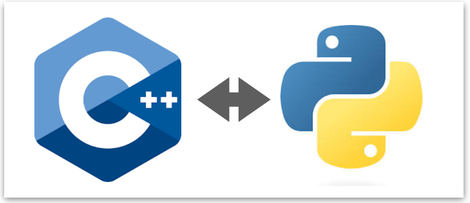
NEUTRINO DATA RELEASES



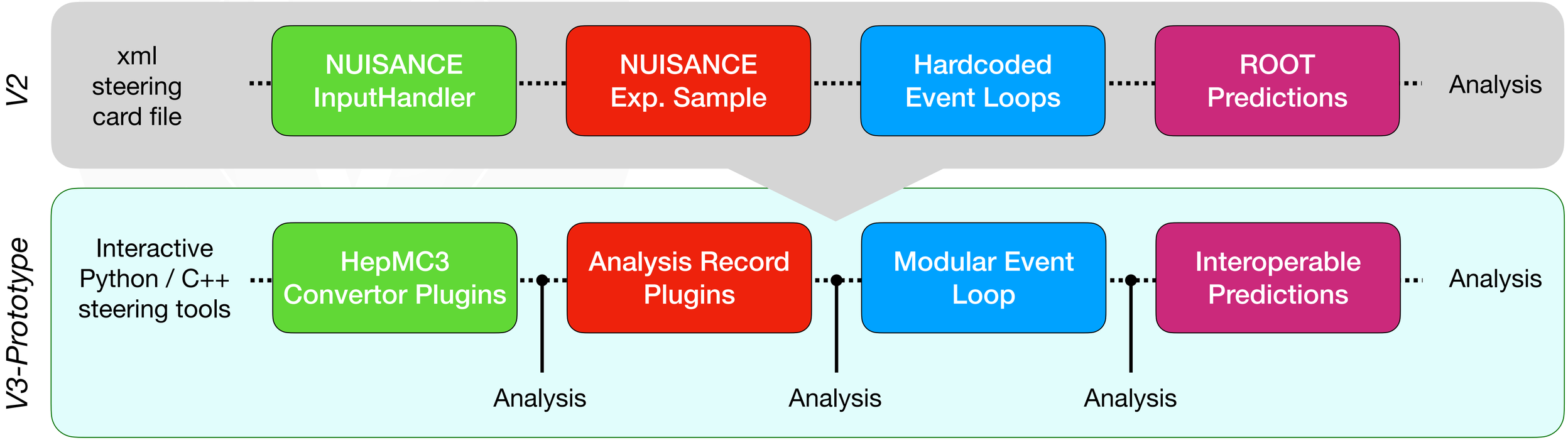


PROTOTYPING EFFORTS : nuisance3

- Exploring how the structure of NUISANCE could be rewritten to make it more accessible.
- Move away from prior monolith structure to modular interfaces.
- Enable external processing/tuning libraries to interface at different stages in the analysis chain through C++ or python.

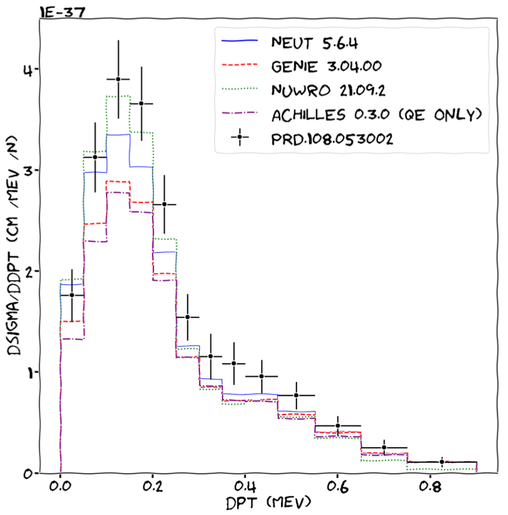


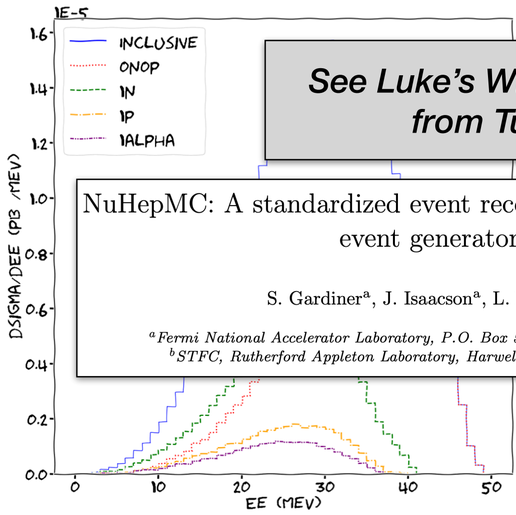
Full C++ to Python Bindings



GENERATOR PLUGINS

- Existing NUISANCE ‘v2 FlatTree’ ROOT format already widely used for generator model studies.
- Testing alternative event source structure focussed on analysis in HepMC3 format.
- Boost plugin structure to isolate generator dependencies whilst allowing passthrough information inside HepMC itself.
- `EventFrame` analysis tools designed to be similar to `RDataFrame`/Pandas interfaces.



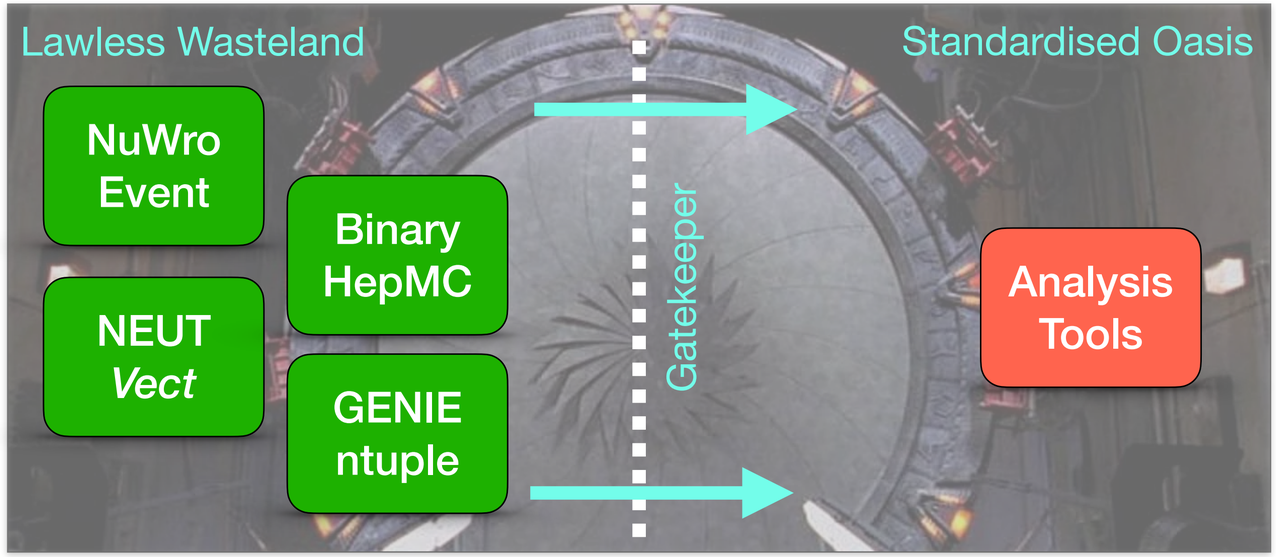


See Luke’s WLCG-HSF Talk from Tuesday!

NuHepMC: A standardized event record format for neutrino event generators

S. Gardiner[a], J. Isaacson[a], L. Pickering[b]

[a] *Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, USA*
[b] *STFC, Rutherford Appleton Laboratory, Harwell Oxford, United Kingdom*



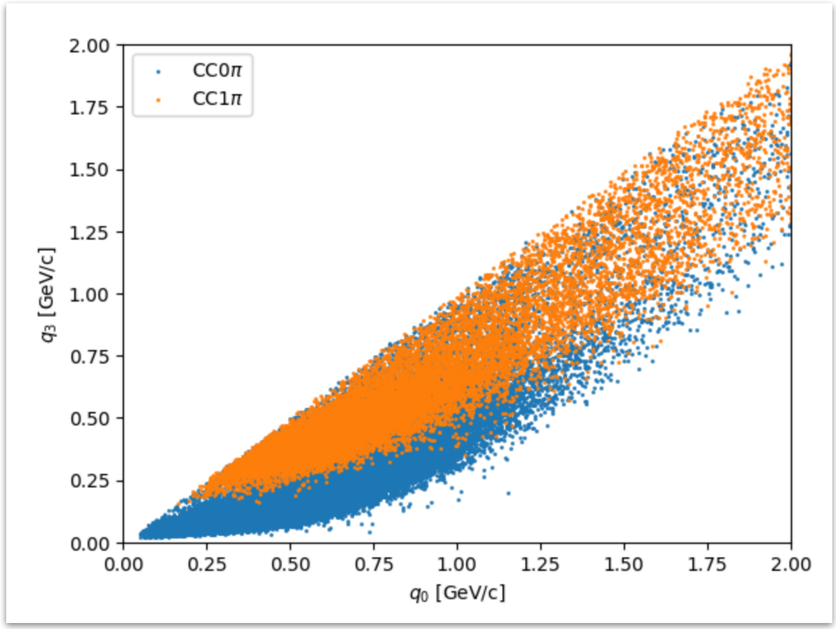
GENERATOR ANALYSIS

- Analysis columns added to event source objects based on user defined event processing hooks.
- Flexibility for complex projections or weighting functions for the column automatically.
- Data frame evaluation outputs using numpy, pandas, Arrow (extendable interoperability).

```
evs.add_column(...)  ⟶  evs.evaluate()  ⟶  evs.scatter("q0","q3")
```

- `WeightCalculator(GenEvent const &)`
- `Projections(GenEvent const &)`
- `Selection(GenEvent const &)`
- `UserDefined(GenEvent const &)`

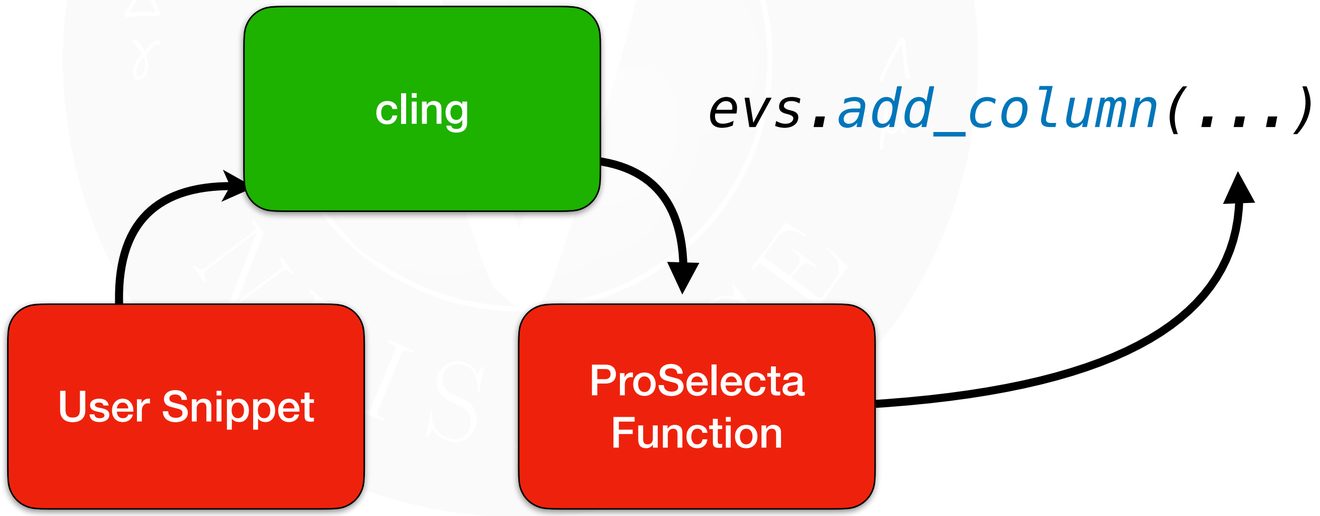
event.number	weight.cv	process.id	q0	q3
0	1	300	263.1	487.6
1	1	452	1731	2036
2	1	200	272.3	684.4
3	1	401	317.7	331.1
4	1	250	163.4	577.3
5	1	200	115.7	503.6
6	1	200	92.75	261.2
7	1	200	353.1	789.7
8	1	200	86.2	352.7
9	1	200	104.1	212
10	1	600	1650	1813
11	1	200	372.3	769.4
12	1	500	2218	2799
13	1	550	1175	1307
14	1	452	755.3	1172
15	1	200	219.9	592.8
16	1	200	116.8	286.5
17	1	400	607.8	647
18	1	200	316.2	782.9
19	1	250	475.7	1058
20	1	402	585.6	962.3
...	...	...	...	...



JIT C++ EVENT ANALYSIS

- **ProSelecta** : a modular NuHepMC3 event analysis tool that supports JIT compiling of analysis functions.

- Returns template projection/selection signatures.



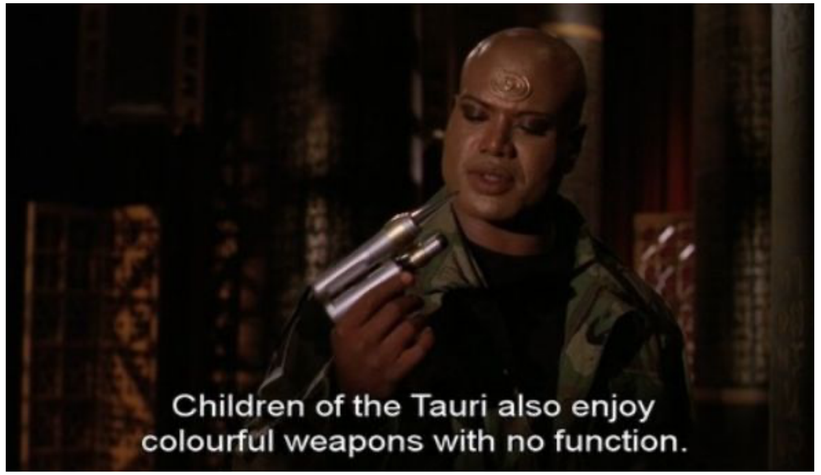
ProSelecta Neutrino Tools

```
// Particle Selectors
Target(HepMC3::GenEvent const &ev)

Beam(HepMC3::GenEvent const &ev, int PID)
BeamAny(HepMC3::GenEvent const &ev,
        std::vector<int> const &PIDs)

OutPartHM(HepMC3::GenEvent const &ev, int PID)
OutPartHMAny(HepMC3::GenEvent const &ev,
             std::vector<int> const &PIDs)



// Standard Projections
double q0(HepMC3::GenEvent const &ev)
double q3(HepMC3::GenEvent const &ev)
double Q2Lep(HepMC3::GenEvent const &ev);


...

double CosThetaLep(HepMC3::GenEvent const &ev) {

  auto pin = ps::sel::BeamAny(ev, pdg::groups::kNeutralLeptons);

  if (!pin) {
    return kMissingDatum;
  }

  int nupid = pin->pid();
  int ccpid = nupid > 0 ? (nupid - 1) : (nupid + 1);

  auto pout = ps::sel::OutPartFirstAny(ev, {ccpid, nupid});

  return parts::CosTheta(pin, pout);
}
```

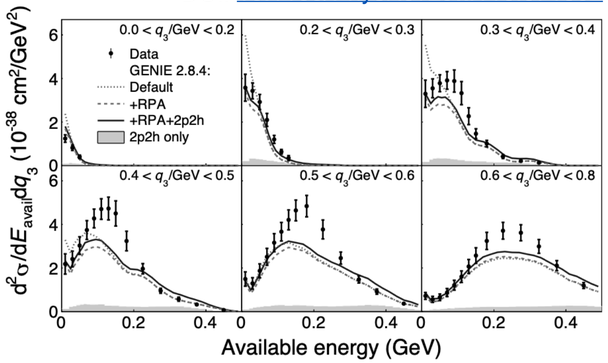
SNIPPET ANALYSES

- Automatic loading of HepMC3 extensions makes development slightly easier but real strength is the ability to prototype analysis data releases and signal selections.

- Directly compare many different experimental projection approaches kept in a single file loaded at runtime.

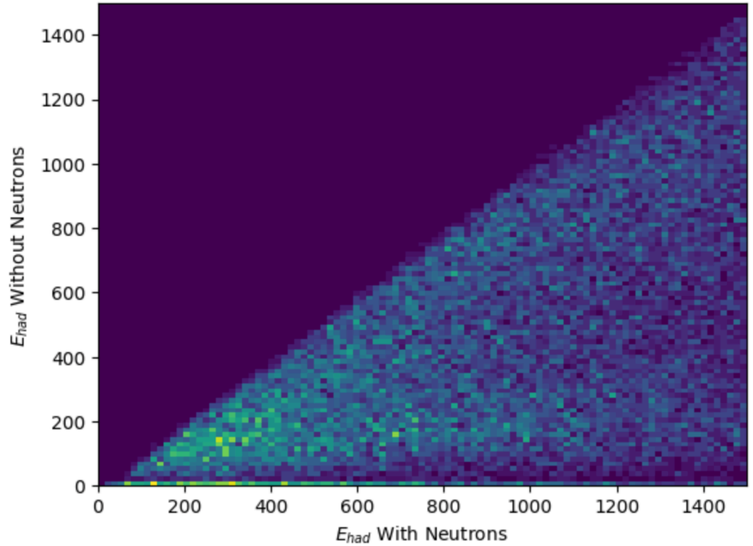


DOI: 10.1103/PhysRevLett.116.071802



MINERvA ν-CH
Low Recoil Measurement

```
// MINERvA CC-inclusive Signal Definition
int MINERvA_CCINC_CCEavq3_Filter(ev) {
  auto nu = ps::sel::Beam(ev, ps::pdg::kNuMu);
  if (!nu) return false;

  auto mu = ps::sel::OutPartHM(ev, ps::pdg::kMuon);
  if (!mu) return false;

  double cosangle = ps::proj::event::CosLep(ev);
  if (cosangle) <  0.93969262078) return false;

  if (ps::proj::event::ELep(ev) < 1.5 * ps::GeV)
    return false;

  return true;
}
```



DATA PLUGINS : HEPDATA ν-A Scattering

- Exploring the use of HEPDATA as a standard data release format for adoption by the neutrino community.
- YAML based data release with associated correlation matrices, supporting flux info.
- Additional information beyond the data tables is needed to reliably preserve the original analysis.
  - Historically most of the effort has been in the "Reanalysis Processor" stage.

HEPData

Measurement of double-differential muon neutrino charged-current interactions on C_8H_8 without pions in the final state using the t2k off-axis beam

The T2K collaboration Abe, Ko ; Andreopoulos, Costas ; Antonova, Maria ; *et al.*

Phys.Rev.D 93 (2016) 112012, 2016.

Inspire Record 1421157 DOI 10.17182/hepdata.77052

Download All

Filter 10 data tables

Table 1
Data from Equation 6
10.17182/hepdata.77052.v1/t1
Total signal cross-section per nucleon integrated over all the muon kinematics phase space in Analysis I.

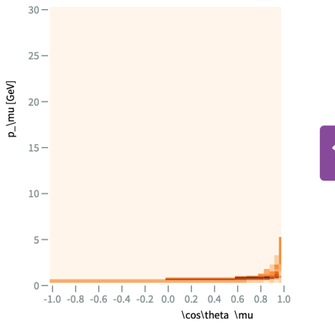
Table 2
Data from Figure 6
10.17182/hepdata.77052.v1/t2
Results of the double differential cross-section measurement bin-by-bin in Analysis I.

Table 3
Data from Figure 6
10.17182/hepdata.77052.v1/t3
Covariance matrix for shape systematics error in Analysis I.

Showing 50 of 67 values Show All 67 values

Analysis		I
RE		NUMU C --> MU- X
$\cos\theta_\mu$	p_μ [GeV]	$d^2\sigma/dp d\cos\theta$ [10^{-38} cm^2 / nucleon / GeV]
-1.0 - 0.0	0.0 - 0.3	0.255394
-1.0 - 0.0	0.3 - 0.4	0.207948
-1.0 - 0.0	0.4 - 30.0	0.000271813
0.0 - 0.6	0.0 - 0.3	0.208457
0.0 - 0.6	0.3 - 0.4	0.724434
0.0 - 0.6	0.4 - 0.5	0.551849
0.0 - 0.6	0.5 - 0.6	0.234113
0.0 - 0.6	0.6 - 30.0	0.000368685
0.6 - 0.7	0.0 - 0.3	0.178667

Visualize



Brushing Enabled?

DATA PLUGINS : HEPDATA ANALYSES

- Suggesting a standard for the community to also release projection/selection functions necessary for a comparison directly to HEPDATA.
- Self-contained snippets based on HepMC3 event selections. Transparent, readable analysis steps that define exact processing needed.
- nuisance3 setup to automatically parse HEPDATA snippets using ProSelecta, fully dynamic processing of comparisons.

HEPDATA : submission.yaml

```yaml
qualifiers:
 - name: likelihood
   value: Poisson
 - name: scaling
   value: EventRateScaleToData
 - name: weighting
   value: Default

 - name: Filter
   value: ANL_167744_CC1npip_Selection
 - name: ppi
   value: ANL_167744_CC1npip_Project_ppi
```

Histogram Processing

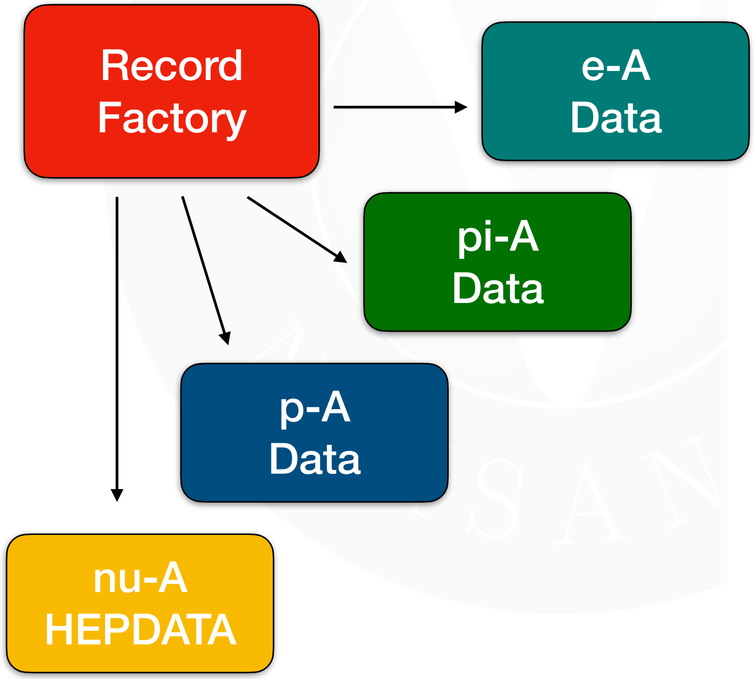
Event Processing

HEPDATA : analysis.cxx

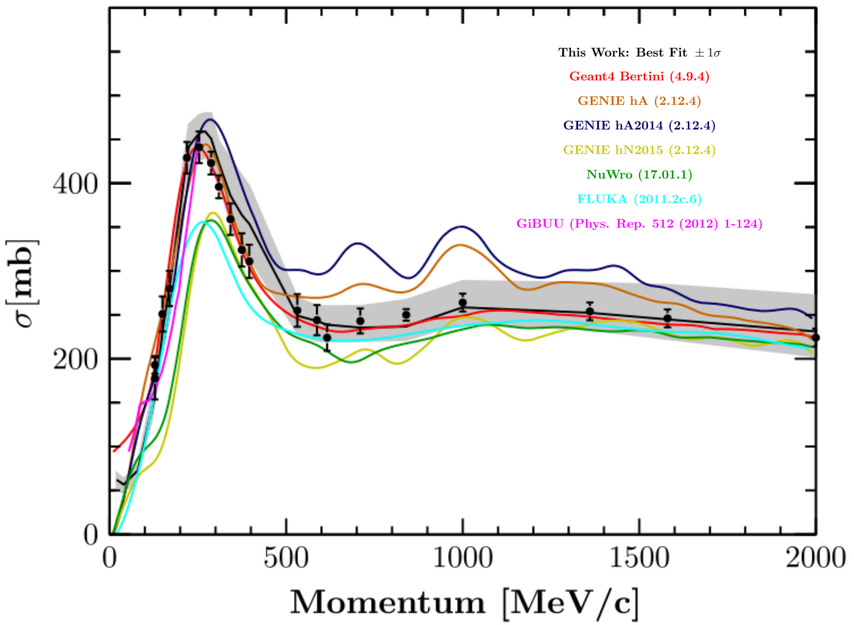
```
double ANL_167744_CC1npip_Project_ppi(ev) {

    if (!ANL_167744_CC1npip_Selection_lowW(ev)) {
        return ps::kMissingDatum;
    }

    auto Pnu  = ps::sel::Beam(ev, ps::pdg::kNuMu)->momentum();
    auto Pn   = ps::sel::OutPartHM(ev, 2112)->momentum();
    auto Ppip = ps::sel::OutPartHM(ev, 211)->momentum();
    auto Pmu  = ps::sel::OutPartHM(ev, 13)->momentum();

    return Ppip.length() * ps::MeV;
}
```

DATA PLUGINS : p-A/π-A/e-A Scattering

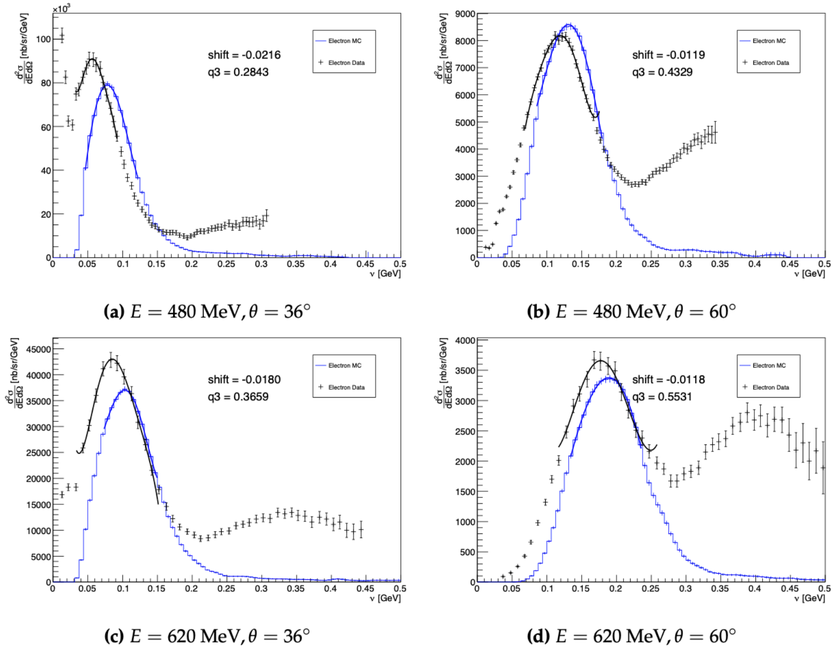
- Long term plan to combine neutrino HEPDATA global analyses with interfaces to existing pion and electron scattering data initiatives.
- **Understanding correlations in nuclear models across multiple classes of data.**



E. Pinzon, π-A Scattering



J. McElwee, e-A Scattering



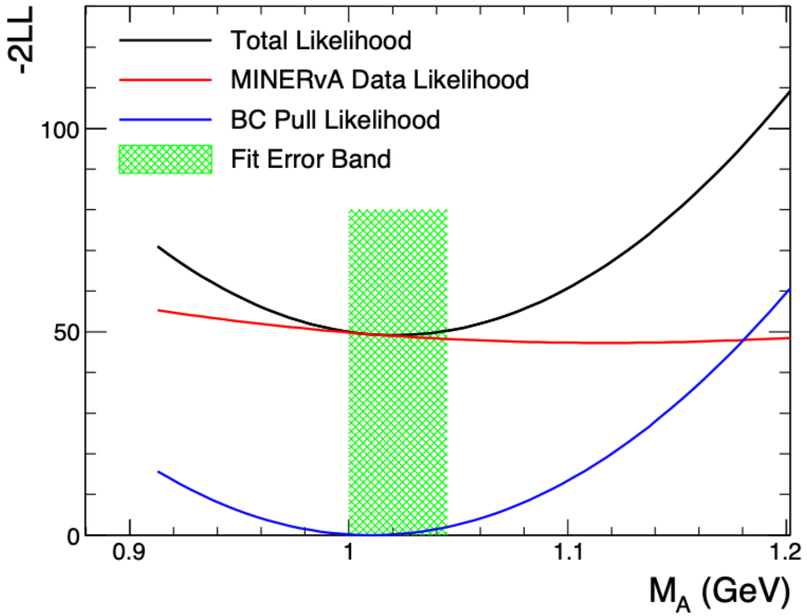
(a) $E = 480$ MeV, $\theta = 36°$ **(b)** $E = 480$ MeV, $\theta = 60°$

(c) $E = 620$ MeV, $\theta = 36°$ **(d)** $E = 620$ MeV, $\theta = 60°$

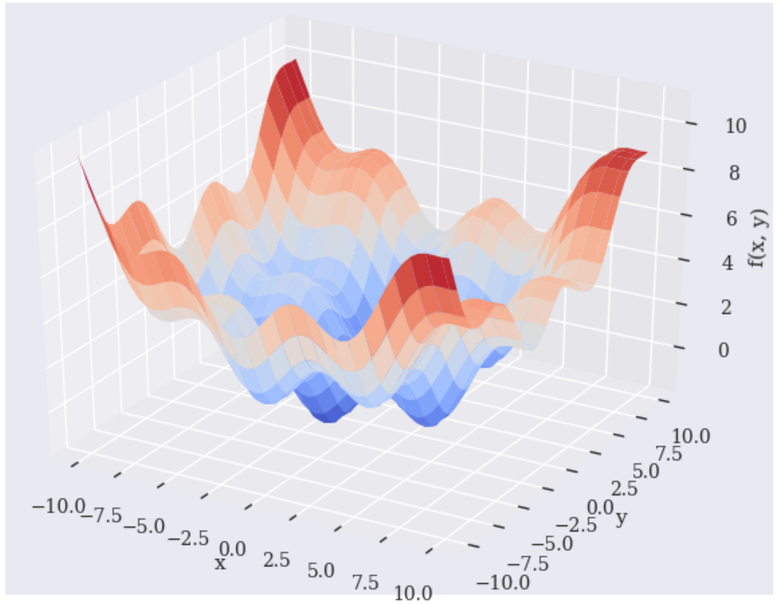
TUNING INTERFACES

- Move to python as a steering language allows previous hard-coded NUISANCE routines and global likelihood functions to be evaluated interactively.

- Looking at potential for direct interface into external tuning and systematic evaluations.



NUISANCE v2
Minuit Result



scipy.optimize
basinhopping

CONCLUSIONS

- NUISANCE provides standardised generator tuning tools for the neutrino community.

- Looking at redeveloping core code structure to use module generator and analysis methods.
  - Moving away from monolithic comparison routines.
  - Exploring use of automated compilation of data release projection/signal operators.
  - Python steering code for interfacing to external tuning tools.

- Processes being considered based on feedback from experience in the neutrino community, but welcome thoughts, suggestions, and comments on lessons learnt in the collider community.

THANKS FOR LISTENING!