Science & Technology Facilities Council

UK Research and Innovation

Report from NuInt 18

Lukas Koch

STFC Rutherford Appleton Laboratory

PPD Seminar

14/11/2018

T2K

[Eligio Lisi]

A new field is born/re-branded

Electroweak Nuclear Physics

Science & Technology Facilities Council

UK Research and Innovation

[Eligio Lisi]

A new field is born/re-branded

Electroweak Nuclear Physics

"multifaceted and interdisciplinary"
= a mess of different communities

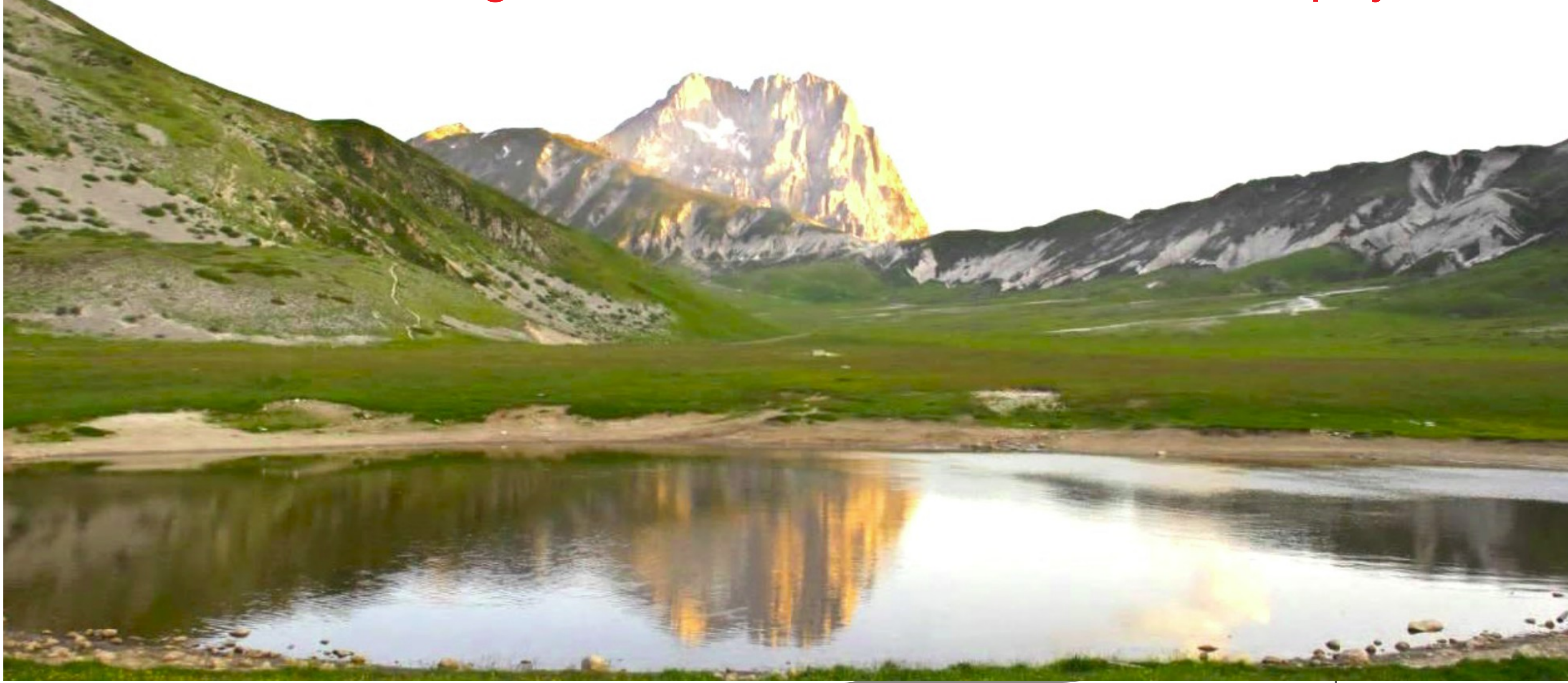
Intersection of neutrino cross-sections, electron scattering, nuclear physics, ...

Learn from Astroparticle Physics

- Particle physics, Astrophysics, Cosmic ray physics, Cosmology were in a similar position at the end of the last century
- Realised that they needed to join communities to tackle important problems
  - dark matter, baryon asymmetry and stability, neutrino masses, …
- Came together under the umbrella term "Astroparticle Physics"
- Dedicated journals, schools soon emerged, leading to a common language, flourishing field

Electroweak Nuclear Physics

electron scattering, neutrino cross-sections, nuclear physics, ...

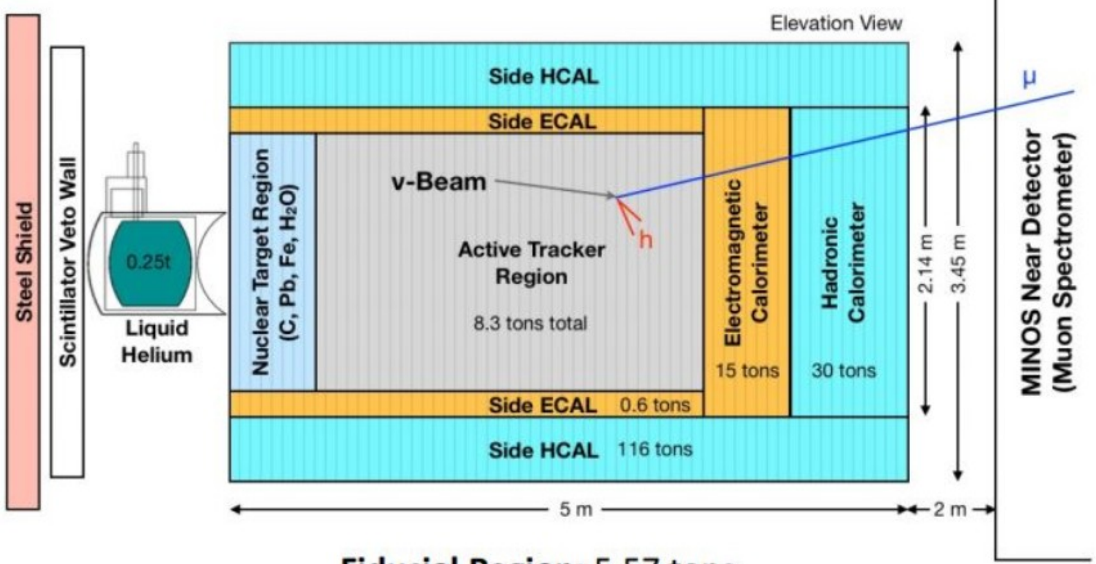


[Aaron Bercellie]

MINERvA

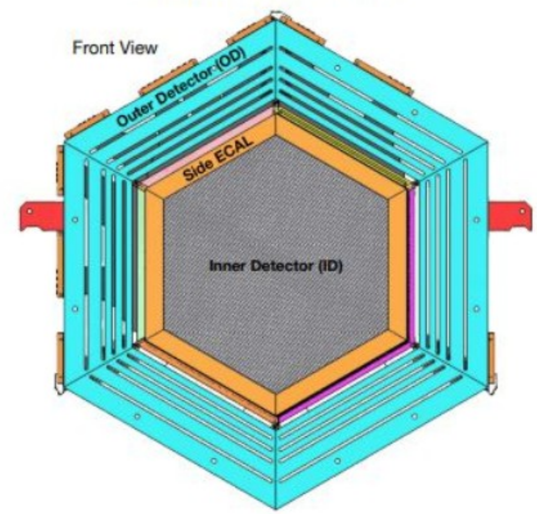
MINERvA Detector

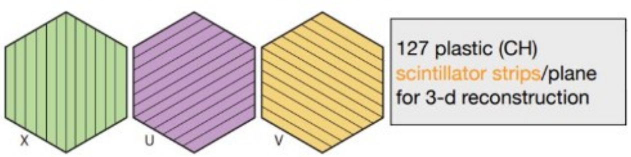
Side View
Complete Detector



Fiducial Region: 5.57 tons

Front View
Single Module

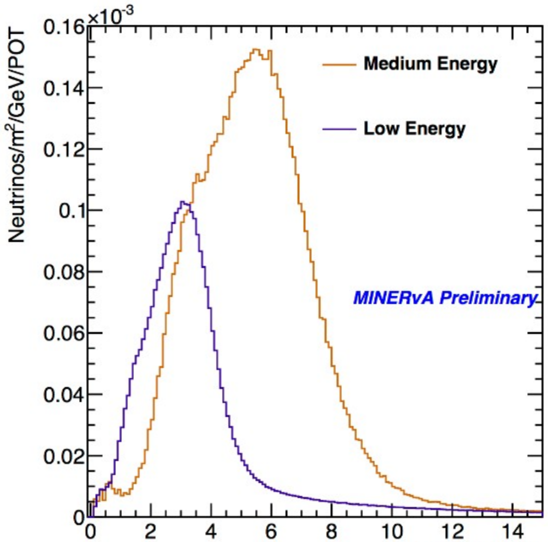


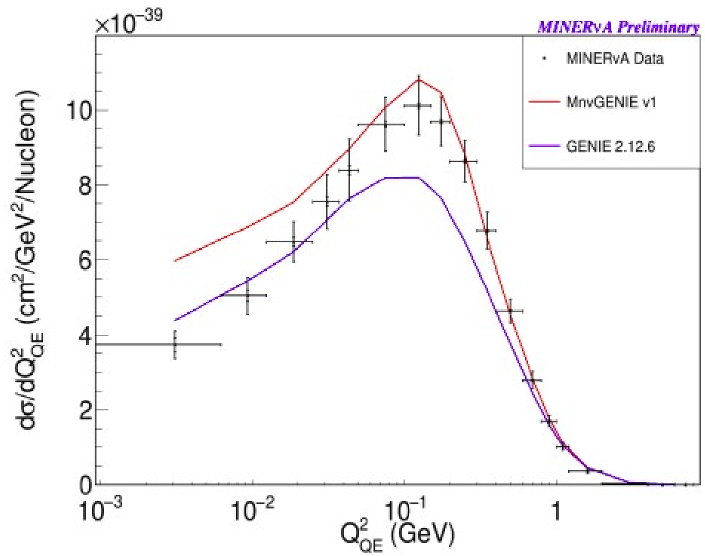


Nucl. Instrum. Methods Phys. Res., Sect. A 743, 130 (2014).
Nucl. Instrum. Methods Phys. Res., Sect. A 789, 28 (2015).

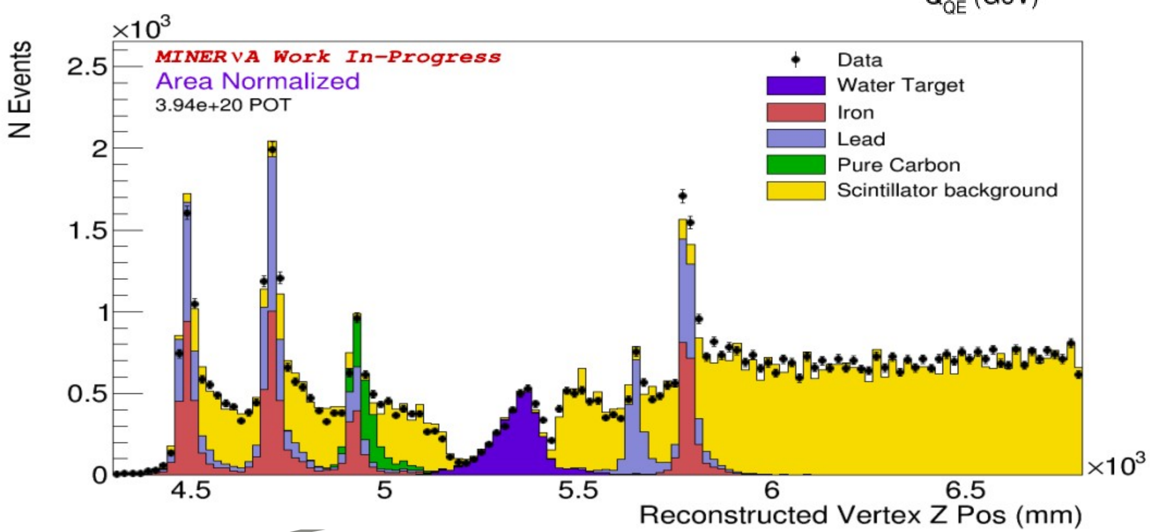
[Jeffrey Kleykamp]

New measurements with new beam energy





CC inclusive (μ + X) on scintillator (carbon)



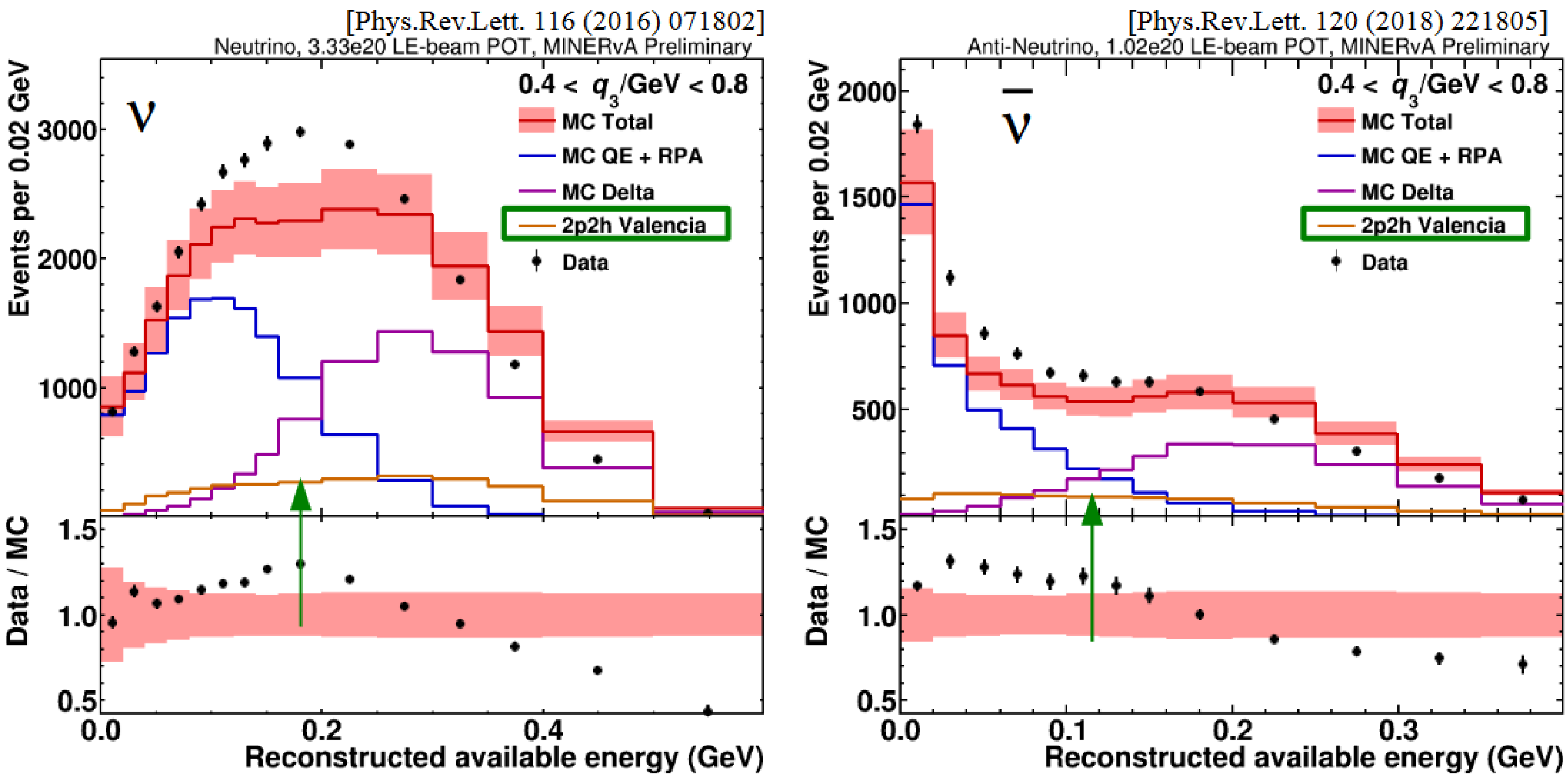
Measurements on nuclear targets upcoming

Science & Technology Facilities Council

UK Research and Innovation

[Xianguo Lu]

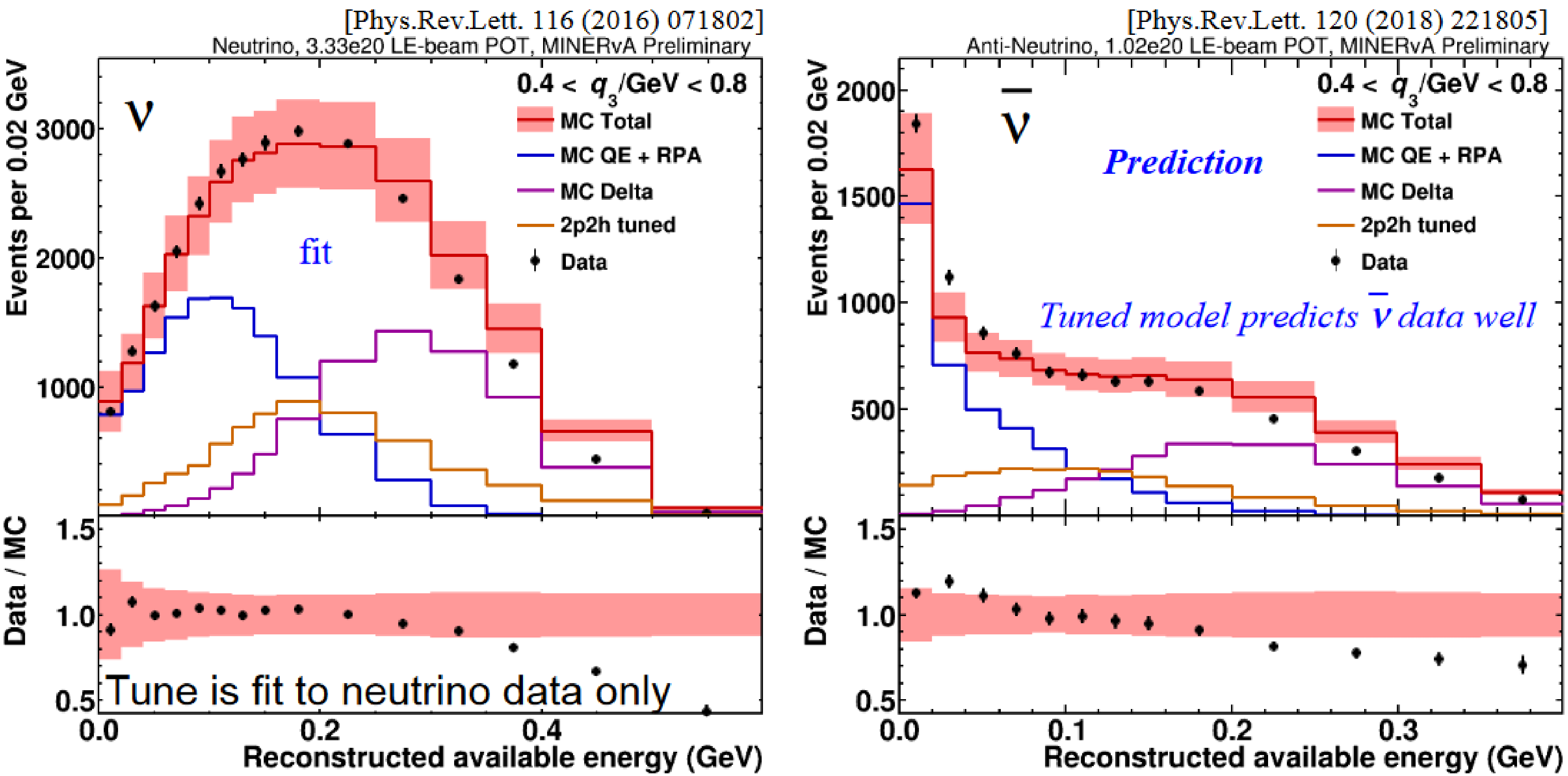
Generator tuning on CC-inc



2p2h = 2 particles 2 holes = correlated nucleon pair as initial state

[Xiangguo Lu]

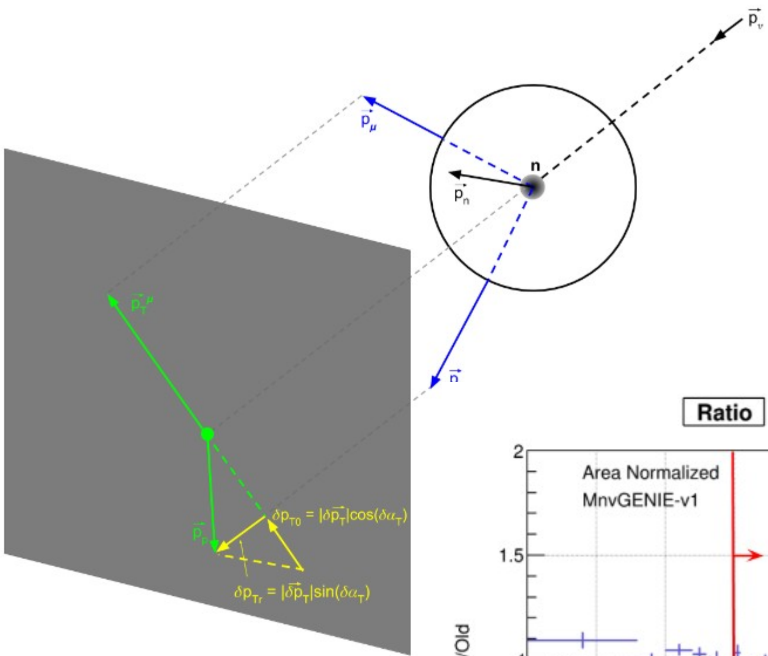
Generator tuning on CC-inc



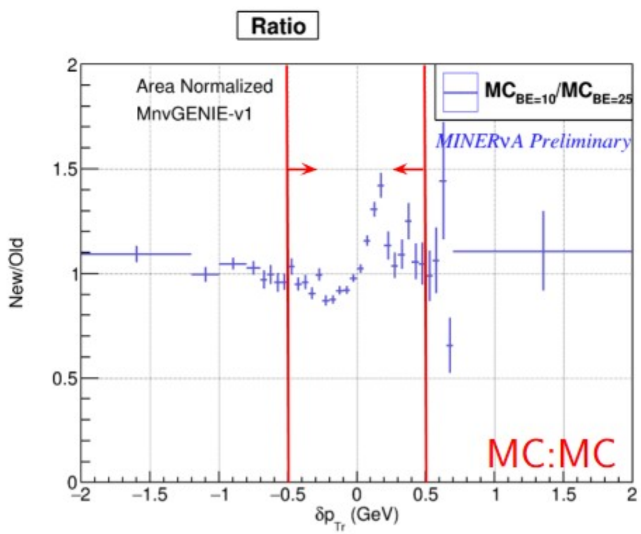
Theorists say this should not work, but it does

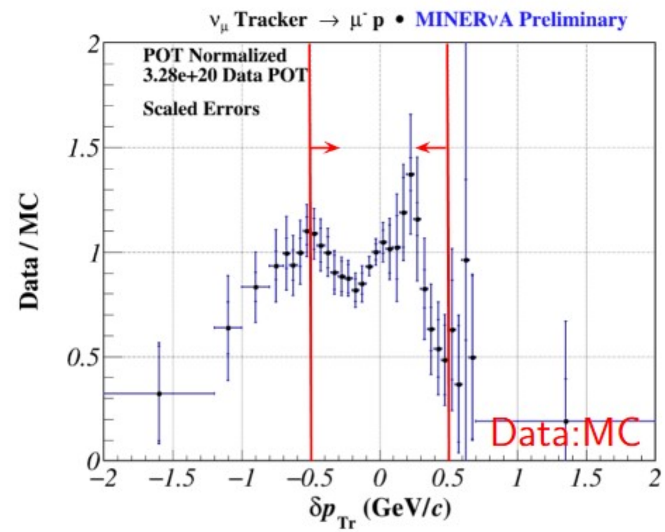
[Rob Fine]

Double transverse variables



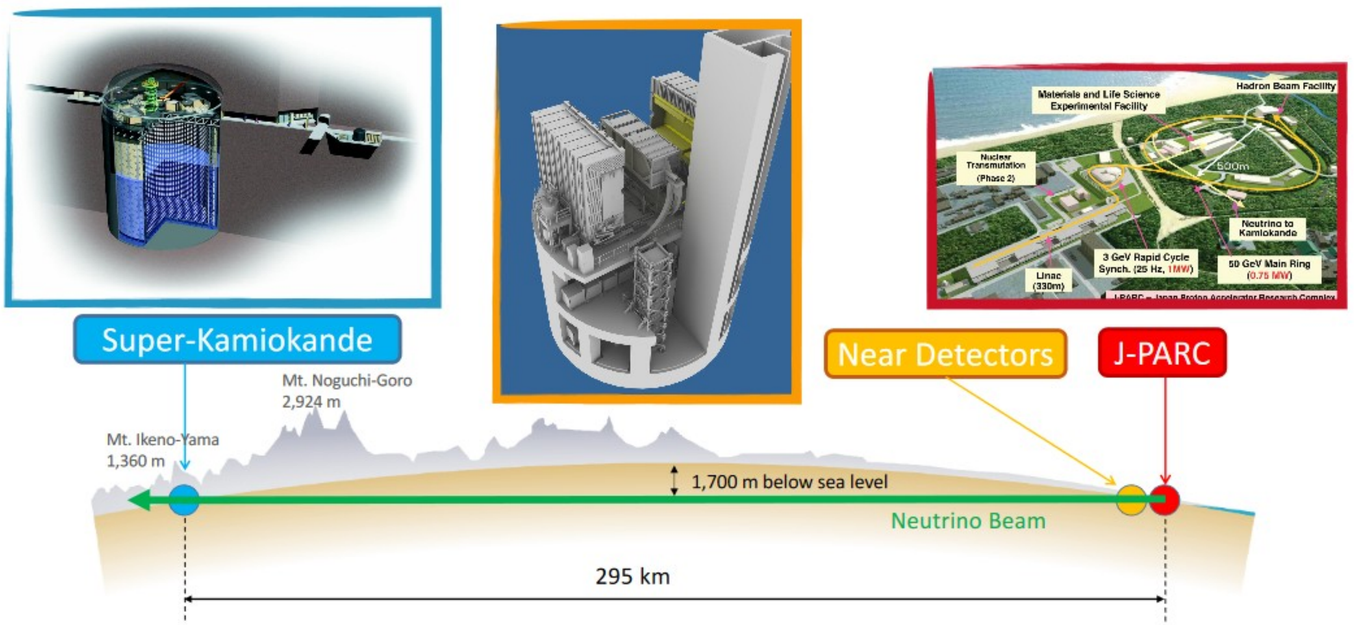
Momentum imbalance transverse to neutrino direction can teach us something about the nuclear interactions, e.g. binding energy

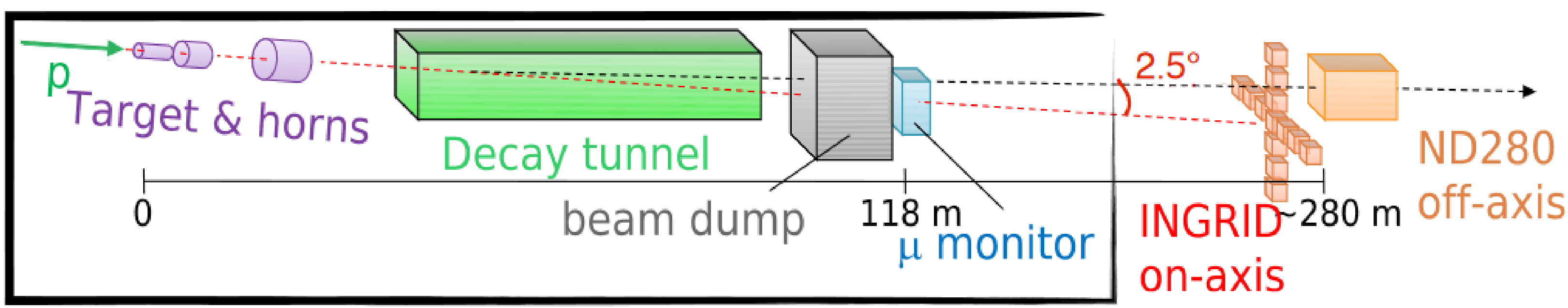




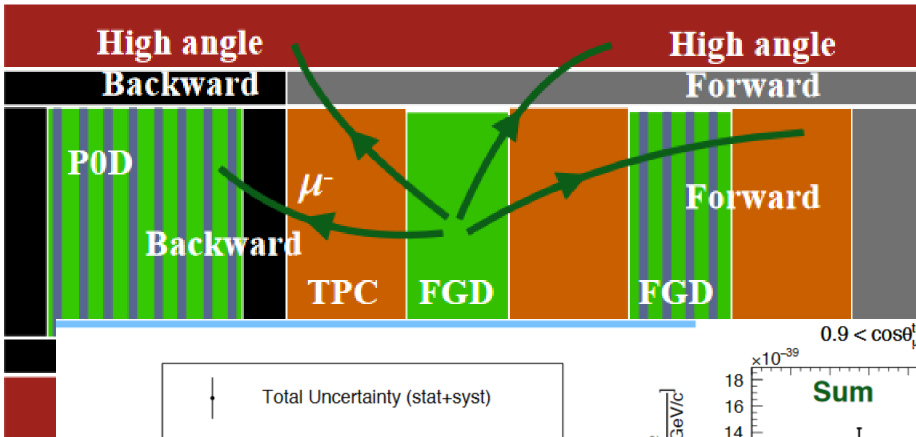
[Ciro Riccio]

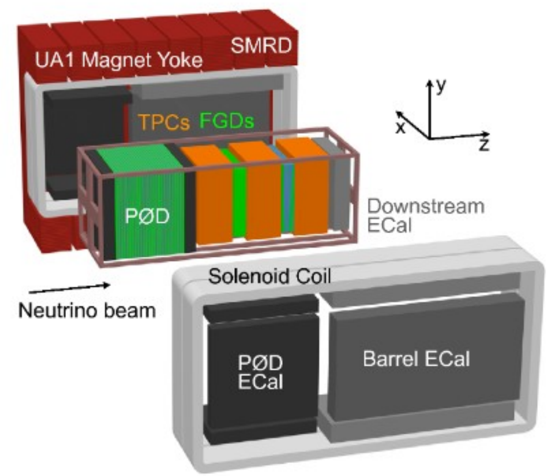
T2K

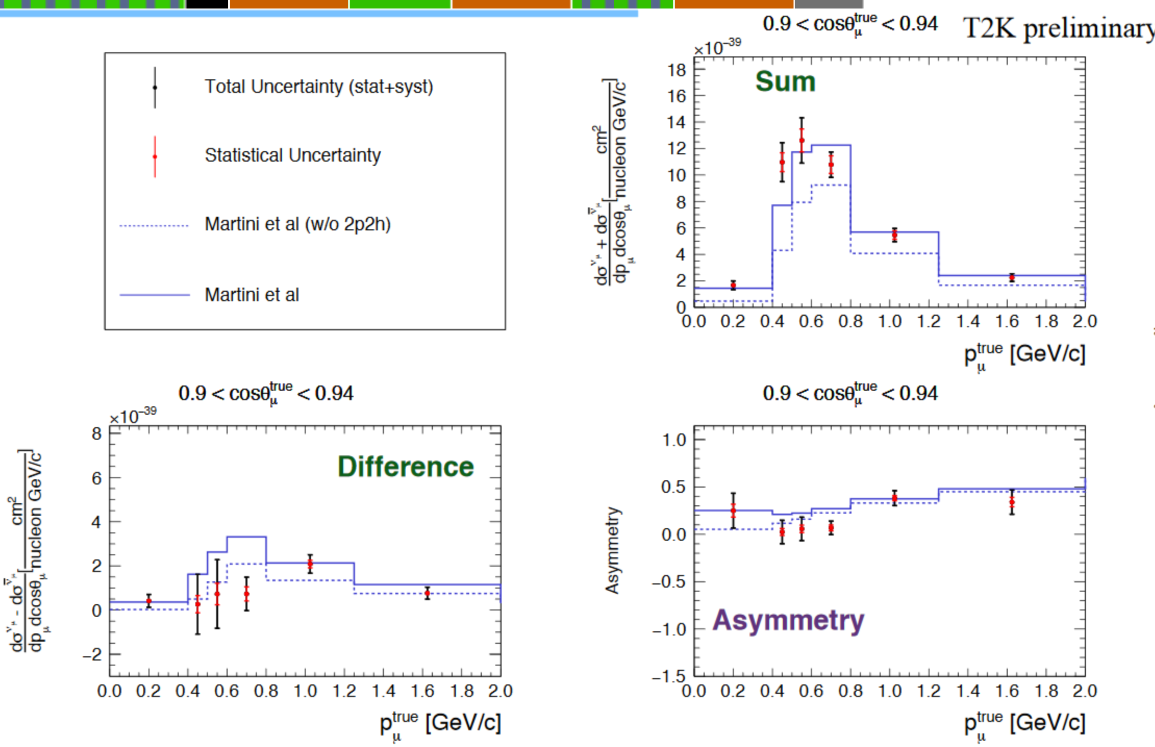


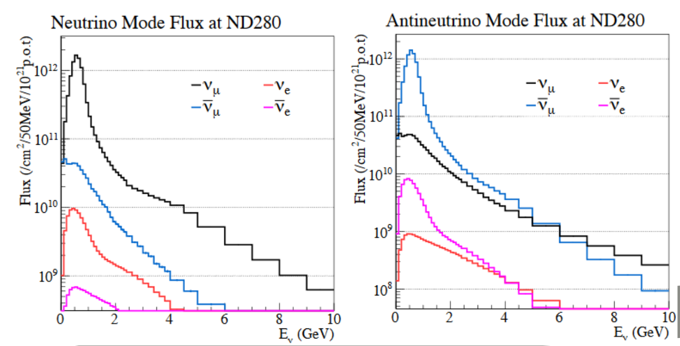


High-acceptance CC0π ν-ν̄ comparison



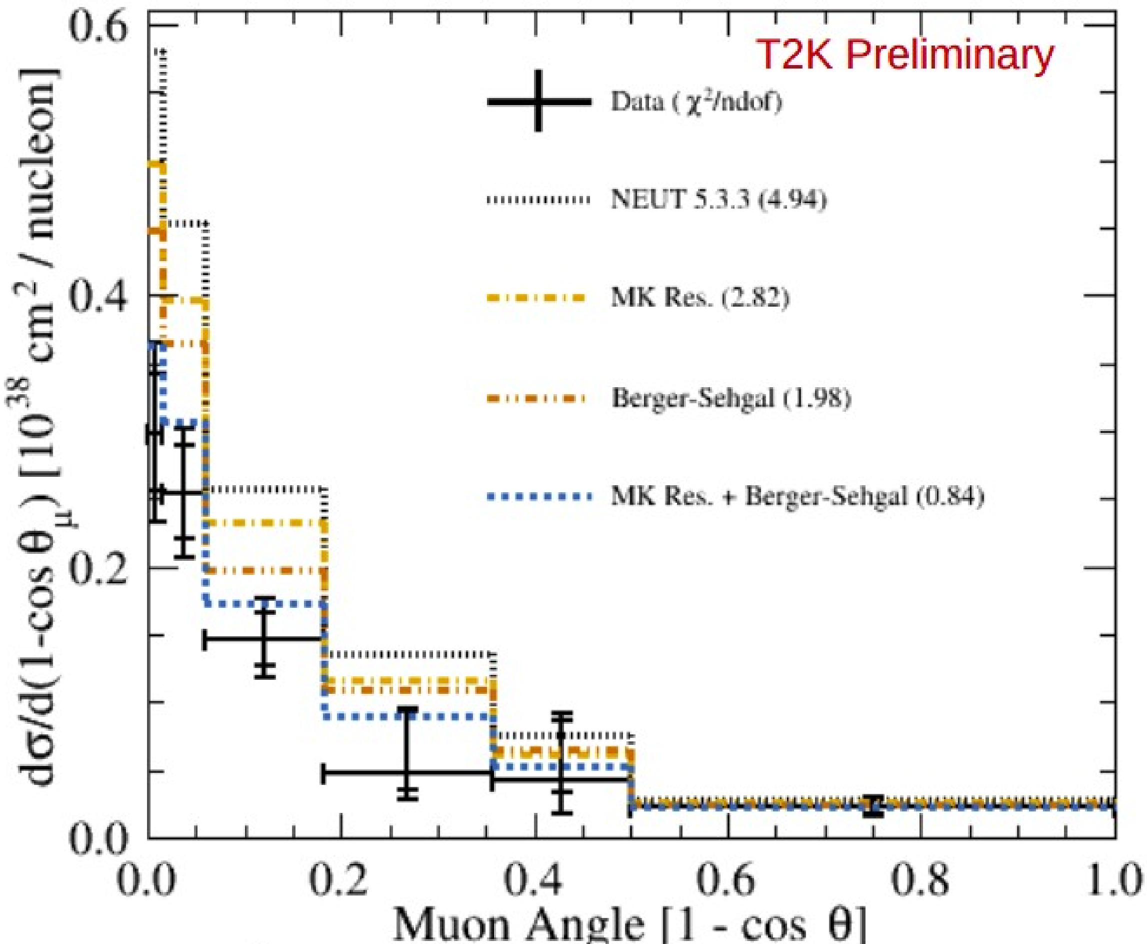


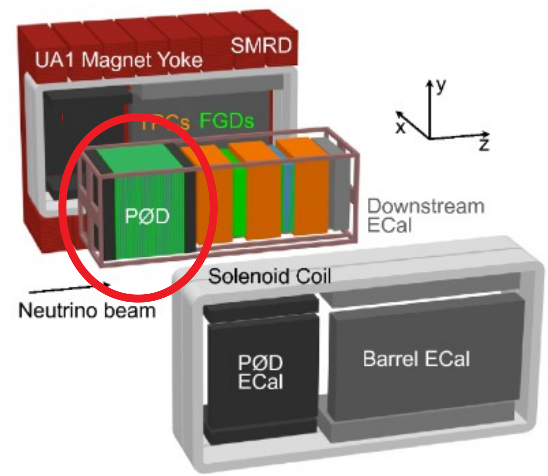




[Daniel Cherdack]

CC1π measurement favours new MK pion model

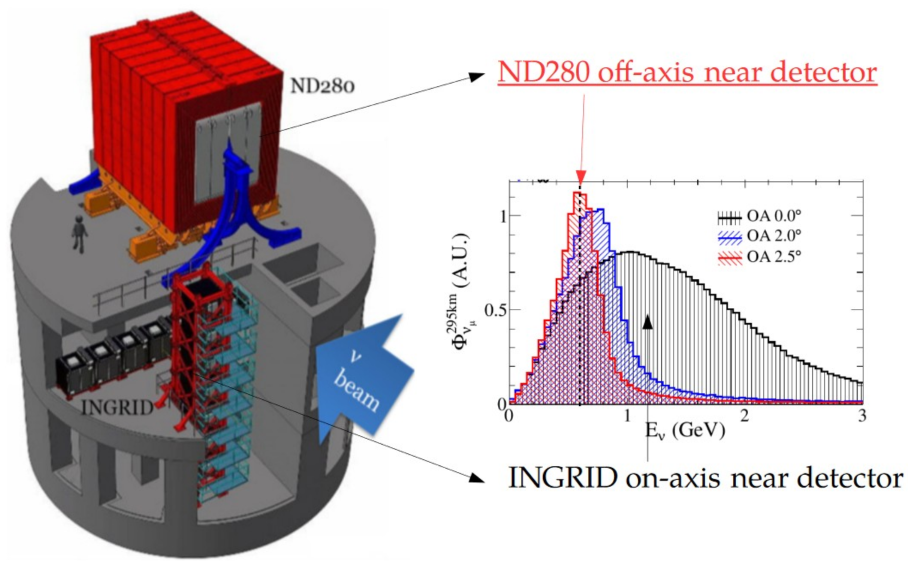


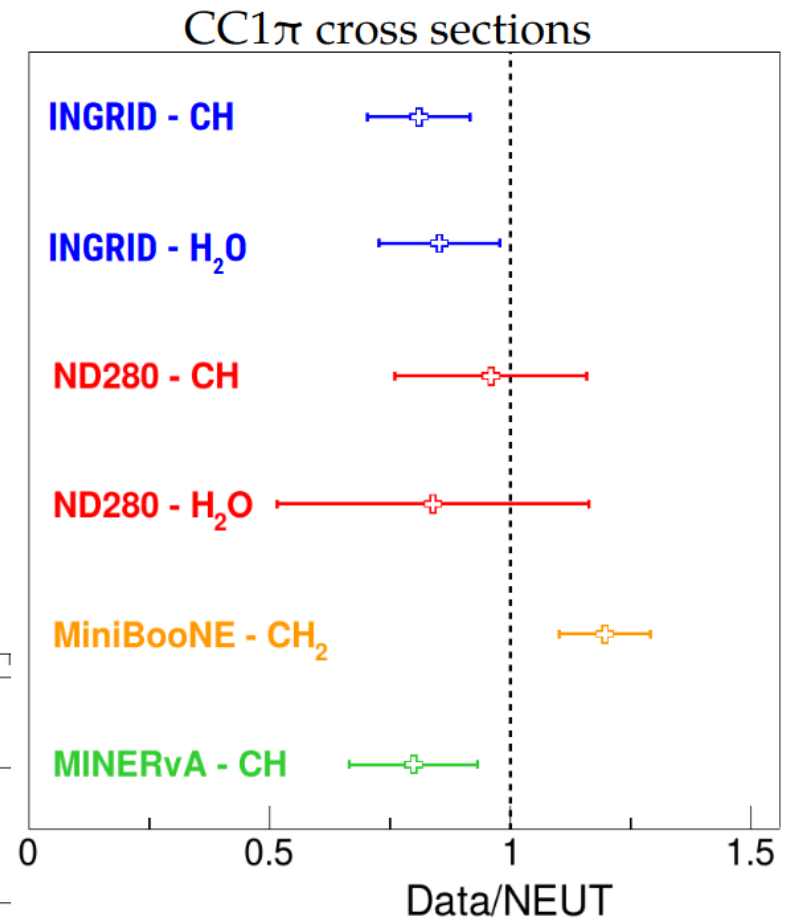


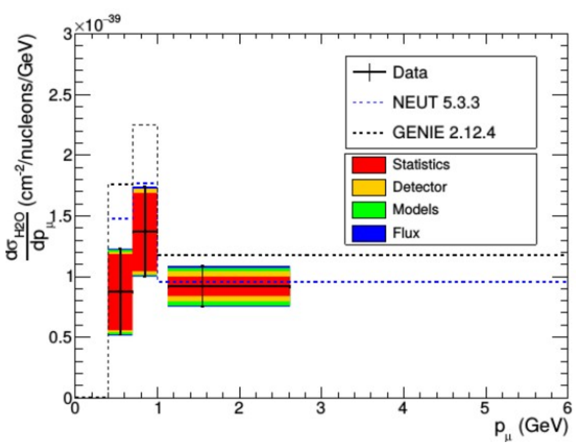
Science & Technology Facilities Council

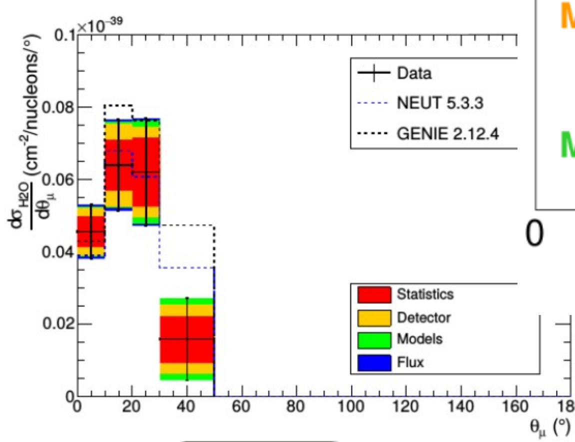
UK Research and Innovation

On-axis CC1π measurements









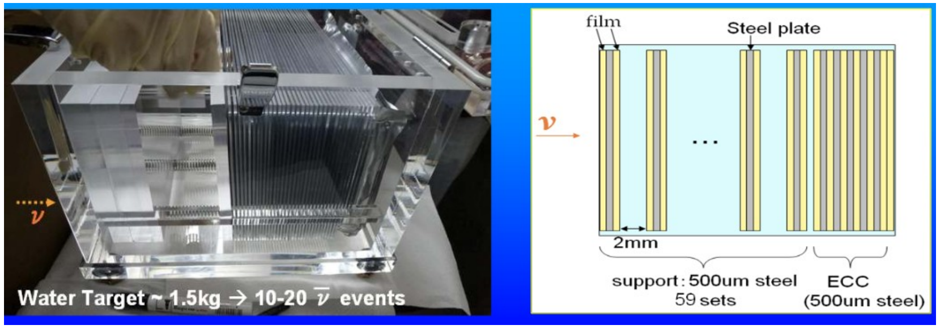
Science & Technology Facilities Council

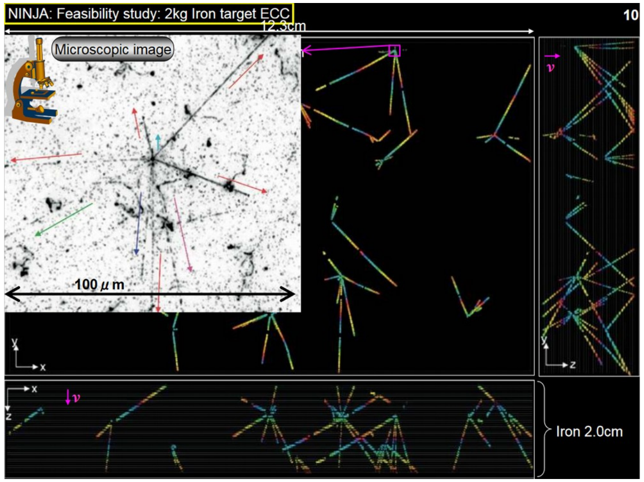
UK Research and Innovation

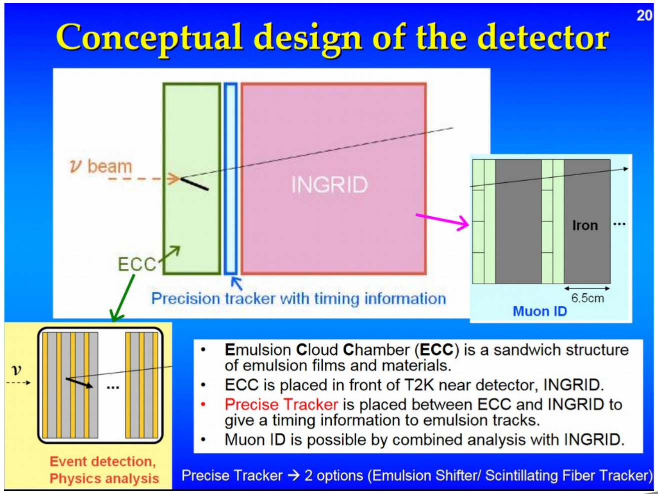
[Tsutomu Fukuda]

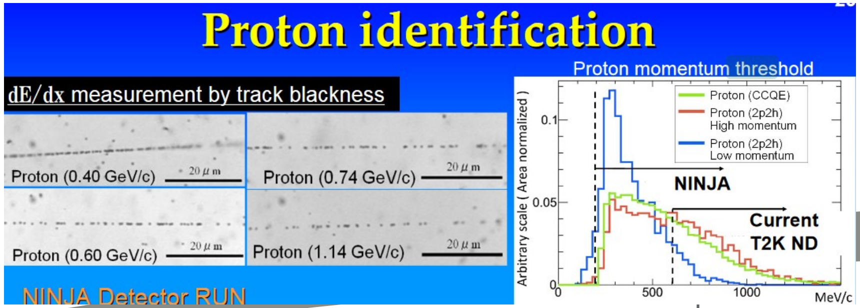
NINJA

- Neutrino Interaction research with Nuclear emulsion and J-parc Accelerator







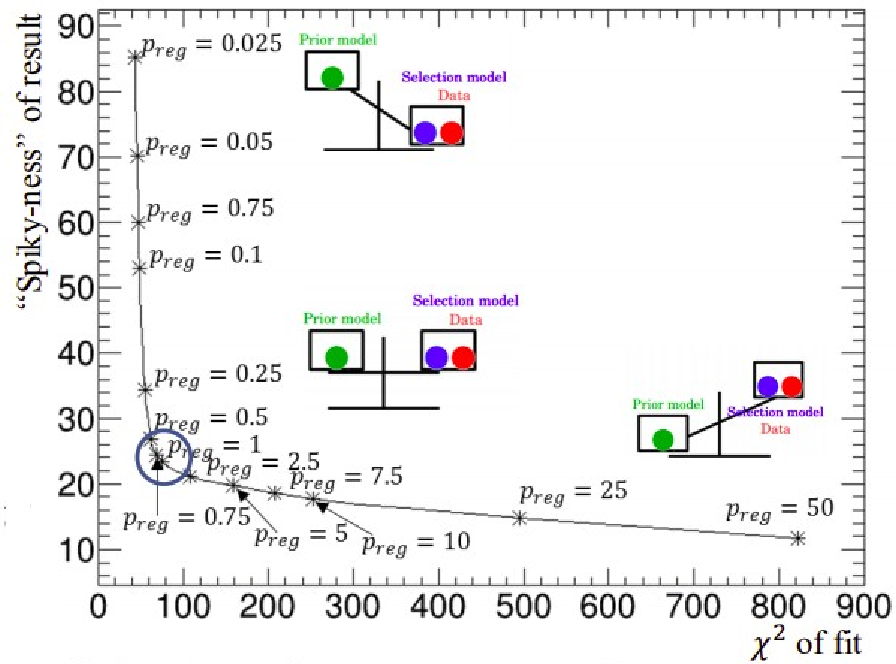


Science & Technology Facilities Council

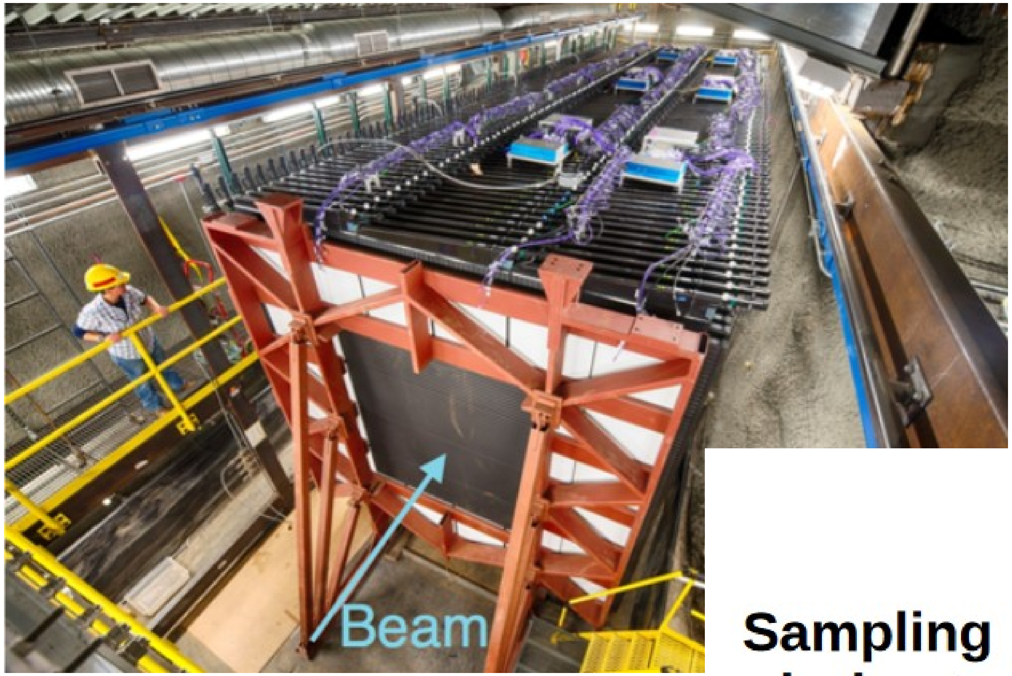
UK Research and Innovation

Method questions

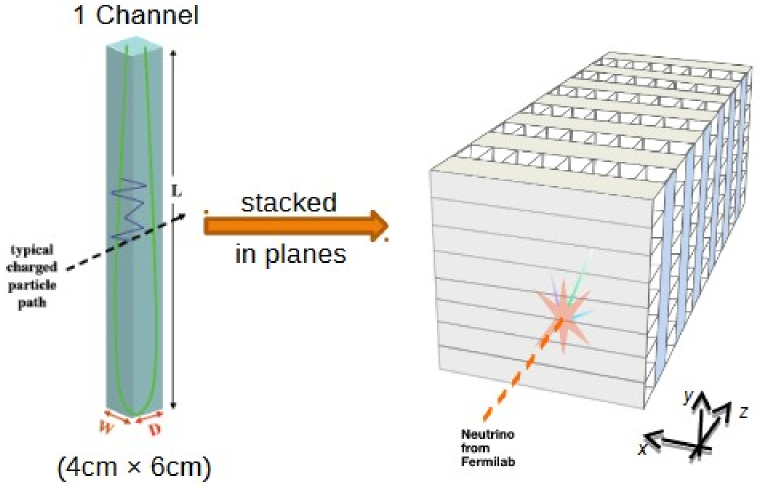
- Unfolded results can produce strongly correlated data
  - Difficult to impossible to interpret by eye
- Regularisation reduces correlation but introduces bias
  - Need to choose a regularisation strength, e.g. data-driven L-curve method
- ALWAYS also provide unregularised results
  - Most correct/useful for global fits
- Alternative to unfolding, provide raw data and forward-folding matrix
  - Response Matrix Utilities (ReMU) software package provides needed functionality



NOvA

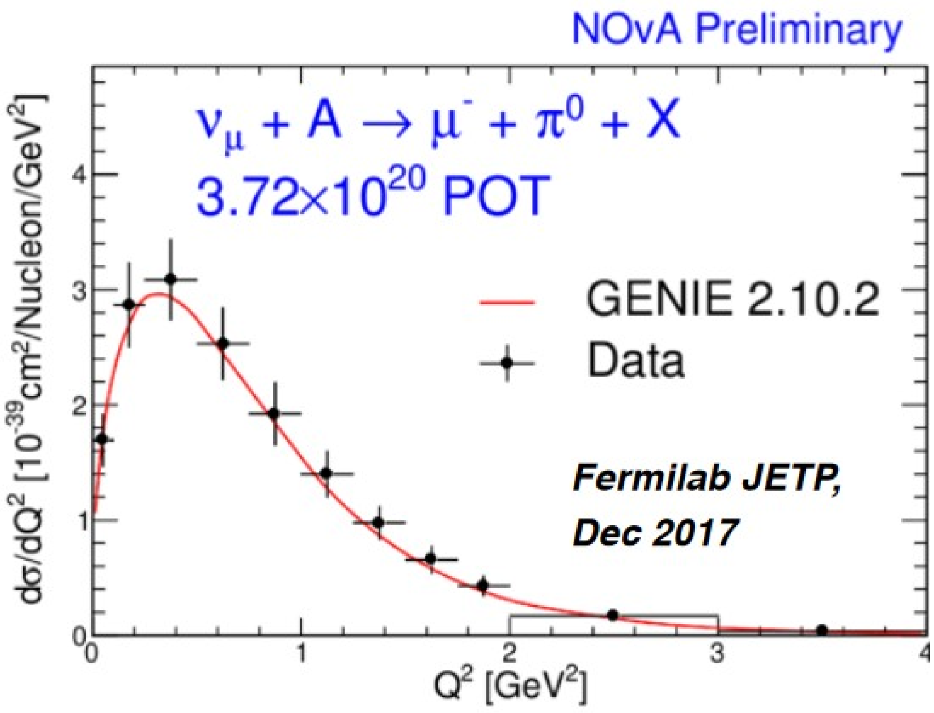


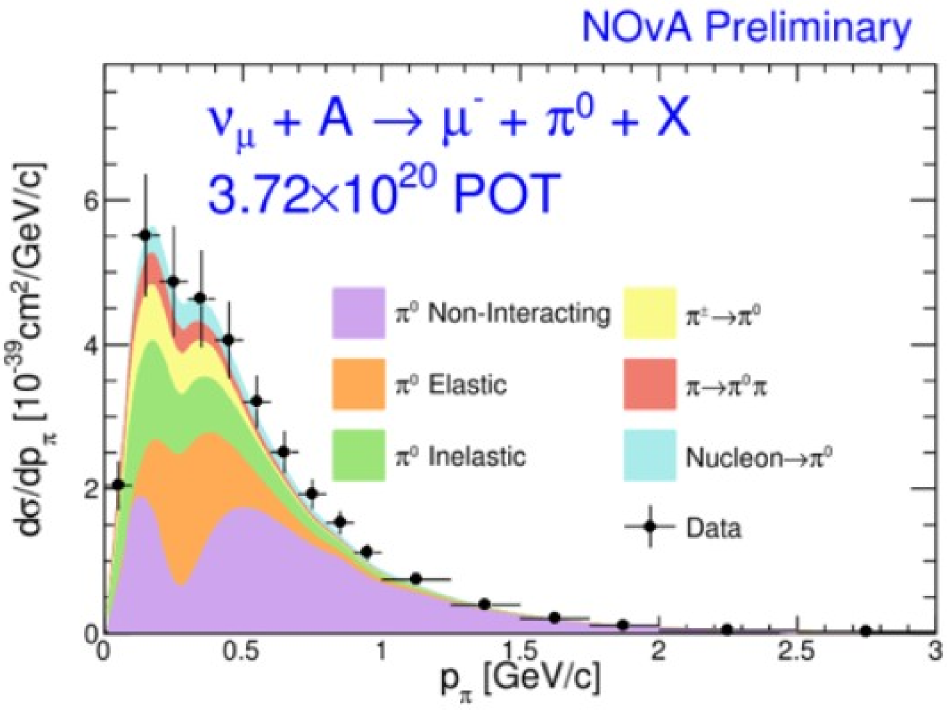
Sampling calorimeter detectors



[Leo Aliagat]

CC π^0 inclusive

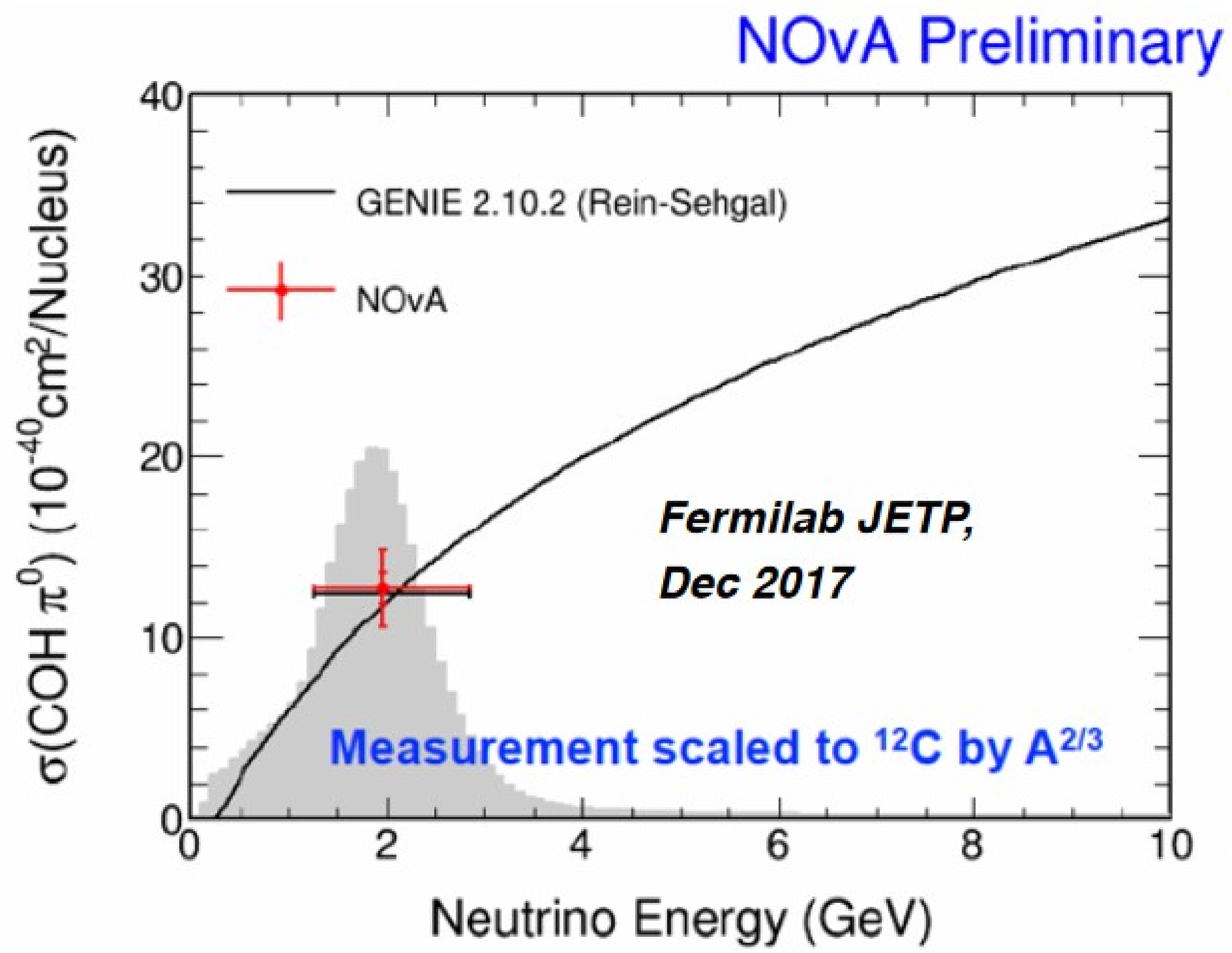


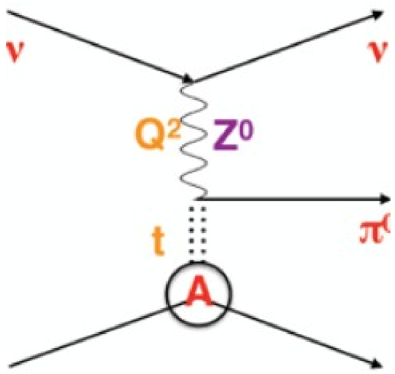


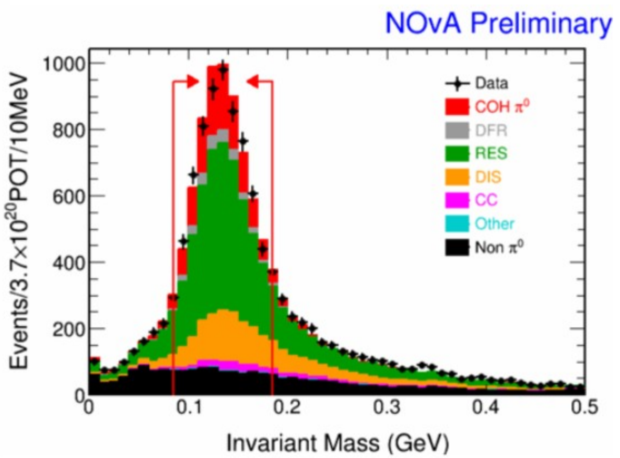
Science & Technology Facilities Council

UK Research and Innovation

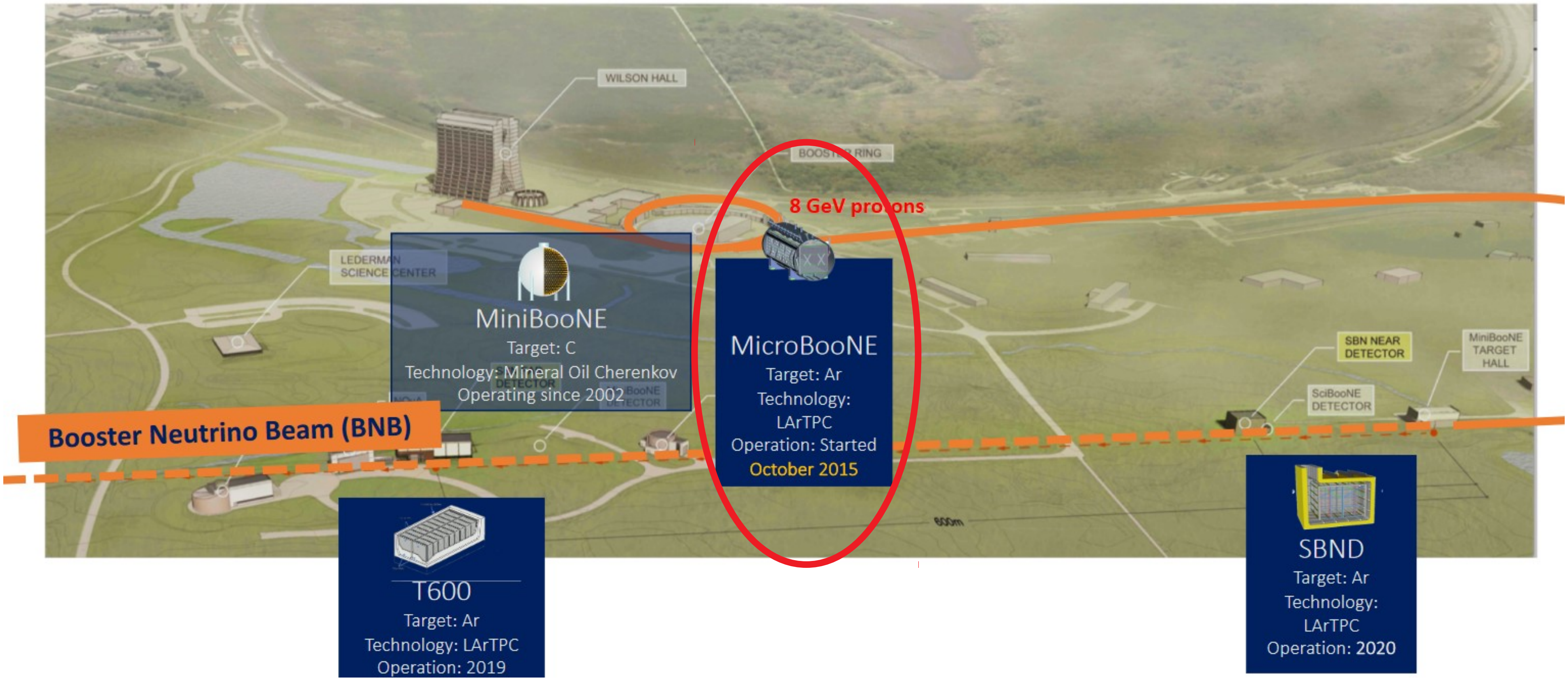
NC coh π^0


NOvA Preliminary
GENIE 2.10.2 (Rein-Sehgal)
NOvA
Fermilab JETP, Dec 2017
Measurement scaled to ¹²C by A^2/3
σ(COH π⁰) (10⁻⁴⁰cm²/Nucleus)
Neutrino Energy (GeV)
0
2
4
6
8
10
0
10
20
30
40


ν
ν
Q²
Z⁰
t
A


NOvA Preliminary
Events/3.7×10²⁰POT/10MeV
Invariant Mass (GeV)
Data
COH π⁰
DFR
RES
DIS
CC
Other
Non π⁰
0
200
400
600
800
1000
0
0.1
0.2
0.3
0.4
0.5

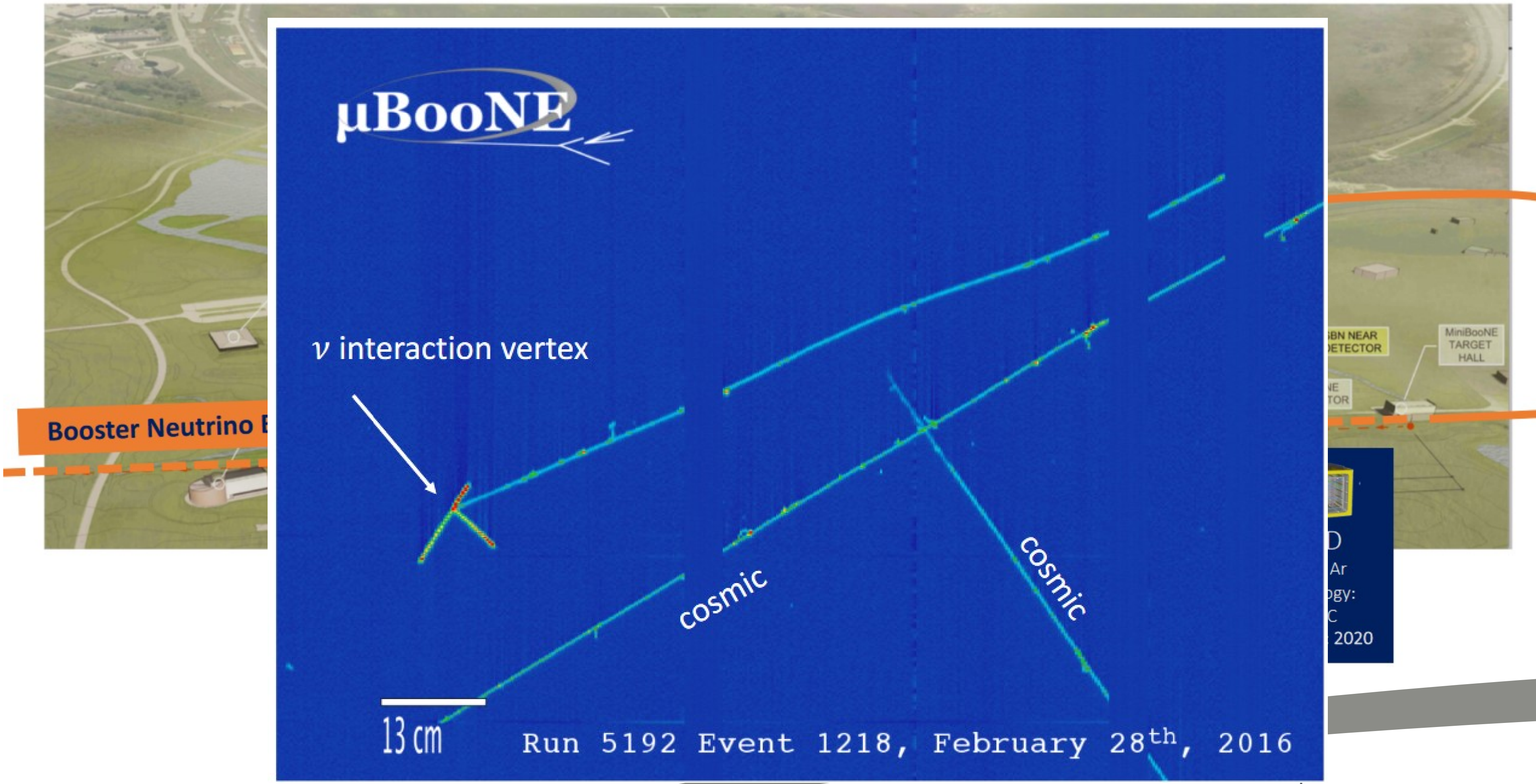
MicroBooNE



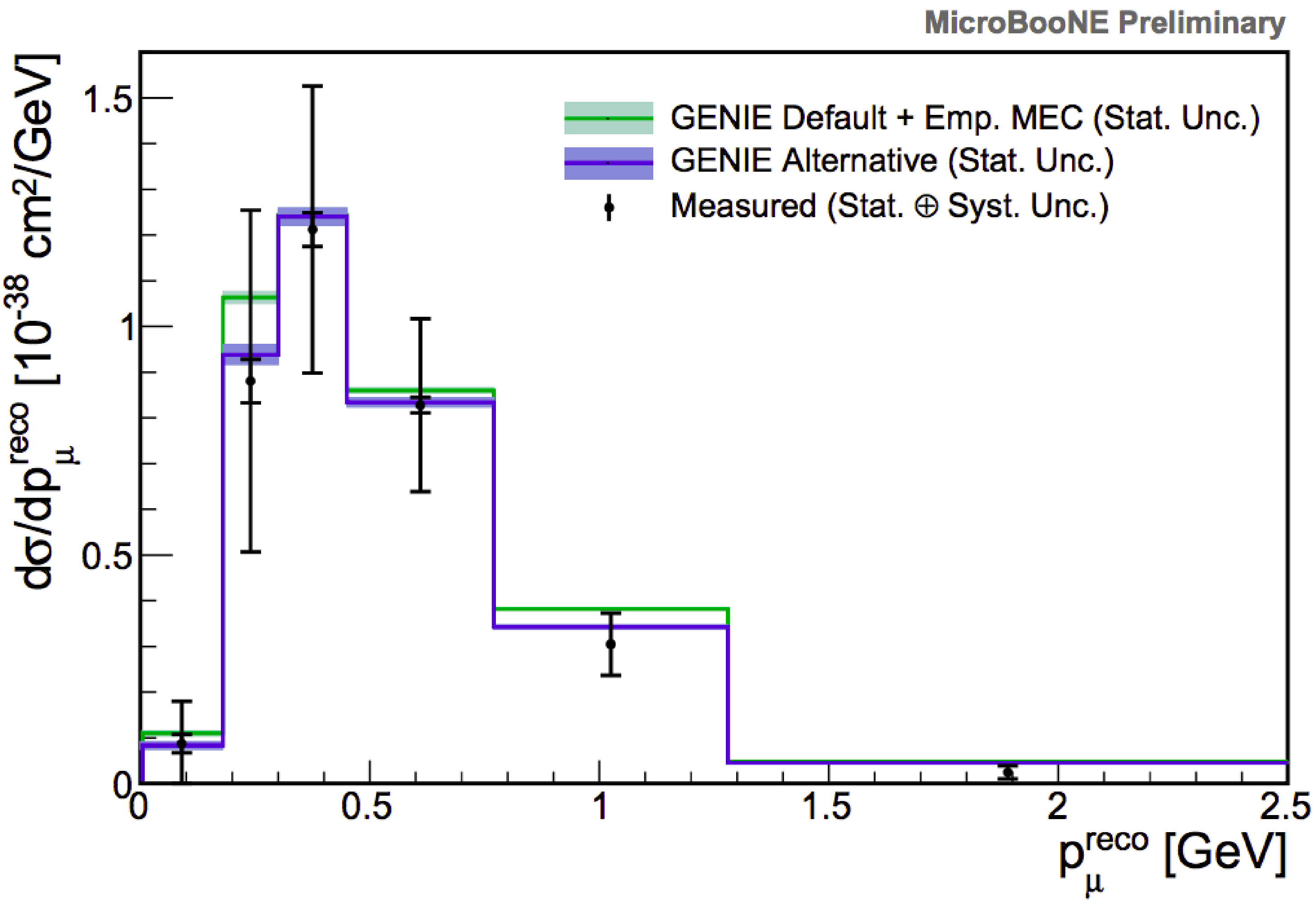
Science & Technology Facilities Council

UK Research and Innovation

MicroBooNE

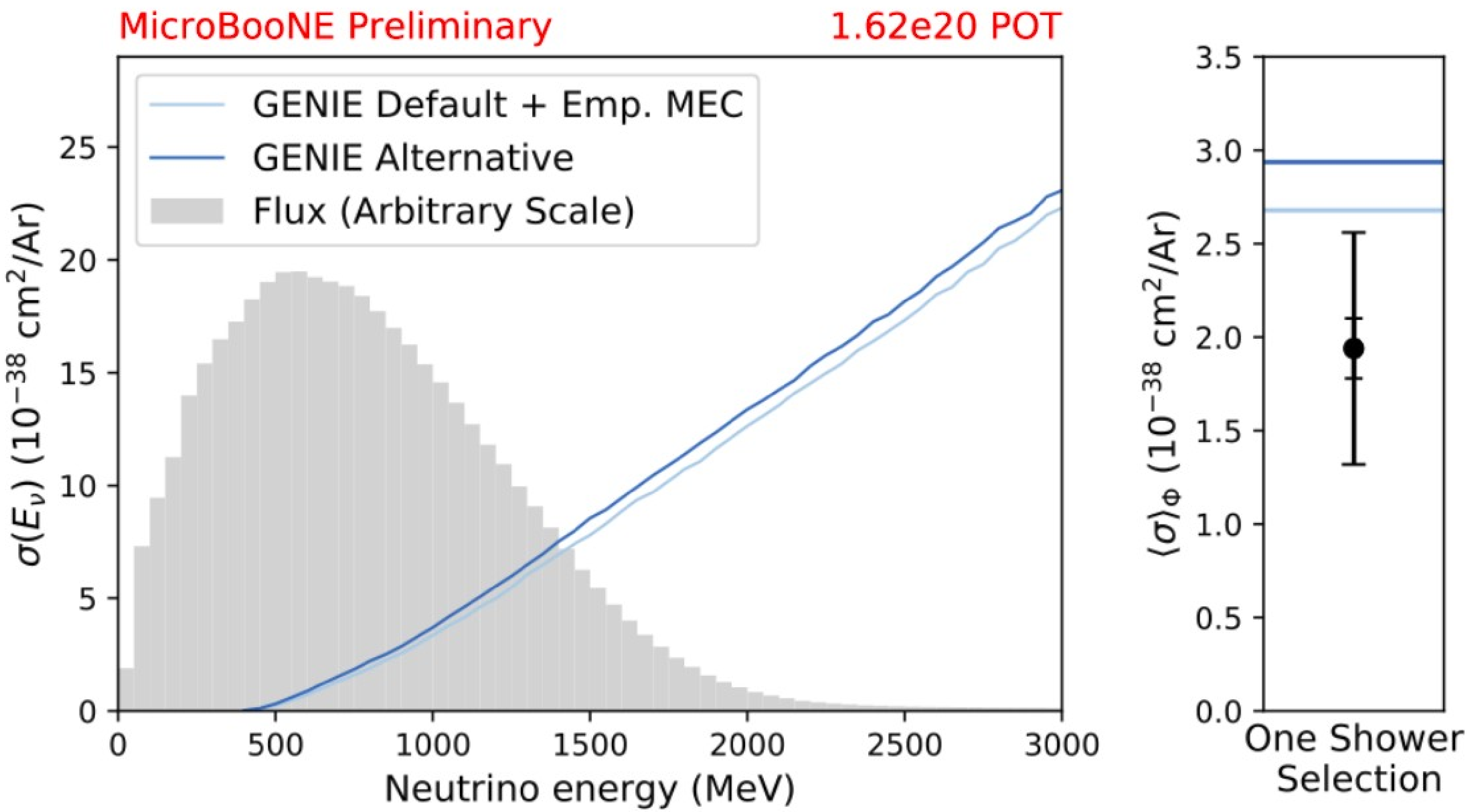

μBooNE
ν interaction vertex
cosmic
cosmic
13 cm
Run 5192 Event 1218, February 28th, 2016
Booster Neutrino
MiniBooNE TARGET HALL

CC inclusive (μ + X)



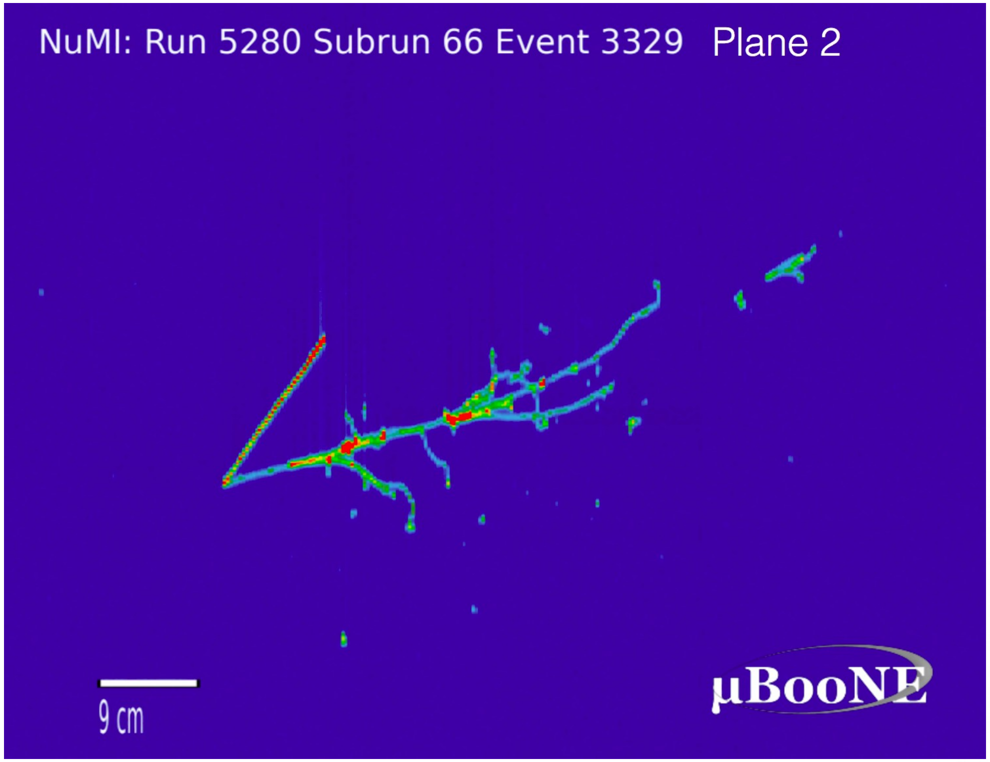
Science & Technology Facilities Council
UK Research and Innovation

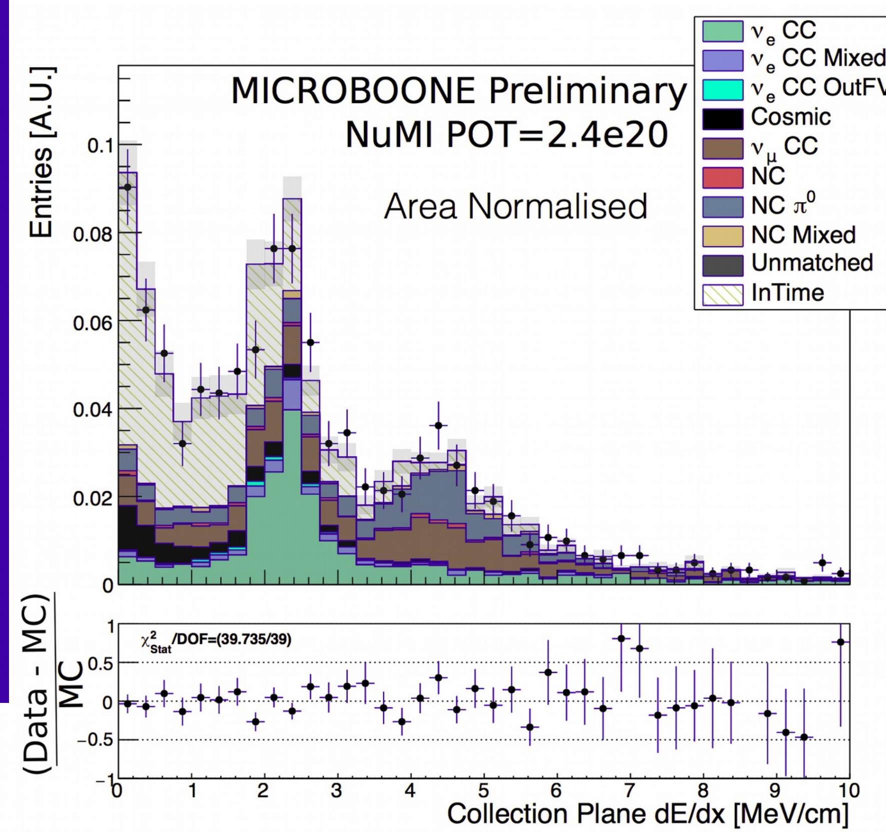
CC π^0 inclusive

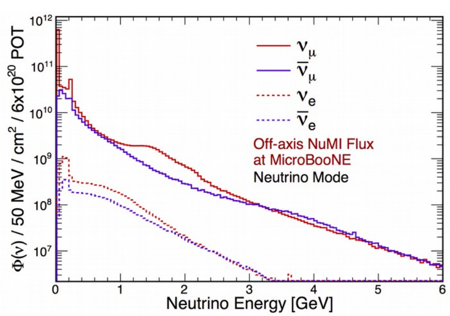


Science & Technology Facilities Council
UK Research and Innovation

Preparation for ν_e measurement





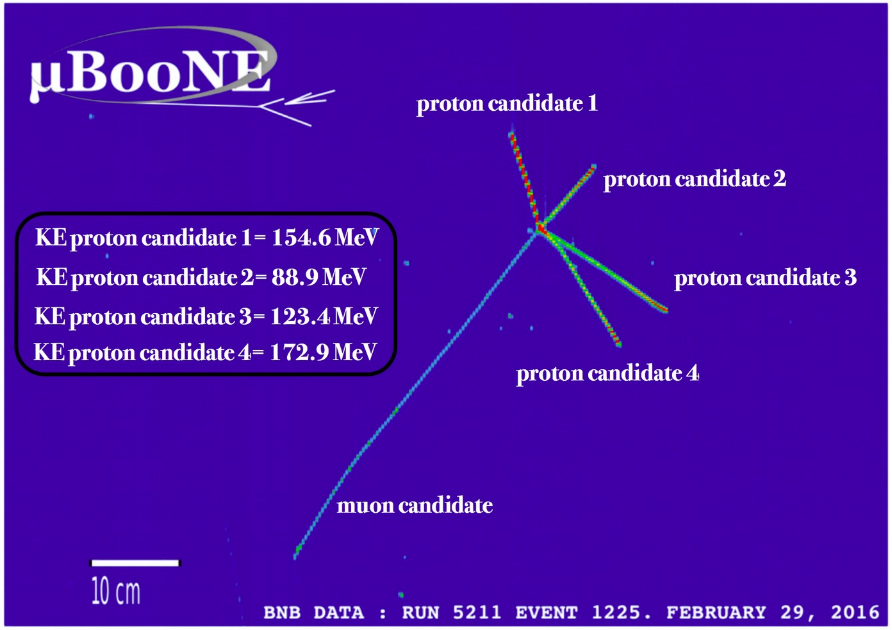


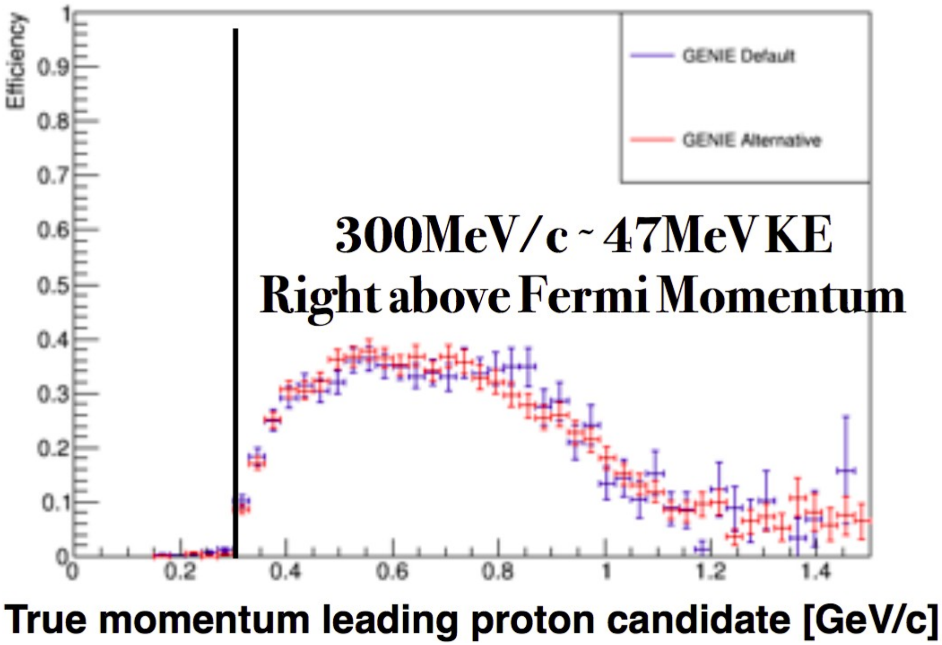
Science & Technology Facilities Council

UK Research and Innovation

[Raquel Castillo Fernandez]

Protons in MicroBooNE





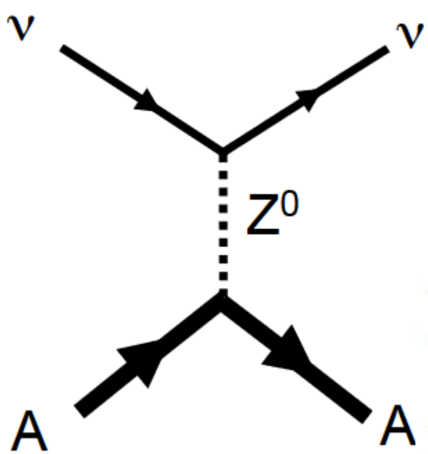
Work is on-going to lower the threshold towards the technical limit of ~20 MeV KE (wire pitch).

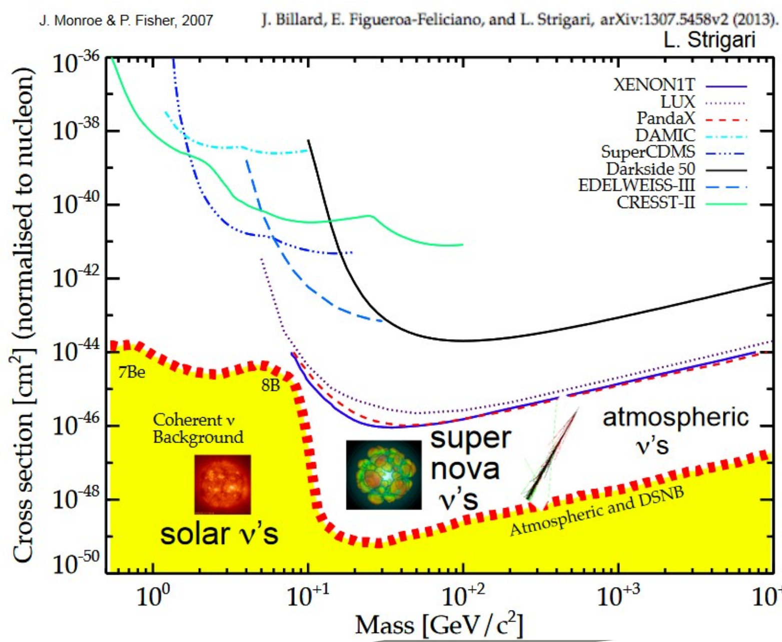
Science & Technology Facilities Council

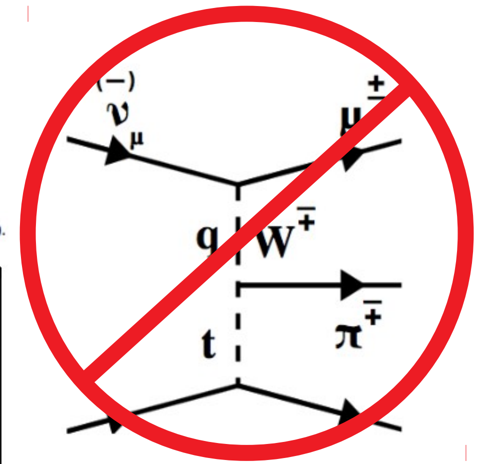
UK Research and Innovation

CEvNS

- Coherent Elastic Neutrino-Nucleus Scattering





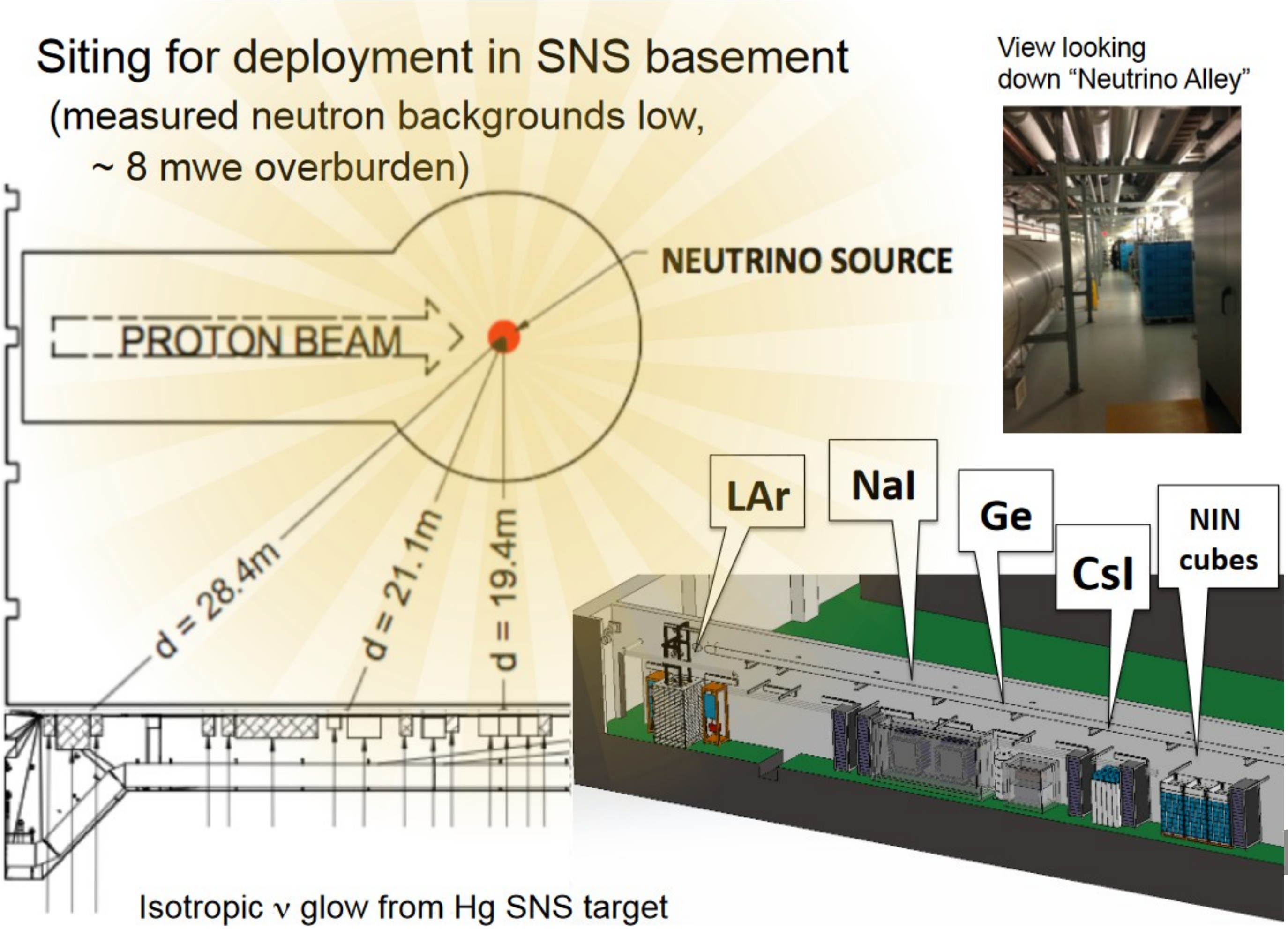


Science & Technology Facilities Council

UK Research and Innovation

[Kate Scholberg]

Stopping pion neutrino beam



Science & Technology Facilities Council

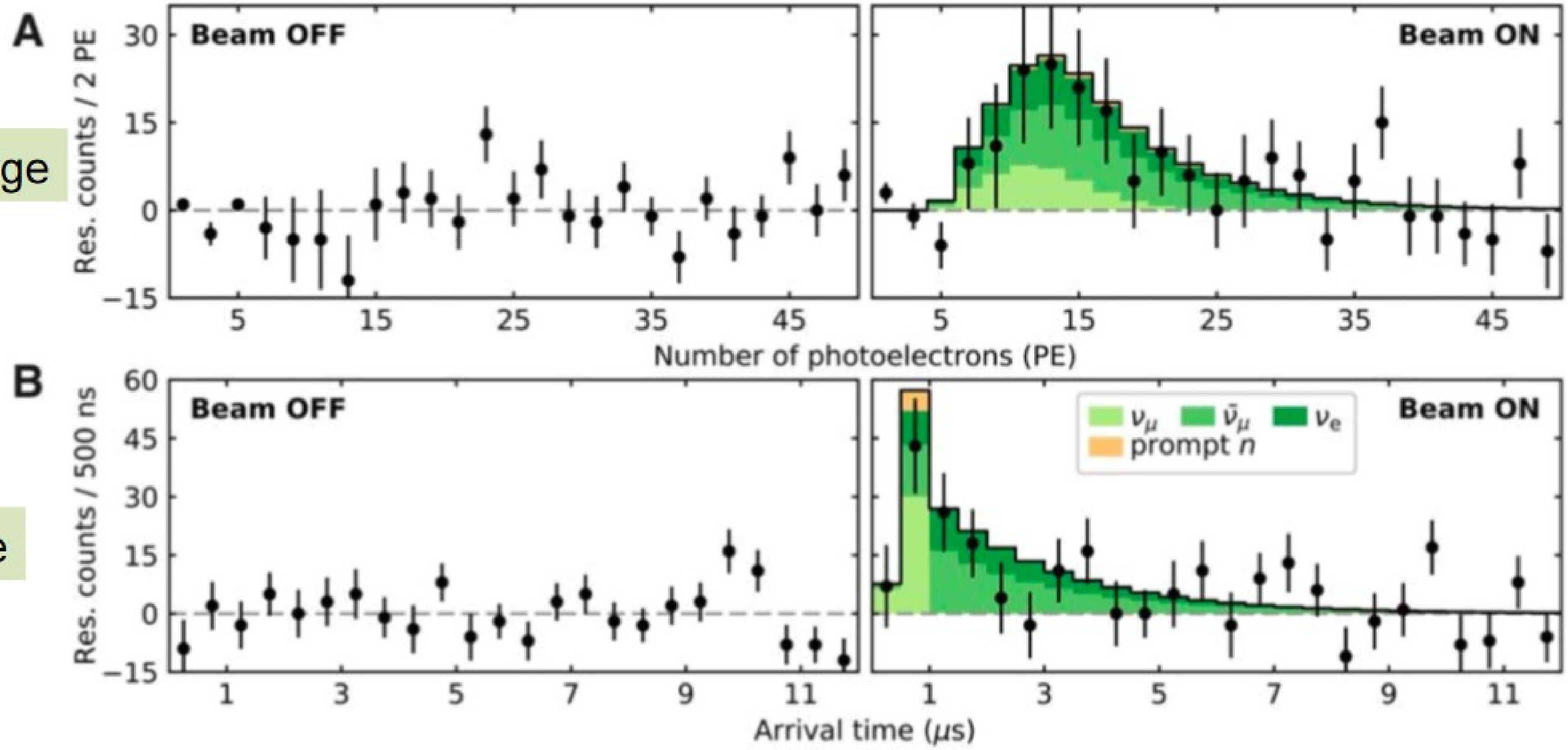
UK Research and Innovation

First light

First light at the SNS with 14.6-kg CsI[Na] detector

Charge

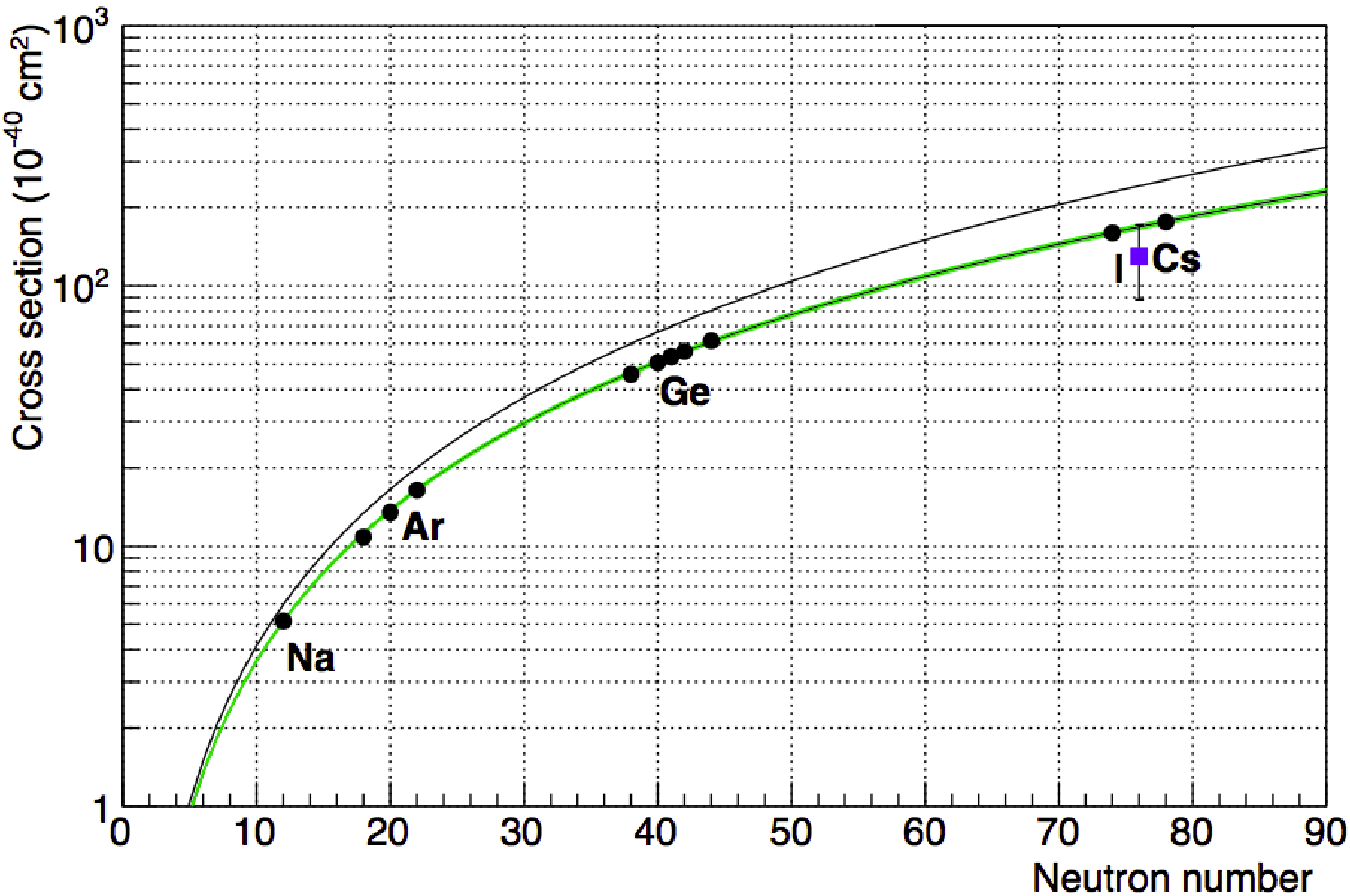
Time



Science & Technology Facilities Council

UK Research and Innovation

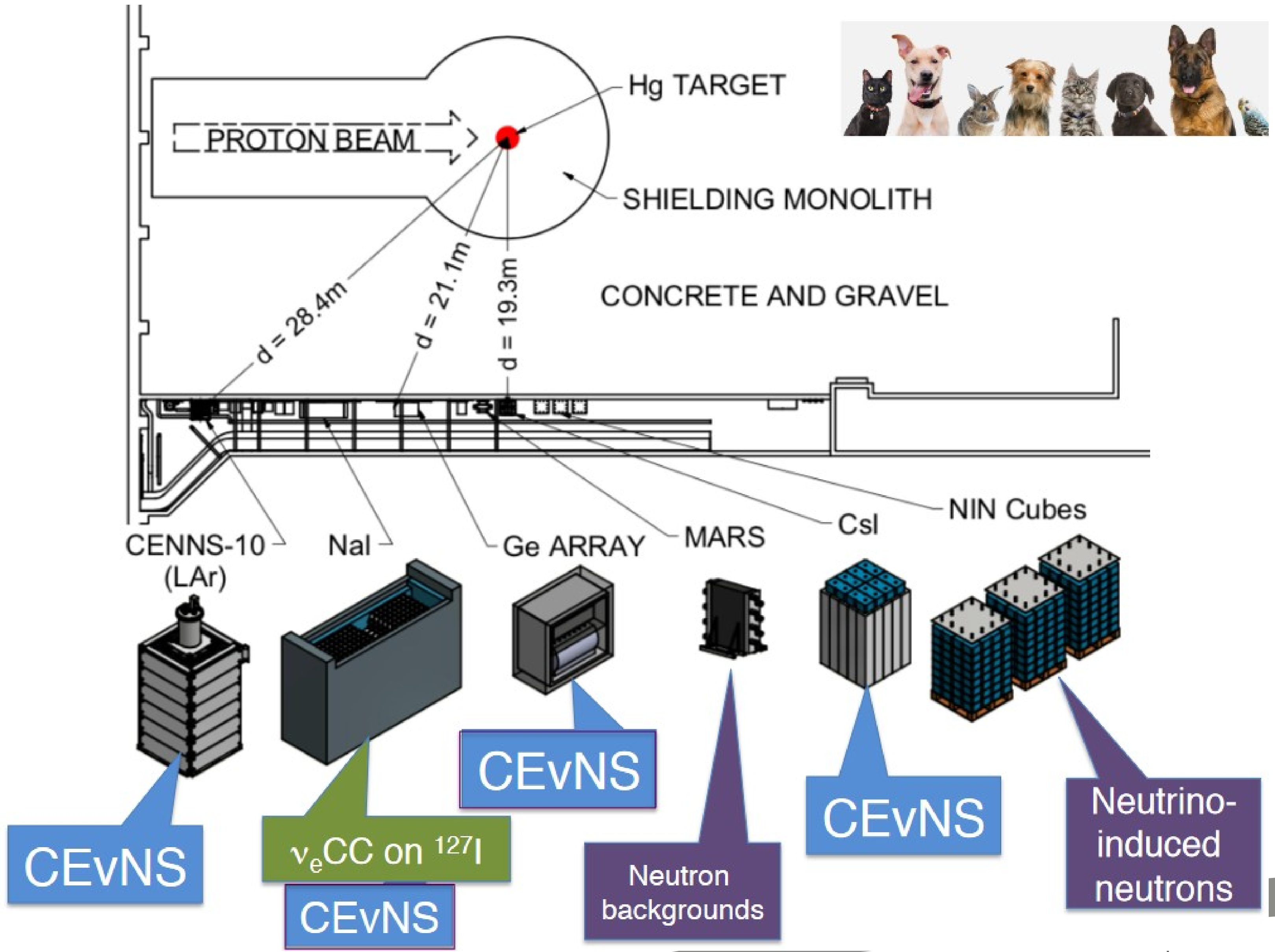
But much more to do



Science & Technology Facilities Council

UK Research and Innovation

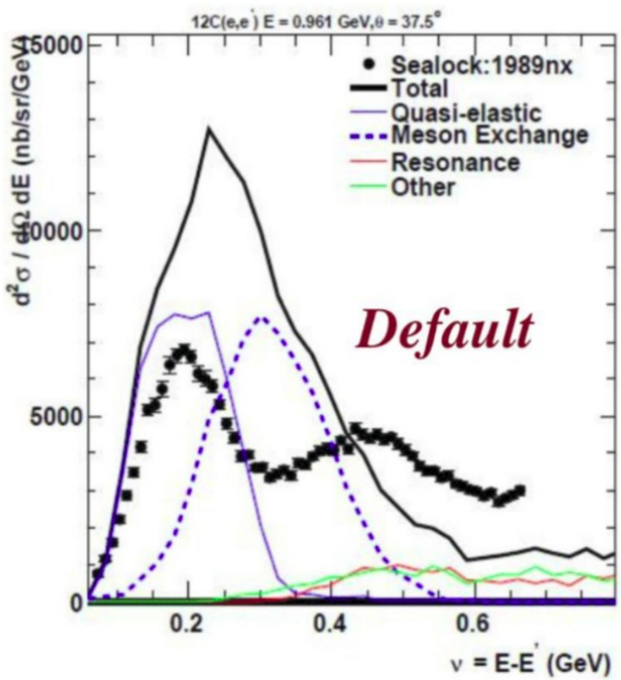
[Kate Scholberg]

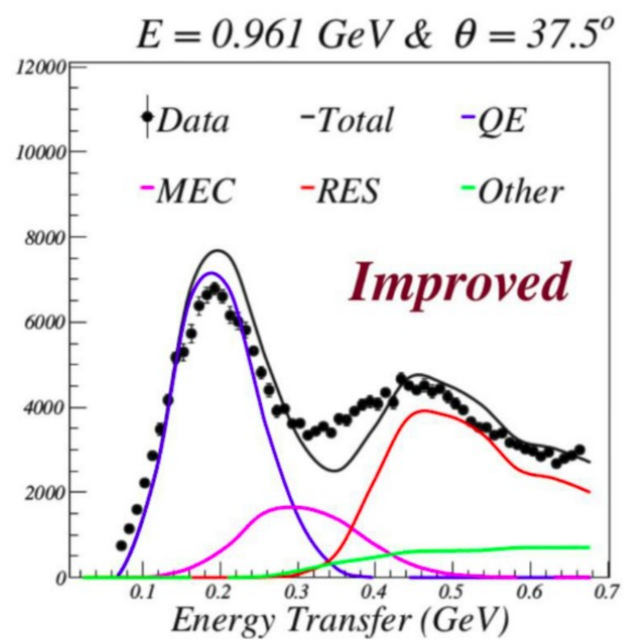


Electron scattering

***Generator vs (e,e') data**

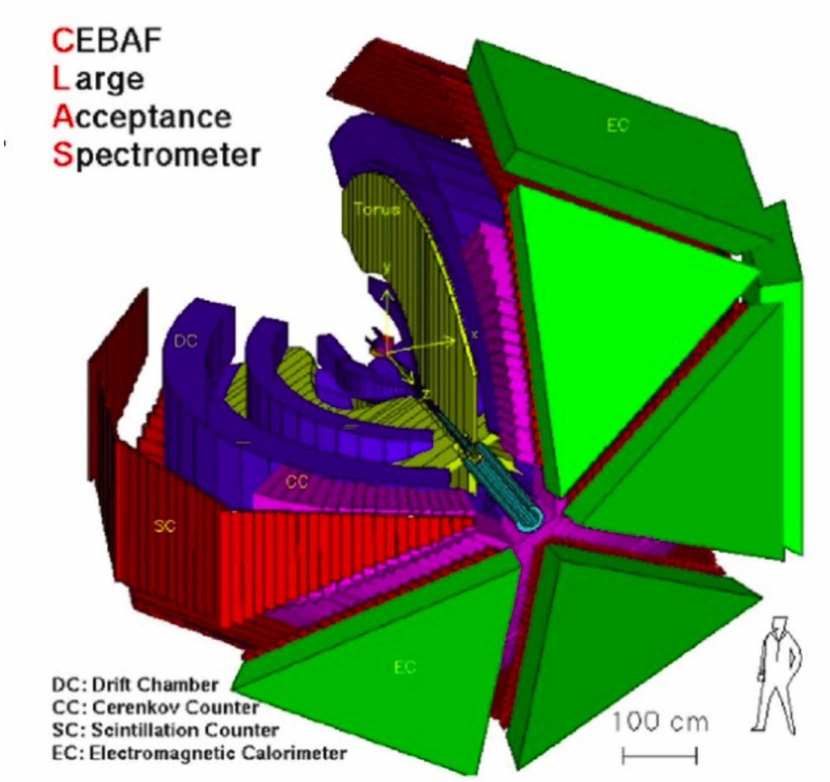
When we got started ...   *... Today*



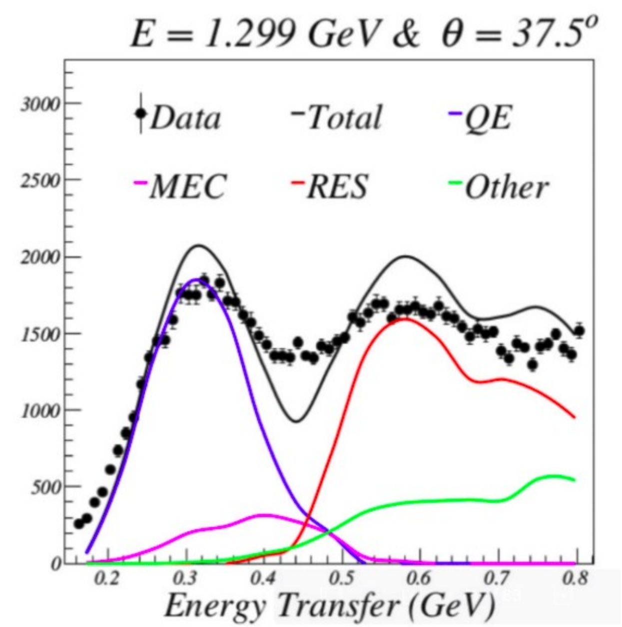


**Genie R-2_12_10*

e4nu



^{12}C



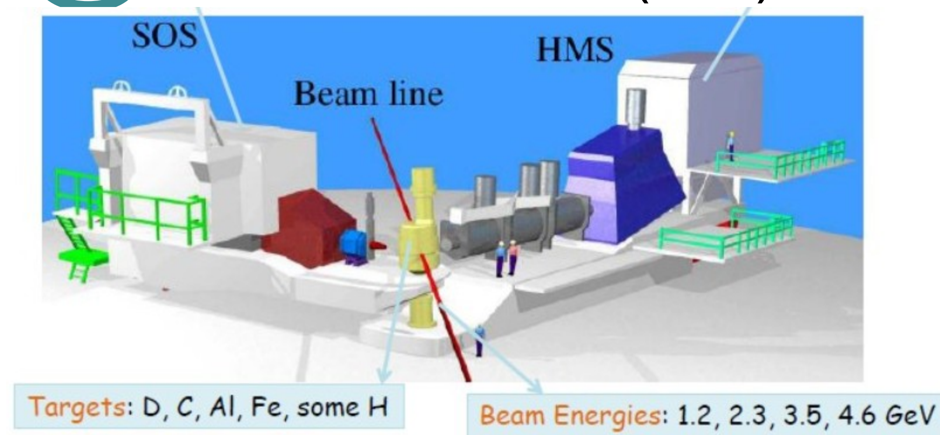
Science & Technology Facilities Council

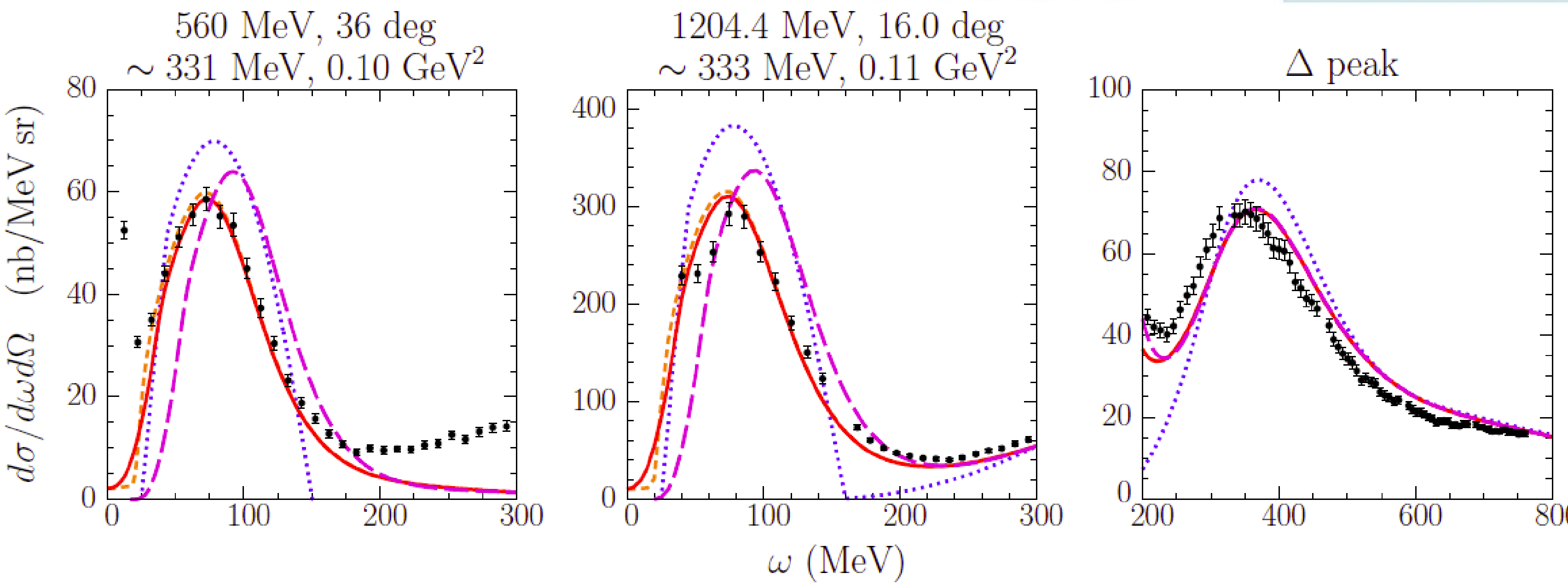
UK Research and Innovation

E04-001 @ JLab

(e,e')

New data covers QE,
resonances,
and beginning of DIS

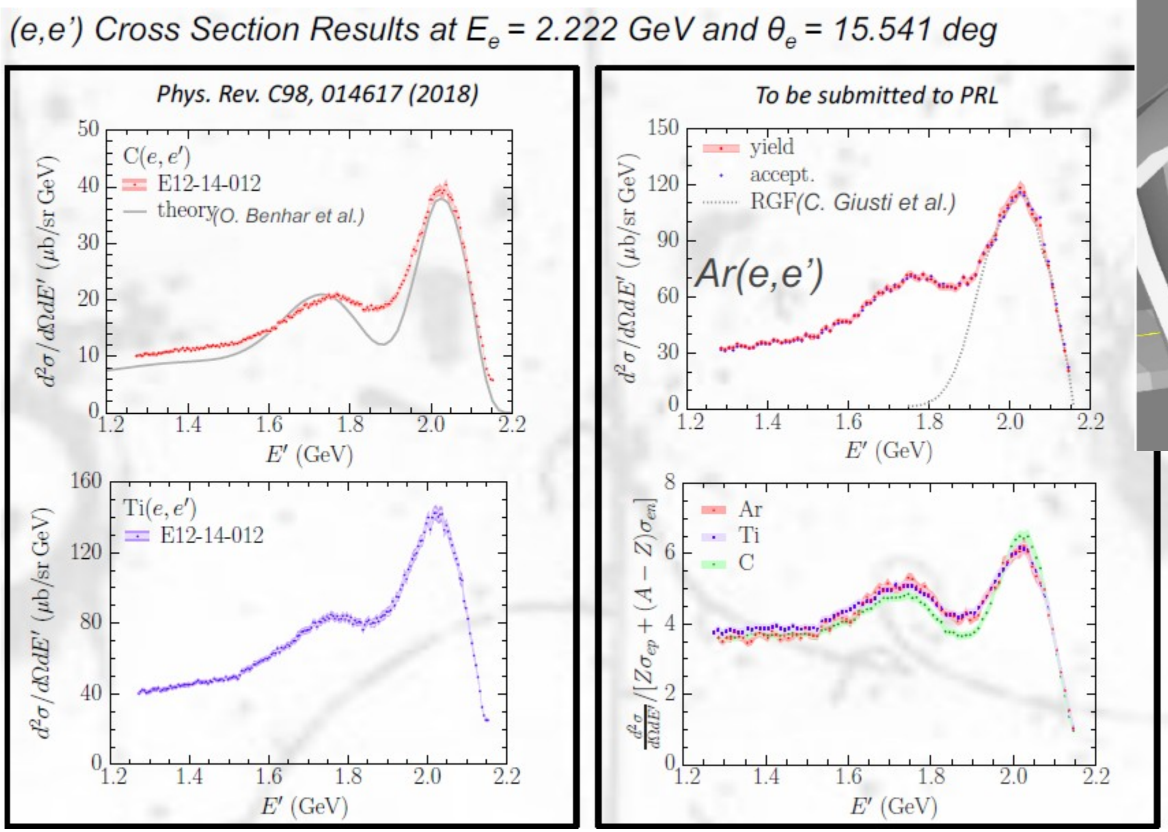


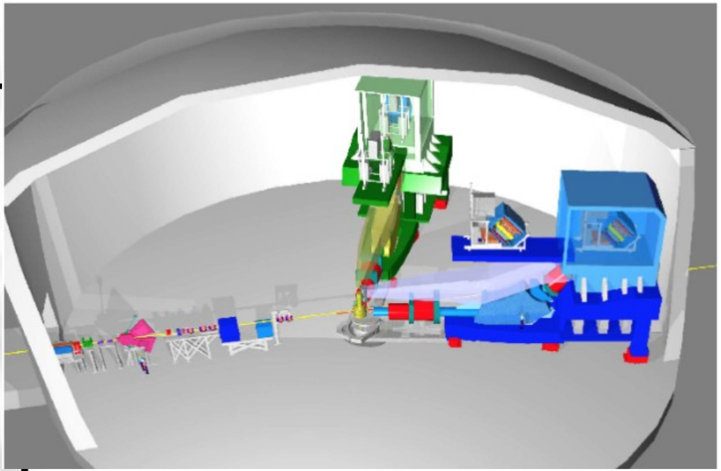


Science & Technology Facilities Council

UK Research and Innovation

E12-14-012 @ JLab





(e,e'p) coming soon

Conclusions

- Electroweak Nuclear Physics is a very active "multifaceted and interdisciplinary" field
  - but we can handle it
- Wide array of experimental endeavours ongoing
  - Model development is driven by experiments
- Will re-branding be successful?
  - Might hear the term "electroweak nuclear physics" more in the future
- Exciting times to be a neutrino/nuclear physicist
- All NuInt 18 presentations are online:

https://indico.cern.ch/event/703880/contributions/

Science & Technology Facilities Council

UK Research and Innovation

Thank you

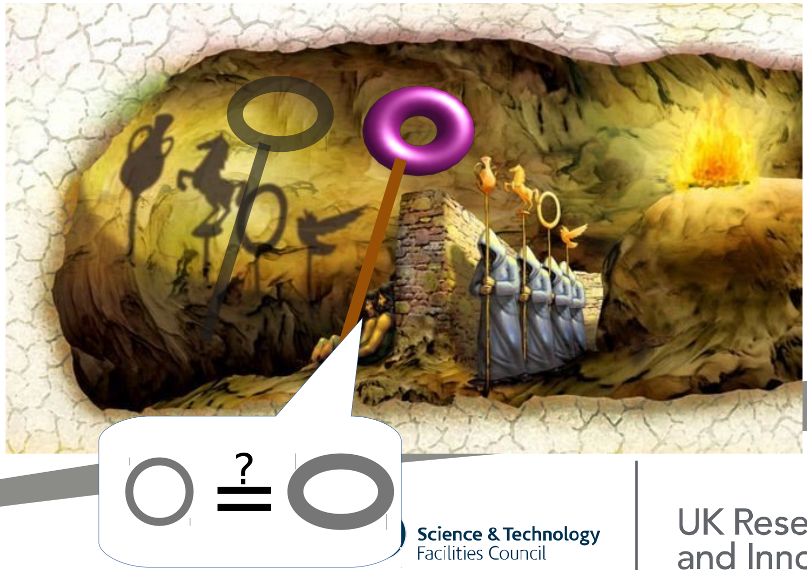
The cross-section cave

- What we see is not what we are interested in
  - Lost events due to efficiency
  - Added events due to background
  - Different event properties due to smearing

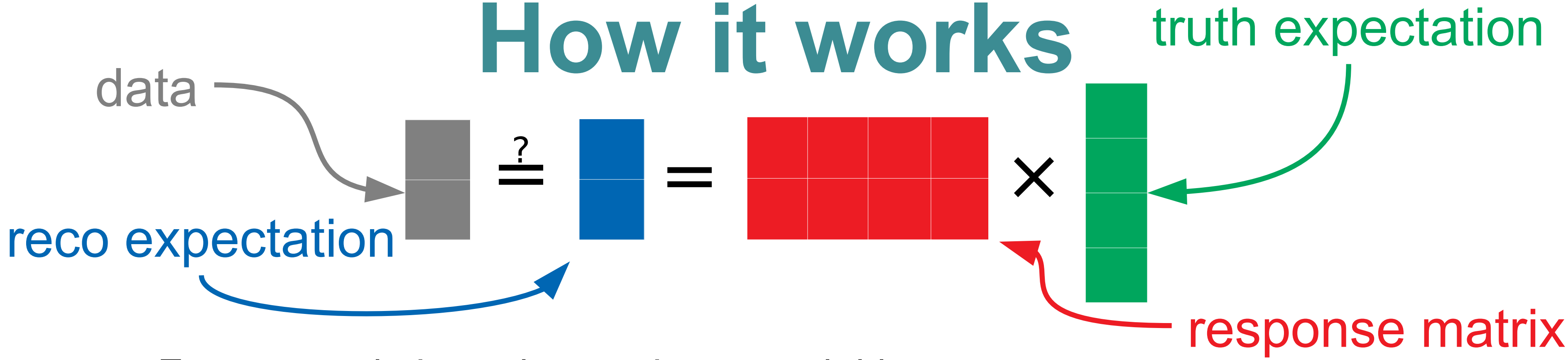
cross-section extraction

- The canonical way: Unfolding
  - “Undo” the detector and selection effects
  - Challenging to do right without introducing bias

- Another way: Forward-folding
  - Apply detector effects to theory
  - Brings its own sets of challenges



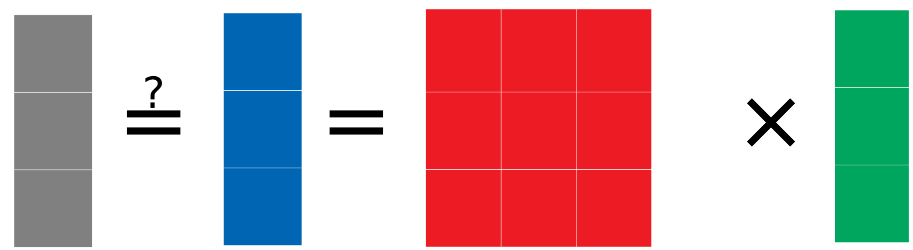
How it works

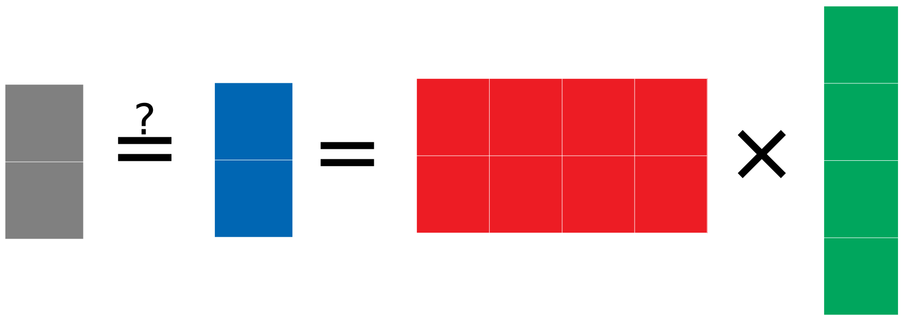


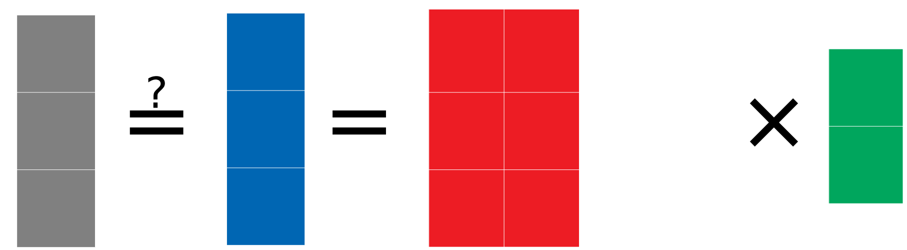
- Every event belongs in exactly one truth bin and up to one reconstructed bin (if it gets reconstructed)
- P(reco bin = i | truth bin = j) = R_{ij} = efficiency × smearing
  - Response matrix describes average detector response to true events
- reco expectation = response matrix × truth expectation
  - Can (and truth usually must) be binned in multiple variables
- The data is the data is the data
  - No uncertainty on the data points, 4 is exactly 4!
  - All systematics in response matrix or physics model
- All comparisons between data and theory (likelihoods, chi-squares, chi-by-eye) are done in reco space.

Reasons to do it: Flexible binning

- Flexible number of bins
- reco bins ≠ truth bins

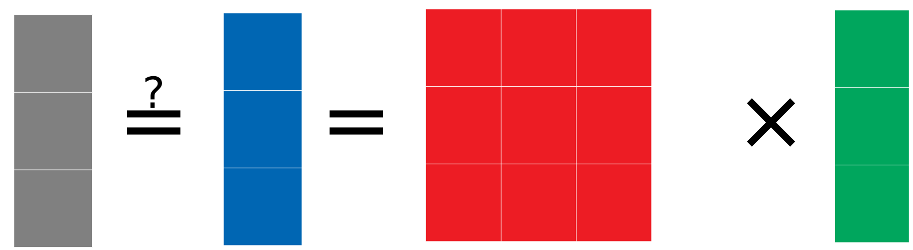


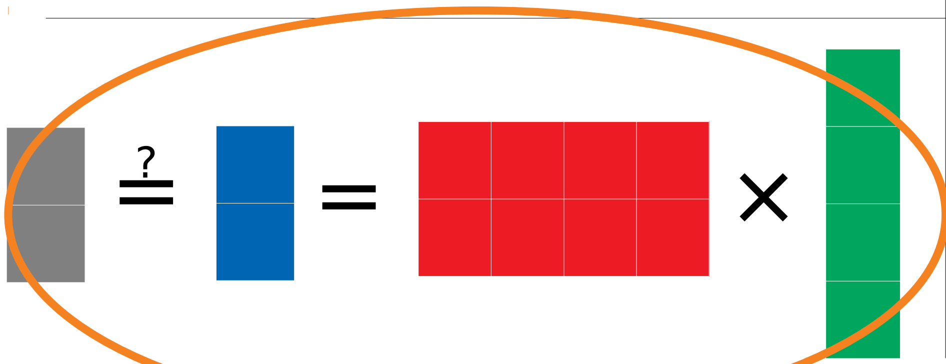


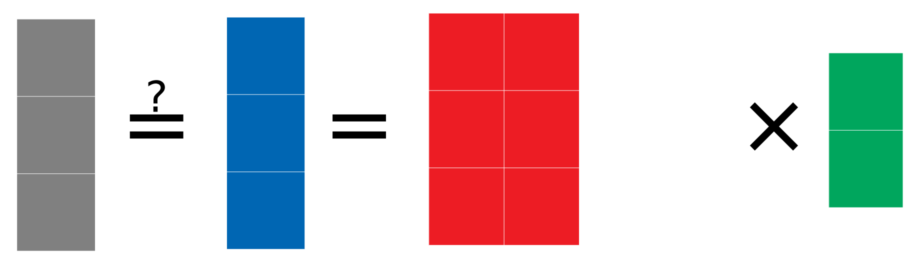


Reasons to do it: Flexible binning

- Flexible number of bins
- reco bins ≠ truth bins







- Combine coarse reco binning with fine truth binning
  - Great for analyses with low statistics
  - Admit we are not able to constrain truth completely

Science & Technology Facilities Council

UK Research and Innovation

Reasons to do it: reco level data

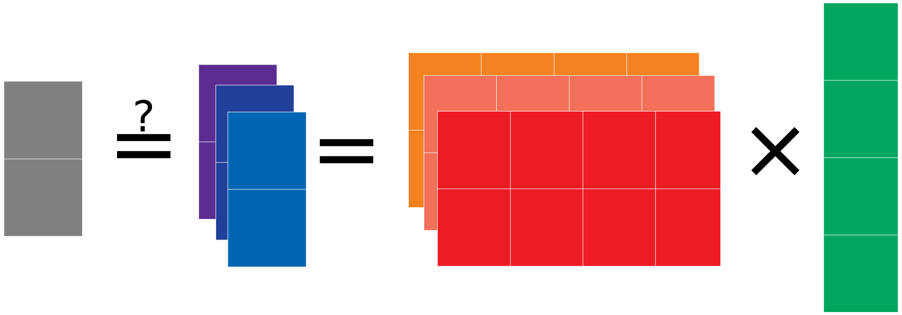
- No data point correlation
  - Theory predictions will be correlated, but probably much less than what unregularised unfolding might do
  - Chi-by-eye
- Robert D. Cousins, Samuel J. May, Yipeng Sun, [arXiv:1607.07038]
  Should unfolded histograms be used to test hypotheses?:

  "It seems remarkable that, even though unfolding by matrix inversion would appear not to lose information, in practice the way the information is used (linearizing the problem via expressing the result via a covariance matrix) already results in some failures of the bottom-line test of GOF. This is *without any regularization or approximate EM inversion*."

"D'Agostini"

Detector uncertainties

- Matrix only describes single possible detector
  - True detector probably behaves slightly differently
- Cover detector uncertainties with “toy simulations”
  - Variations and weights of same events
- Each toy yields own response matrix
- Each response matrix yields own reco prediction



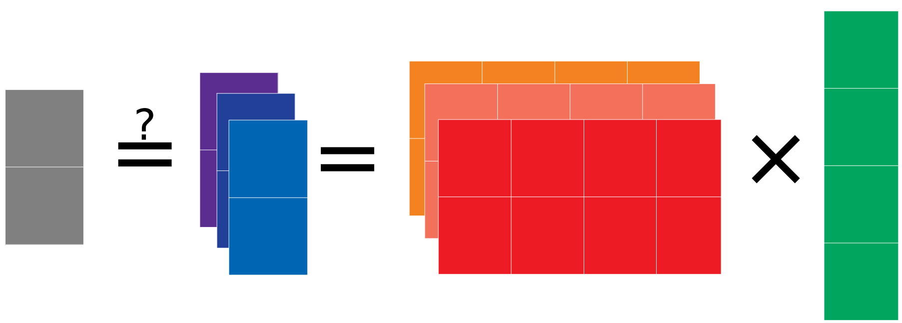
- Compare to data w/ marginal, i.e. average, likelihood

The data usage cycle

Experiment takes data, builds response matrix

Use for likelihood tests to reject models / fit parameters

Inside experiment collaboration



Publish parameter fits as well as raw data and matrix

Fits inform building of new models

Use published data and matrix for likelihood tests to reject models / fit parameters https://remu.readthedocs.io https://github.com/ast0815/remu

Response Matrix Utilities

- Implements all of this (and then some)
- Input:
  - Toy variations of selection (detector systematics)
  - Truth and reco binning
- Provides methods to:
  - Build matrix
  - Evaluate statistical detector uncertainty
  - Forward-fold truth (i.e. model)
  - Compare to data (e.g. compute likelihoods, p-values, MCMC)
- Pure python (+ standard scientific packages numpy, etc)
  - Easy to install and use
  - $ pip install remu
- Tell me what you expect/want/need!

The hard part for the analyser

- Make the response matrix model-independent!
- What:
  - Matrix elements depend *only* on detector properties
- Why:
  - A model-dependent matrix defeats the purpose of being able to test arbitrary models with it
- How:
  - Understand your detector and analysis
  - Choose an appropriate truth binning (variables to bin in, granularity of binning)
  - See backup slides

To conclude

- Forward folding is every bit as challenging as unfolding
  - Need to really understand the detector to decide which variables to bin in and how
  - High MC statistics requirements
- ReMU implements the necessary machinery
- Method promises some advantages over unfolding
  - Works with low real data statistics
  - Best model separation power in reco space [arXiv:1607.07038]
- Method paper in preparation

A few things to think about

- How to best handle backgrounds
  - Backgrounds are just another set of truth bins
  - Can be handled organically (simultaneous fits)
- Provide background templates
- Provide experiment/analysis specific convenience functions
- Plot release → data release → algorithm release
  - $ pip install t2k-results

    >>> t2k_results.thisorthat_xsec_result.fit(my_model)
  - No more manual overlaying plots copied from papers

Science & Technology Facilities Council

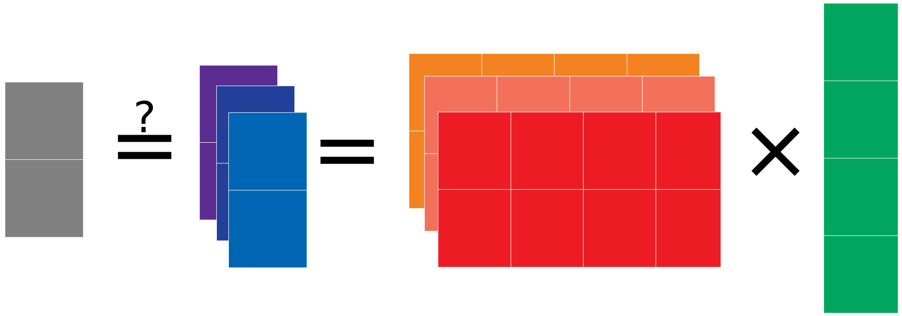
UK Research and Innovation

Thank you!

https://remu.readthedocs.io https://github.com/ast0815/remu

Statistical uncertainty

- Generating MC costs time and money
  - In theory we could have arbitrarily precise matrices
  - In practice we don't
- Quantify statistical uncertainty of matrix elements
- Generate random matrices according to stats
- Handle just like systematic uncertainties
  - In a way the statistical uncertainty is just another detector systematic



Three step matrix building model

- R_{ij} = $eff_j \times smear_{ij} \times weight_{ij}$
- Efficiency
  - Binomial process
  - Parameters ~ conjugate distribution: Beta
- Smearing
  - Multinomial process
  - Parameter ~ conjugate distribution: Dirichlet
- Weighting
  - What matters are the average weights
  - Use standard error of the mean: Normal

What to bin in

- Ideal:
  - Bin in *all* truth variables that affect reconstruction
- This goes beyond the variables of physical interest, i.e reco variables!
  - Measuring muon momentum distribution, but true cos(theta) affects efficiency? You *must* bin in true cos(theta)!
  - Might lead down some weird rabbit holes (angular separation of tracks, total particle multiplicity, ...)
- Realistic:
  - Bin in *most important* variables that affect reconstruction

Science & Technology Facilities Council

UK Research and Innovation

What *not* to bin in

- Never ever use truth variables that need a “physics” model to propagate to the reco level!
- Neutrino energy? Bad choice.
  - Measurable effect depends on interaction model, nuclear model, FSI…
- Muon momentum? Good choice.
  - Directly accessible by detector (track curvature)
- HMN momentum? Even better choice!
  - Do you assume the muon to be selected as HMN?
  - What about confusion with high-momentum pions?
- Rule of thumb:
  - Bin in variables as “close” to low-level reconstructed quantities
  - “Could you see it in an event display?”

The exponential #bins problem

- #bins = (#bins/variable) ^ (#variables)
- MC stats are cheap (compared to data) but not free
- Need to compromise
  - Bin coarsely (but beware in-bin variations!)
  - Concentrate on most important truth variables
  - Reduce #reco bins
  - #response matrix bins = #truth bins × #reco bins
- Aim: Reduce model-dependence to a negligible level
  - Will never remove it completely
- Constraint: Sufficient MC events in bins

The bitter truth

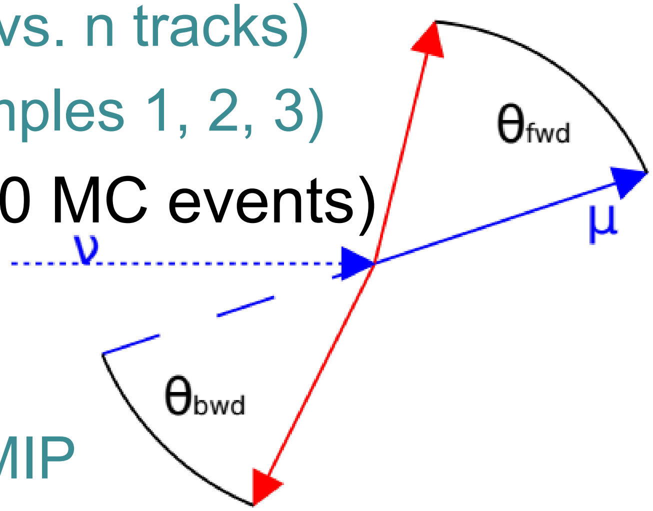
- There will be truth bins with not enough events
- Constrains the phase space of testable models
  - n_{test} < ($n_{generated}$ / safety factor) in all truth bins
- Best way to avoid this:
  - Build response matrix with MC covering the full phase space (“particle parties”)
- Realistic way to mitigate this:
  - Build response matrix with MC from multiple generators, turn dials to widest possible phase space
- Response matrix depends only on detector properties
  - Mix and merge all the models!

Gas interaction example (WIP)

- Reco binning: 16 bins
  - 2 bins in main MIP (muon or pion) angle (forwards vs. backward)
  - 2 bins in particle multiplicity (1 track vs. n tracks)
  - 4 bins in selection (main, control samples 1, 2, 3)

Gas interaction example (WIP)

- Reco binning: 16 bins
  - 2 bins in main MIP (muon or pion) angle (forwards vs. backward)
  - 2 bins in particle multiplicity (1 track vs. n tracks)
  - 4 bins in selection (main, control samples 1, 2, 3)
- Truth binning: 11760 bins (5353 w/ >0 MC events)
  - 7 bins in true MIP momentum
  - 7 bins in true MIP cos(theta)
  - 5 bins in true forward separation of MIP
  - 6 bins in true backwards separation of MIP
  - 8 bins in event category (4 in FV + 4 out of FV)



Science & Technology Facilities Council

UK Research and Innovation

The other hard part

- Getting everyone on board to use this
  - This will mean extra work for theorists/model-builders
- But it is worthwhile
  - Better model separation power
  - Works with low statistics
  - Endorsed by actual statisticians!
- This is not just dumping work on theorists
  - This is hard for experimentalists too!
  - Have to work together for better physics results
- Make this as painless as possible
  - There will be *some* pain…
  - Tell me what you want/need/expect!