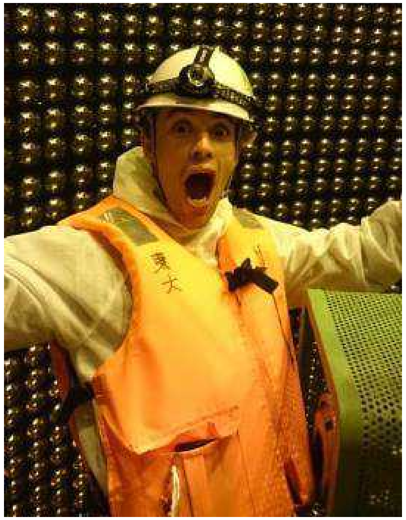
Neutrinos, data, and generator tuning



Clarence Wret

UNIVERSITY OF OXFORD

Luke Pickering

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Callum Wilkinson

BERKELEY LAB
Lawrence Berkeley National Laboratory

Stephen Dolan

CERN

clarence.wret@physics.ox.ac.uk

HSF Event Generator Tuning Workshop
28 June 2023

Who am I?

- Started my PhD with implementing neutrino interaction models, fitting them to various data (now 9 years ago...)
- Wrote a framework which compares neutrino generators and fits their predictions to data (NUISANCE)
- Tuned single pion production model to external data, and worked with theorists to develop and implement new models
- Have worked with experiments to test and perform their own generator tunes to specific data
  - T2K, HK, DUNE, NOvA, MicroBooNE, MINERvA
- Now supporting effort as much as possible, with users from neutrino experiments and phenomenology

- **Will need to massively simplify some topics here**, feel free to email me if you have questions!

clarence.wret@physics.ox.ac.uk

Neutrino oscillation 101

*Massive simplification, there are actually 3 states, matter interactions and CP violation, and this equation gets much much more complicated!

- Neutrino oscillations have E_ν dependence:

$$P_{\alpha\to\beta,\alpha\neq\beta} = \sin^2(\boxed{2\theta})\,\sin^2\left(1.27\frac{\boxed{\Delta m^2}\boxed{L}}{\boxed{E}}\frac{[\mathrm{eV}^2]\,[\mathrm{km}]}{[\mathrm{GeV}]}\right)^*$$

Nonetheless, the take-home message remains

- L is the distance the neutrinos travel, E is their energy
  - L/E is the central parameter in experiment design
- L/E determines what ranges of θ and Δm^2 your experiment is sensitive to, roughly speaking:
  - L/E ~ 1 km/MeV → measure θ_{13} and Δm^2_{13} → short baseline reactor experiment
  - L/E > 100 km/MeV → measure θ_{12} and Δm^2_{12} → long baseline reactor experiments, solar experiments
  - L/E ~ 400-500 km/GeV → measure θ_{23} and Δm^2_{23} → long baseline accelerator and atmospheric experiments
- Baseline is fixed, neutrino energy is not

Neutrino oscillation 101

- Neutrino oscillations have E dependence:

$$P_{\alpha\to\beta,\alpha\neq\beta} = \sin^2(2\theta)\sin^2\left(1.27\frac{\Delta m^2 L}{\boxed{E}}\frac{[\mathrm{eV}^2]\,[\mathrm{km}]}{[\mathrm{GeV}]}\right)$$

- L is the distance the neutrinos travel, E is their energy
  - L/E is the central parameter in experiment design
- L/E determines what ranges … iment is sensitive to, roughly speak…
  - L/E ~ 1 km/MeV → me… actor experiment
  - L/E > 100 km/MeV → measu… eline reactor experiments, solar experiments
  - **L/E ~ 400-500 km/GeV → measure θ_{23} → long baseline accelerator and atmospheric experiments**
- Baseline is fixed, neutrino energy is not

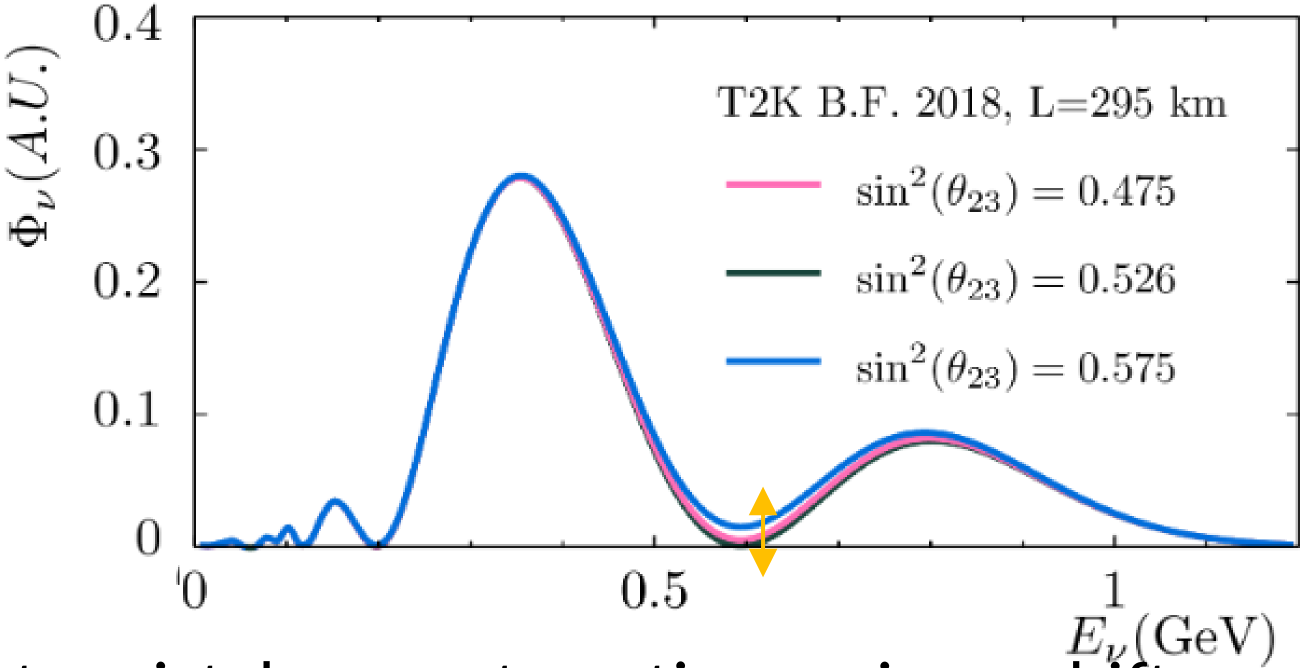
The focus of this talk, where most of the tuning and generator work happens

Neutrino oscillation 101

- Neutrino oscillations have E_ν dependence

$$P_{\alpha\to\beta,\alpha\neq\beta} = \sin^2(\boxed{2\theta})\sin^2\left(1.27\frac{\Delta m^2 L}{E}\frac{[\mathrm{eV}^2]\,[\mathrm{km}]}{[\mathrm{GeV}]}\right)$$

- Shift in amplitude biases measurement of θ, shift in frequency biases Δm^2



- Possible to mistake a systematic causing a shift as an oscillation parameter value

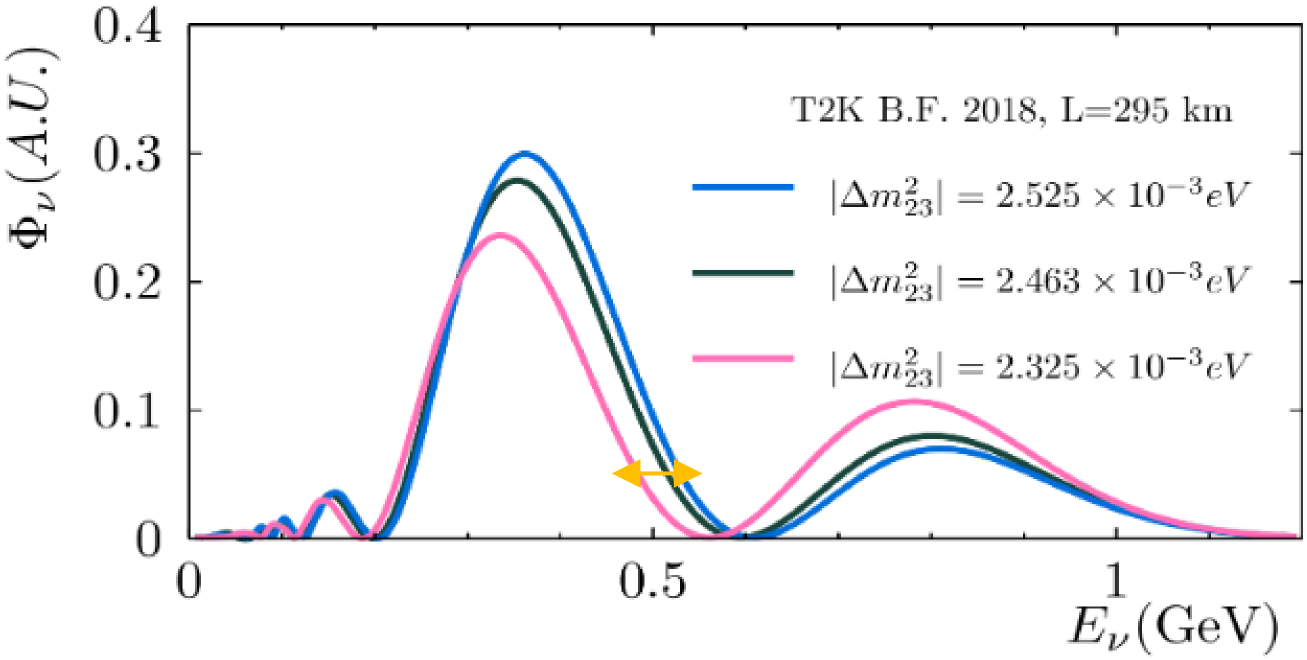
L. Pickering NuFact2019

Neutrino oscillation 101

- Neutrino oscillations have E_ν dependence

$$P_{\alpha\to\beta,\alpha\neq\beta} = \sin^2(2\theta)\sin^2\left(1.27\frac{\boxed{\Delta m^2}L}{E}\frac{[\mathrm{eV}^2]\,[\mathrm{km}]}{[\mathrm{GeV}]}\right)$$

- Shift in amplitude biases measurement of θ, shift in frequency biases Δm^2

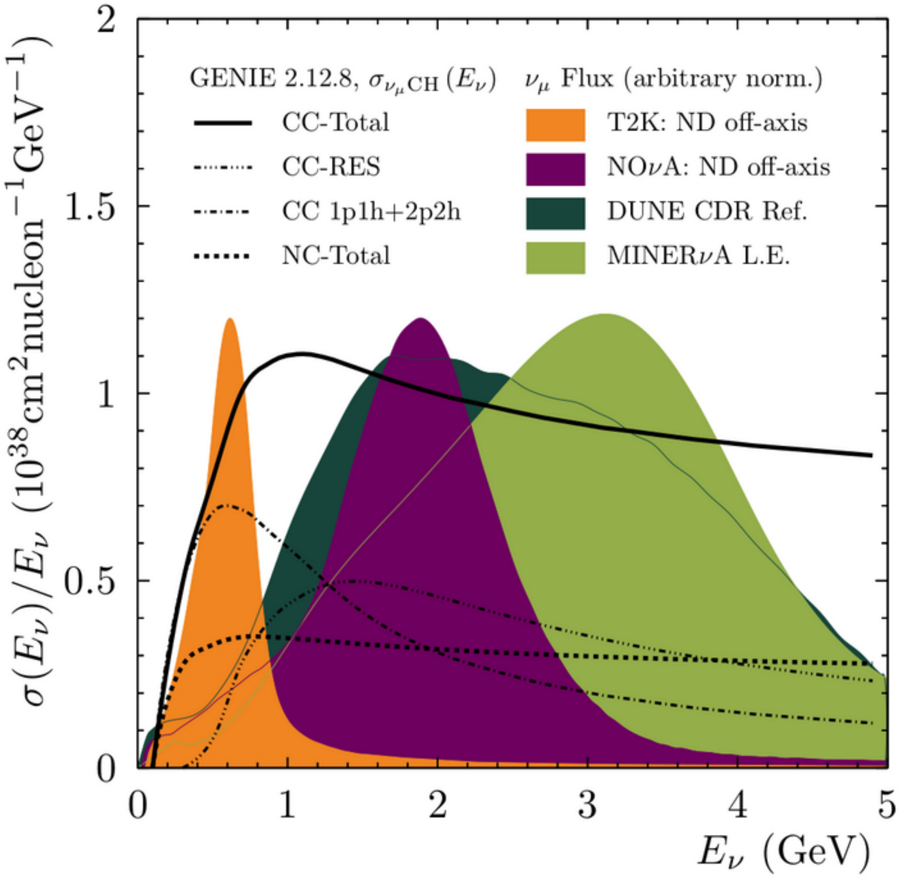


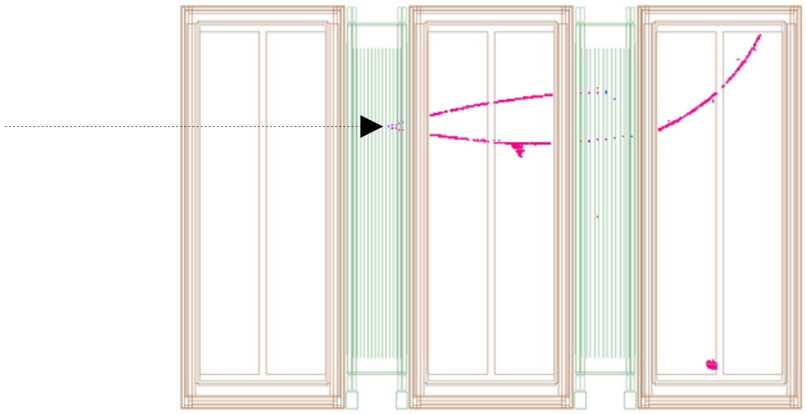
- Possible to mistake a systematic causing a shift as an oscillation parameter value

L. Pickering NuFact2019

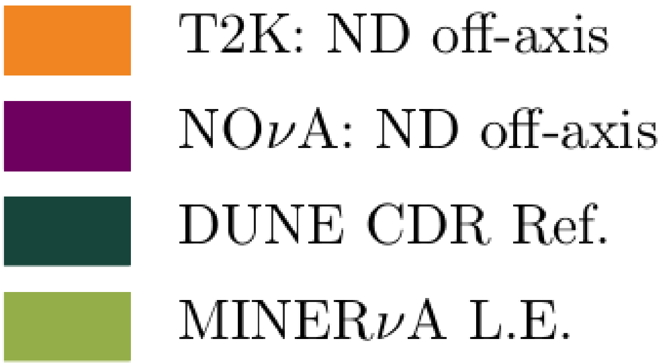
Introduction

- Neutrino energy is on similar scale to **nuclear effects**
- The neutrinos do not have a fixed energy: the neutrino energy is a distribution
  - Not precisely measured event-by-event, instead **inferred from interaction products**





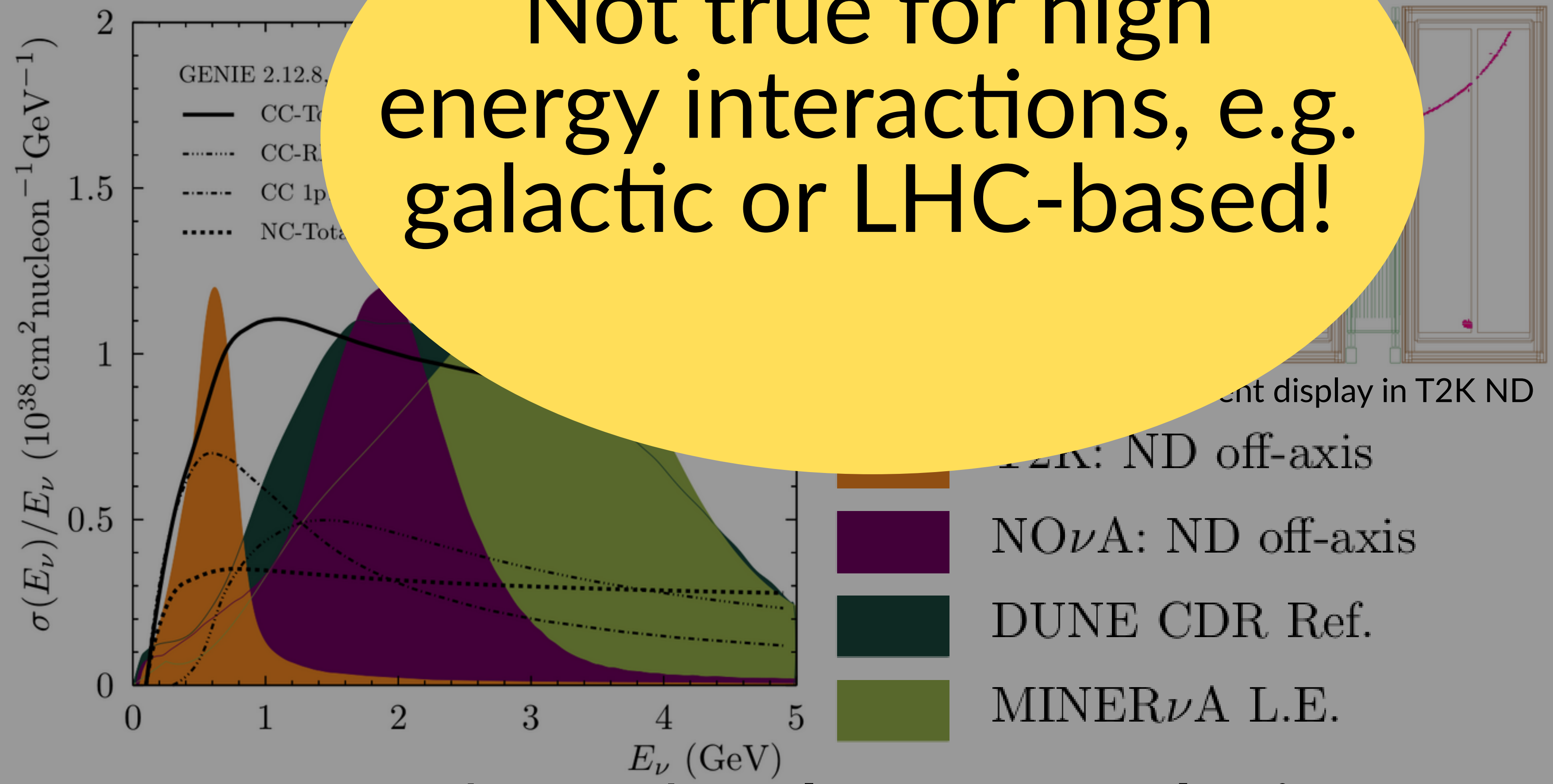
Event display in T2K ND



- Not safe from the perils of **low energy physics, energy transfers on 20-300 MeV scale**

Introduction

- Neutrino energy is on similar scale to **nuclear effects**
- The neutrinos do not have a fixed energy: the neutrino energy is a distribution
  - Not precisely m d **inferred from interaction r**



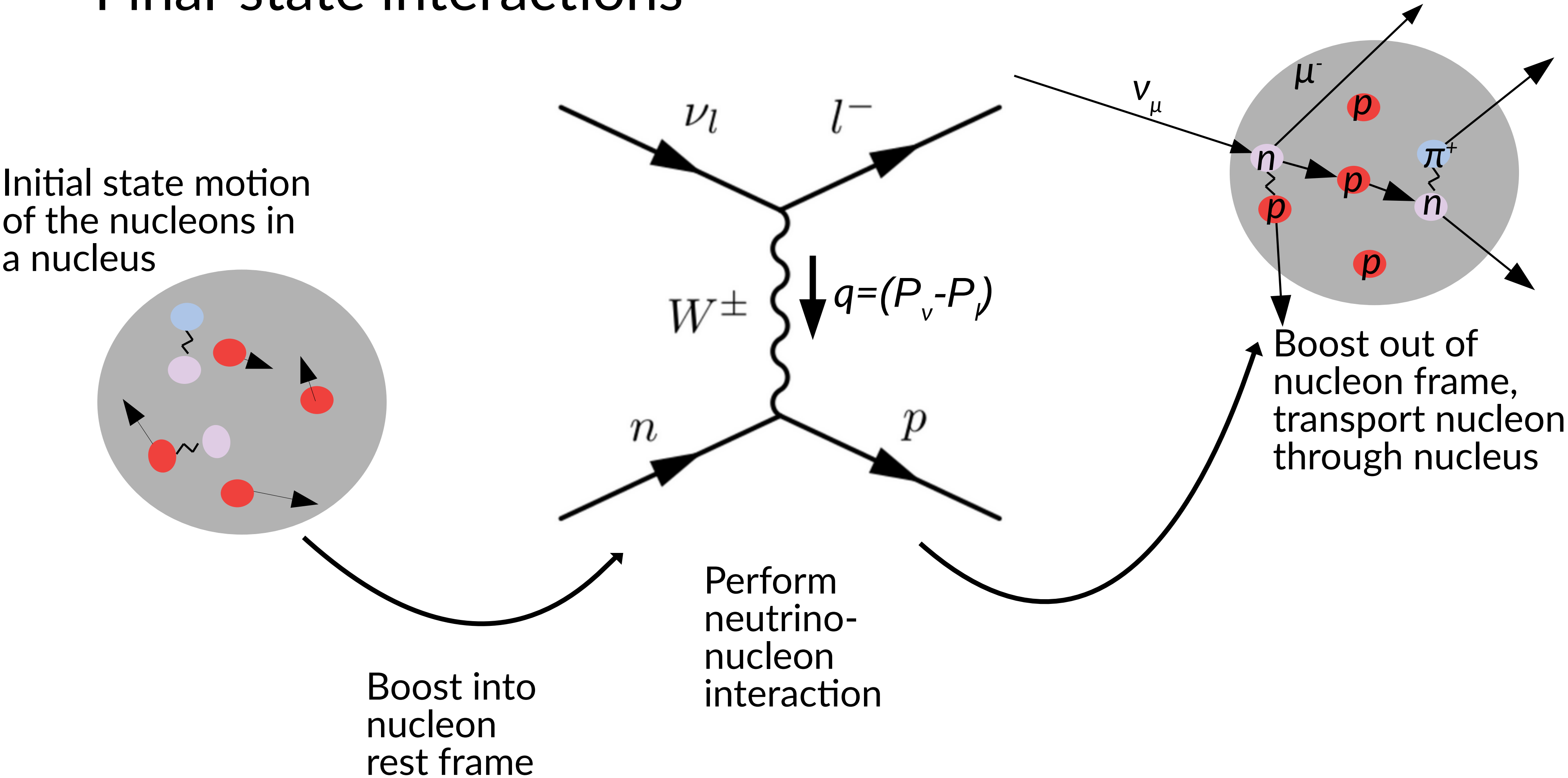
nt display in T2K ND

Not true for high energy interactions, e.g. galactic or LHC-based!

- Not safe from the perils of **low energy physics**

Theory landscape

- Historically factorised problem into three stages:
  - Initial state motion
  - Nucleon-level interaction (hard scatter)
  - Final-state interactions



Initial state motion of the nucleons in a nucleus

Boost into nucleon rest frame

Perform neutrino-nucleon interaction

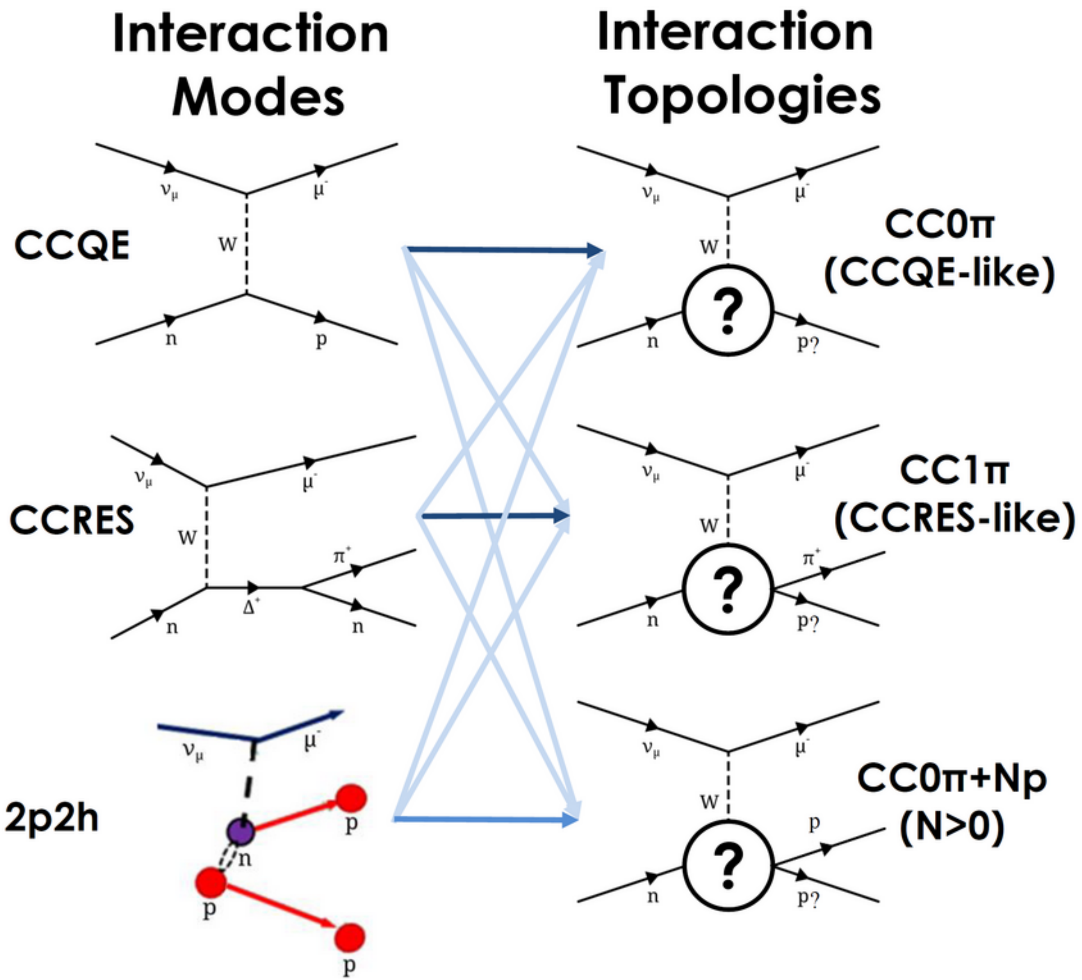
Boost out of nucleon frame, transport nucleon through nucleus

Theory landscape

- Active theory community, especially in the recent 10 years
  - Moving away from factorisation approach
- Theory groups often work outside of generator community, and generator groups contact theorists
  - This is significantly improving with more direct contact and dedicated workshops
  - More theory groups are becoming aware of the problem regarding neutrino interactions
  - Securing long-term funding has been troublesome, sitting between experimental physics and nuclear physics, outside LHC
    - This is improving **significantly** with the advent of the high-statistics experiments DUNE and HK
- Lots of recent effort on nuclear effects and the simplest CC interaction without any pions, and outgoing nucleon(s)
- Some efforts on nucleon/quark level too, e.g. non-resonant backgrounds, DIS transition
- Some work on integrating more sophisticated nucleon-nuclear transport models, e.g. INCL++

Theory landscape

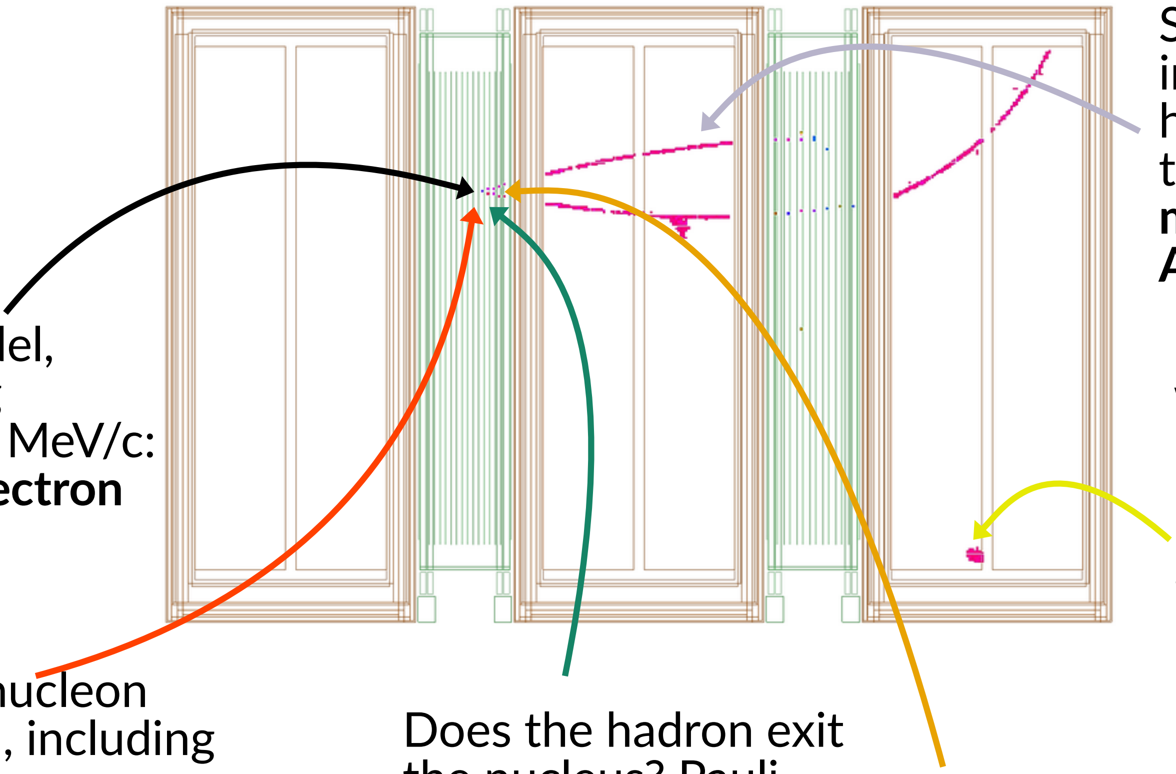
- The observable topology in a detector comes from a vast array of physics: challenging theory
  - It is often not enough to “simply” write a model for a specific Feynman diagram or process, because the observable detected final state has contributions from many different such diagrams



Theory landscape

- The observable topology in a detector comes from a vast array of physics: challenging theory

We see a single muon, single pion event in our detector, what is the physics, where is the model from?



Secondary interactions of hadrons inside the detector: **models from π-A scattering**

Initial state model, nucleon moving between 0-250 MeV/c: **models from electron scattering**

Was this a neutron? **Model from nucleon interaction and hadron propagation**

Neutrino-nucleon interaction, including multiple resonances and non-resonant (DIS?) backgrounds: **models from neutrino-nucleon data (bubble chambers) and electron scattering**

Does the hadron exit the nucleus? Pauli blocked? Removal energy? **Models from electrons and neutrinos**

Hadronic final state interactions inside the nucleus: **models from π-A scattering**

Generator landscape

- The generator market is vast, and expanding
  - GENIE, NEUT, NuWro, GiBUU, Achilles, NUANCE, …
  - **No clear winner for experiments**: some generators have excellent integration into experiments, others have very detailed nuclear model implementations but less developed uncertainty model, and so on
- Have tools to compare these, and experiments will often devise systematics based on generator or theory differences
  - This isn't ideal, but a stop-gap solution until a clearer picture emerges
- Implementing models into generators takes significant time
  - Different generators have different approaches here: some working directly with theorists, others ask theorists to implement their models
  - Working towards a more general framework which can be shared across generators
- Computational aspects are becoming a problem
  - Complex precision nuclear physics is not currently feasible
  - Effort needs to be spent on improving numerical aspects, or effective approaches

Data landscape

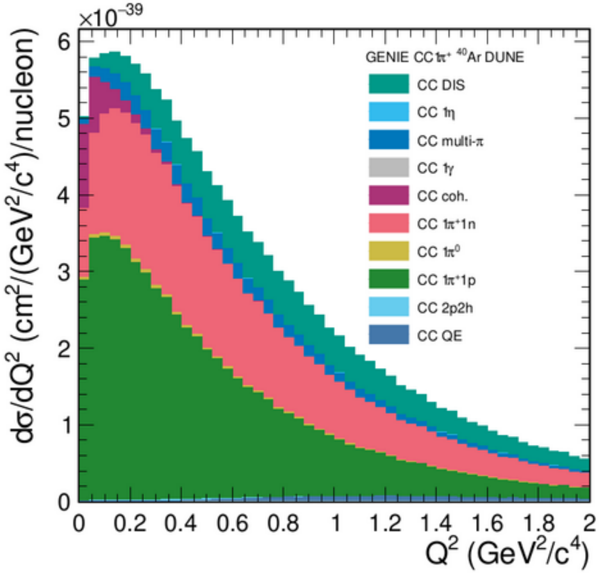
- Neutrino-nucleon data
  - Bubble chamber, 1950s-1980s: no covariance matrices, low statistics, not always clear if data is corrected for detector effects or not; **difficult!**
  - Recently some efforts in accessing nucleon physics with CH target; not yet clear what impact of model dependence is, or value of result (need more stats?)
- Nuclear target data
  - Initially poor quality: unreliable model-dependent background subtraction, missing covariance matrices, unclear and model-dependent signal definitions, unfolding issues, model-dependent corrections
  - Has improved dramatically in the last 10 years; some forward-folding, moving towards only reporting what is seen (less model-dependent correction), awareness of problems related to unfolding, efforts towards data preservation
- Integrity of data is evaluated on case-by-case basis
  - Some measurements are not suitable for generator tuning, e.g. model dependent cuts directly impacting physics conclusions
  - Some measurements have missing or corrupt covariance matrices
- Integrity of tuning
  - You might disagree with choices made by your generator's tuning effort; what's the solution?
- Experiment-specific, sometimes measurement-specific, tuning commonplace

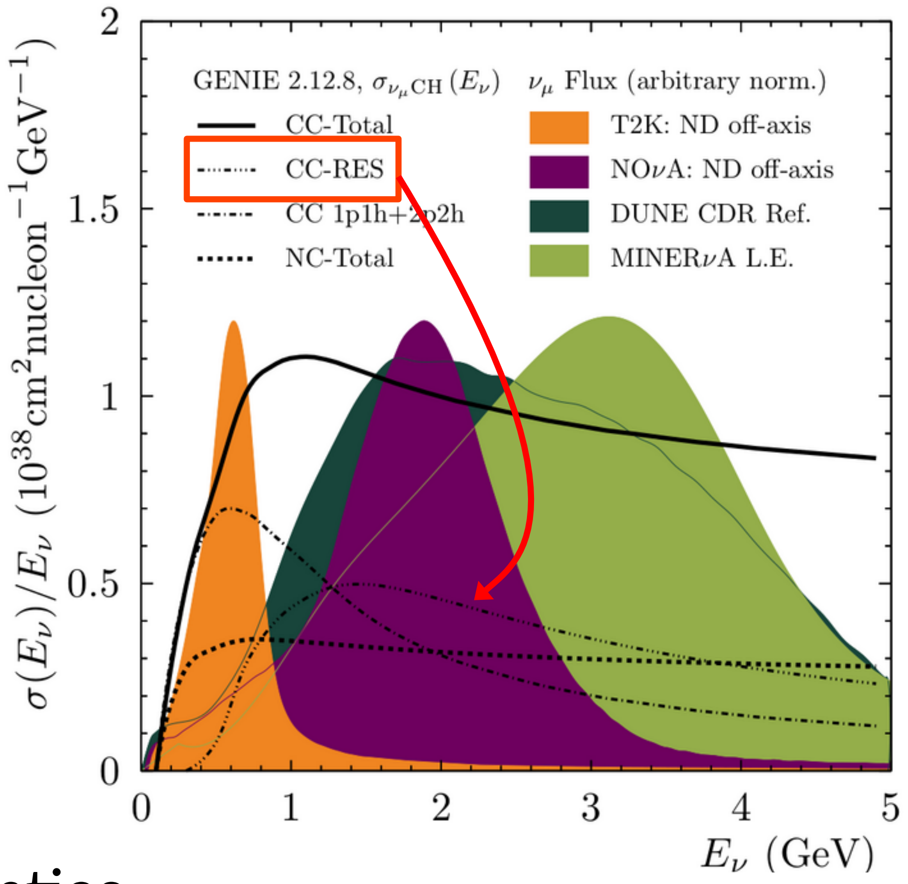
Data landscape

- Charged lepton scattering data
  - Historically used by theorists to constrain vector components of interaction
  - e4nu group is actively working with GENIE and CLAS data for e.g. Ar40 nuclei (DUNE target), amongst others
  - GiBUU historically emphasised multiple probes: significant important work
  - NuWro has electron scattering routine, explored somewhat
  - NEUT work in progress, explored somewhat
- Hadron scattering data
  - Generally constrains the pion and nucleon FSI
  - However, a particle colliding with a target is not necessarily equivalent to a particle moving out of a nucleus
  - All generators have used this data to varying extent
- Photon scattering data
  - Primarily used by theorists and GiBUU, little work done by other generators (at least to my knowledge)
- Some effort towards unifying data releases on HEPdata, but not commonplace at the moment

Example of tuning, CC1π⁺

- Say you want to tune the resonant CC1π⁺ model on DUNE; important contributor to NOvA and DUNE's oscillation analysis

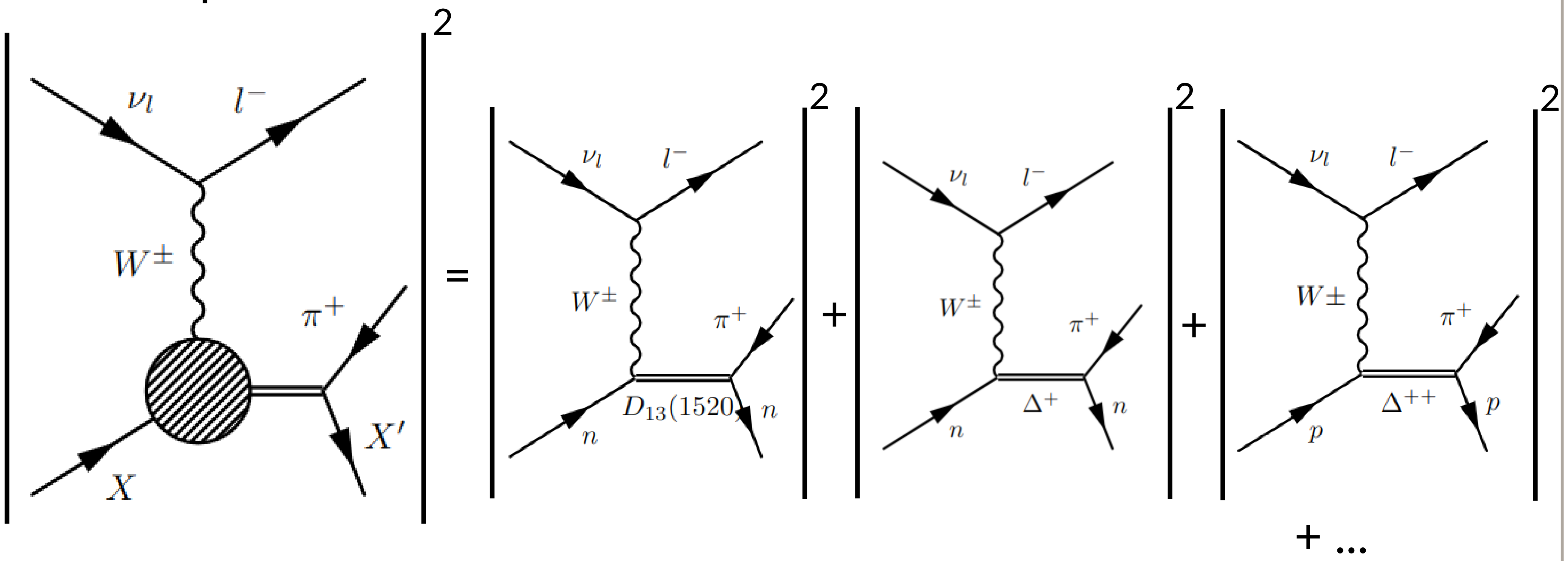




- Complex nuclear target (^{40}Ar)
  - Choose initial state model
  - Tune nucleon model
  - Add in nuclear effects and uncertainties
  - Add in non-resonant contributions and uncertainties
  - Tune to relevant data on a nuclear target
  - Likely inflate uncertainties

Example of tuning, CC1π^+

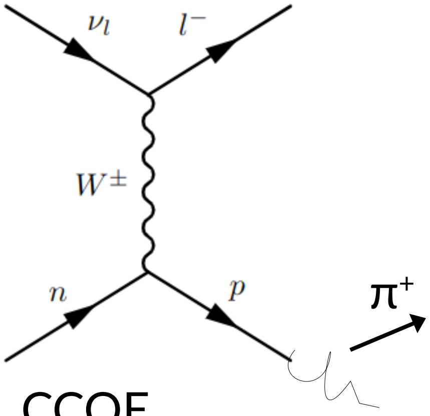
- Tune nucleon-neutrino model to old bubble chamber experiments for a specific interaction



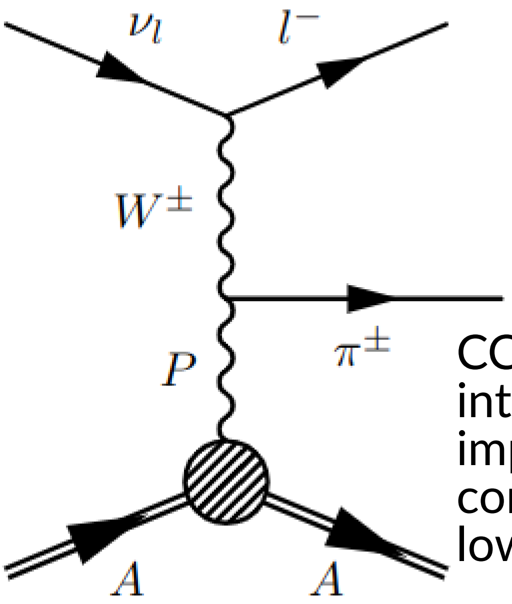
- Select bubble chamber data in relevant neutrino energy range
  - If you’re lucky, the bubble chamber data has a hadronic mass cut, attempting to isolate the resonance you’re interested in
  - For T2K, this is the Δ(1232); for DUNE, many resonances play an important part and there are no experiments in the corresponding energy range: out of luck!

Example of tuning, CC1π^+

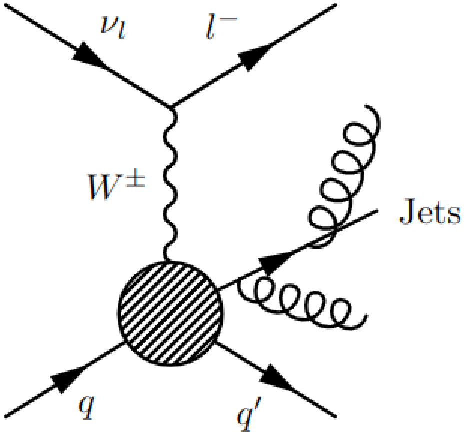
- Use output from nucleon level fit as input to nuclear level fit
  - Significantly inflate and/or invent reasonable systematics based on nuclear physics or empirical observations
- There is Pauli blocking for low momentum nucleons, removal energy, pion final state interactions, and other effects (some not even known!)
- New diagrams contribute, since the π^+ can come from many places



CCQE interaction with nucleon FSI producing charged pion



CC coherent interaction, important contributor at low Q^2



DIS interactions where a multi-pi interactions has a pion absorbed in the nucleus

<u>Need a lot more uncertainty when tuning to nuclear data, possibly constrained by other fits</u>

Example of tuning, CC1π^+

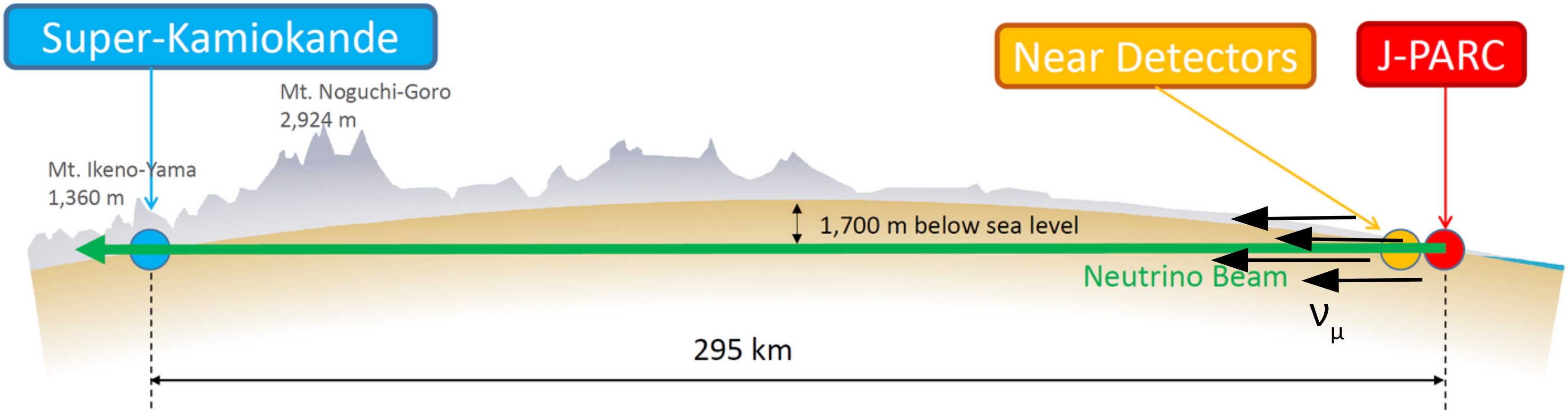
- Select neutrino-nucleus data relevant for the experiment's neutrino energy range
  - For example, MINERvA data at E_ν~3.5 GeV might not be suitable to T2K or MicroBooNE (E_ν~0.5-1 GeV)
  - Measurements from T2K (with a CH or H_2O target) might not be suitable for MicroBooNE or DUNE (Ar^{40} target)
    - The nuclear physics may have fundamentally different implications!
- These tunes are almost always never complete; require parameter uncertainty inflation so that result can reasonably cover a range of data
  - Often the muon kinematics are better described than the pion kinematics: the hadronic part is trickier
- If you're an oscillation experiment with a near detector, you almost certainly also tune your model using your own near-detector data
  - Tune to other experiments to set reasonable input priors, and check for tensions in parameters

Summary

- Progress on neutrino interactions has accelerated significantly in the last ~10 years
- A wide array of neutrino interaction data is amounting on nuclear targets, although the emerging picture is not clear (yet?)
  - Old data largely considered obsolete now, favouring better techniques and analysis methods; model-dependence in data still critical to assess however
- Nucleon target data is old and unreliable: requires significant scrutiny before using!
  - Programme at FNAL is looking into a modern bubble chamber experiment, but future is uncertain for now
- Theory is moving away from impulse approximation and factorisation, focussing on nuclear effects
- Generator programme increasingly vast: experiment-specific generators and general-purpose generators are available, with their own tunes
- Tuning in neutrino interaction physics uses many sources of data: neutrino scattering, electron scattering, hadron scattering, and more
  - Maturing programme of tuning for all generators
- Tuning to nuclear data is not straightforward: many theoretical contributions leads to many free parameters, often leading to an effective model and experiment-specific tuning
  - Subjectivity in data choice, and knowledge of modelling is critical
- Still have many lessons to learn from LHC community!

Why use external data over ND?

- Often use near detector to constrain systematics “before oscillations”



- Rate at both detectors have common ingredients

$$R(\vec{\mathbf{x}}) = \underbrace{\overbrace{\Phi(E_\nu) \times \boxed{\sigma(E_\nu, \vec{\mathbf{x}})} \times \epsilon(\vec{\mathbf{x}})}^{\text{Near}} \times P(\nu_A \to \nu_B)}_{\text{Far}}$$

- Your ND isn’t perfect → Use external data!