How we use (and don’t use) neutrino cross sections

Callum Wilkinson

BERKELEY LAB

Who am I?

T2K cross-section working group convener

Developer of the NUISANCE open-source package for data-MC comparisons and tuning

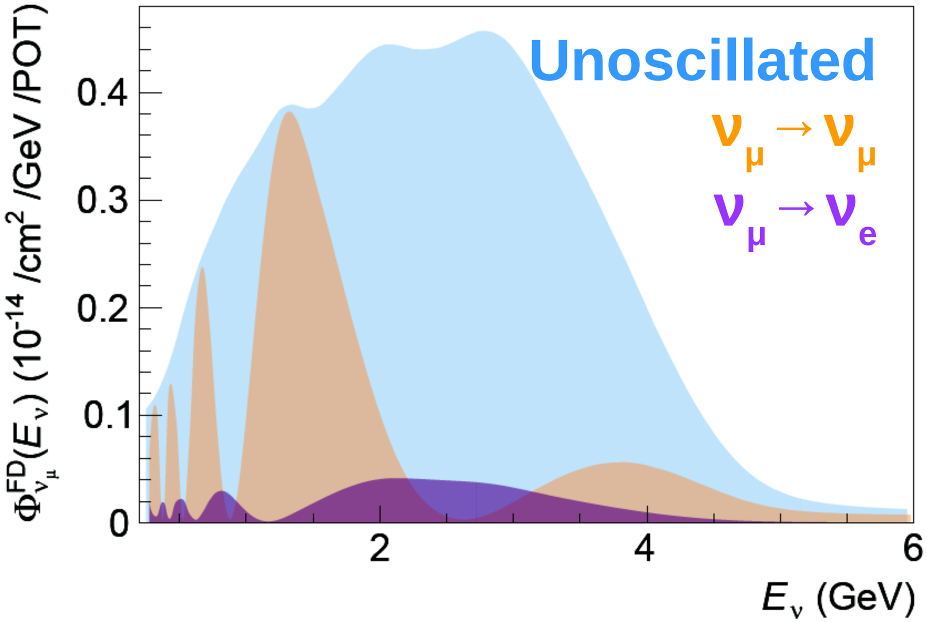
DUNE long-baseline working group convener

Importance for oscillation analyses

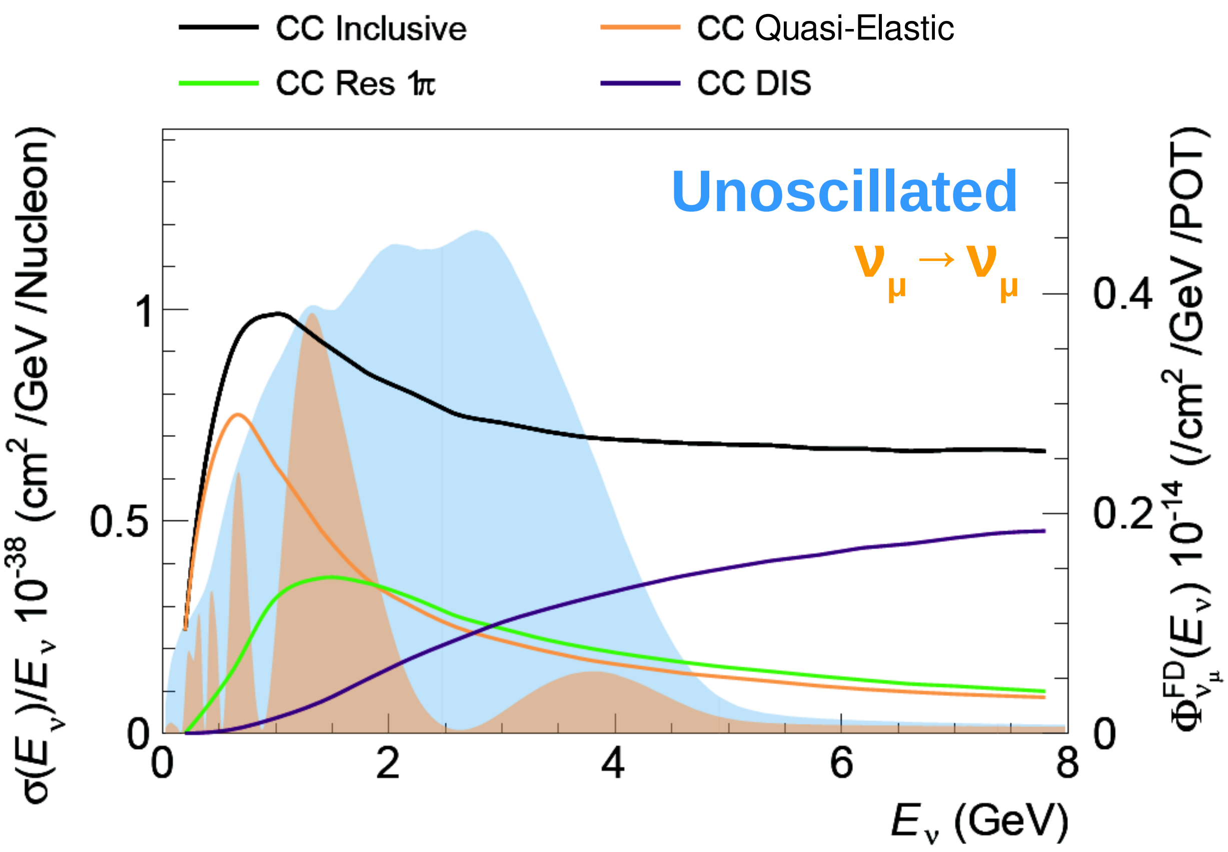
Event rate
Neutrino flux
Cross section
Detector smearing
Oscillation probability

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

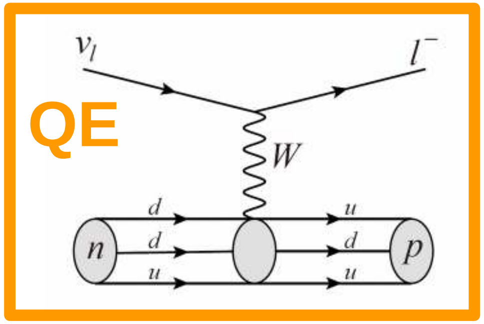
- Dramatic E_ν or flavor change
- Near/far ratios don’t fully cancel systematics
- **Critical dependence on the cross section**



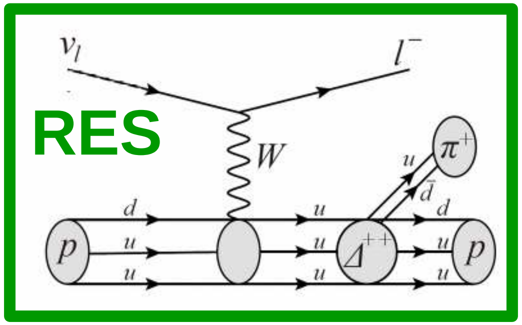
Why are cross sections such a challenge?



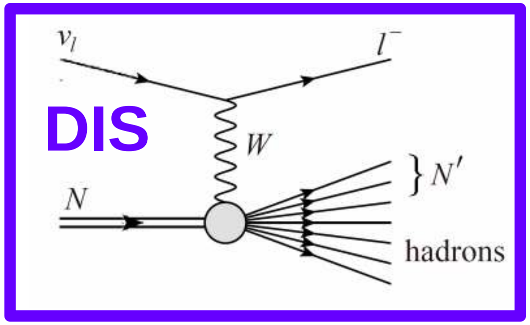
Region of interest spans a number of distinct interaction modes, with different physical descriptions!



QE



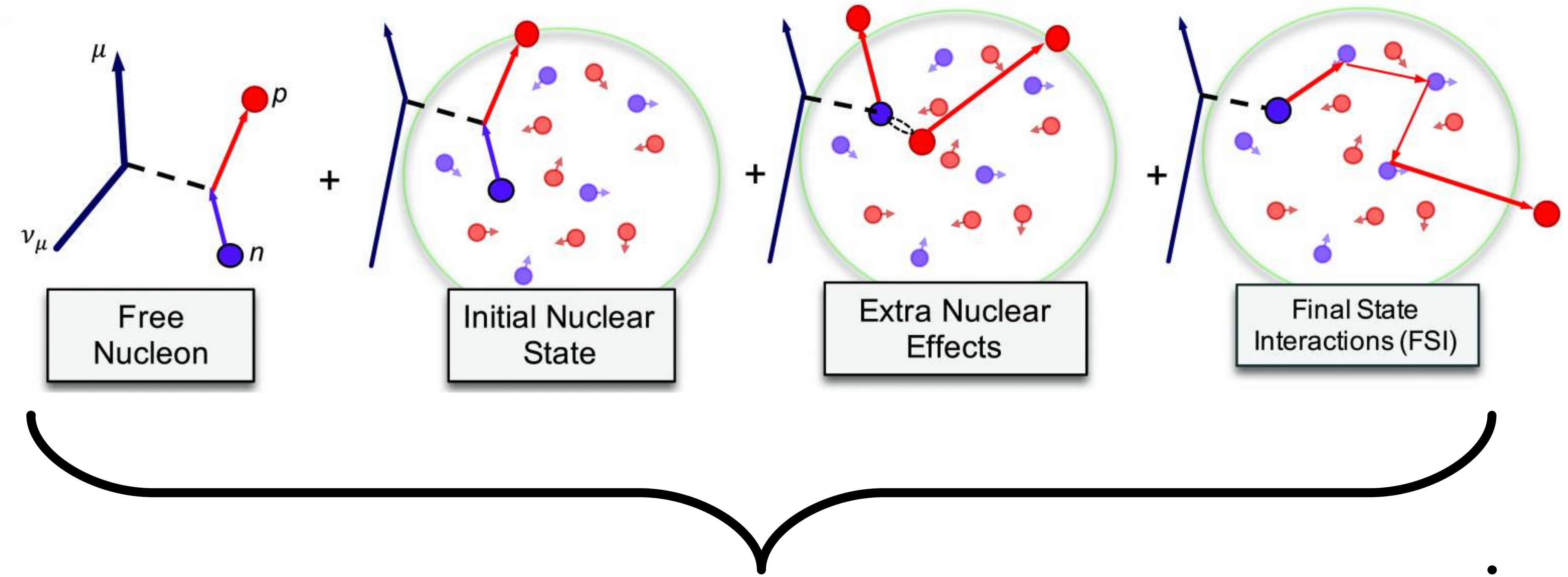
RES



DIS

Energy transfer

The plot thickens



$$R(\vec{\mathbf{x}}) = \int dE \;\; \Phi(E_\nu) \times \sigma(E_\nu, \vec{\mathbf{x}}) \times \epsilon(\vec{\mathbf{x}}) \times P(E_\nu; \nu_A \to \nu_B)$$

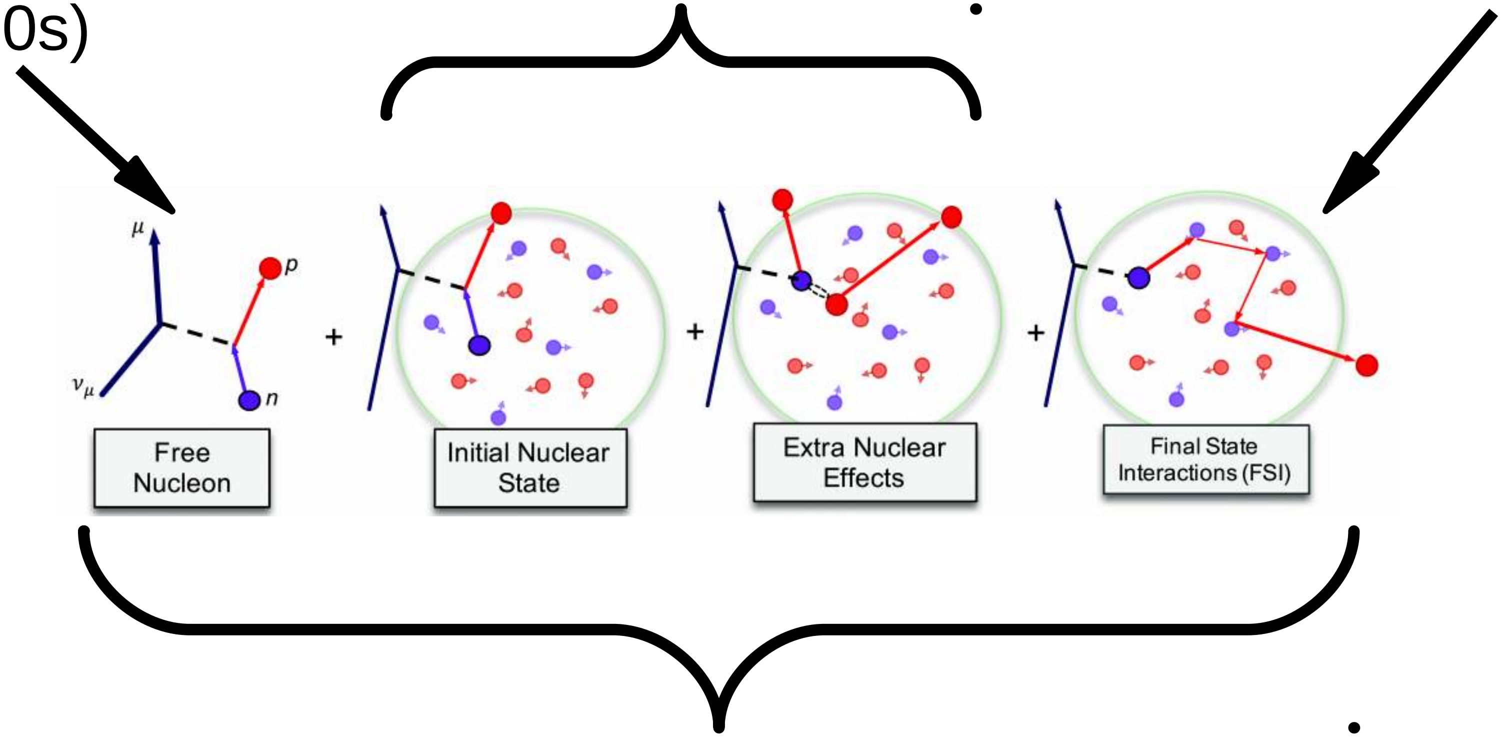
Many confounding nuclear effects to model

What relevant information is there?

ν-H_2/D_2 bubble chamber (1970s-1980s)

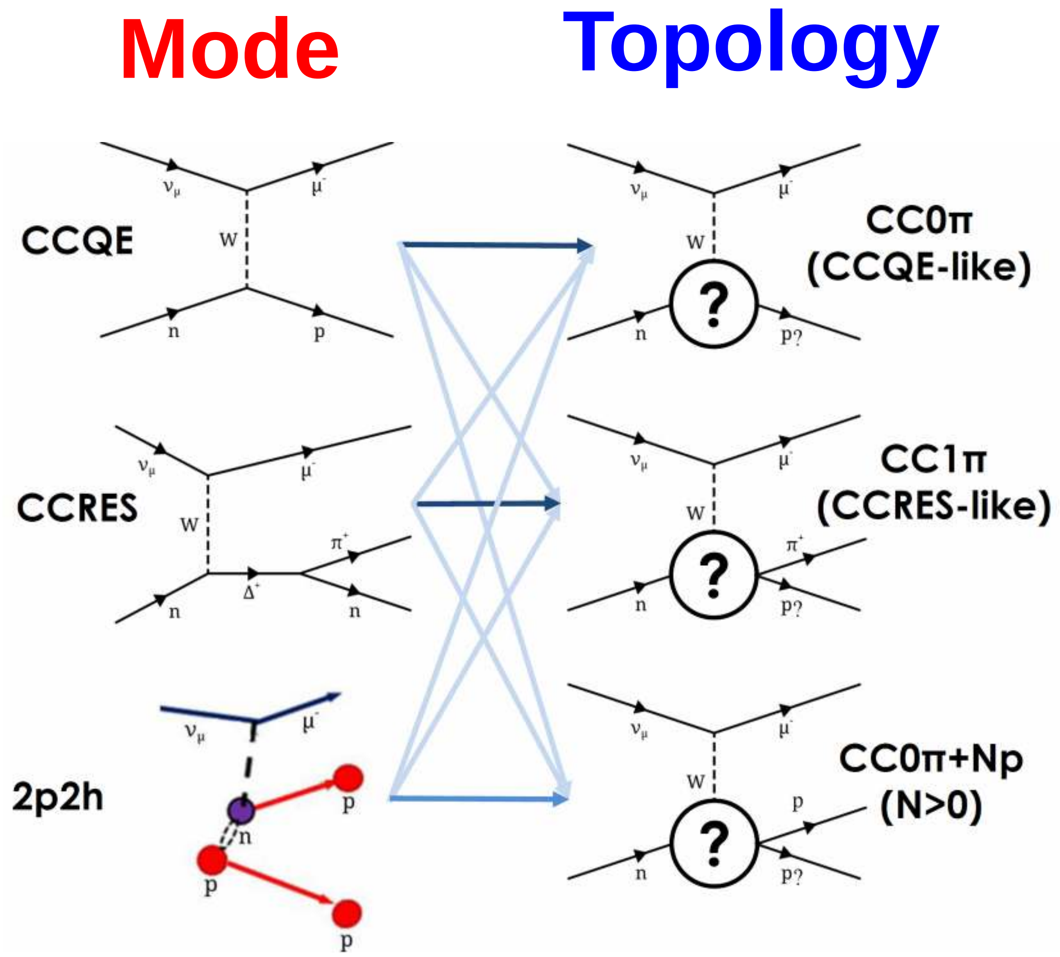
e-A scattering (1970s-present)

π-A, γ-A



Neutrino-nucleus experiments

Neutrino-nucleus scattering measurements



Many modes contribute to any measurement

Complicated FSI effects

Integrated over broad E_ν/ω region

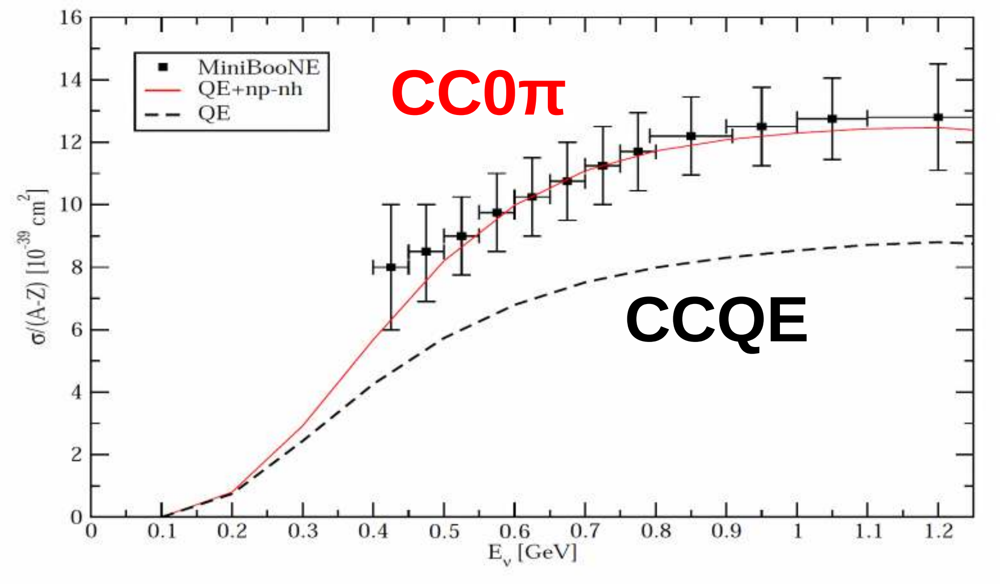
$$\tilde{\sigma}_k(\vec{\mathbf{y}}) = \sum_i \int_{E_{\min}}^{E_{\max}} \sigma_i(E_\nu, \vec{\mathbf{x}}) \times \mathrm{FSI}(\vec{\mathbf{x}}, \vec{\mathbf{y}}) dE_\nu$$

Tuning models to data

- Tuning σ_i parameters requires multiple post-FSI datasets to break degeneracies!
  - Multiple fluxes
  - Different selections
  - Different acceptance
  - Detector technologies
  - Multiple targets

- This necessity has motivated a lot of work measuring neutrino cross sections in a lot of different ways – **a vital first step!**

CC0π data model discrepancies

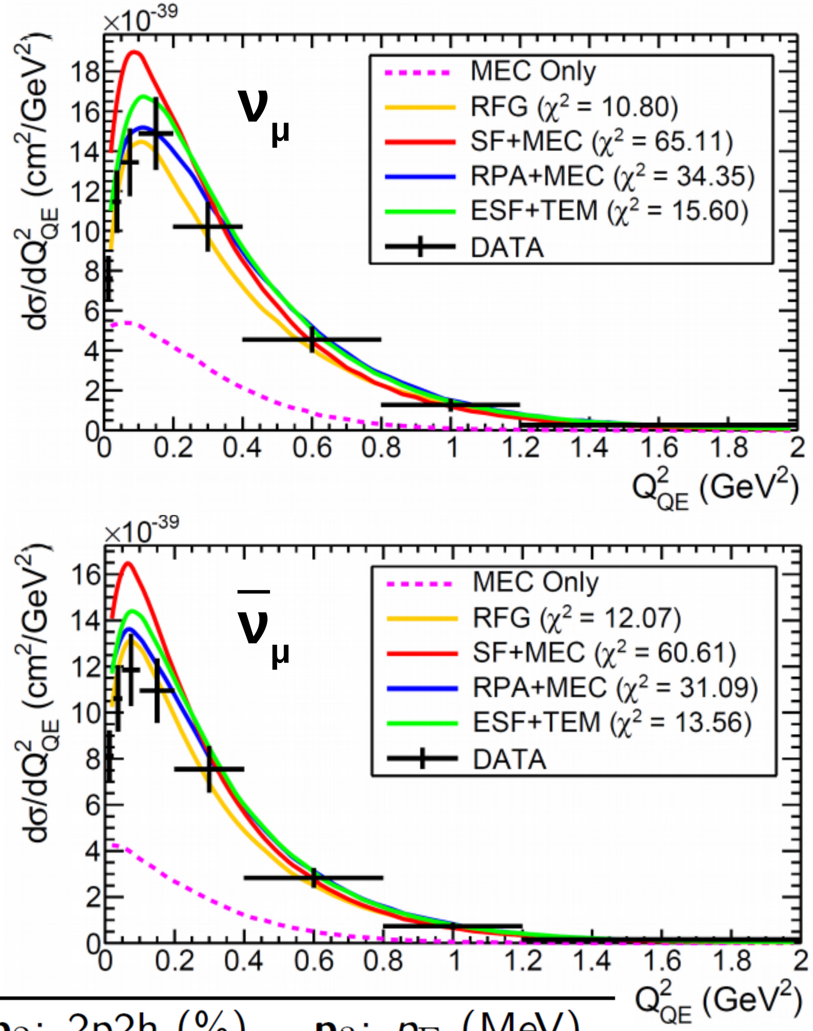


PRC 80 (2009) 065501

- That’s not to say that a single dataset can’t show that “the model” is insufficient to explain data!
- **Example:** MiniBooNE CC0π results sparked a number of theory models which all sought to explain the huge data-MC difference
- Many were broadly successful...

CC0π fit ~2015

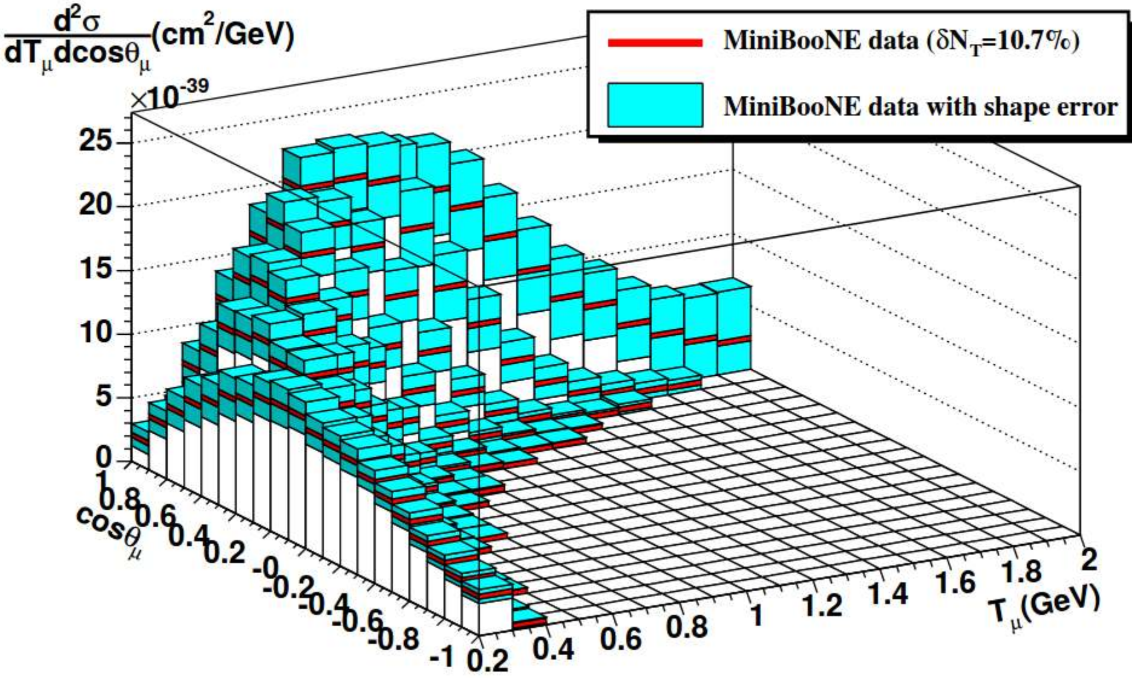
- Attempt to fit all CC0π data to two (semi-complete) models in NEUT:
  - MiniBooNE T_μ-cosθ_μ ν_μ
  - MiniBooNE T_μ-cosθ_μ $\bar{\nu}_\mu$
  - MINERvA Q^2 ν_μ & $\bar{\nu}_\mu$ (with corr.)
- Used to select CC0π model for T2K osc. analysis
- Unable to fit data well, **surprising** and **unsatisfactory** results.



Model	χ^2/DOF	$\mathbf{p}_1$: M_A (GeV)	$\mathbf{p}_2$: 2p2h (%)	$\mathbf{p}_3$: p_F (MeV)
Theory expectation		∼1.00	∼100	∼217
A: SF+2p2h	97.5/228	1.33±0.02	0 (at limit)	234±4
B: RFG+RPA+2p2h	97.8/228	1.15±0.03	27±12	223±5

Gory details in PRD 93 (2016) 072010

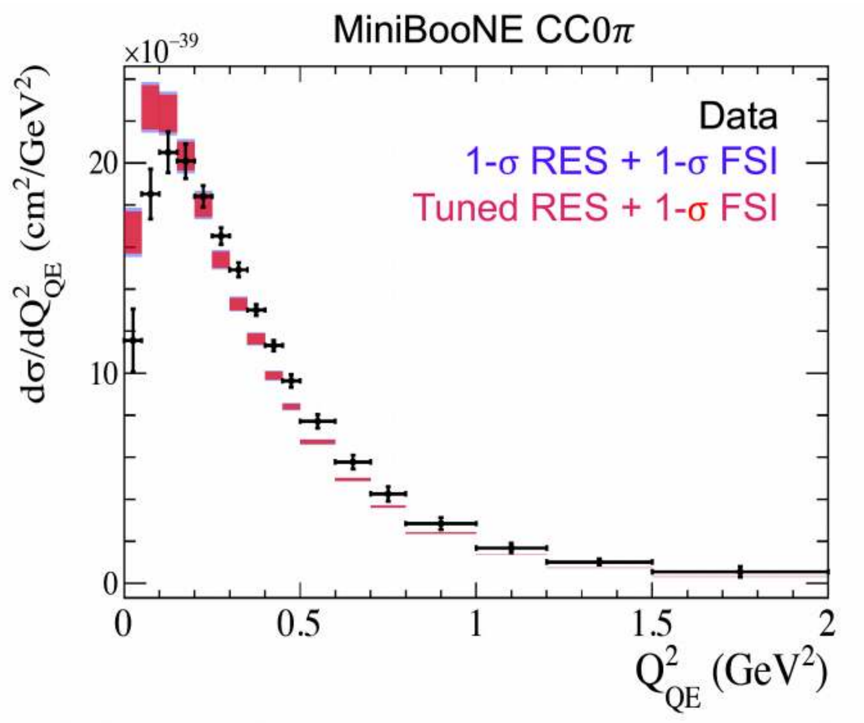
Challenge #1: missing information

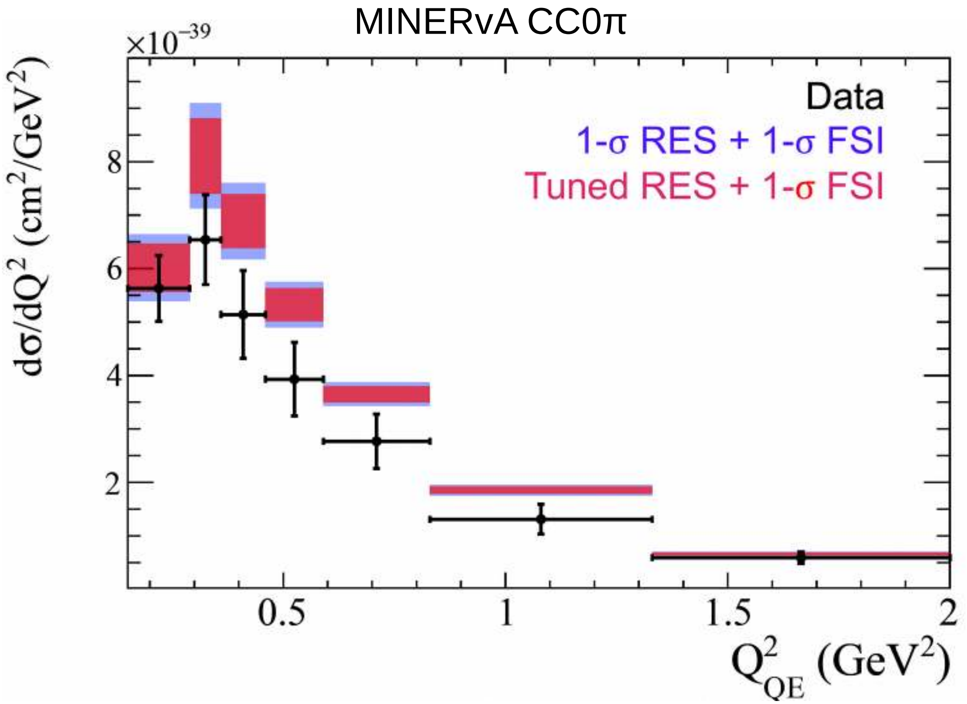


PRD 81 (2010) 092005

- Some older datasets are missing vital information
- **Example:** MiniBooNE CC0π data – bin-to-bin correlations are obviously strong, but no covariance provided
- Naively using the information provided yields χ^2/DOF ~0.1…

Challenge #2: difficult to factorize problem

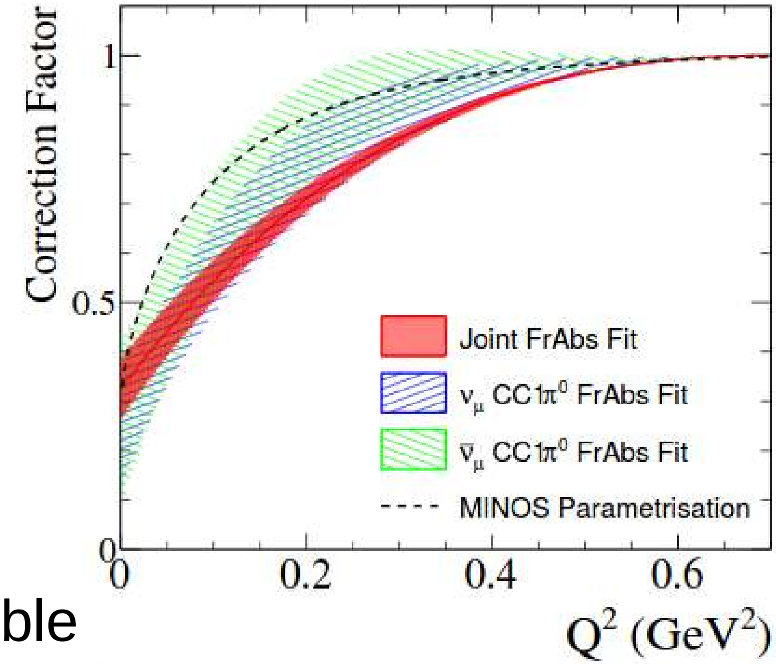




- Difficult to isolate regions of the model to tune with data
- **Example:** pion production + pion FSI uncertainties shown for the two CC0π samples used in 2015 fits
- Can’t tune the CCQE/2p2h models alone without making assumptions about these!

CC1π fit ~2019

- Joint MINERvA-NUISANCE tuning of GENIE model to published MINERvA pion production data [1]:
  - ν_μ-CC1$\pi^\pm$ [2]
  - ν_μ-CCN$\pi^\pm$ [3]
  - ν_μ-CC1π^0 [4]
  - $\bar{\nu}_\mu$-CC1π^0 [2]
  - p_μ, T_π, θ_μ and θ_π available for each



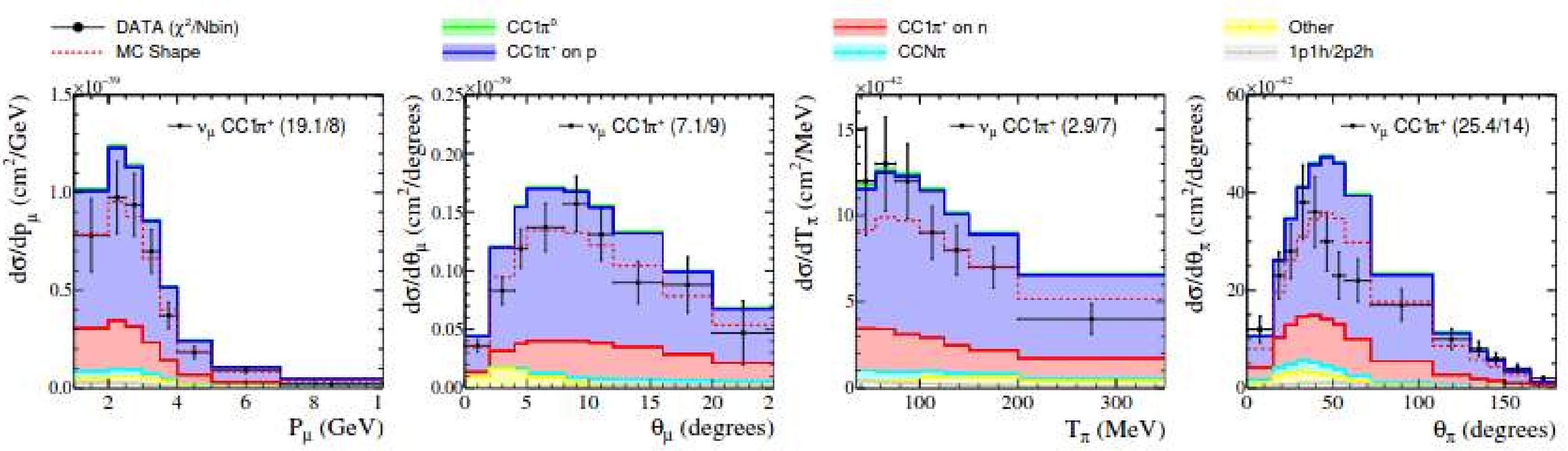
- Tensions found between MINERvA and bubble chamber data, but a low-Q^2 suppression helped

- Used to look for biases in DUNE sensitivity and T2K data analysis

[1] PRD 100 (2019) 072005
[2] PRD 92 (2015) 092008

[3] PRD 94 (2016) 052005
[4] PRD 96 (2017) 072003

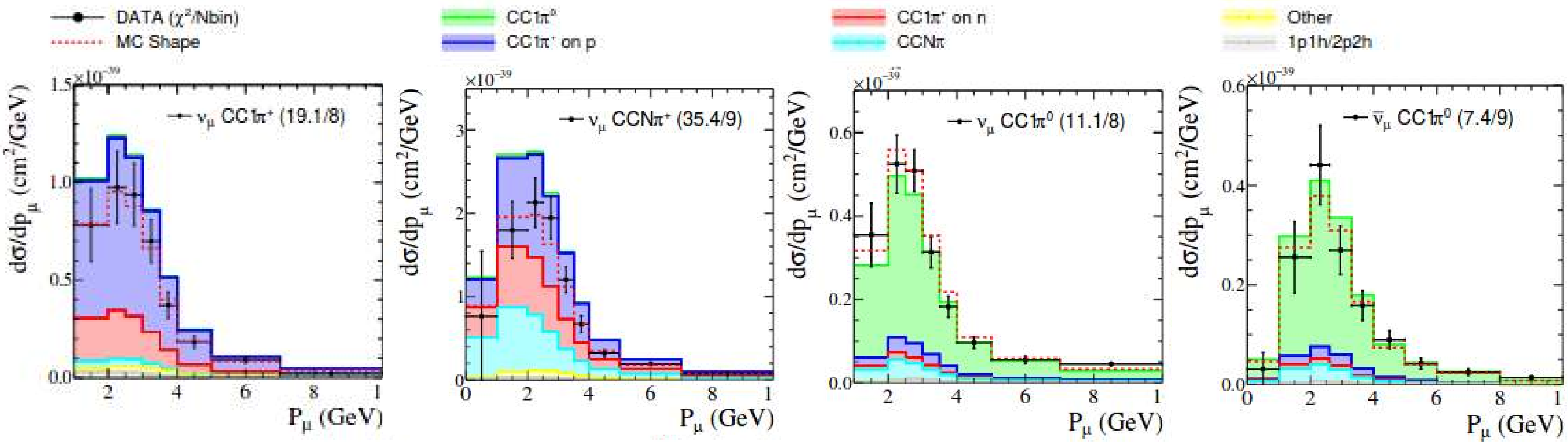
Challenge #3: cross-correlations



PRD 100 (2019) 072005

- Experiments often produce multiple projections of the same data, which may help constrain different aspects of the model
- However, without correlations *between* projections, including all of these pieces of information is challenging → ad hoc solutions
- Typically, correlations are large for all experiments due to flux normalization uncertainties of 5-15%

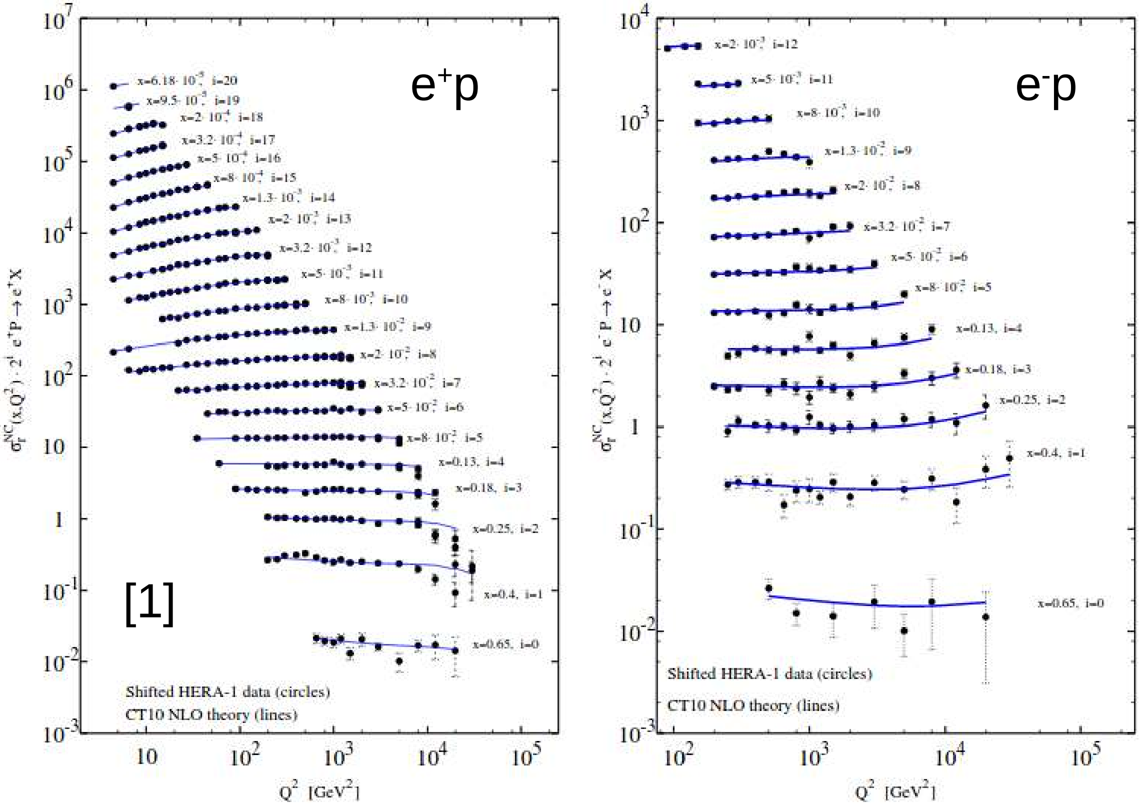
Challenge #3: cross-correlations



PRD 100 (2019) 072005

- Different analyses from the same experiment are also correlated, at least through the flux, and may contain a subset of the same events
- Experiments in the same beamline may complicate this issue further, all leading to underestimated tension in global fits…
- Large covariances matrices help, but quickly become unwieldy. Producing correlations with old datasets may not be possible

Challenge #3: lessons from parton PDFs?



- Parton PDF fits had similar missing correlation issues [1]
- Final $e^{\pm}p$ HERA reanalysis presented the impact of 114 systematics separately, for each bin [2]
- Avoids large covariances, and can see the effect of each systematic in the fit

- However, significant effort (and a consistent set of systematics) required to make this sort of move
- Control samples complicate this sort of approach as they introduce correlations between systematics...

[1] CTEQ, PRD82 (2010) 074024
[2] H1 and ZEUS, JHEP (2010) 1001:109

Recommendations #1: correlations

- ***Recommendations just intended as a starting point for discussion!***

- Correlations between datapoints in a distribution are essential for a usable result

- Correlations between multiple distributions are necessary for fits to large ensembles of data to be successful

- We should seek to learn from fields which tackled similar issues to optimize our “legacy measurements” for the future

Challenge #4: reliability of data

$$\frac{d\sigma}{dx_i} = \frac{\sum_j \widetilde{U}_{ij}^{-1}\left(N_j - B_j\right)}{\Phi_\nu T \Delta x_i \epsilon_i}$$

- Common assumption that cross-section extraction takes the measured rate and presents it in a slightly massaged form
- “The data is the data”, right?

Challenge #4: reliability of data

Some unfolding methods introduce bias

The signal definition and background subtraction can be model dependent

$$\frac{d\sigma}{dx_i} = \frac{\sum_j \widetilde{U}^{-1}_{ij}\left(N_j - B_j\right)}{\Phi_\nu T \Delta x_i \epsilon_i}$$

The choice of variables can rely on an implicit model correction

Efficiency corrections couple to model in complex ways

- Common assumption that cross-section extraction takes the measured rate and presents it in a slightly massaged form
- “The data is the data”, right?

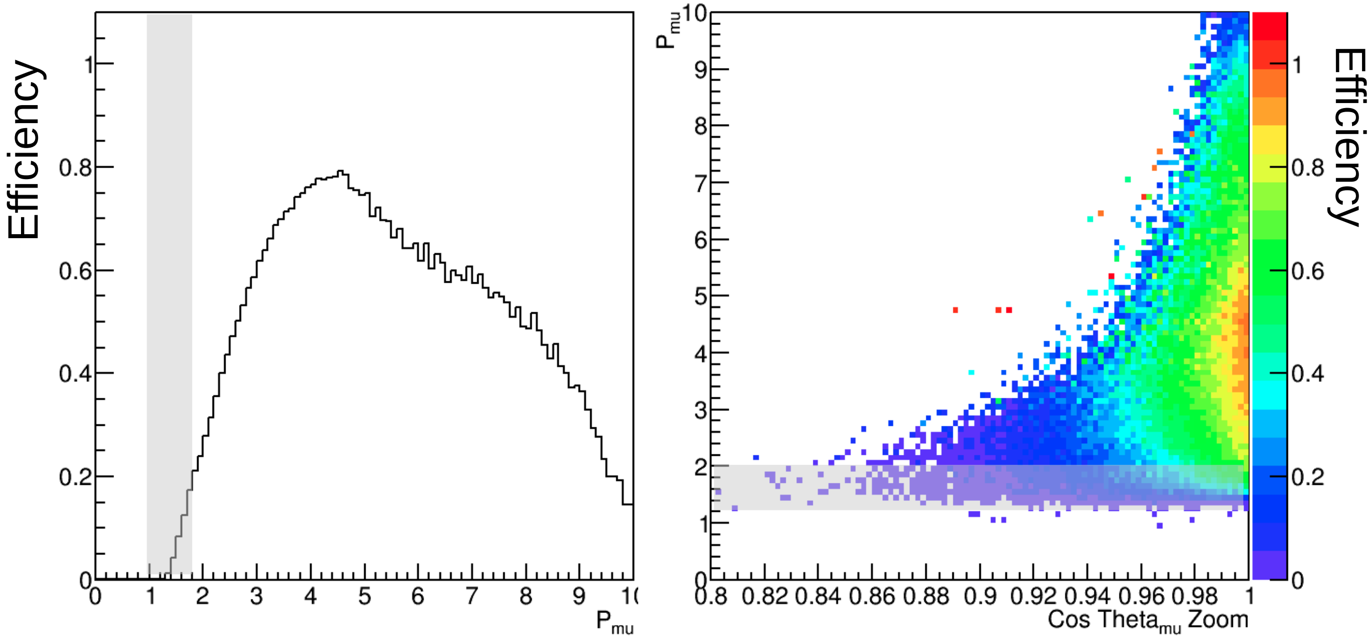
Challenge #4: reliability of data

- A problem for all fitting efforts (e.g., parton PDFs) is the need to decide which data to exclude or “deweight” in fits
- As XSEC data producers, we want to ensure our work is as “future proof” as possible, and will be used
- Conversely, as XSEC consumers, we need to ensure that we don’t allow imperfect data to bias our analyses!

Example problem: 1D efficiency correction

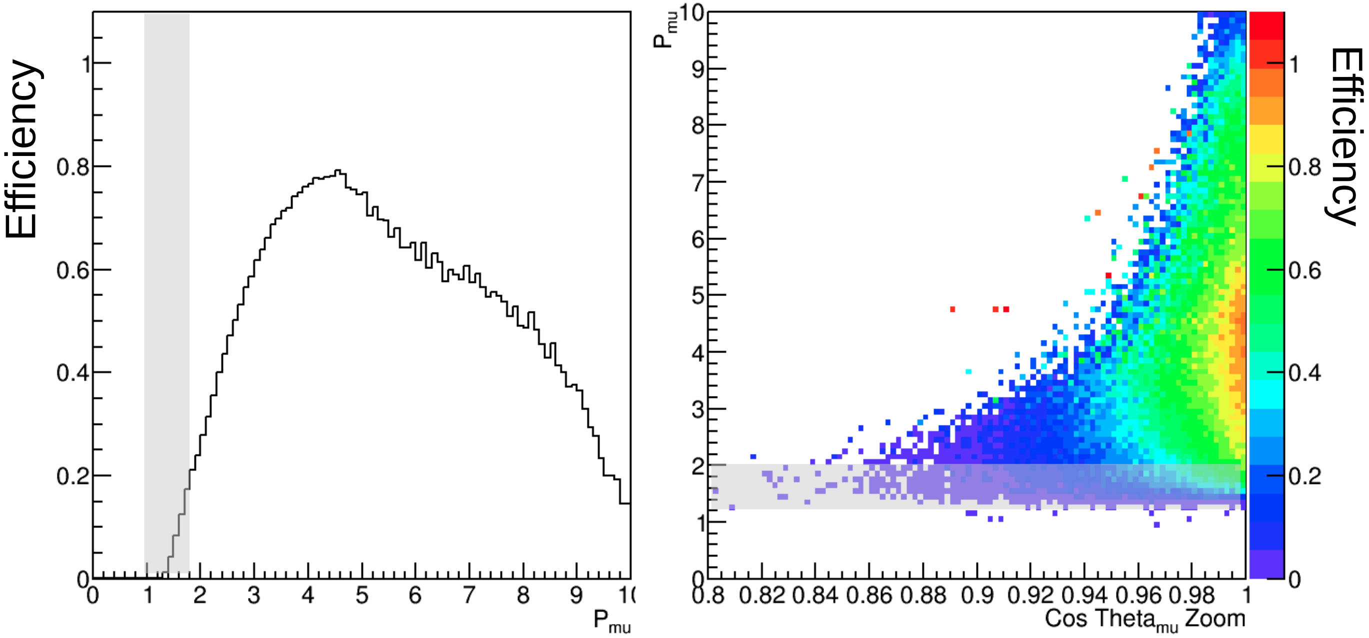
- Efficiency corrections often done on a bin-by-bin basis, with all other degrees of freedom integrated out, but can lead to model dependence

- E.g., making a CC-inc measurement differential in p_μ integrates out $\cos\theta_\mu$ variations. All events in a p_μ bin are given the same correction
  → **implicit bias to simulated $\cos\theta_\mu$ distribution**



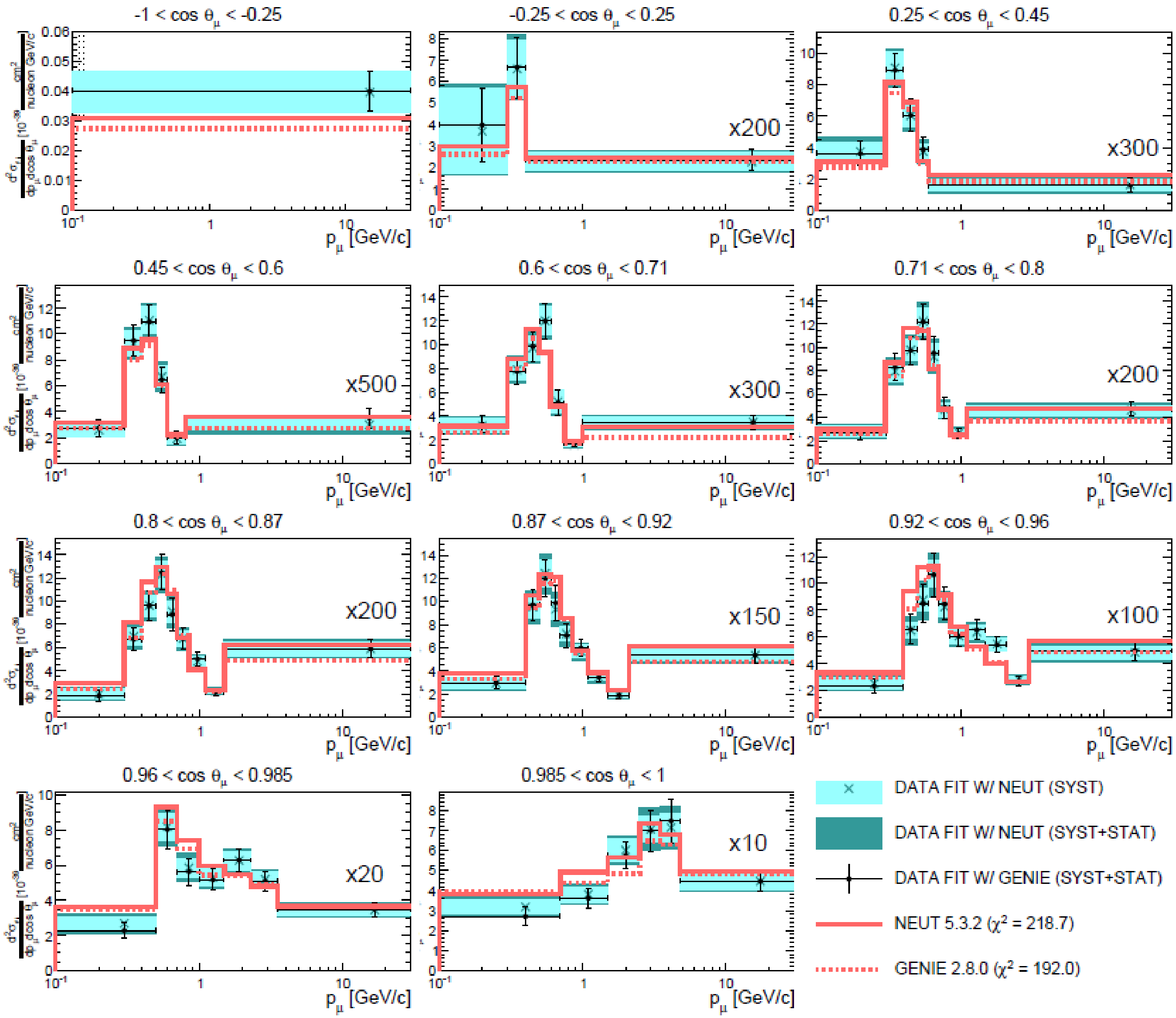
Mitigation

- In this simple case, the solution is to extract the cross section in 2D then collapse it onto 1D (or make the efficiency is flat in $\cos\theta_{\mu}$)
- But what happens to other degrees of freedom? Is the efficiency flat as a function of all other particle kinematics in the event? Do they matter?
- The **hope** is often that these will be covered by XSEC modelling systematics…



Diagnosing efficiency issues

- In this T2K CC-INC analysis, GENIE (2.8.0) and NEUT were used
- Despite extensive care to add robust modelling systematics, results differ
- Also worth remembering that these simulations have a lot of similarities!
- Multiple generators are a good (but not exhaustive) way to look for issues



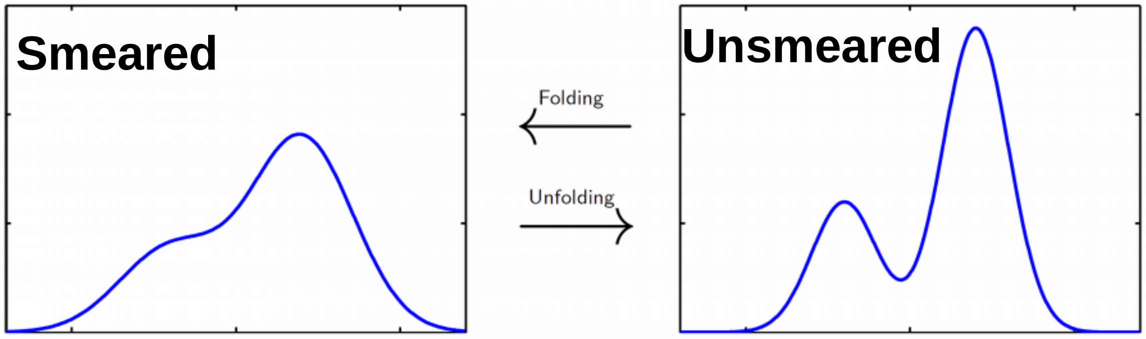
PRD 98 (2018) 012004

Recommendations #2: efficiency issues

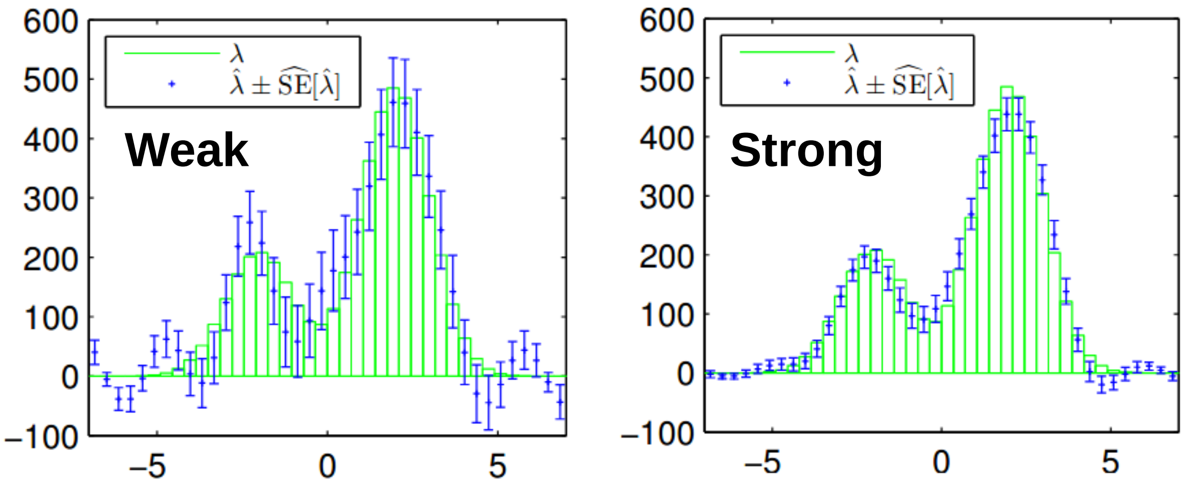
- Phase space restrictions should be applied to signal definition to avoid correcting for unsampled regions of phase space
- Care should be taken to ensure that variables which have a large impact on the efficiency are not implicitly integrated over
- Limitations of simulation and assumptions made should be acknowledged in papers so data consumers don’t have to guess
- **Independent** MC package/tune/versions should be pushed through the detector simulation and analysis results compared
- Fake data studies to test robustness of analysis should be carried out, where reweighting is not limited to analysis bins

Unfolding methods

- Attempt to go from the smeared (measured) space to the true space



- Many detailed methods available, each with benefits and pitfalls
- A key issue is “regularization”, e.g., how strongly the results are smoothed to avoid statistical uncertainties from blowing up



Figures from M. Kuusela

Recommendations #3: unfolding

- Model comparisons should be carried out in both reco and true spaces with an appropriate goodness of fit metric for each

- Comparisons to control samples used in the analysis should also be shown in the reco space for completeness

- Regularization of results should be performed using data-driven methods where possible

- Methods which produce regularized results can generally also produce unregularized results, both are useful!

- Not unfolding at all is a useful option, but can make it harder to propagate some uncertainties

Summary

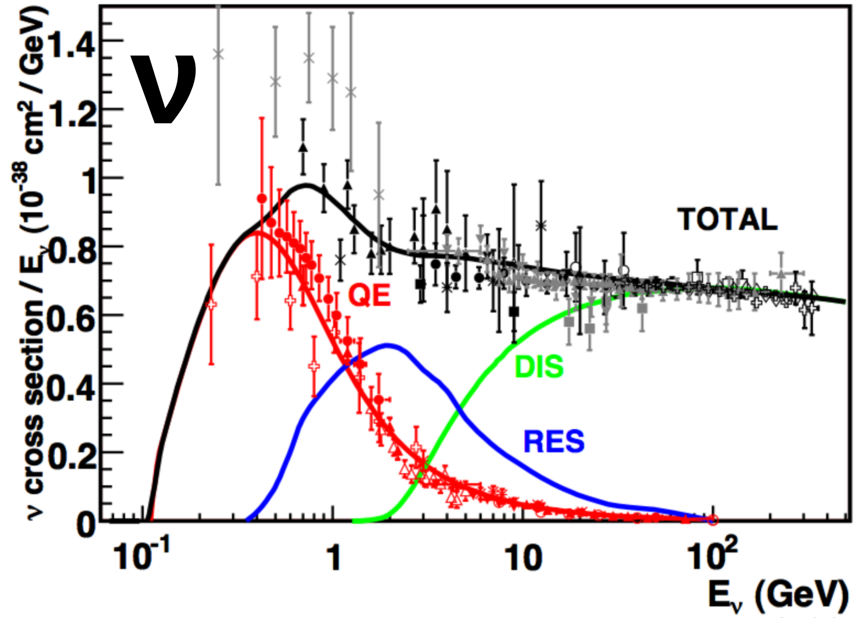
- Neutrino XSEC measurements are important to build systematic uncertainty models for oscillation experiments
- A number of issues limit the utility of some older datasets
- Need to “future-proof” new and recent data to ensure its continued use in the precision oscillation era
- Some lessons we could learn from other communities, e.g., parton PDF fitters, but some problems unique to our community
- I’ve made a few recommendations for ways to ensure that the hard work that goes into all XSEC analyses is fully utilized
- We should gather suggestions from the community and come up with a set of recommendations (but not requirements)

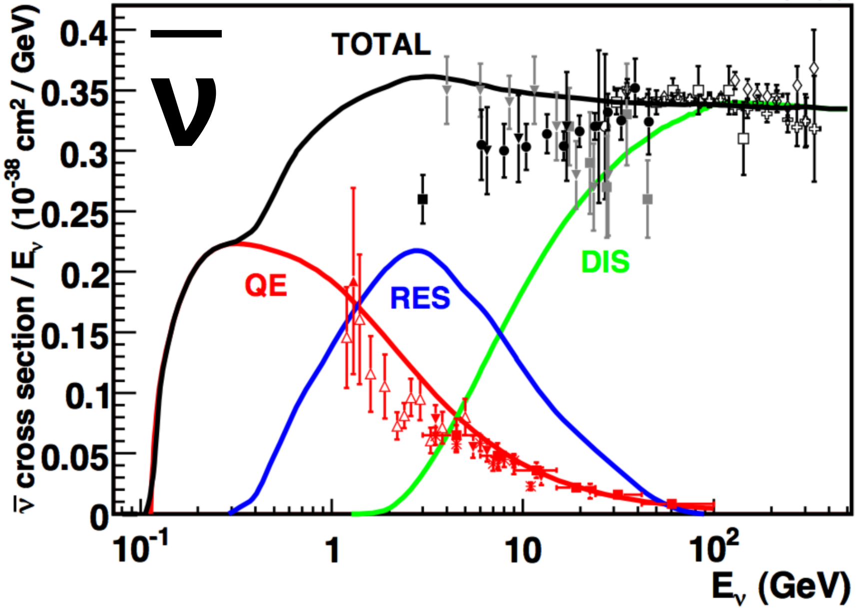
Backup

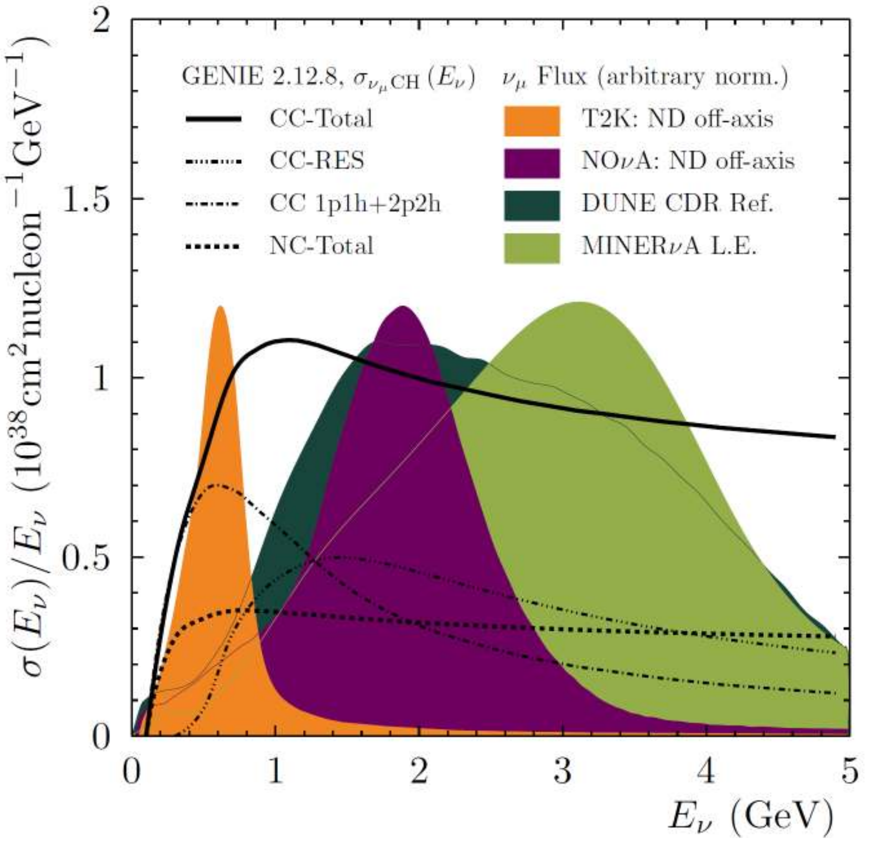
Why do we want to measure XSECs?

All oscillation experiments live in the 0.1-10 GeV transition region.

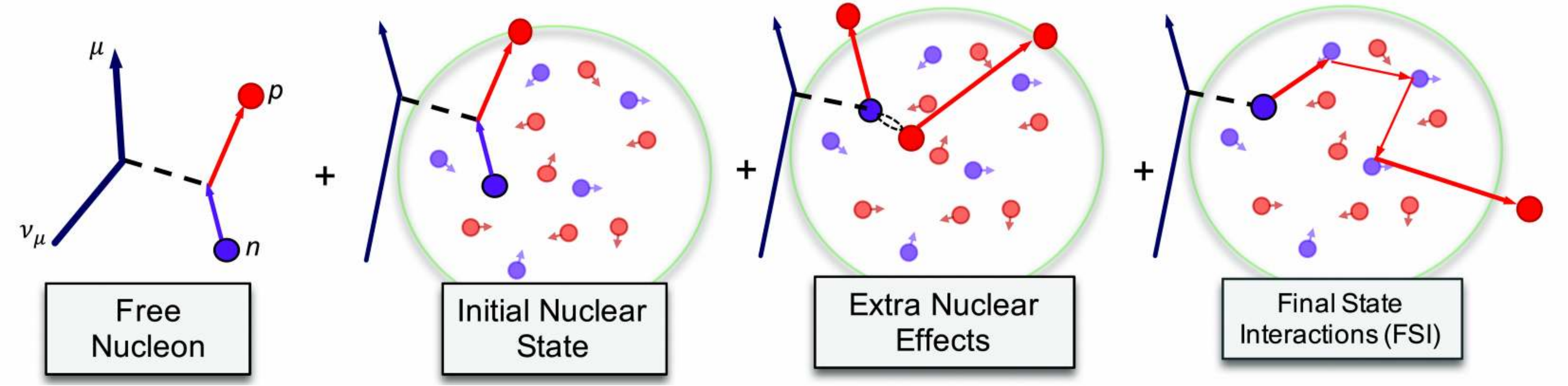
Multiple models required with different physical assumptions.





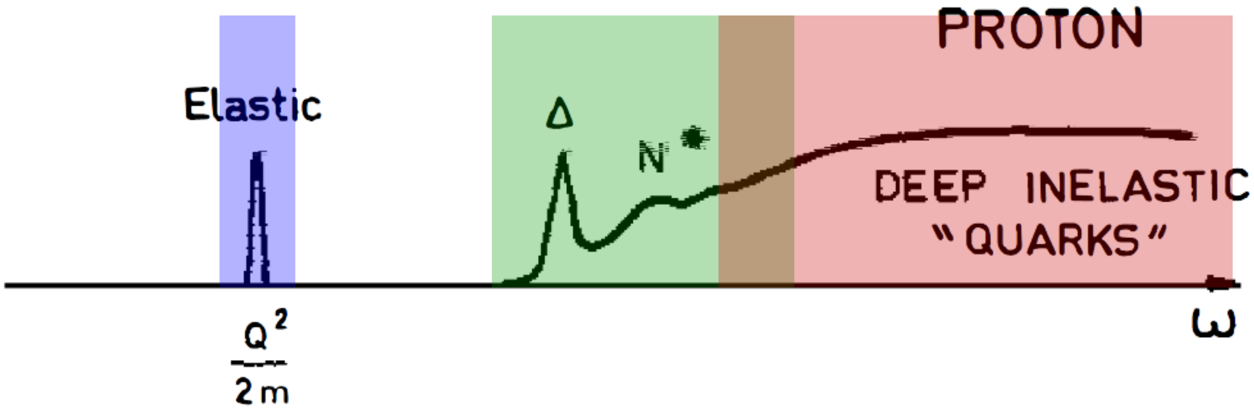


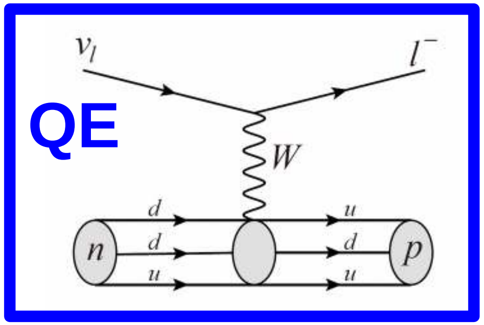
Nuclear targets

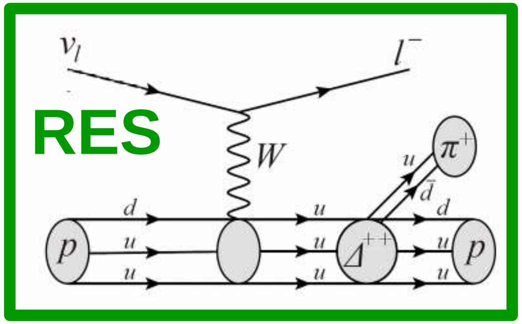


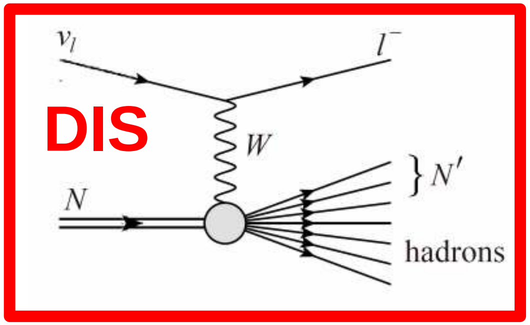
- **Free nucleon:** the interaction level cross section, including hadronization at high energy transfer
- **Initial nuclear state:** how nucleons behave inside the nucleus. E.g., Relativistic Fermi Gas.
- **Nuclear effects:** additional effects due to the presence of multiple nucleons. E.g. np-nh interactions.
- **Final State Interactions:** subsequent interactions before interaction products exit the nucleus.

Free nucleon response

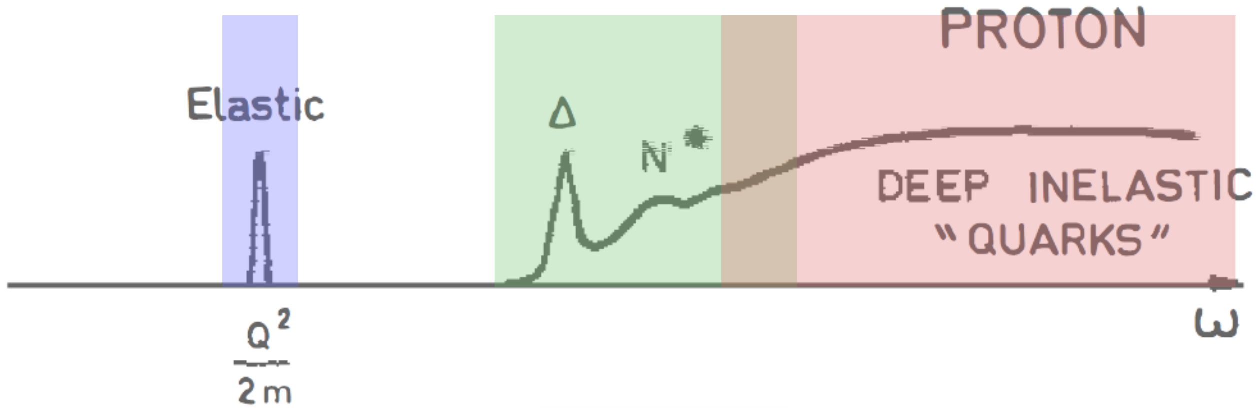


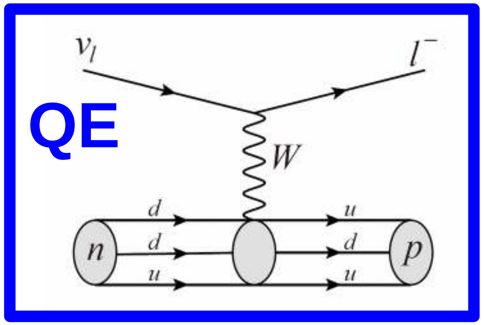


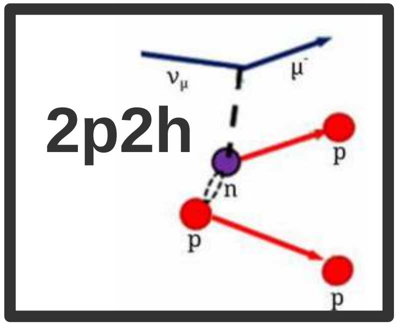


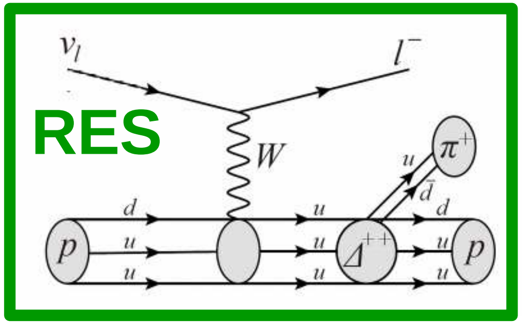


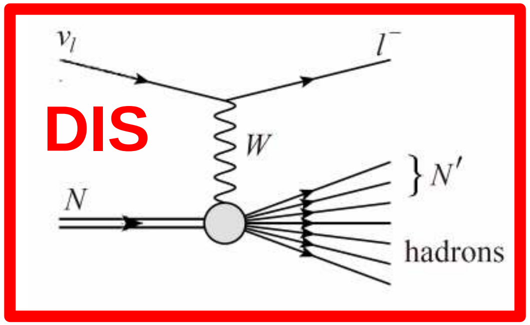
Nuclear response

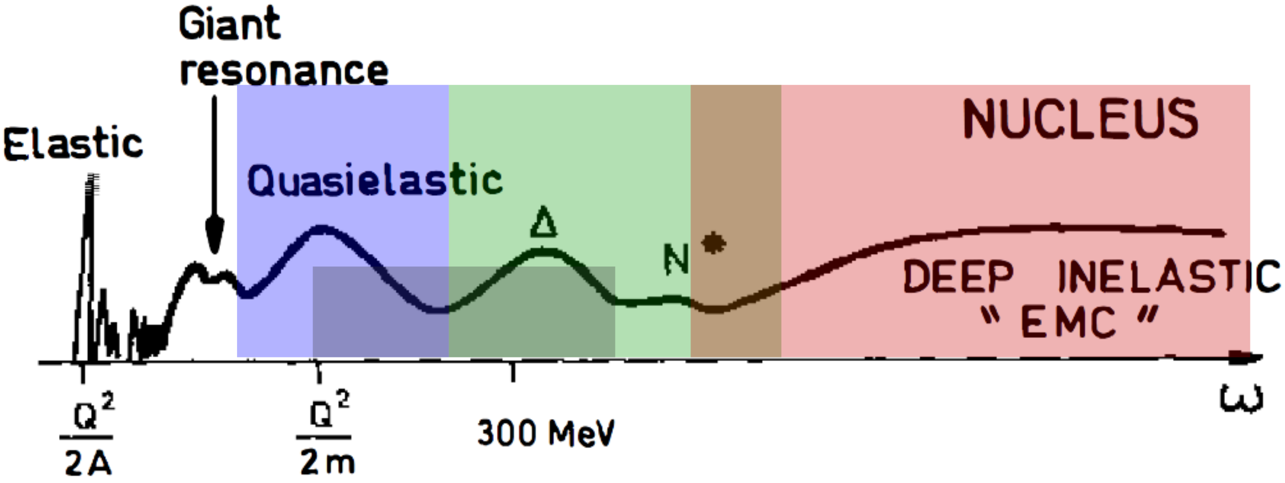






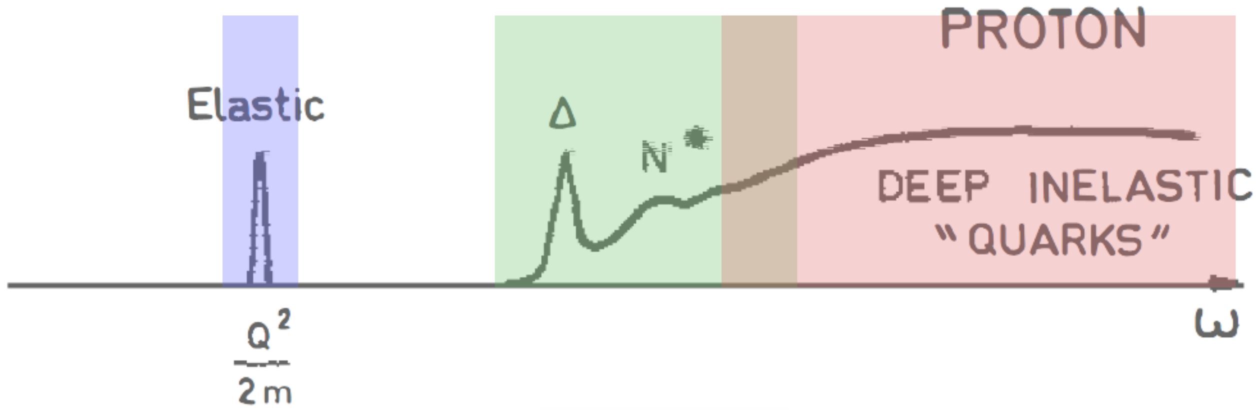


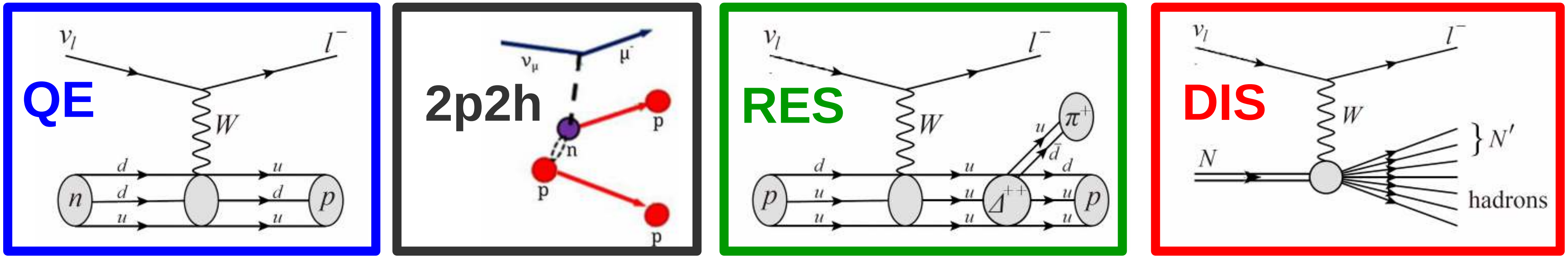


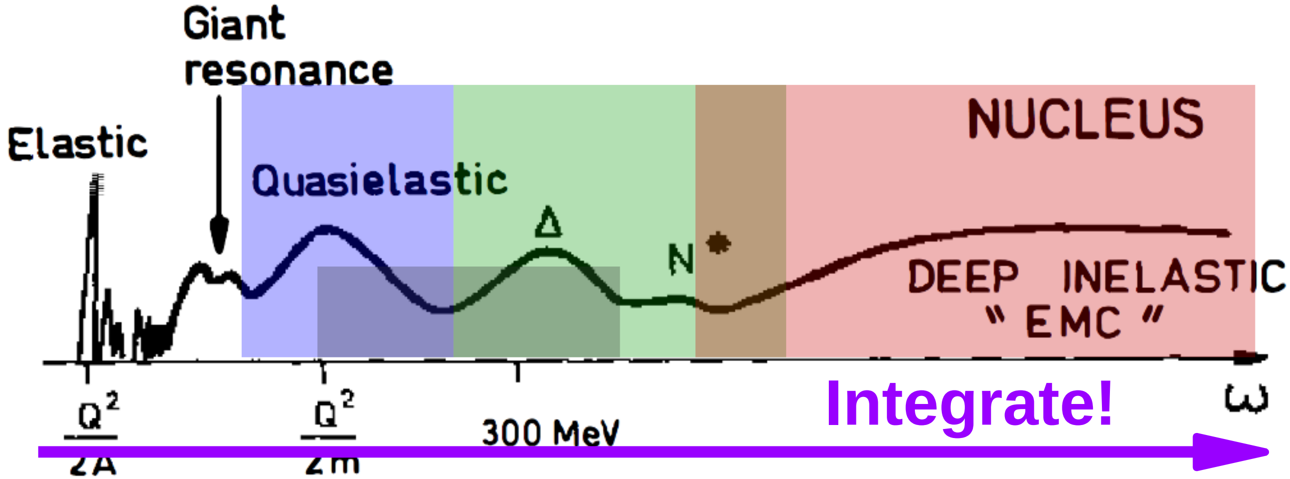


Interactions with more than one nucleon contribute

Nuclear response

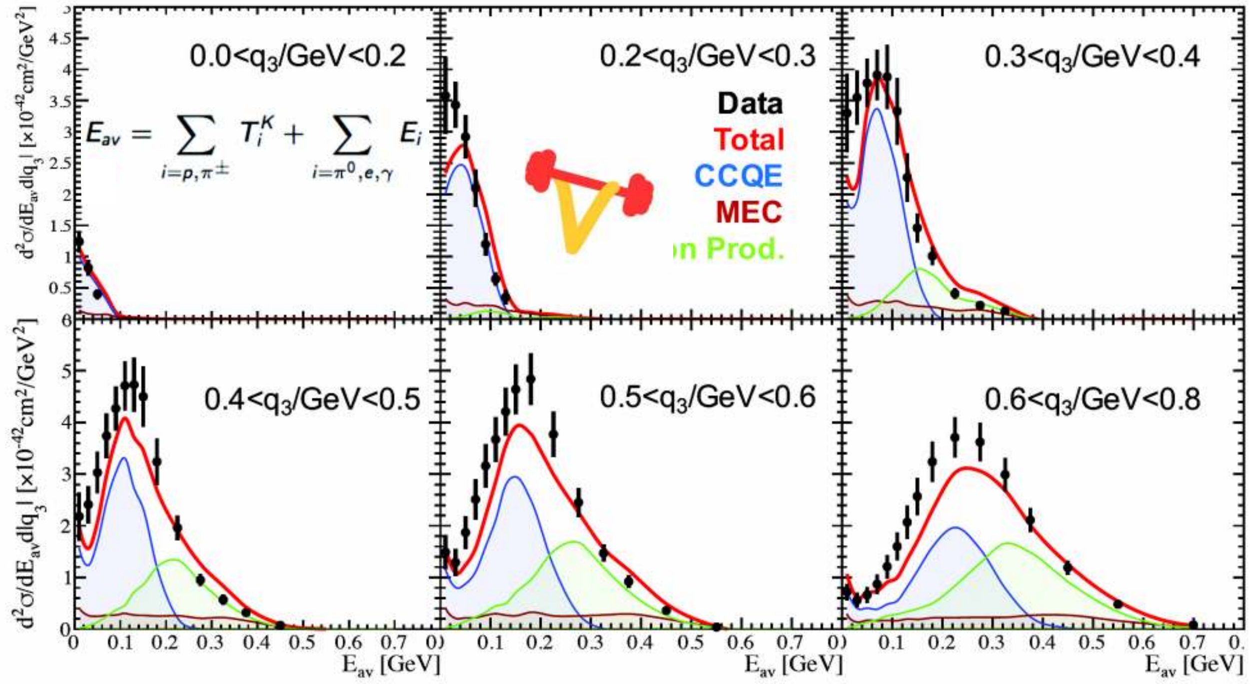






Can't reconstruct ω, so no way to avoid poorly modelled regions!

Counter example: MINERvA CC-inclusive

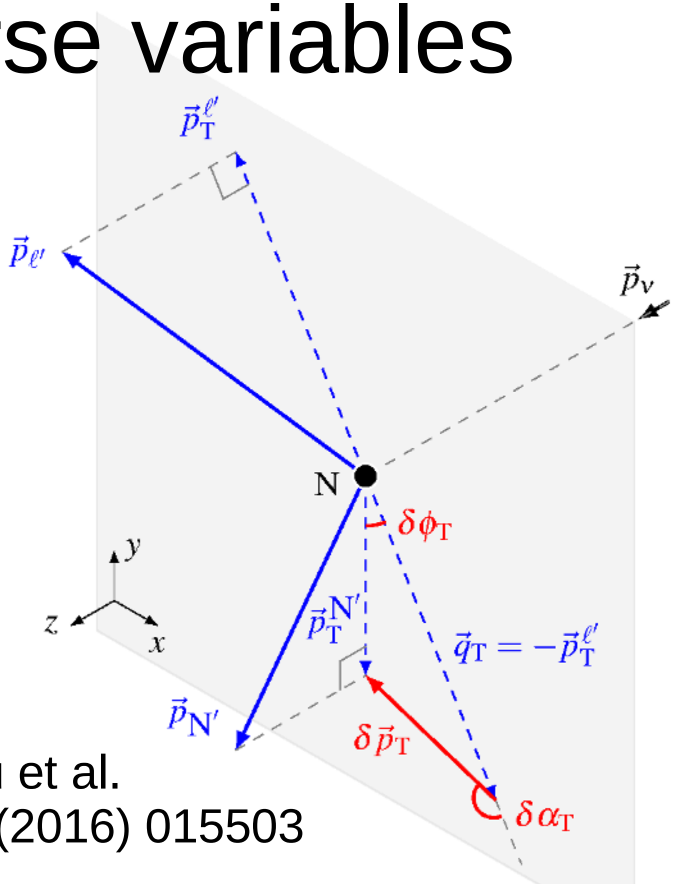


PRL 116 (2016) 071802

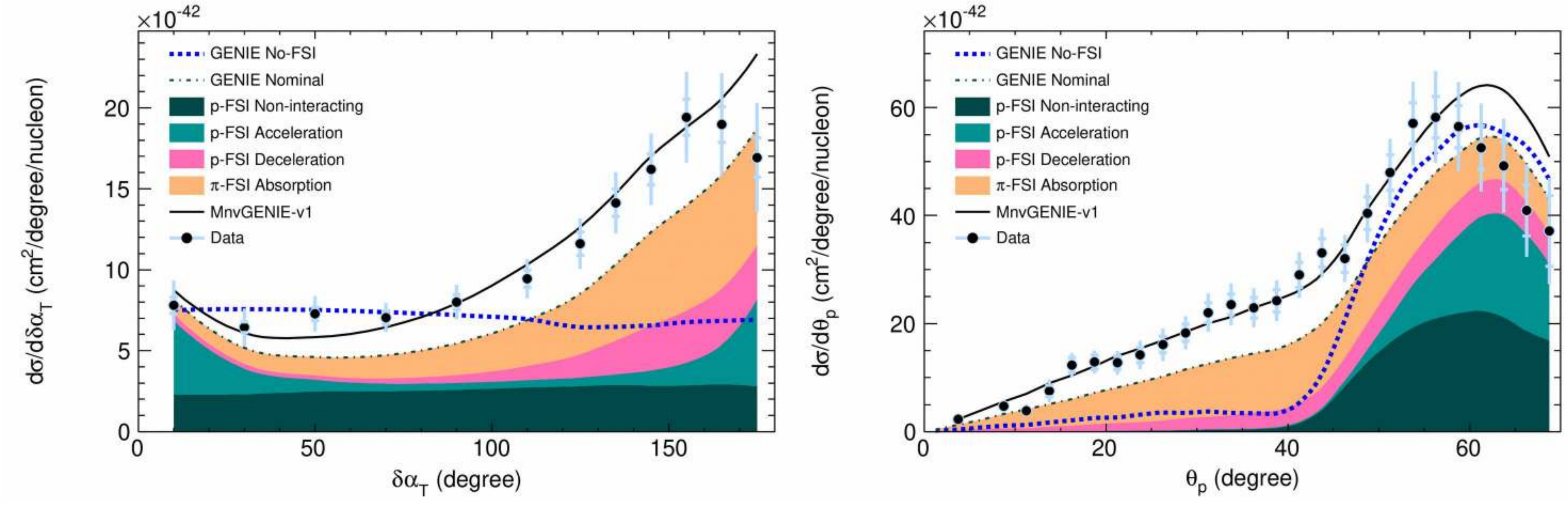
- Effort to measure low momentum transfer events in *interaction level* kinematics
- Goes some way to isolating “dip” region between QE and Delta (pion production)
- Clear data excess in this area, but still difficult to really isolate the problematic region

Counter example: transverse variables

- Some interesting new variables used in CC0π analyses aim to break degeneracies between effects
- Stringent model tests, and *suggestive* of underlying problems in some cases
- Other thoughtful variables may help in future



X -G. Lu et al.
PRC94 (2016) 015503



PRL121 (2018) 022504

CC0π reconstructed kinematics

Reconstructed neutrino energy and reconstructed four-momentum transfer are given by:

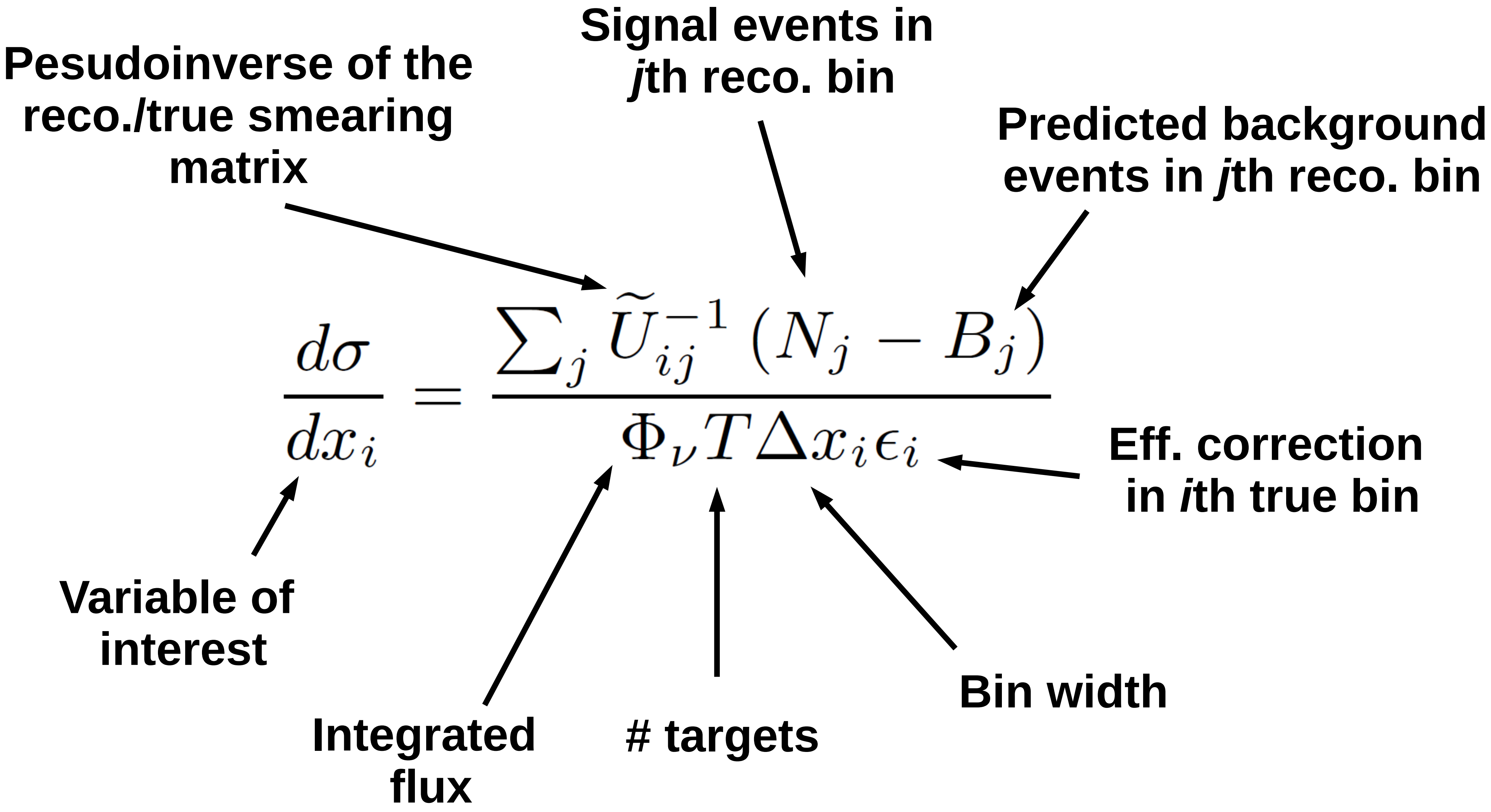
$$E_\nu^{QE} = \frac{2M'_n E_\mu - (M'^2_n + m_\mu^2 - M_p^2)}{2(M'_n - E_\mu + \sqrt{E_\mu^2 - m_\mu^2}\cos\theta_\mu)}$$

$$Q^2_{QE} = -m_\mu + 2E_\nu^{QE}(E_\mu - \sqrt{E_\mu^2 - m_\mu^2}\cos\theta_\mu)$$

Where $E_\mu = T_\mu + m_\mu$, and $M'_n = M_n - E_B$.

E_B = 34 MeV for all datasets except MINERvA anti-neutrino, which uses E_B = 30 MeV.

Cross section extraction



$$\frac{d\sigma}{dx_i} = \frac{\sum_j \widetilde{U}_{ij}^{-1}\left(N_j - B_j\right)}{\Phi_\nu T \Delta x_i \epsilon_i}$$

Pesudoinverse of the reco./true smearing matrix

Signal events in *j*th reco. bin

Predicted background events in *j*th reco. bin

Eff. correction in *i*th true bin

Variable of interest

Bin width

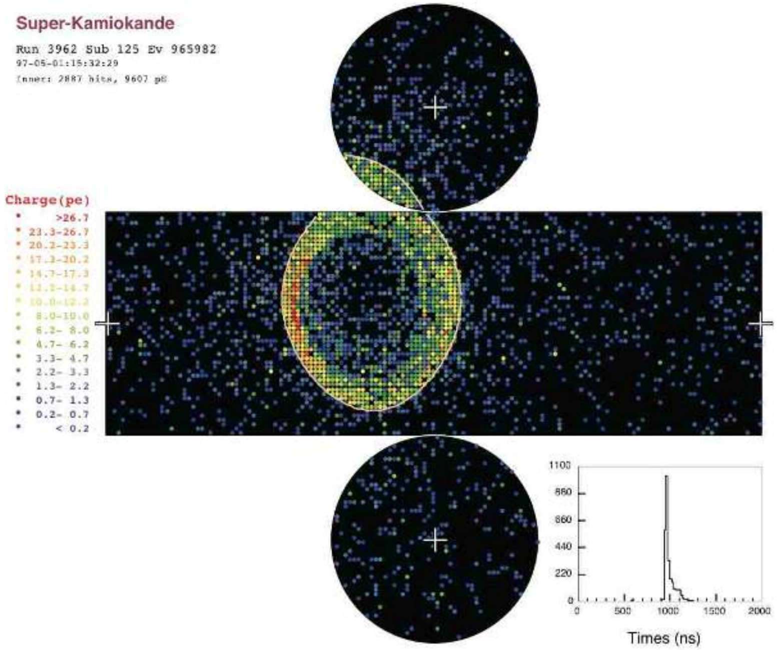
Integrated flux

 # targets

E_ν reconstruction methods

(1) **Leptonic** variables only:

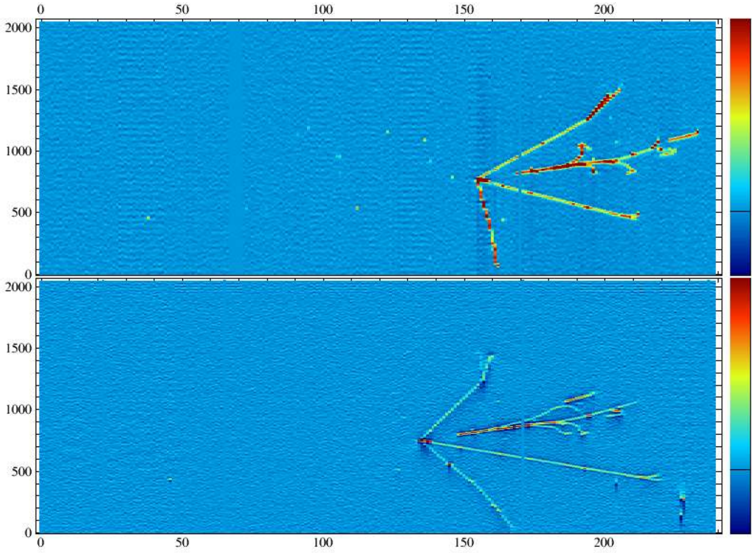
$$E_\nu^{QE} = \frac{m_p^2 - {m'}_n^2 - m_\mu^2 + 2{m'}_n E_\mu}{2({m'}_n - E_\mu + p_\mu \cos\theta_\mu)}$$



Water Cherenkov: T2K, Hyper-K

(2) **Leptonic** and **hadronic** information:

$$E_\nu = E_\mu + E_{\text{had}}$$



Tracking calorimeter: NOvA; **Liquid Argon TPCs:** DUNE

So what do we need to know?

(1) **Leptonic** variables only:

$$E_\nu^{QE} = \frac{m_p^2 - {m'}_n^2 - m_\mu^2 + 2{m'}_n E_\mu}{2({m'}_n - E_\mu + p_\mu \cos\theta_\mu)}$$

- E_ν → muon kinematic spectrum
- Pion production below threshold
- Pion production + absorption rate
- Smearing from nuclear model

(2) **Leptonic** and **hadronic** information:

$$E_\nu = E_\mu + E_{\text{had}}$$

- Pion production rate below experimental threshold
- Neutral energy fraction (and dependence on E_ν)
- Smearing from nuclear model

Cross section model dependence

The signal definition and background subtraction can be model dependent

Some unfolding methods introduce bias

$$\frac{d\sigma}{dx_i} = \frac{\sum_j \widetilde{U}^{-1}_{ij}\left(N_j - B_j\right)}{\Phi_\nu T \Delta x_i \epsilon_i}$$

Choice of variables

Efficiency corrections couple to model in complex ways

Choice of variables

- **Model-independent:** final state particle kinematics, or some combination of them (e.g., Q^2_{QE}). *Combinations are prone to subtle efficiency issues!*

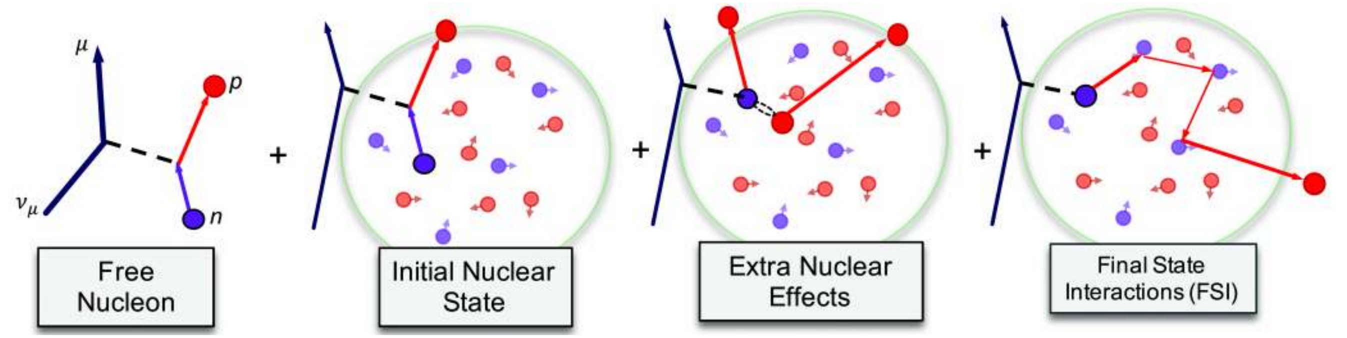
$$Q^2_{\mathrm{QE}} = -m_\mu^2 + 2E_\nu^{\mathrm{QE,\ RFG}}(E_\mu - \sqrt{E_\mu^2 - m_\mu^2}\cos\theta_\mu)$$

- **Model-dependent:** interaction-level kinematics, Q^2, E_ν W…

- Perception that theorists will prefer interaction-level variables because they are easier to use… but this risks them not being used in the future!

- **Recommendation:** produce results in final state particle kinematic variables *as well as* anything more exotic

Model-independent signal definitions



⇒ **Visible particles**

- Experiments can only measure final state particles, e.g., 1μ 0π:

CC0π = CCQE + npnh(0π) + CC1π(+abs) + …

- Many previous measurements try to correct for irreducible backgrounds to make the result easier to use...

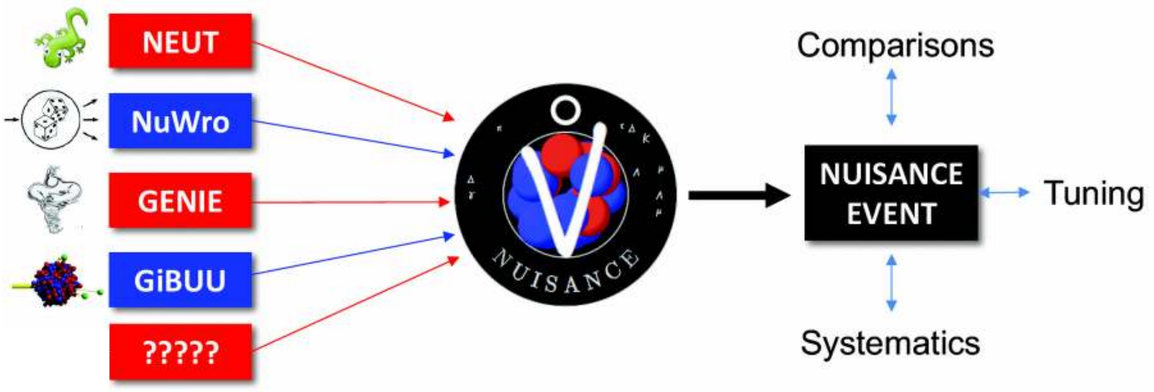
… but trying to recover CCQE introduces model dependence

$$\textbf{CCQE} = \underbrace{\textbf{CC0}\pi}_{\textbf{Data}} - \underbrace{\textbf{npnh(0}\pi\textbf{) - CC1}\pi\textbf{(+abs) -} \ldots}_{\textbf{Generator}}\ \textbf{???}$$

NUISANCE

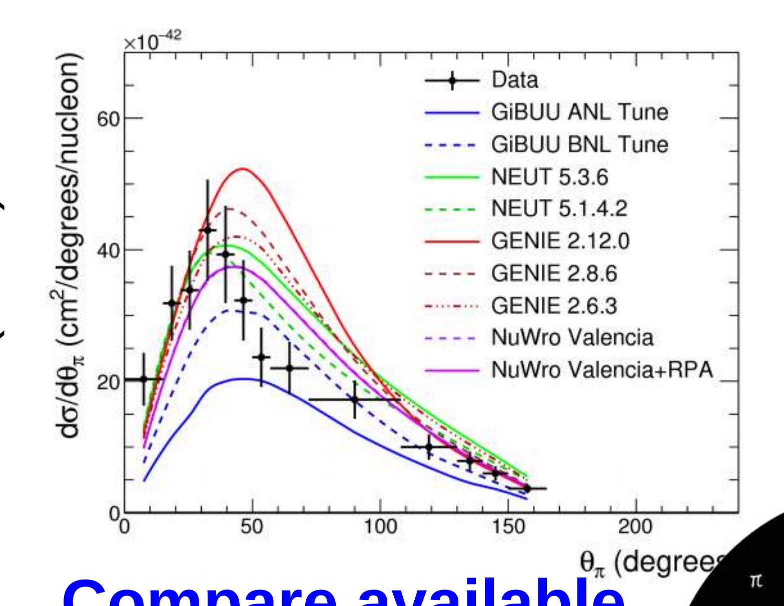
- General purpose cross section comparison and tuning framework:
  - Large collection of datasets included (~130)
  - Support for multiple Monte Carlo generators

- ***Open source (***GNU GPLv3): **nuisance.hepforge.org**



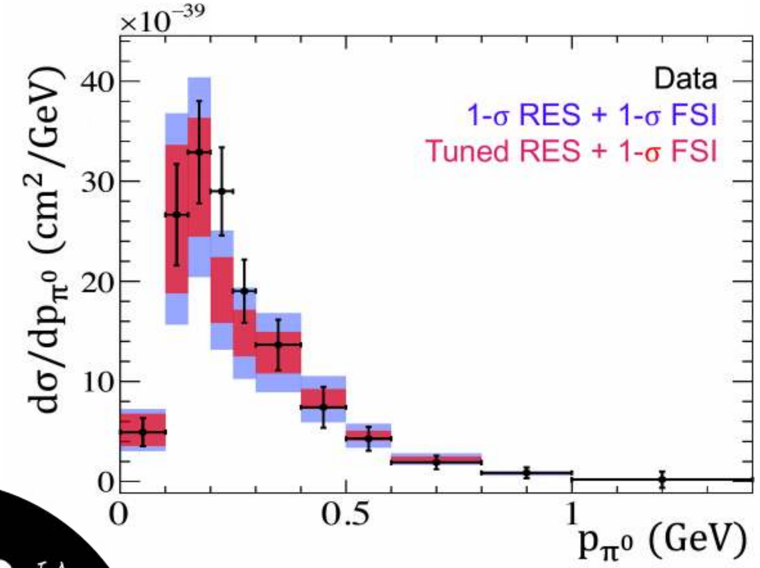
P. Stowell, C. Wret, C.W. et al. JINST 12 (2017) P01016

MINERvA ν_μ-CH CC1π^+
PRD 92 (2015) 092008



Compare available models

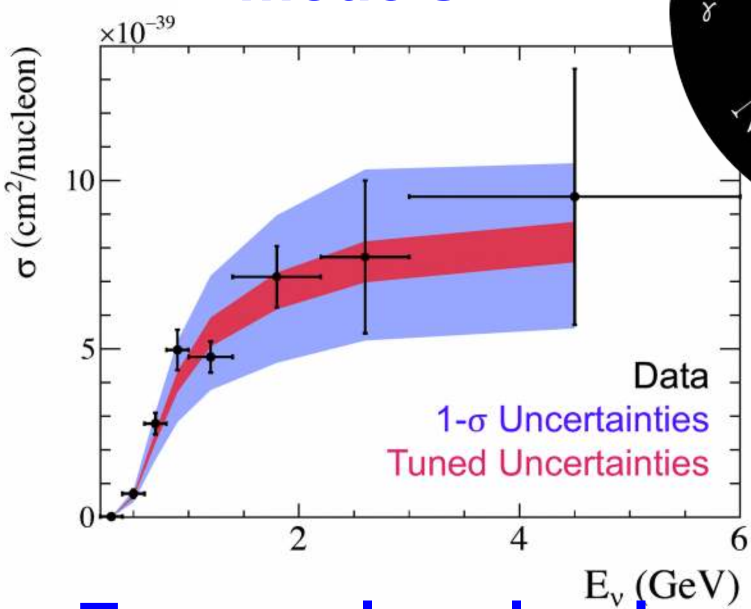
MB ν_μ-CH$_2$ CC1π^0
PRD 83 (2011) 052009



Use data to validate nuclear uncertainties

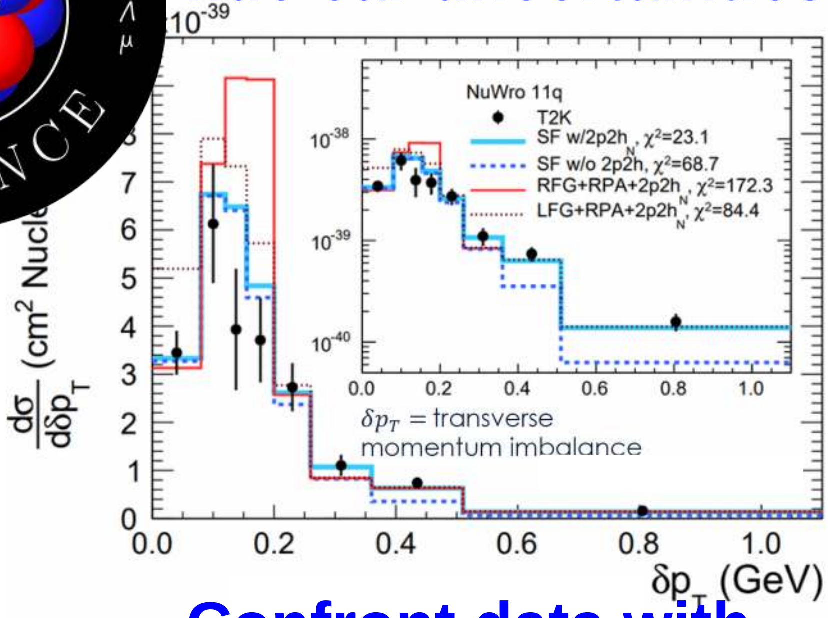
NUISANCE

ANL CC1p1π^+
PRD25 (1982) 1161-1173



Tune nucleon level uncertainties (QE, 1π)

T2K CC0π
PRD 98 (2018) 032003



Confront data with various models