SIS, DIS and hadronization uncertainties

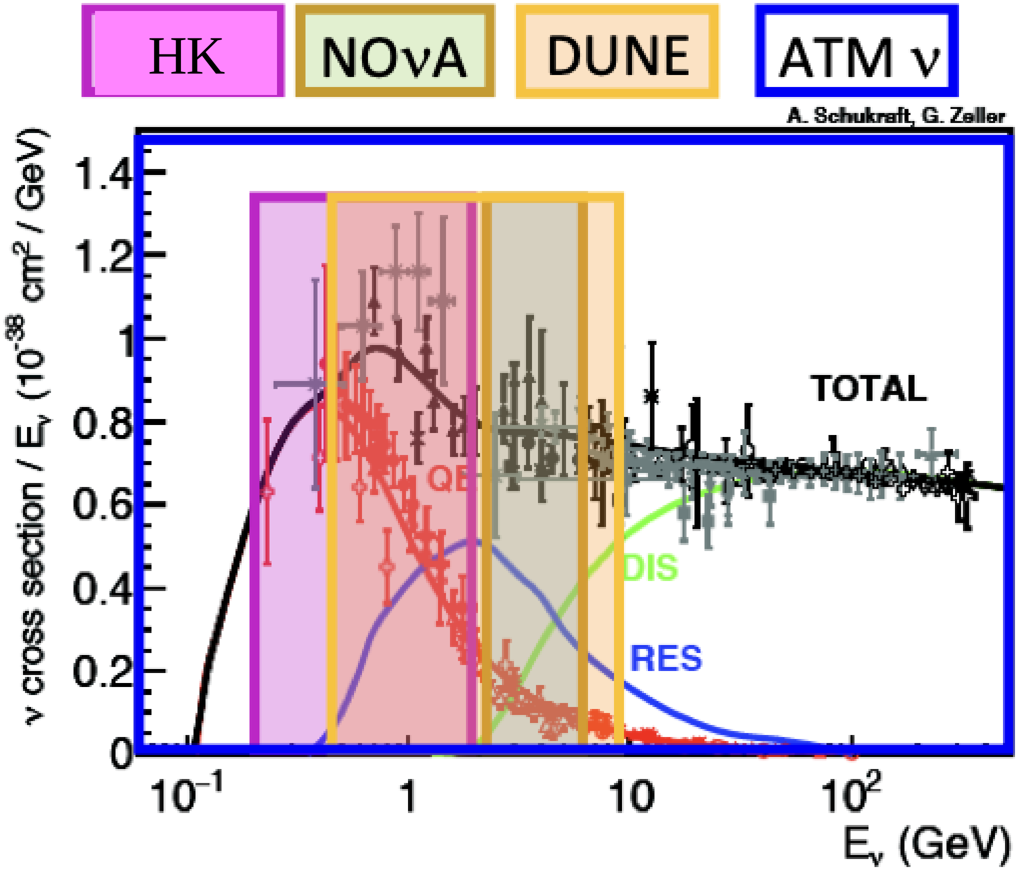
"Neutrino Interactions and the next generation of neutrino experiments"

ECT* Workshop, Torino

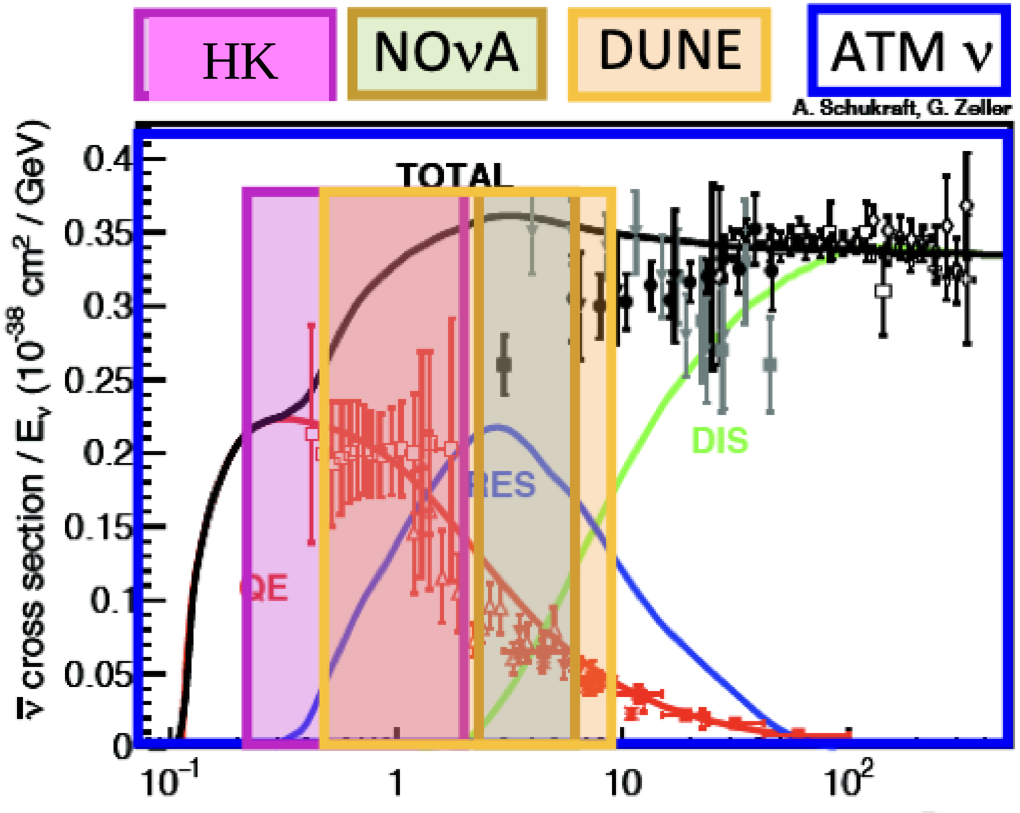
Julia Tena Vidal at Tel Aviv University

Processes of interest

Neutrino cross-section

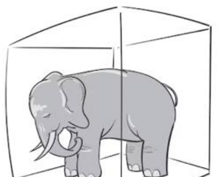


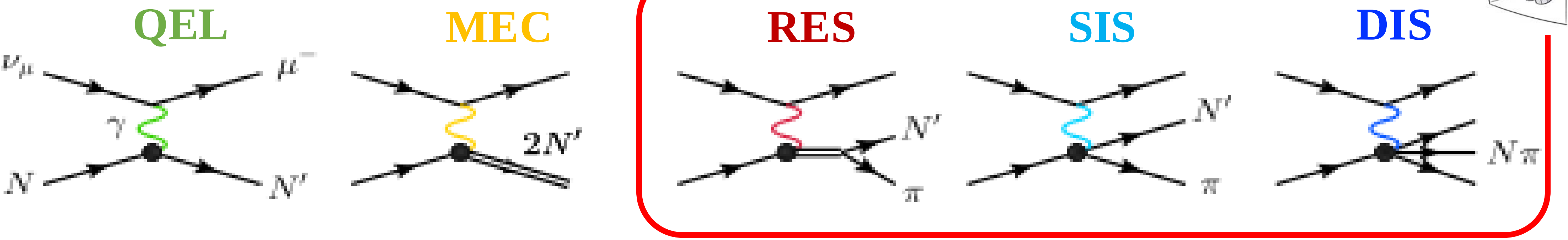
Anti-neutrino cross-section



J. A. Formaggio, G. P. Zeller, Rev. Mod. Phys. 84 (2012) 1307 E_ν (GeV)

Processes of interest





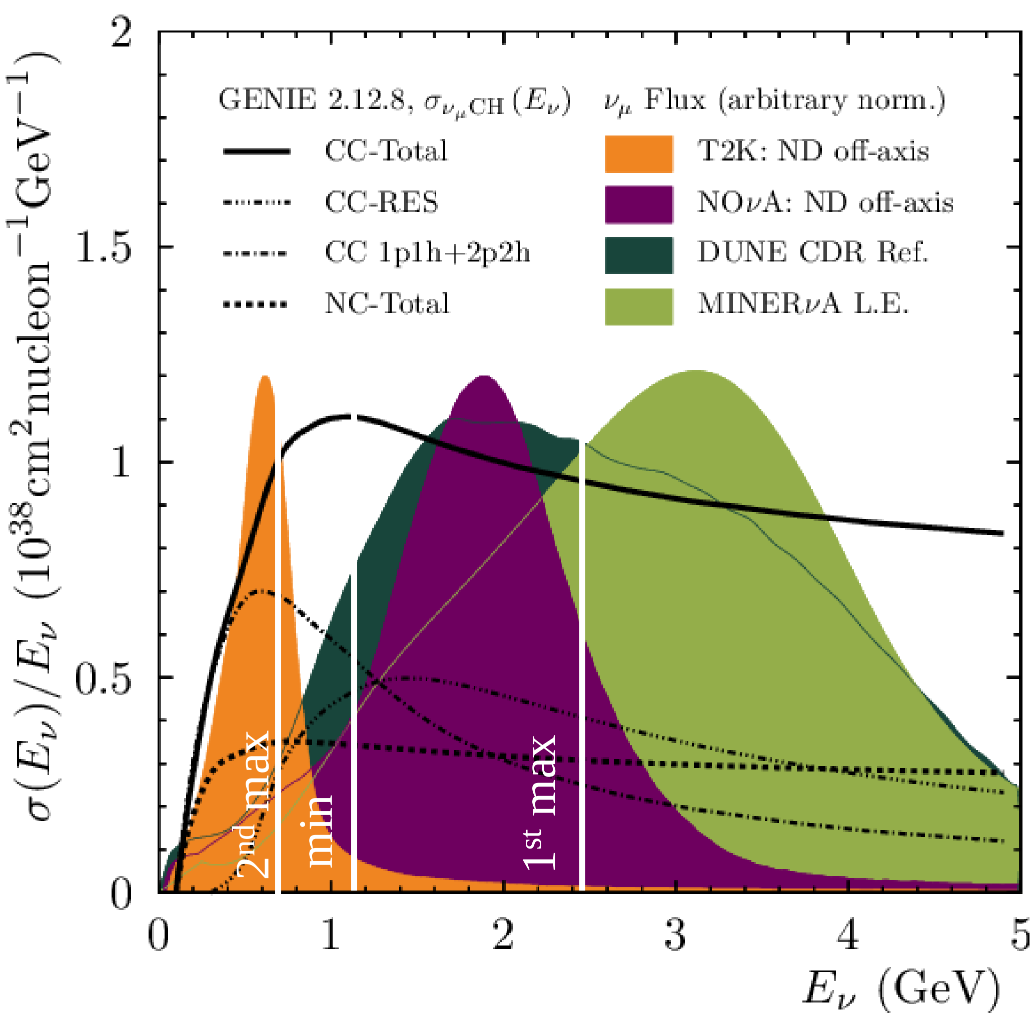
Process	SBN [%]	T2K/HK [%]	DUNE [%]
CCQE	45.9	42.0	20.0
CCMEC	10.5	8.3	5.5
CCRES	38.0	35.5	**41.0**
CCSIS/DIS	5.0	13.5	**30.1**

Table from S. Dolan

Processes of interest - DUNE

DUNE's oscillation sensitivity will be driven by SIS/DIS Systematics

- Pion production events dominate the event rate
- What are the relevant uncertainties?
- Are nuclear effects well understood?
- Can we reconstruct their energy?
    - See Adi Ashkenazi's talk
- Can we propagate the relevant uncertainties to the oscillation analysis?



Processes of interest - SBND

SBND will have unprecedented statistics

Physical Process	**Model Configurations (CC QE & CC 2p2h)**			
	Nieves	**SuSAv2**	**LS-E**	**SM**
	Charged Current			
QE	2,152,939	2,373,600	2,078,485	2,776,187
MEC	483,859	552,181	543,418	0
RES	1,160,763	1,160,763	1,160,763	1,160,763
DIS /SIS	220,335	220,335	220,335	220,335
Coherent	6950	6950	6950	6950
Other	2384	2384	2384	2384

Table from R. Johnes
Predictions from GENIE

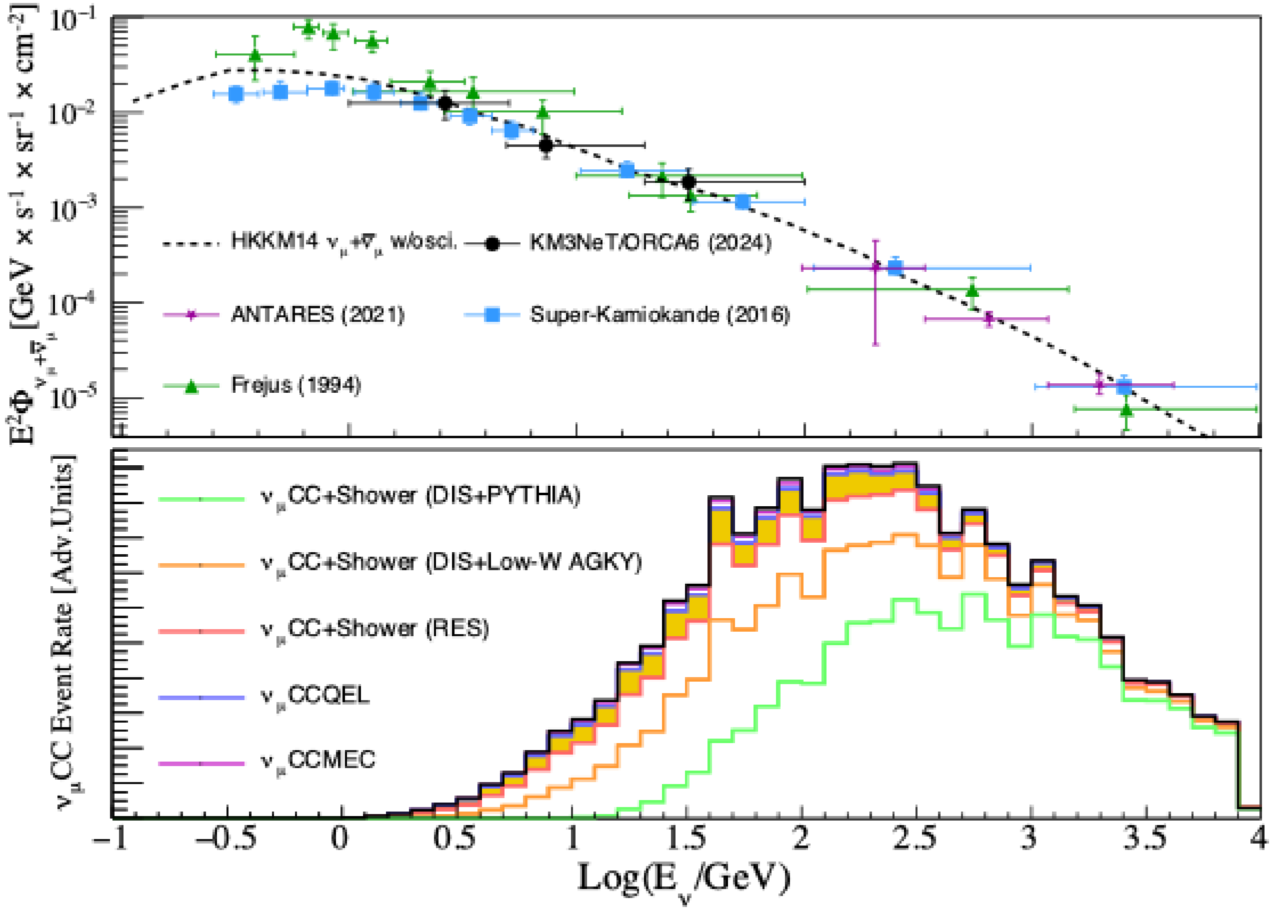
Processes of interest – KM3NeT-ORCA

KM3NeT-ORCA will be **dominated** by SIS/DIS events

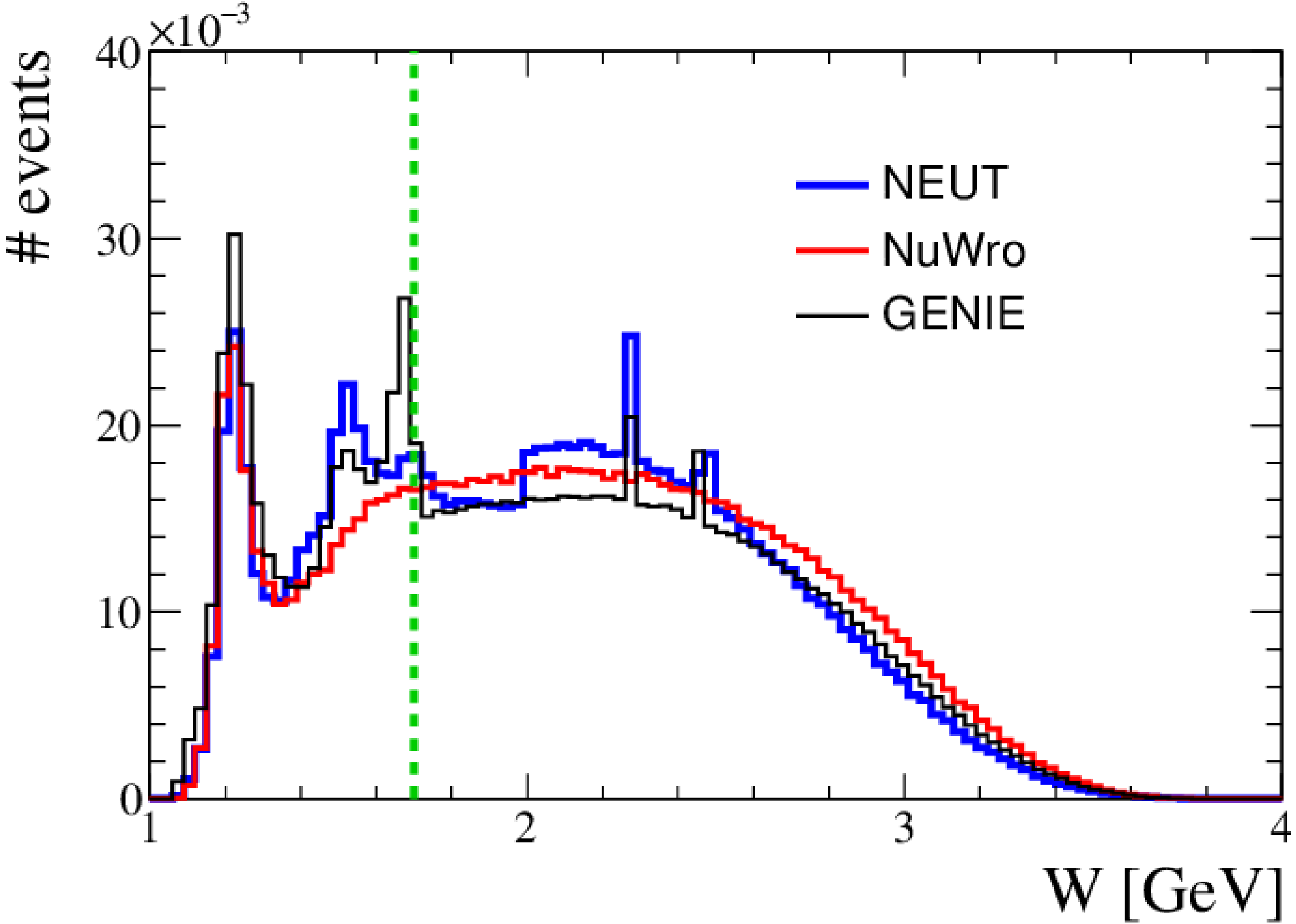
- PYTHIA systematics are most relevant
- Impact on oscillation analysis unknown

External constraints needed

- No Near Detector
- Down-going sample contaminated by cosmic muons

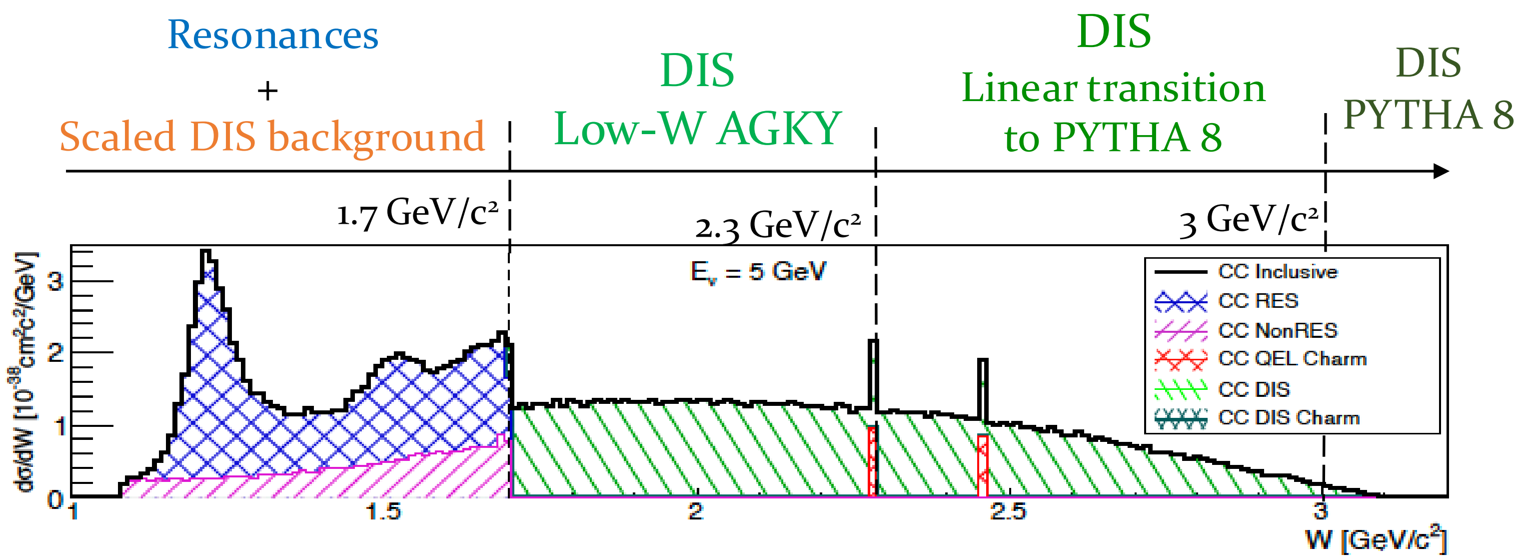


The Shallow Inelastic Scattering Region
Generators perspective



By C. Bronner

The Shallow Inelastic Scattering Region GENIE perspective


Resonances
+
Scaled DIS background

DIS
Low-W AGKY

DIS
Linear transition
to PYTHA 8

DIS
PYTHA 8

1.7 GeV/c²

2.3 GeV/c²

3 GeV/c²

GENIE's Shallow-Inelastic Scattering model

RES

- Rein-Sehgal or Bergher-Sehgal are the starting point
- Added additional resonances
- Dipole Parameterization

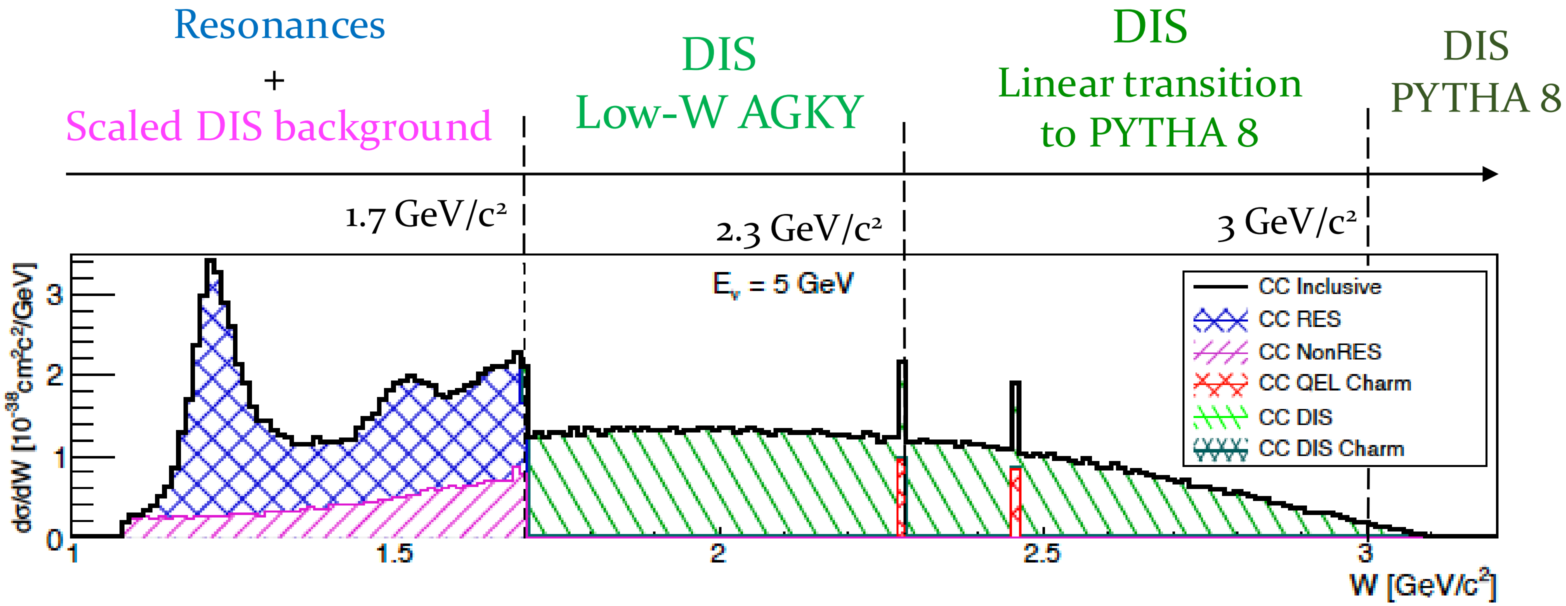
Non-resonant bkg

- Duality-based approach
- Scaled Bodek-Yang model
- Scaling factors depend on initial state and hadron multiplicity
- Coupled to low-W AGKY model

DIS

- Bodek-Yang model
- Cross-section calculation at partonic level
- AGKY hadronization model

The Shallow Inelastic Scattering Region GENIE perspective



Resonances
+
Scaled DIS background

DIS
Low-W AGKY

DIS
Linear transition
to PYTHA 8

DIS
PYTHA 8

1.7 GeV/c²

2.3 GeV/c²

3 GeV/c²

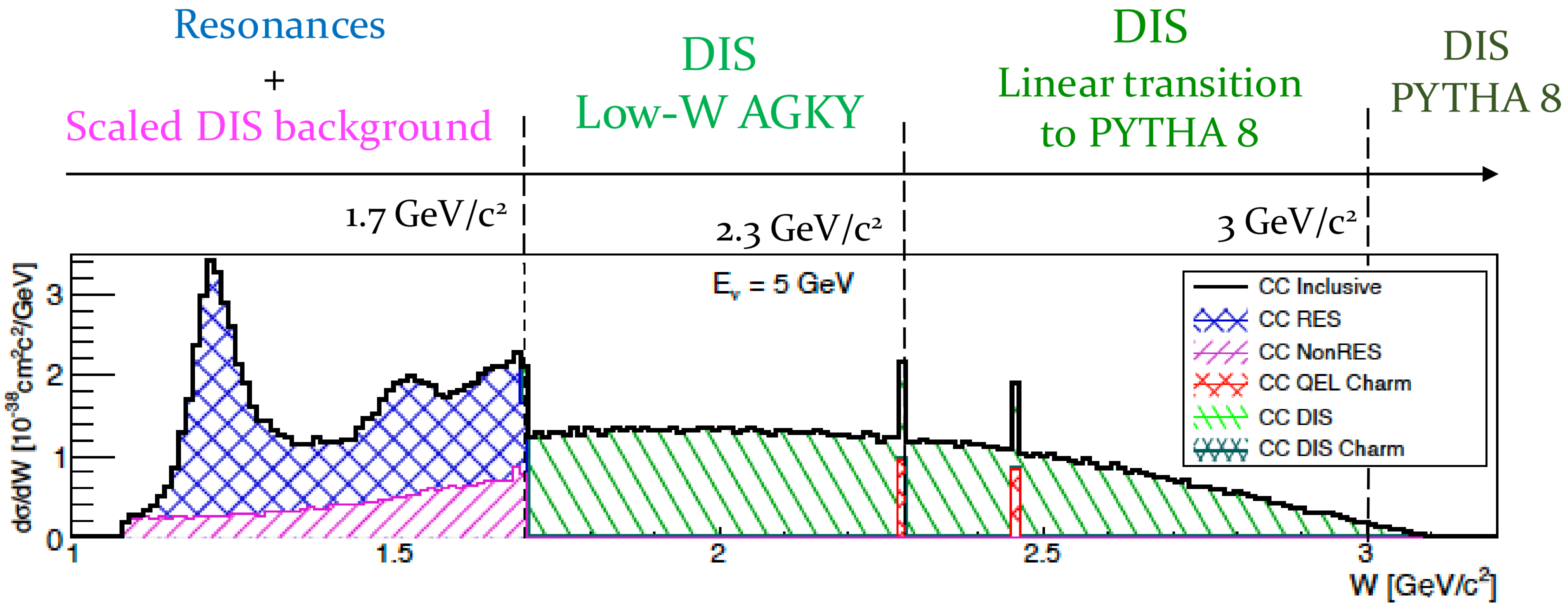
$$\frac{d^2\sigma^{inel}}{dQ^2dW} = \begin{cases} \dfrac{d^2\sigma^{RES}}{dQ^2dW} + \dfrac{d^2\sigma^{Non-RES}}{dQ^2dW} \; for\ W < \boldsymbol{W_{cut}} \\ \dfrac{d^2\sigma^{DIS}}{dQ^2dW} \; for\ W \geq \boldsymbol{W_{cut}} \end{cases}$$

Free parameters

RES is modelled with Rein-Sehgal or Berger-Sehgal models

- Resonances are added coherently
- Not full kinematical models
- RES models do not account for NRB

The Shallow Inelastic Scattering Region GENIE perspective



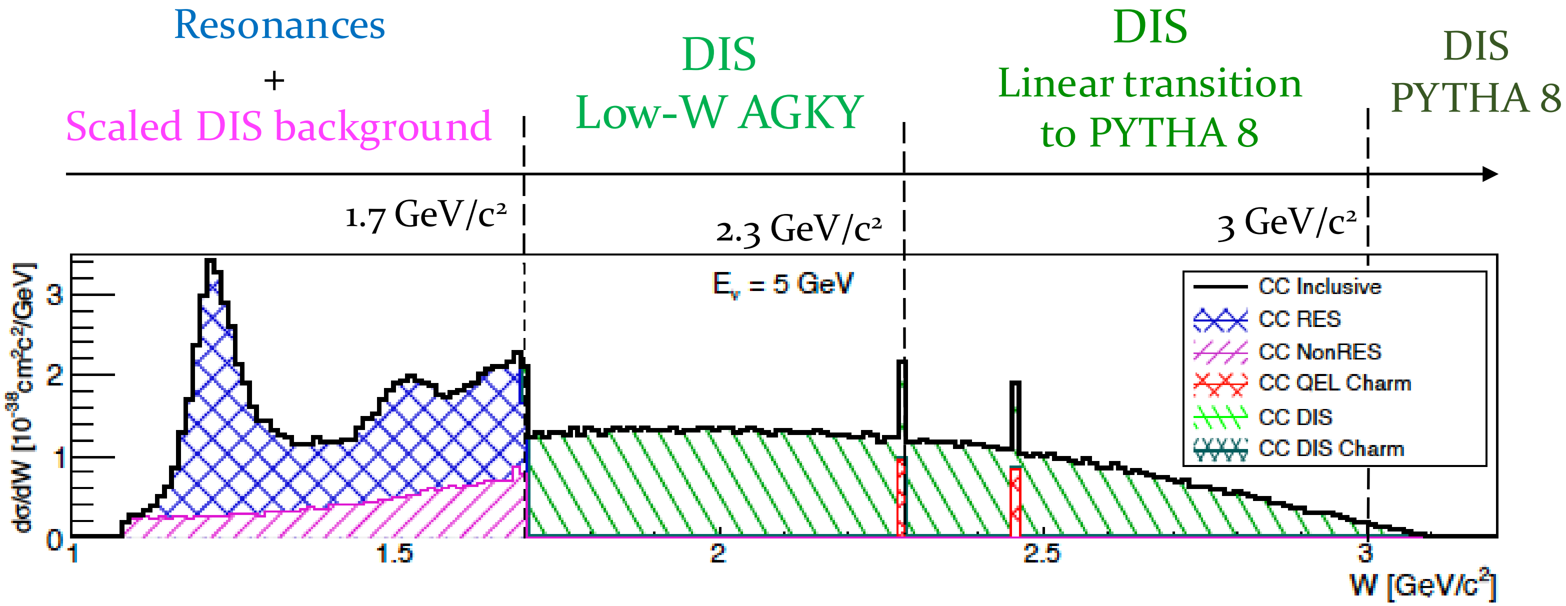
$$\frac{d^2\sigma^{inel}}{dQ^2dW} = \begin{cases} \dfrac{d^2\sigma^{RES}}{dQ^2dW} + \dfrac{d^2\sigma^{Non-RES}}{dQ^2dW} & for\ W < W_{cut} \\ \dfrac{d^2\sigma^{DIS}}{dQ^2dW} & for\ W \geq W_{cut} \end{cases}$$

Free parameters

RES is modelled with Rein-Sehgal or Berger-Sehgal models

- Resonances are added coherently
- Not full kinematical models
- RES models do not account for NRB

The Shallow Inelastic Scattering Region GENIE perspective


Resonances
+
Scaled DIS background

DIS
Low-W AGKY

DIS
Linear transition
to PYTHA 8

DIS
PYTHA 8

1.7 GeV/c²

2.3 GeV/c²

3 GeV/c²

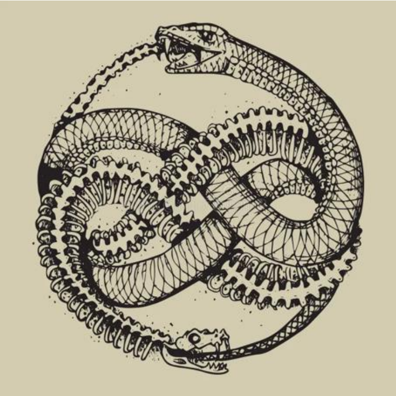
$$\frac{d^2\sigma^{Non-RES}}{dQ^2dW} = \frac{d^2\sigma^{DIS}}{dQ^2dW}\cdot\Theta(\boldsymbol{W_{cut}}-W)\cdot\sum_m \boldsymbol{f_m}(Q^2,W)$$

Free parameters

- NRB modelled with Bodek and Yang extrapolated at $W < W_{cut}$
- f_m parameters couple with the AGKY hadronization model

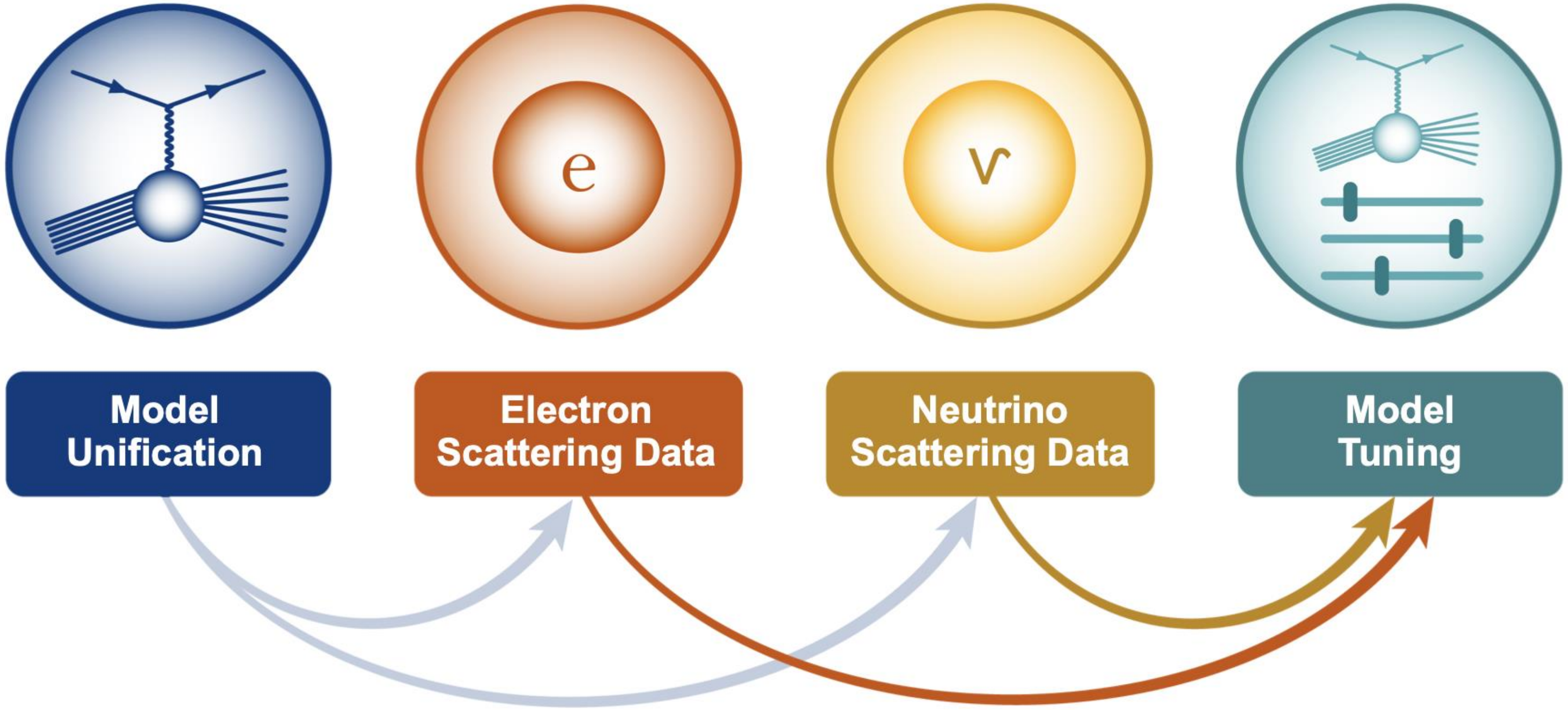
The Shallow Inelastic Scattering Region
GENIE perspective

- **In GENIE, RES, SIS, DIS and Hadronization models must be tuned altogether**



  - RES + SIS will determine the total cross-section at the SIS region
  - The DIS model determines the SIS cross-section (before scaling)
  - Hadronization determines the pion multiplicity of non-RES and DIS events
  - The event multiplicity determines the overall scaling of SIS
- **Pion production systematics already important at the free nucleon level**
  - Electron-scattering inclusive data widely used
  - Neutrino-scattering – bubble chamber data

Towards a global tune



Model Unification

Electron Scattering Data

Neutrino Scattering Data

Model Tuning

Towards a global tune

Model unification

- Ideally, models have clear V-A separation, with specific parameters
  - Not available in all event generators

e

Tune your generator against eA data

- High precision constraints on V-parameters
- Nuclear models – Spectral functions
- Final-State interaction models

ν

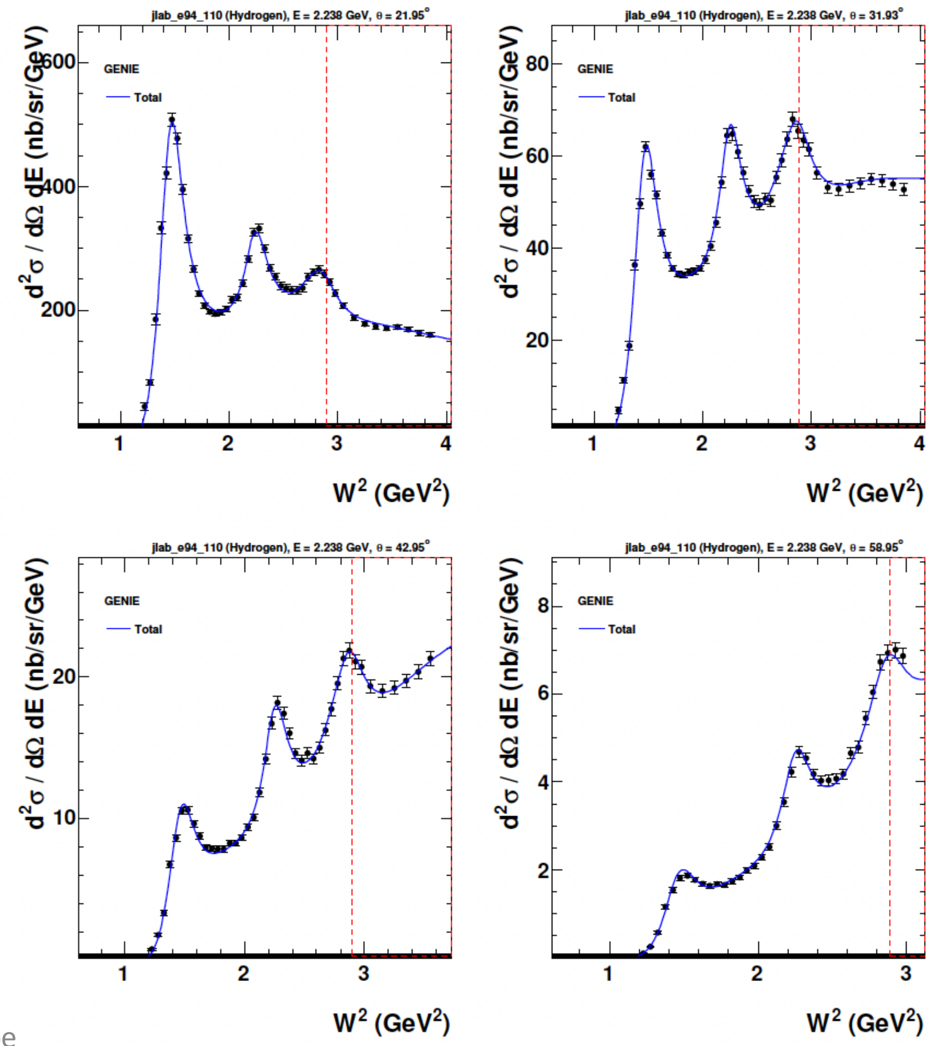
Propagate tune results to neutrino tune

- Results from the electron tune can be imposed as priors to avoid bias
- Constrain FSI and nuclear model with electron data – break degeneracies
- Ideally, also axial part, but this might be tricky for some models

Tuning of the Shallow-Scattering Inelastic region Datasets available – electron scattering

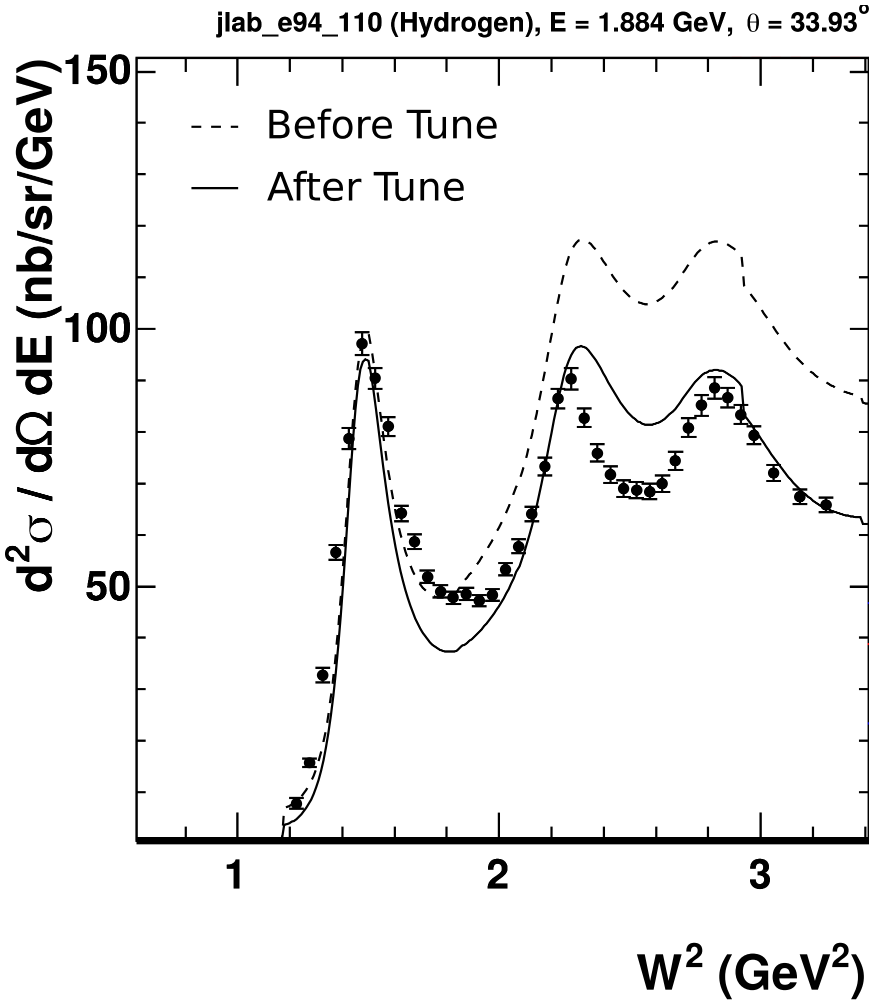
- Inclusive data from JLAB and SLAC
- as a function of W^2
  - Computed using e^- information
- For different beam energies and angles
- on hydrogen and deuterium targets

(*) Data is compared against Boosted-Christy prediction



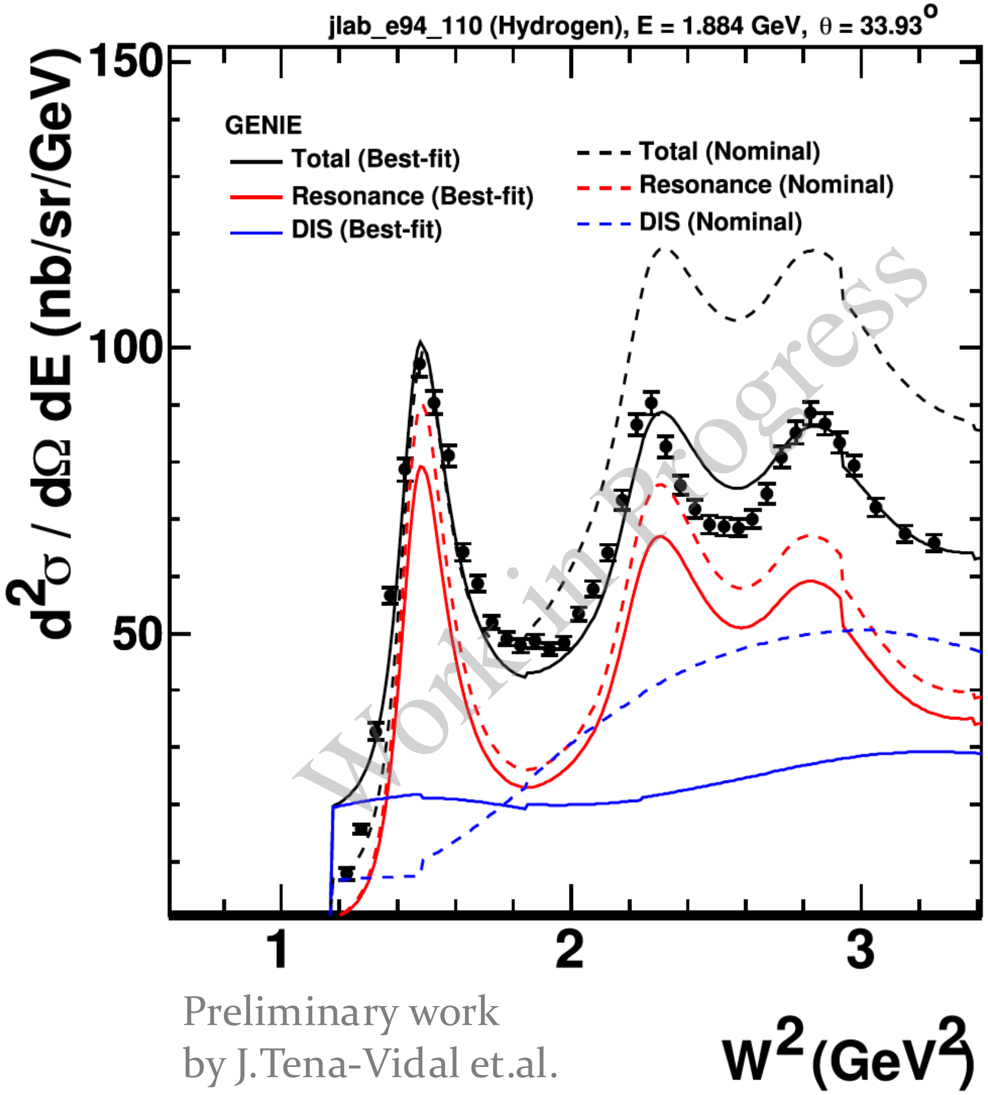
Tuning of the Shallow-Scattering Inelastic region Inclusive electron scattering tunes

- **Can tune directly cross section models**
- **e-N inclusive** can be calculated directly using GENIE cross-section algorithms
  - Known beam energy, probe and target type (nucleon)
  - No need to generate events
  - **Using Berger-Sehgal model but same concept applies to all models**



Tuning of the Shallow-Scattering Inelastic region with *e*-N data

- Non-resonant background parameters never tuned to electron data
  - Double counting is guaranteed
  - Model overpredicts data above the delta region
- Excellent inclusive data available from JLAB and SLAC
  - Fine W binning allows tuning of RES/SIS/DIS
  - Delta peak constrains RES Scaling
  - Multiplicity 2 and 3 non-resonant parameters can be constrained using fine W binning

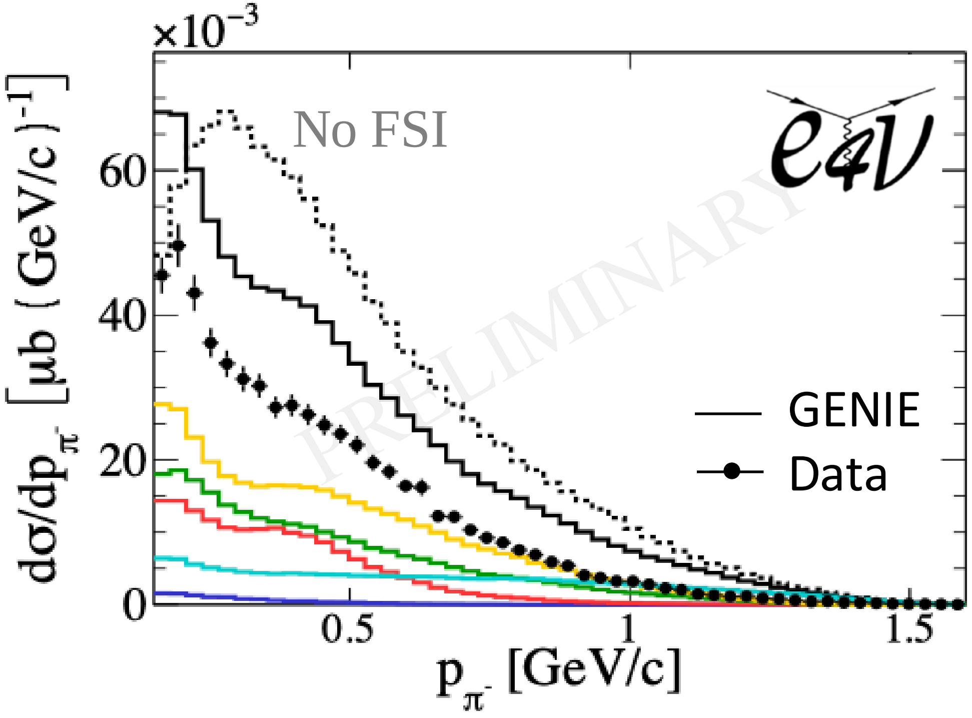


Preliminary work
by J.Tena-Vidal et.al.

Propagating pion production uncertainties from electron tunes

- All generators use an ad-hoc non-RES production model
  - Data driven parameters obtained from tunes to data
  - Cannot easily apply electron constraints to neutrinos
- **But excellent free nucleon description isolates nuclear effects!**
  - Currently over-predicting data on ^{12}C

C(e,e'1p1π^-) @ 2.2 GeV



see Adi Ashkenazi's talk

Neutrino data – bubble chamber experiments

- Bubble chamber experiments provided with the first flux-unfolded integrated cross-section measurements
- Mostly inclusive measurements, few exclusive (one-, two-pion, QEL..)
- Measurements as a function of E_ν, Q2…
  - Big bias on neutrino energy
- Statistically limited, ~ 100 events
- Poor neutrino flux knowledge
- MC-based data-corrections
  - Model dependent cuts
- Missing systematic uncertainties
  - Not quantified by experiments
  - Large normalization uncertainties lead to inconsistencies between experiments
    - Re-analysis of ANL/BNL data [PhysRevD.90.112017]

Limitations of historical neutrino bubble chamber data

- Bubble chamber experiments provided with the first flux-unfolded integrated cross-section measurements
- Mostly inclusive measurements, few exclusive (one-, two-pion, QEL..)
- Measurements as a function of E_ν, Q2…
  - Big bias on neutrino energy
- Statistically limited, ~ 100 events
- Poor neutrino flux knowledge
- MC-based data-corrections
  - Model dependent cuts
- Missing systematic uncertainties
  - Not quantified by experiments
  - Large normalization uncertainties lead to inconsistencies between experiments
    - Re-analysis of ANL/BNL data [PhysRevD.90.112017]

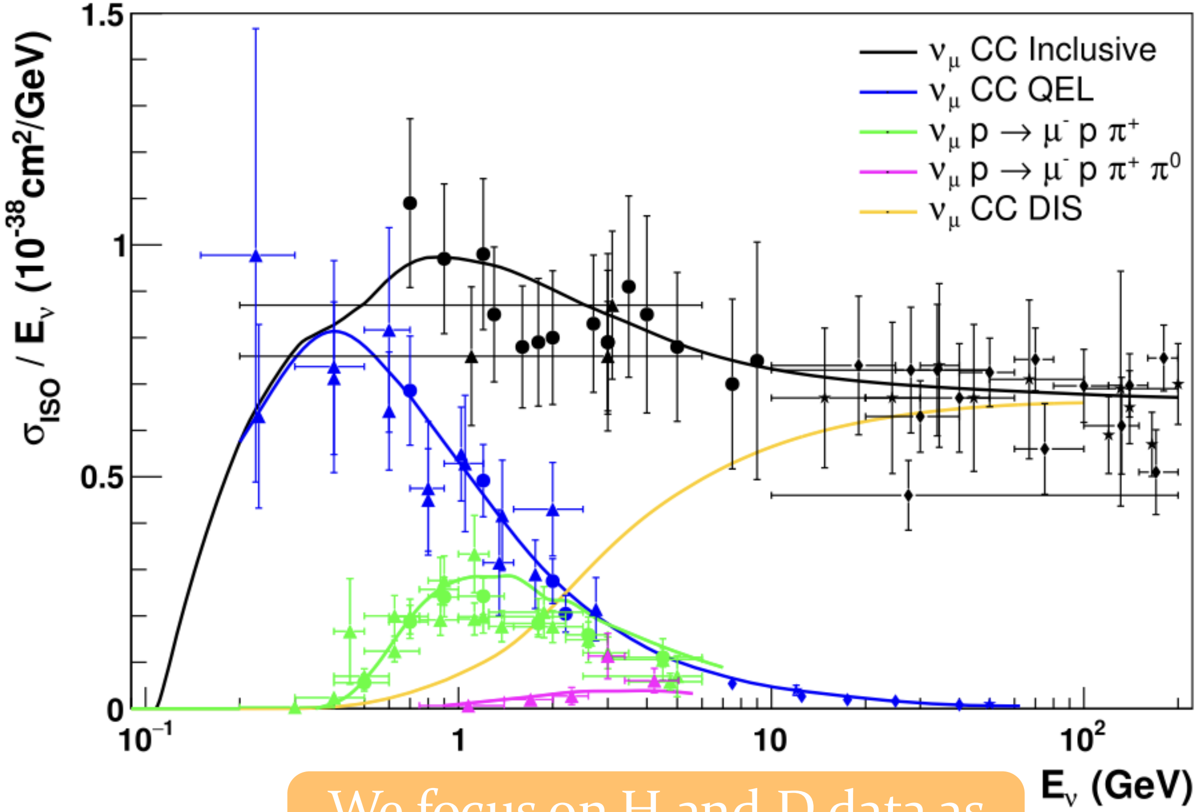
Many reasons to not use these datasets…

… only data available on hydrogen and deuterium for neutrinos!

Neutrino bubble chamber datasets

PhysRevD.104.072009

- ν_μ and anti- ν_μ CC inclusive
- ν_μ and anti- ν_μ CC QEL
- ν_μ and anti- ν_μ CC single-pion
  - $\nu_\mu + n \to \mu^- + n + \pi^+$
  - $\nu_\mu + p \to \mu^- + p + \pi^+$
  - $\nu_\mu + n \to \mu^- + p + \pi^0$
  - $\bar{\nu}_\mu + p \to \mu^+ + p + \pi^-$
  - $\bar{\nu}_\mu + n \to \mu^+ + n + \pi^-$
- ν_μ and anti- ν_μ CC two-pion
  - $\nu_\mu + p \to \mu^- + n + 2\pi^+$
  - $\nu_\mu + p \to \mu^- + p + \pi^+ + \pi^0$
  - $\nu_\mu + p \to \mu^- + n + \pi^+ + \pi^-$

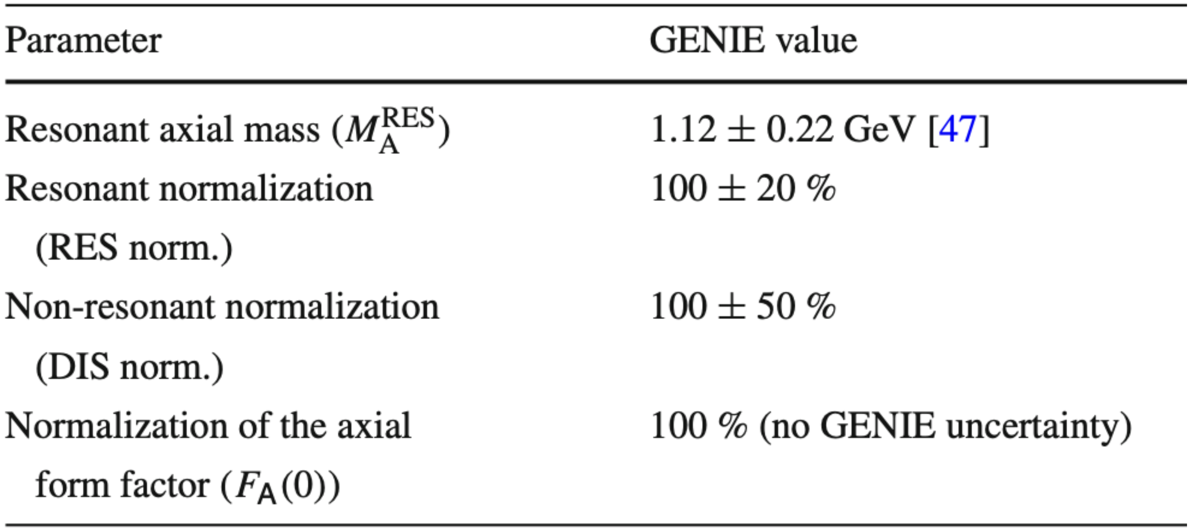


We focus on H and D data as a function of E_ν

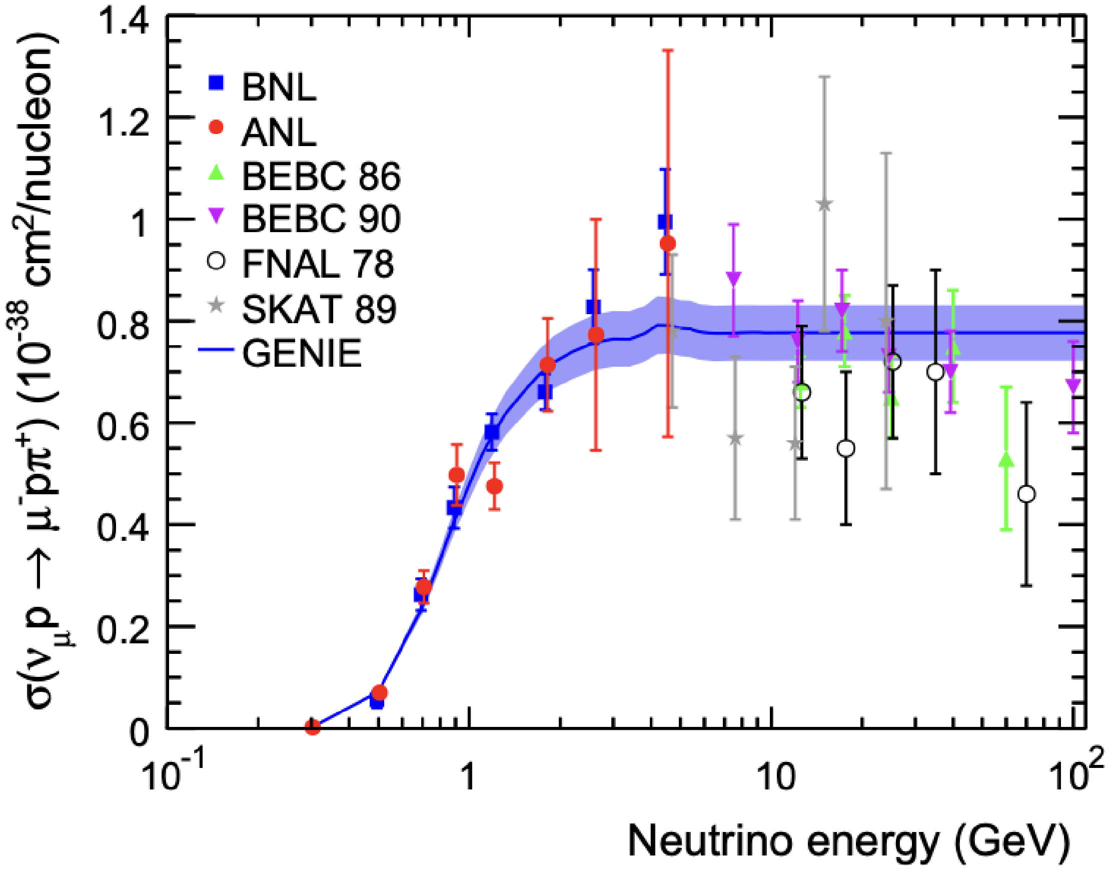
Constraining GENIE model of neutrino induced pion production using reanalized bubble chamber data

Work by P. Rodrigues et. al. exploits Reweight

- Fit with single pion production data only
- Data as a function of E_ν and Q^2



Parameter	GENIE value
Resonant axial mass (M_A^{RES})	1.12 ± 0.22 GeV [47]
Resonant normalization (RES norm.)	100 ± 20 %
Non-resonant normalization (DIS norm.)	100 ± 50 %
Normalization of the axial form factor ($F_A(0)$)	100 % (no GENIE uncertainty)



(a) $\nu_\mu p \rightarrow \mu^- p \pi^+$

Tuning the Shallow-Scattering Inelastic region
Parameters of interest

GENIE tune uses Professor (*PhysRevD.104.072009*)

RES model parameters:

- M_A^{RES}: global fit result applied as prior - $M_A^{RES} = 1.014 \pm 0.014\ GeV$
- S_{RES}: overall scaling factor for RES cross-section

NRB model parameters:

- W_{cut} to determine the end of the SIS region
- R_m parameters for proton and neutron, multiplicity 2 and 3
- *Simplification:* we neglect the AGKY low-W parameters

DIS model parameters:

- S_{DIS}: overall scaling factor for DIS cross-section
- Prior of 1±0.5 to preserve agreement with high E data (>100GeV)

QEL model parameters:

- M_A^{QEL}: global fit result applied as prior - $M_A^{RES} = 1.12 \pm 0.03\ GeV$

Normalization uncertainty:

- Nuisance parameters per experiment to account for missing flux normalization uncertainties

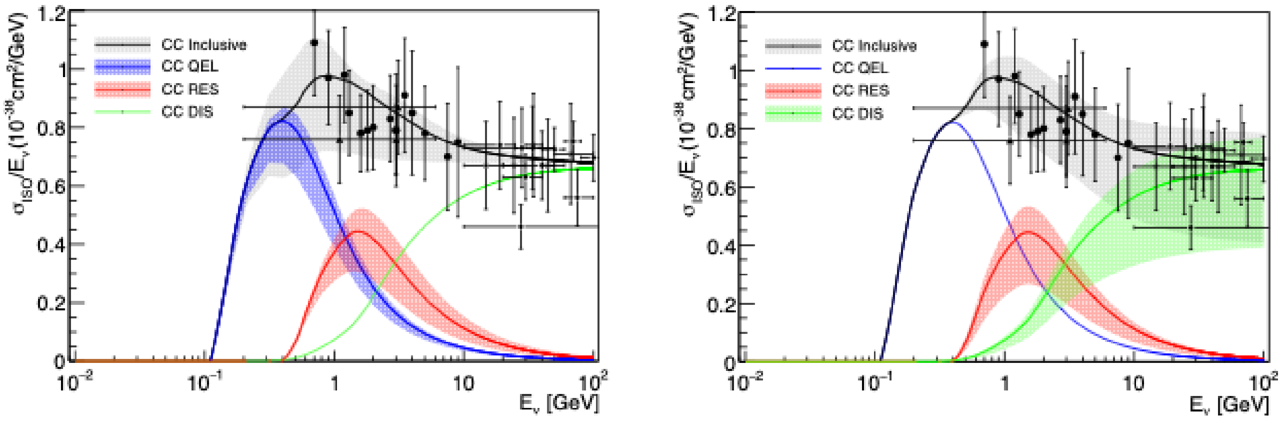
Tuning the Shallow-Scattering Inelastic region Parameters of interest

PhysRevD.104.072009

Your parameter choice might lead to a degenerate result

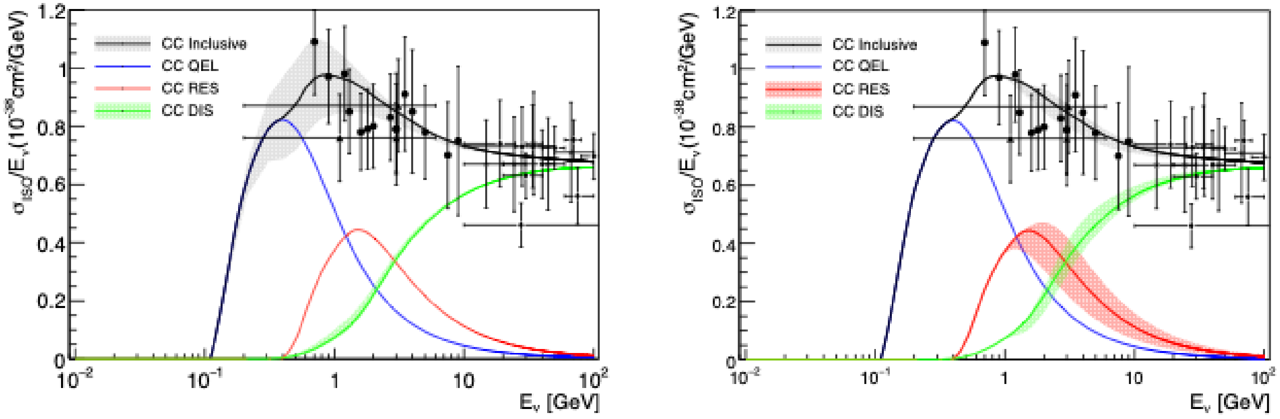
Ways to address it:

- Include in the tune additional data, i.e $\sigma(Q^2)$
- Priors from previous global analysis/tunes



(a) $M_A^{\rm RES}$ and $M_A^{\rm QE}$ impact.

(b) $S_{\rm RES}$ and $S_{\rm DIS}$ impact.



(c) $R_{\nu p}^{\rm CC1\pi}$, $R_{\nu p}^{\rm CC2\pi}$, $R_{\nu n}^{\rm CC1\pi}$ and $R_{\nu n}^{\rm CC2\pi}$ impact.

(d) $W_{\rm cut}$ impact.

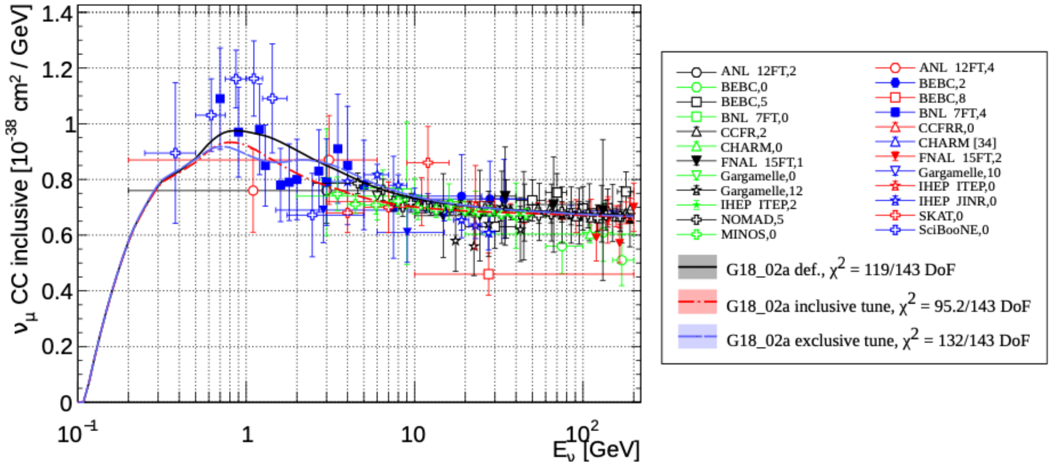
Inconsistencies between datasets

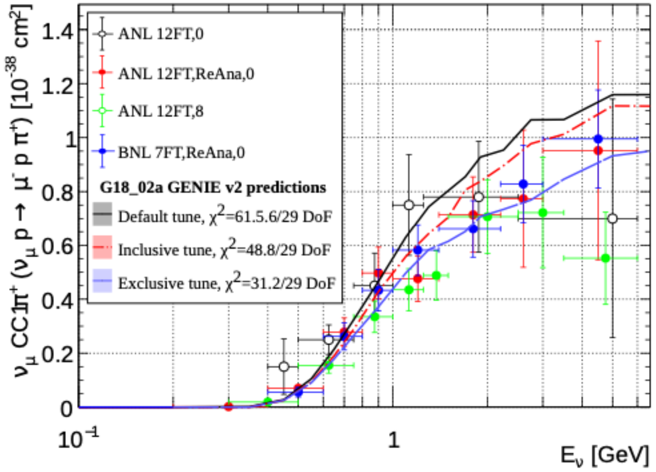
Free-nucleon tune example:
Partial tune to inclusive data has opposite behavior to exclusive tune

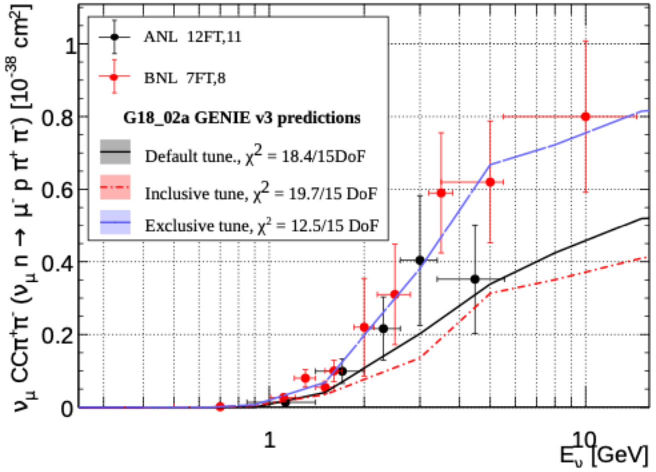
Consequence of the incorrect flux normalization used in the data analysis

Approach:

- Added QEL data
  - Well known $\sigma_{\nu N}^{QEL}(E_\nu)$
- Nuisance parameters



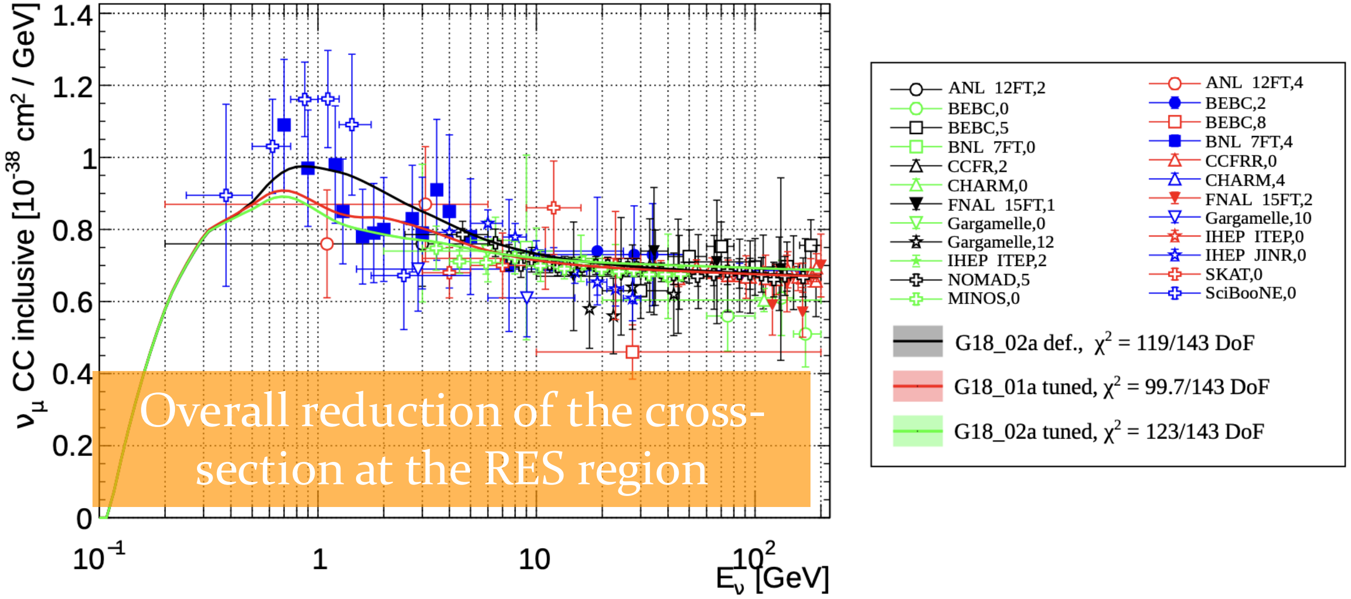




Tune Results

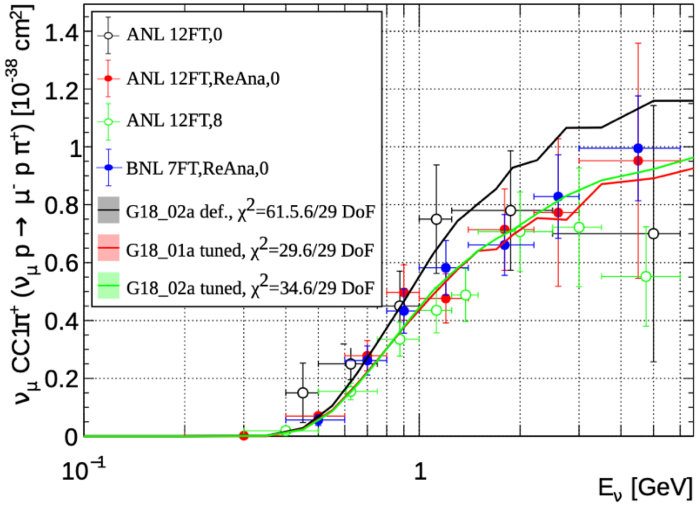
PhysRevD.104.072009

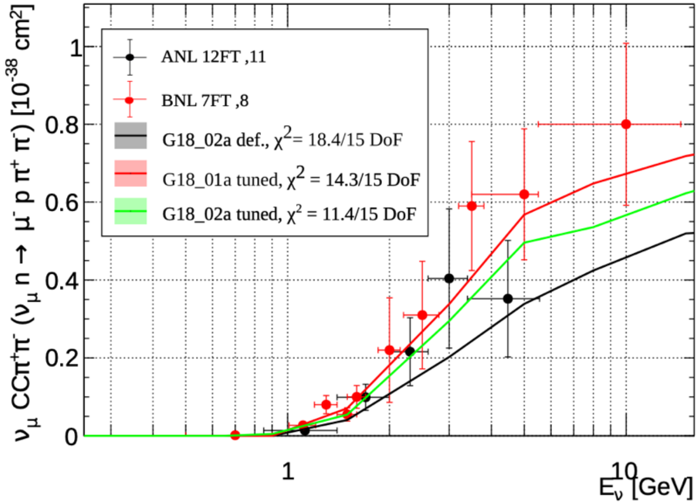
Parameter	Default	G18_02a
S_{RES}	1.00	0.84±0.03
S_{DIS}	1.032	1.06±0.01
$R_{\nu p}^{CC1\pi}$	0.10	0.008
$R_{\nu n}^{CC1\pi}$	0.30	0.03±0.01
$R_{\nu p}^{CC2\pi}$	1.00	0.94±0.08
$R_{\nu n}^{CC2\pi}$	1.00	2.3±0.1
M_A^{QEL}	0.999	1.00±0.013
M_A^{RES}	1.12	1.09±0.014
W_{cut}	1.7	1.81
$\chi^2/157DoF$		1.64



Overall reduction of the cross-section at the RES region

(a) Comparison of ν_μ CC Inclusive cross-section data against against the *default* and tuned CMC.





Supression of 1π production cross-section

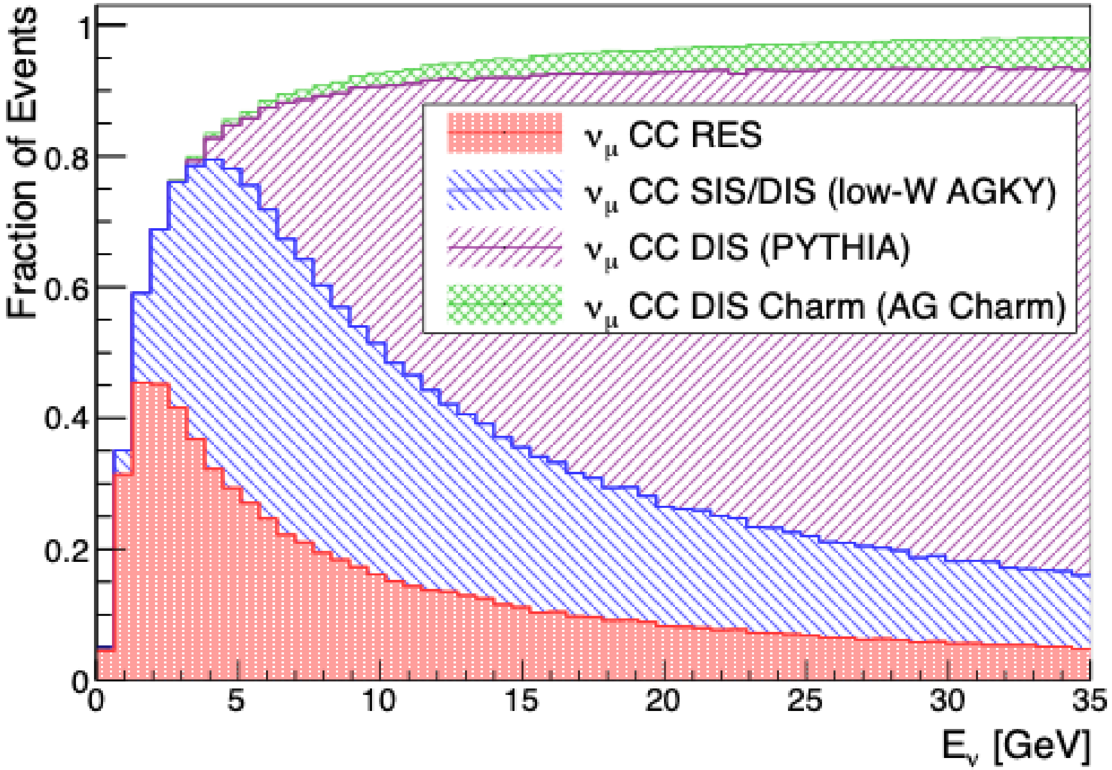
Enhancement of 2π production cross-section

Tuning the **AGKY Hadronization** model
Tuning non-reweightable models

Hadronization models provide with final-state hadrons properties after a (SIS) DIS interaction

Crucial for experiments:

- Experiments like DUNE expect a **large fraction of SIS and DIS events ∽ 30%**
- It determines the number of hadrons, hadronic shower shape, EM fraction of hadronic shower, hadronic shower energy reconstruction…



Tuning the **AGKY Hadronization** model
Tuning non-reweightable models

Most neutrino event generators use (AGKY+) PYTHIA

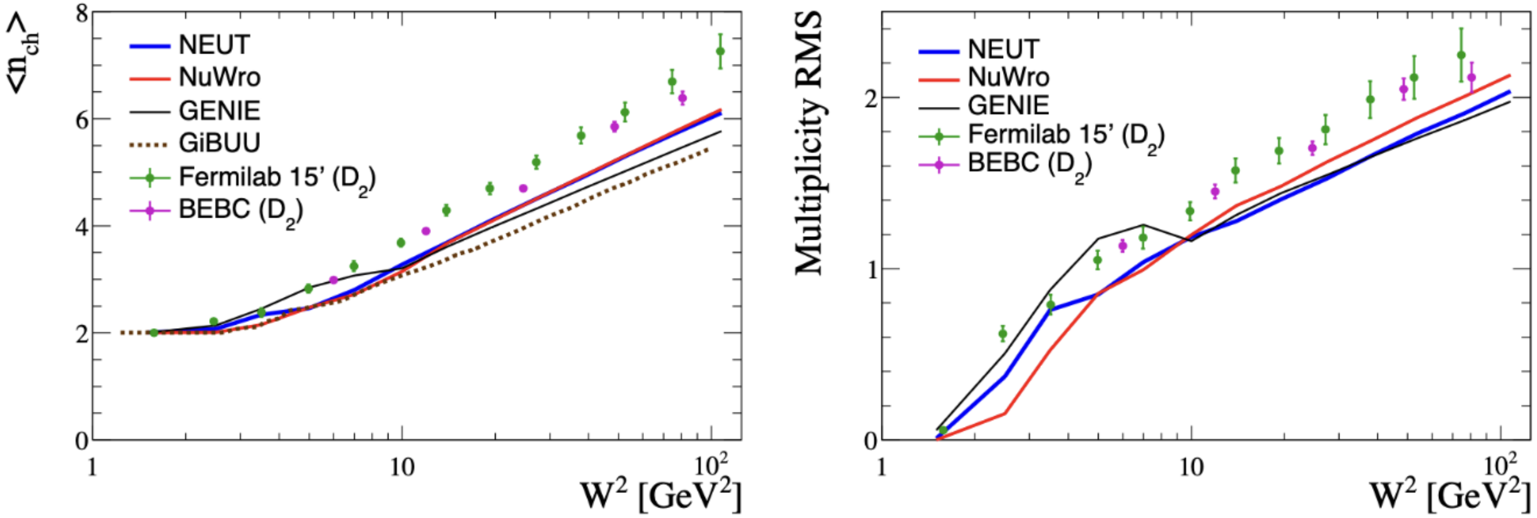


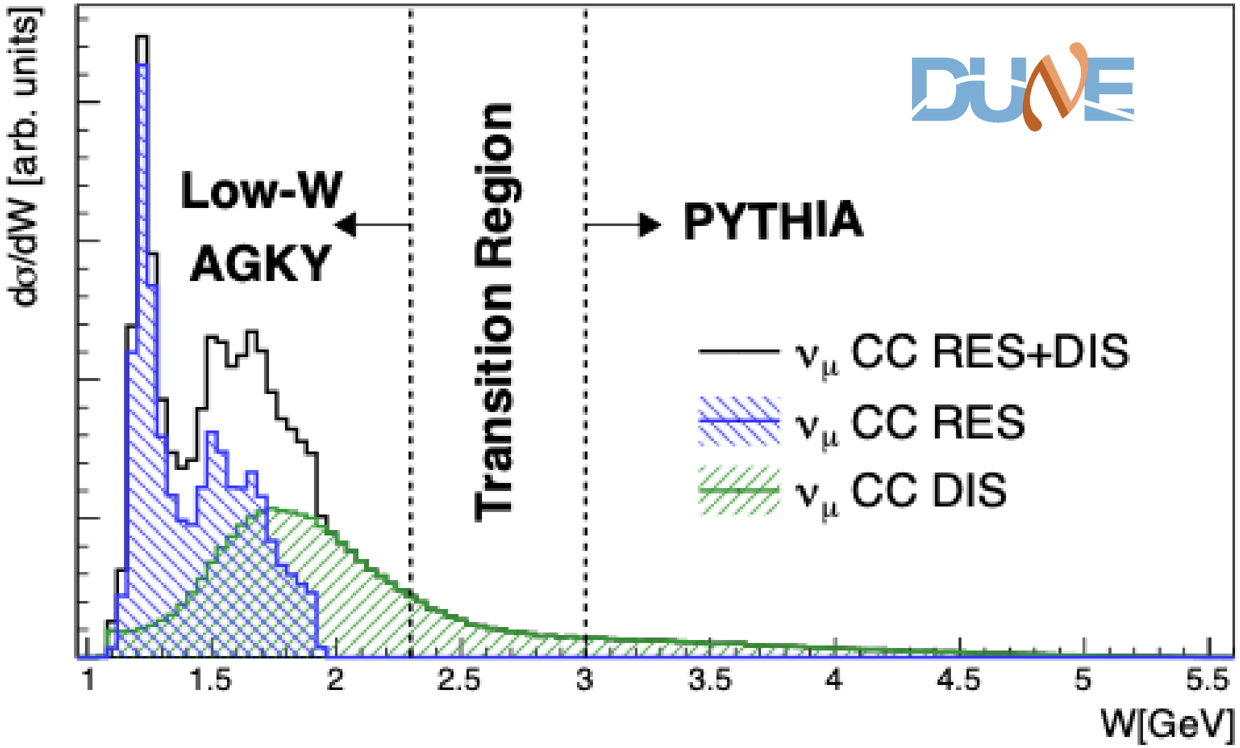
Fig. 2. Mean value (left plot) and dispersion (right plot) as a function of the square of the hadronic invariant mass W^2 of the number of charged hadrons produced in the interactions of neutrinos with protons.

By C. Bronner

Tuning the **AGKY Hadronization** model
Tuning non-reweightable models

Modeling:

- At low-W, model is anchored to bubble chamber data
- Linear transition to PYTHIA
- PYTHIA for W>3 GeV
- In GENIE it is also used to determine the SIS pion multiplicity at the SIS region

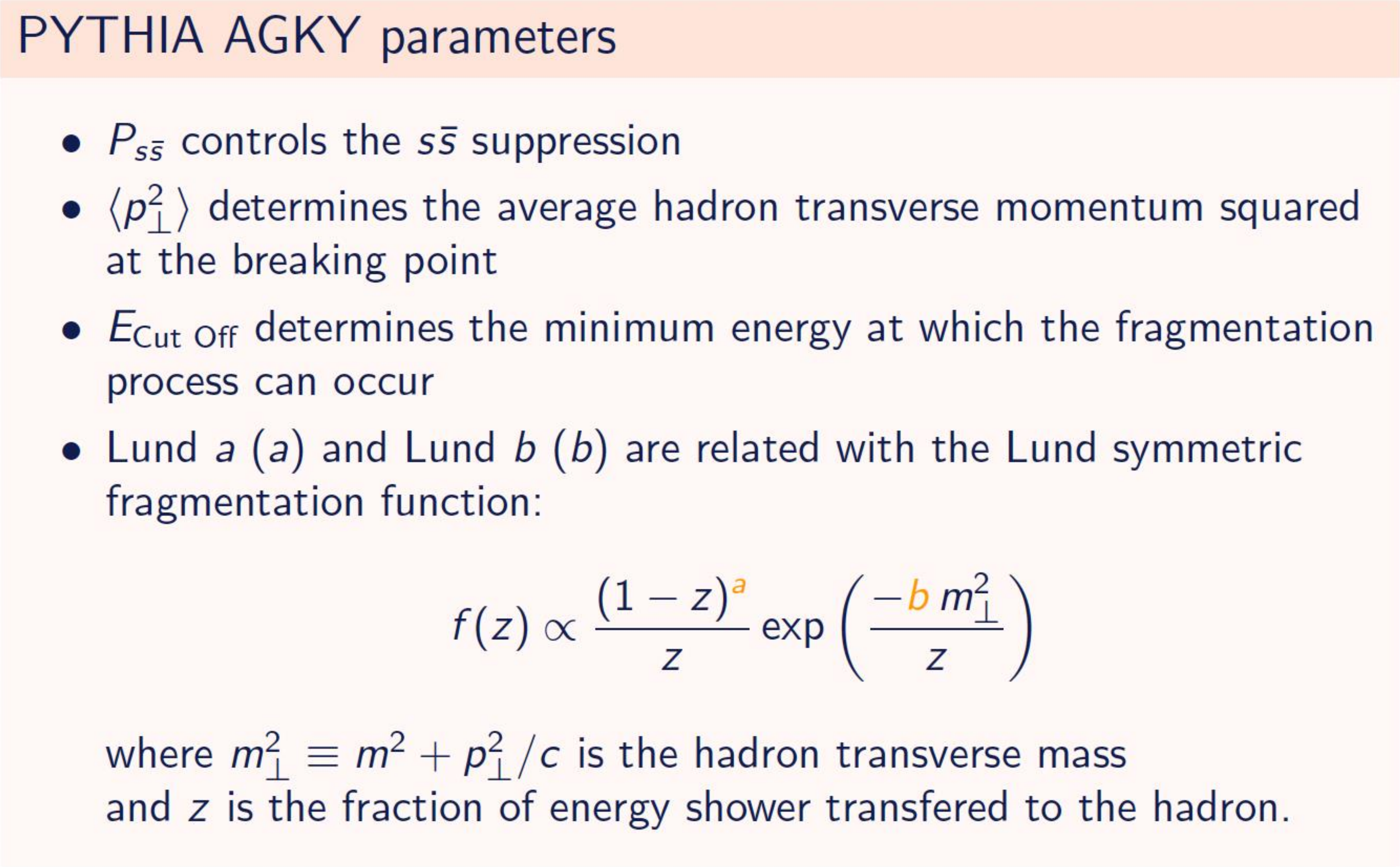


PYTHIA MC

- Based on the Lund String Fragmentation function
- The Generator has many parameters available
- The default parameters are tuned to high energy pp and e^-e^+ experiments
- In GENIE, not all are directly available in the configuration files
  - But adding more is very easy to do
- Tuned with Professor

https://pythia.org

PYTHIA parameters relevant for charged multiplicity tuning



PYTHIA AGKY parameters

- $P_{s\bar{s}}$ controls the $s\bar{s}$ suppression
- $\langle p_\perp^2 \rangle$ determines the average hadron transverse momentum squared at the breaking point
- $E_{\text{Cut Off}}$ determines the minimum energy at which the fragmentation process can occur
- Lund a (a) and Lund b (b) are related with the Lund symmetric fragmentation function:

$$f(z) \propto \frac{(1-z)^a}{z} \exp\left(\frac{-b\, m_\perp^2}{z}\right)$$

where $m_\perp^2 \equiv m^2 + p_\perp^2 / c$ is the hadron transverse mass
and z is the fraction of energy shower transfered to the hadron.

Low-W AGKY

- Data-driven model aimed at describing showers below W<3 GeV
  - where PYTHIA is no longer valid
- Most crucial for accelerator neutrino experiments

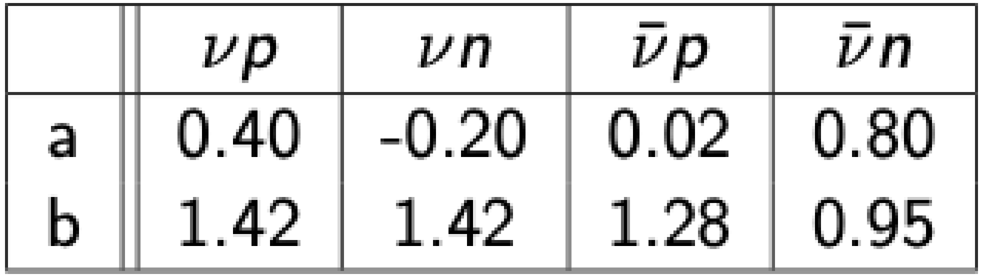
  Aims to describe:
  - Averaged charged multiplicity
  - Averaged neutral multiplicity
  - Total multiplicity
  - Baryon multiplicity
  - Shower kinematics
- **Anchored to bubble chamber data**

Low-W AGKY: How many hadrons?

- **We use an empirical law extracted from data**

$$< n_{ch} > = a + b \cdot \ln(W^2/GeV^2)$$

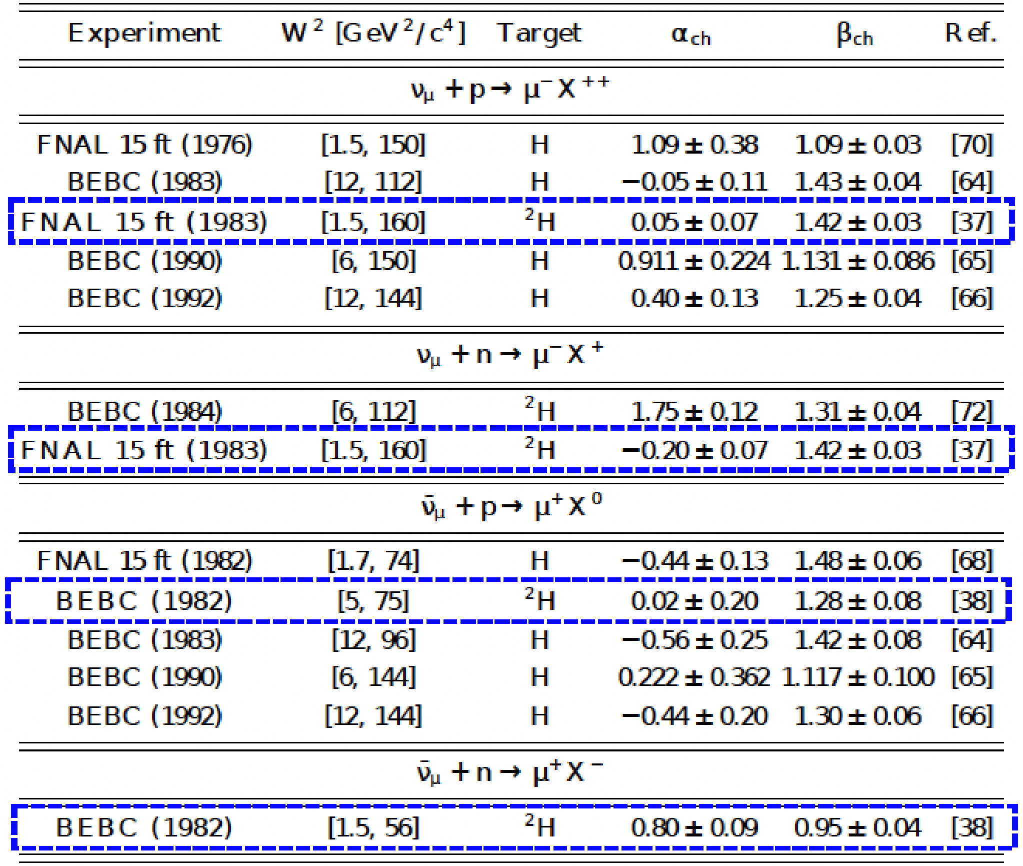
Default GENIE values:



	νp	νn	$\bar{\nu} p$	$\bar{\nu} n$
a	0.40	-0.20	0.02	0.80
b	1.42	1.42	1.28	0.95

Not really coming from a consistent fit to data

- Extracted from Deuterium fits only
- From different datasets (not compatible)

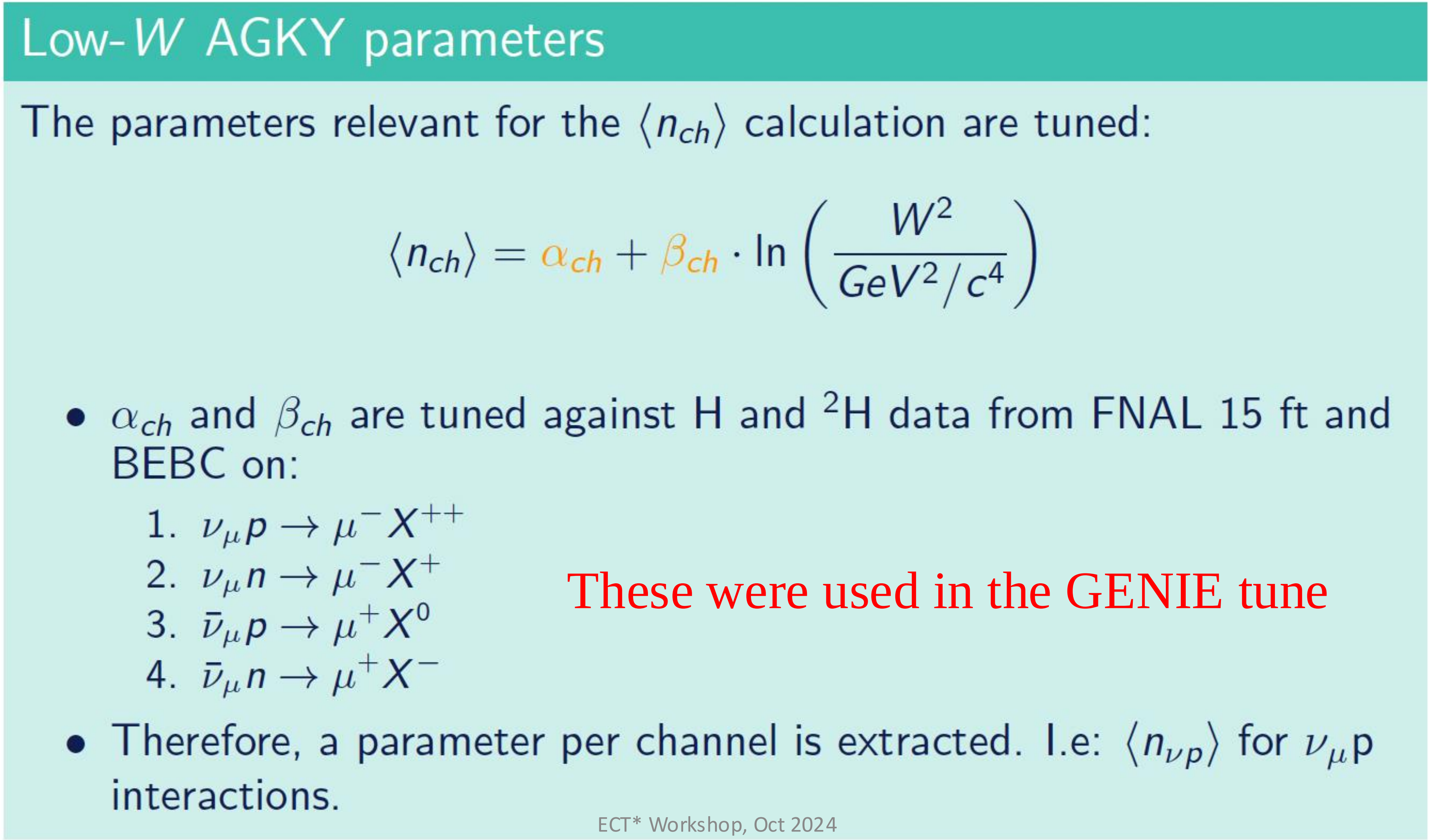


Experiment	W^2 [GeV^2/c^4]	Target	α_{ch}	β_{ch}	Ref.
$\nu_\mu + p \rightarrow \mu^- X^{++}$					
FNAL 15 ft (1976)	[1.5, 150]	H	1.09 ± 0.38	1.09 ± 0.03	[70]
BEBC (1983)	[12, 112]	H	-0.05 ± 0.11	1.43 ± 0.04	[64]
FNAL 15 ft (1983)	[1.5, 160]	^{2}H	0.05 ± 0.07	1.42 ± 0.03	[37]
BEBC (1990)	[6, 150]	H	0.911 ± 0.224	1.131 ± 0.086	[65]
BEBC (1992)	[12, 144]	H	0.40 ± 0.13	1.25 ± 0.04	[66]
$\nu_\mu + n \rightarrow \mu^- X^{+}$					
BEBC (1984)	[6, 112]	^{2}H	1.75 ± 0.12	1.31 ± 0.04	[72]
FNAL 15 ft (1983)	[1.5, 160]	^{2}H	-0.20 ± 0.07	1.42 ± 0.03	[37]
$\bar{\nu}_\mu + p \rightarrow \mu^+ X^{0}$					
FNAL 15 ft (1982)	[1.7, 74]	H	-0.44 ± 0.13	1.48 ± 0.06	[68]
BEBC (1982)	[5, 75]	^{2}H	0.02 ± 0.20	1.28 ± 0.08	[38]
BEBC (1983)	[12, 96]	H	-0.56 ± 0.25	1.42 ± 0.08	[64]
BEBC (1990)	[6, 144]	H	0.222 ± 0.362	1.117 ± 0.100	[65]
BEBC (1992)	[12, 144]	H	-0.44 ± 0.20	1.30 ± 0.06	[66]
$\bar{\nu}_\mu + n \rightarrow \mu^+ X^{-}$					
BEBC (1982)	[1.5, 56]	^{2}H	0.80 ± 0.09	0.95 ± 0.04	[38]

Default GENIE Parameters

References given in ArXiv:2106.05884.

Average multiplicity parameters



Low-W AGKY parameters

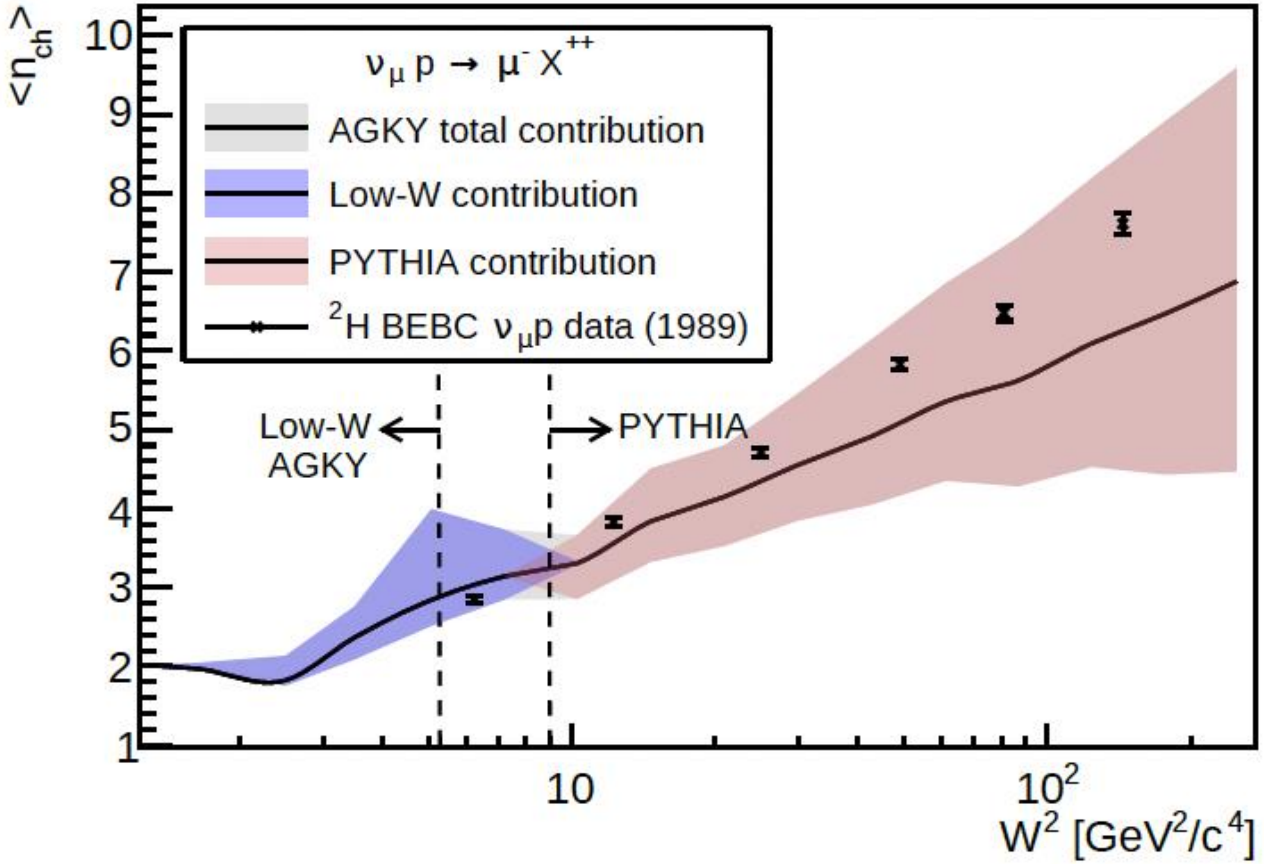
The parameters relevant for the $\langle n_{ch} \rangle$ calculation are tuned:

$$\langle n_{ch} \rangle = \alpha_{ch} + \beta_{ch} \cdot \ln \left(\frac{W^2}{GeV^2/c^4} \right)$$

- α_{ch} and β_{ch} are tuned against H and ^{2}H data from FNAL 15 ft and BEBC on:
    1. $\nu_\mu p \rightarrow \mu^- X^{++}$
    2. $\nu_\mu n \rightarrow \mu^- X^{+}$
    3. $\bar{\nu}_\mu p \rightarrow \mu^+ X^{0}$
    4. $\bar{\nu}_\mu n \rightarrow \mu^+ X^{-}$
- Therefore, a parameter per channel is extracted. I.e: $\langle n_{\nu p} \rangle$ for ν_μp interactions.

These were used in the GENIE tune

Effect of low-W and PYTHIA parameters



Relevant datasets

A total of 154 data points on $\langle n_{ch} \rangle (W^2)$ are available:

$\langle n_{ch} \rangle$ vs W^2 data from FNAL 15 ft and BEBC

Initial state	Target	W^2 range [GeV^2/c^4]	Year
$\nu_\mu p$	H	[1.5,160]	[1976, 1992]
$\nu_\mu n$	H and ^{2}H	[1.5,160]	[1983,1989]
$\bar{\nu}_\mu p$	H	[1.7,144]	[1981,1992]
$\bar{\nu}_\mu n$	H and ^{2}H	[1.5,56]	[1982,1989]

- The data is obtained from (anti)neutrino CC interactions
- The data used is from the 70's and 80's
- The same analysis requirements are implemented to the corresponding predictions
- The $\langle n_{ch} \rangle$ data is independent of the cross section
- Most of the datapoints have $W > 3$ GeV$/c^2$

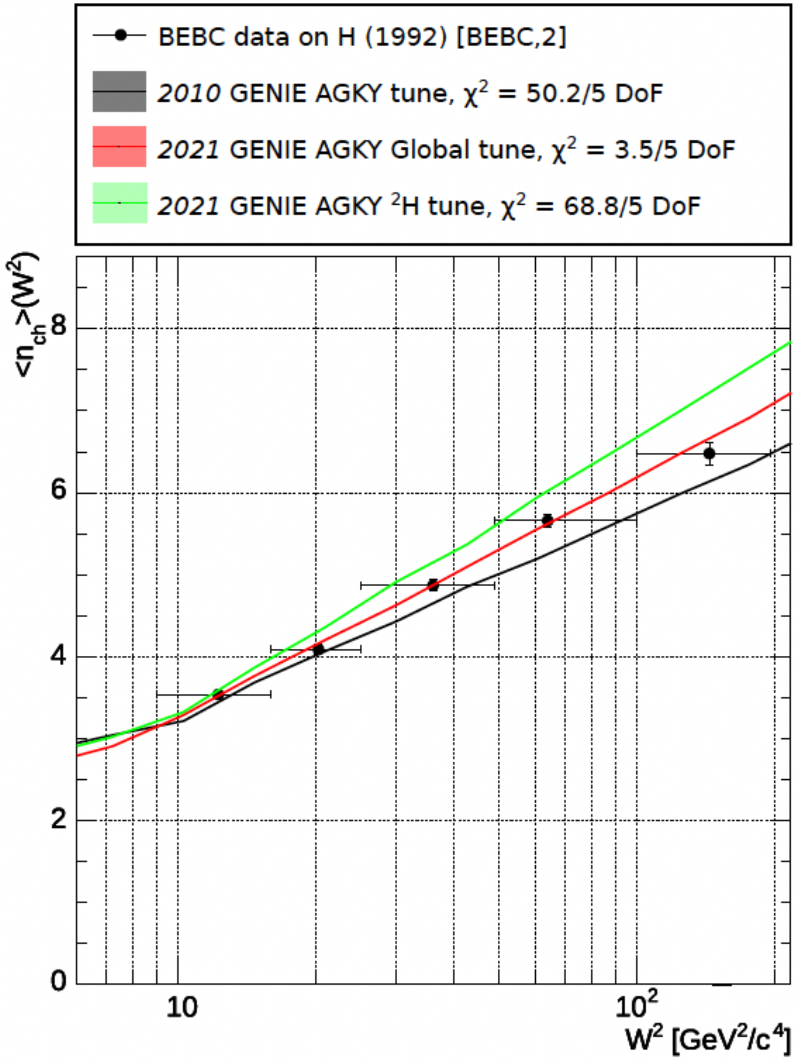
GENIE AGKY tune results

The main effect of the tunes is observed at the PYTHIA region:

→ The tunes increase $\langle n_{ch} \rangle$

This is a consequence of the **increase** in Lund *a* and Lund *b*

Parameter	*2010* GENIE	*2021* Global Fit	*2021* ^{2}H Fit
		Low-W empirical model	
$\alpha_{\nu p}$	0.40	1.1 ± 0.3	1.2 ± 0.4
$\alpha_{\nu n}$	-0.20	$1.75^{+0.14}_{-0.11}$	-0.58 ± 0.07
$\alpha_{\bar{\nu} p}$	0.02	$1.32^{+0.16}_{-0.14}$	1.9 ± 0.08
$\alpha_{\bar{\nu} n}$	0.80	1.11 ± 0.09	1.07 ± 0.3
$\beta_{\nu p}$	1.42	0.79 ± 0.15	0.9 ± 0.3
$\beta_{\nu n}$	1.42	0.5 ± 0.1	1.9 ± 0.3
$\beta_{\bar{\nu} p}$	1.28	0.8 ± 0.1	0.3 ± 0.1
$\beta_{\bar{\nu} n}$	0.95	$0.88^{+0.09}_{-0.08}$	0.9 ± 0.2
		PYTHIA	
$P_{s\bar{s}}$	0.30	0.27 ± 0.04	0.29 ± 0.05
$\langle p_\perp^2 \rangle$ [GeV^2/c^2]	0.44	0.43 ± 0.05	0.43 ± 0.04
E_{CutOff} [GeV]	0.20	0.30 ± 0.04	0.24 ± 0.05
Lund *a*	0.30	1.53 ± 0.13	1.85 ± 0.15
Lund *b* [c^4/GeV^2]	0.58	1.16 ± 0.09	1.0 ± 0.2
	$\chi^2 =$	87.9/62 DoF	29.5/32 DoF



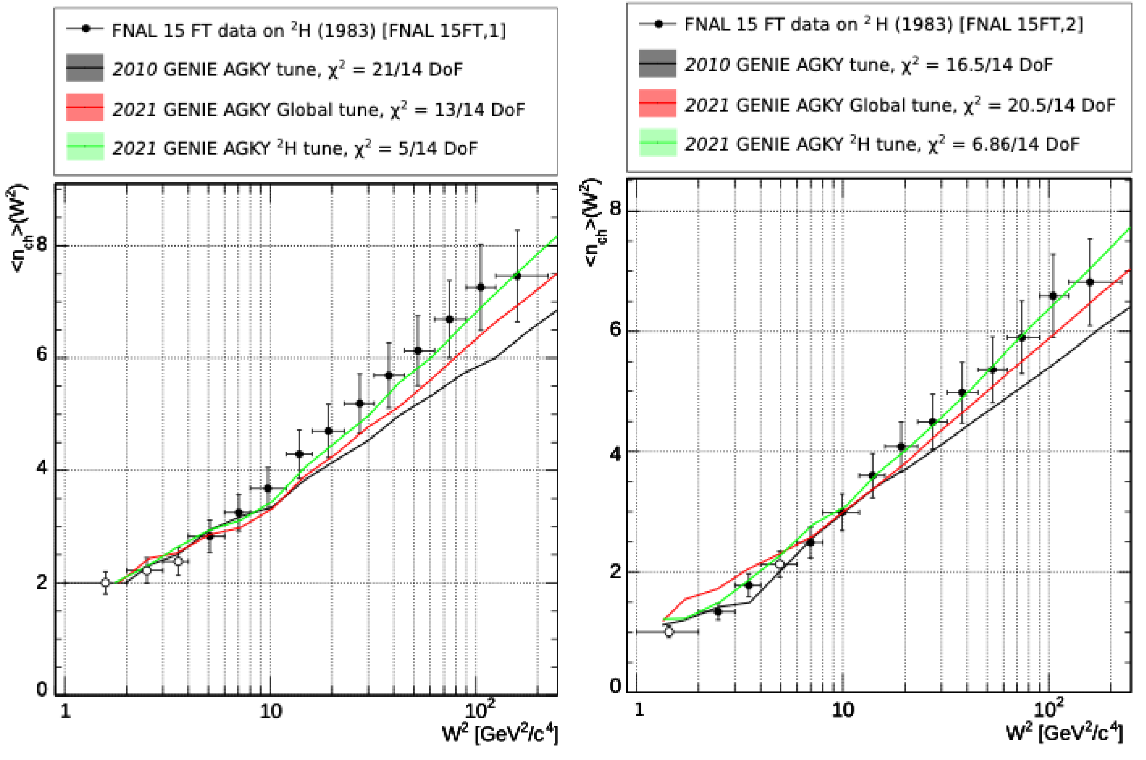
$\nu_\mu + p$ on hydrogen.

Tuning the AGKY Hadronization tune

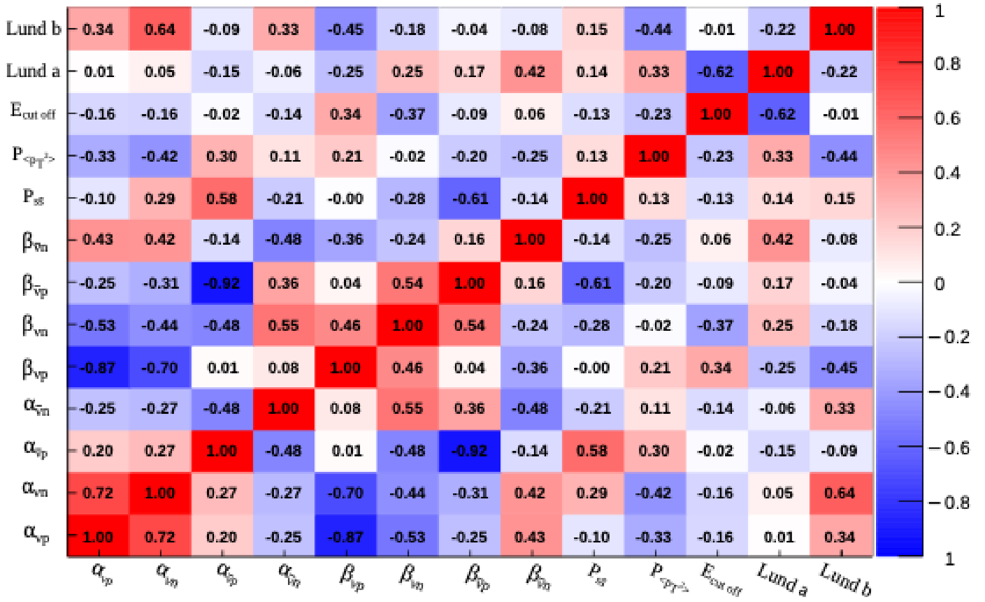
Fully exploiting the GENIE tuning machinery

- First global AGKY tune
  - Tunning the low-W AGKY + PYTHIA altogether
- Focus on averaged charged multiplicity data
- Data-driven constrains to 13 **non-reweightable parameters**
  - Improved description of H+D data
  - Best-fit parameter estimations
  - Uncertainty estimations

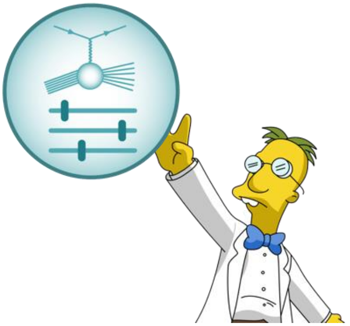
    (*) How can we propagate this uncertainties?



(b) $\nu_\mu + p \rightarrow \mu^- X^{++}$ (e) $\nu_\mu + n \rightarrow \mu^- X^+$

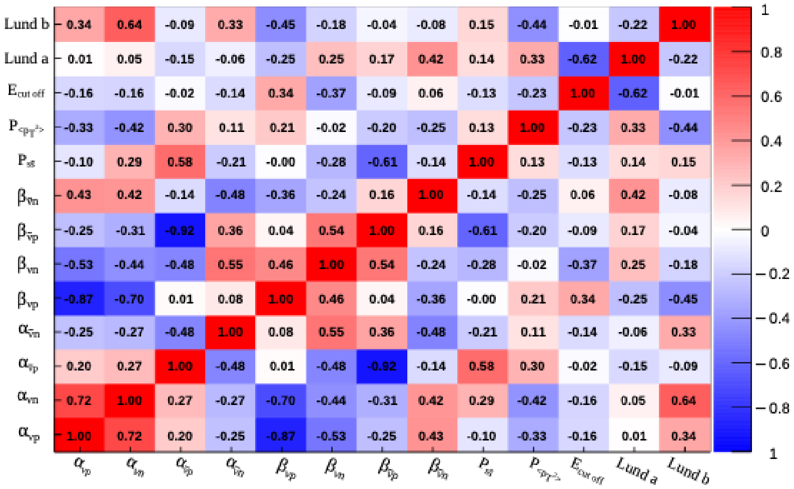


	$\alpha_{\nu p}$	$\alpha_{\nu n}$	$\alpha_{\bar{\nu} p}$	$\alpha_{\bar{\nu} n}$	$\beta_{\nu p}$	$\beta_{\nu n}$	$\beta_{\bar{\nu} p}$	$\beta_{\bar{\nu} n}$	P_{ss}	$P_{<p_T^2>}$	$E_{cut\ off}$	Lund a	Lund b
Lund b	0.34	0.64	-0.09	0.33	-0.45	-0.18	-0.04	-0.08	0.15	-0.44	-0.01	-0.22	1.00
Lund a	0.01	0.05	-0.15	-0.06	-0.25	0.25	0.17	0.42	0.14	0.33	-0.62	1.00	-0.22
$E_{cut\ off}$	-0.16	-0.16	-0.02	-0.14	0.34	-0.37	-0.09	0.06	-0.13	-0.23	1.00	-0.62	-0.01
$P_{<p_T^2>}$	-0.33	-0.42	0.30	0.11	0.21	-0.02	-0.20	-0.25	0.13	1.00	-0.23	0.33	-0.44
P_{ss}	-0.10	0.29	0.58	-0.21	-0.00	-0.28	-0.61	-0.14	1.00	0.13	-0.13	0.14	0.15
$\beta_{\bar{\nu} n}$	0.43	0.42	-0.14	-0.48	-0.36	-0.24	0.16	1.00	-0.14	-0.25	0.06	0.42	-0.08
$\beta_{\bar{\nu} p}$	-0.25	-0.31	-0.92	0.36	0.04	0.54	1.00	0.16	-0.61	-0.20	-0.09	0.17	-0.04
$\beta_{\nu n}$	-0.53	-0.44	-0.48	0.55	0.46	1.00	0.54	-0.24	-0.28	-0.02	-0.37	0.25	-0.18
$\beta_{\nu p}$	-0.87	-0.70	0.01	0.08	1.00	0.46	0.04	-0.36	-0.00	0.21	0.34	-0.25	-0.45
$\alpha_{\bar{\nu} n}$	-0.25	-0.27	-0.48	1.00	0.08	0.55	0.36	-0.48	-0.21	0.11	-0.14	-0.06	0.33
$\alpha_{\bar{\nu} p}$	0.20	0.27	1.00	-0.48	0.01	-0.48	-0.92	-0.14	0.58	0.30	-0.02	-0.15	-0.09
$\alpha_{\nu n}$	0.72	1.00	0.27	-0.27	-0.70	-0.44	-0.31	0.42	0.29	-0.42	-0.16	0.05	0.64
$\alpha_{\nu p}$	1.00	0.72	0.20	-0.25	-0.87	-0.53	-0.25	0.43	-0.10	-0.33	-0.16	0.01	0.34



Professor-based Reweight

How can experiments further exploit the GENIE data-driven systematics in their analysis?



	$\alpha_{\nu p}$	$\alpha_{\nu n}$	$\alpha_{\bar{\nu} p}$	$\alpha_{\bar{\nu} n}$	$\beta_{\nu p}$	$\beta_{\nu n}$	$\beta_{\bar{\nu} p}$	$\beta_{\bar{\nu} n}$	P_{ss}	$P_{<p_T^2>}$	$E_{cut\ off}$	Lund a	Lund b
Lund b	0.34	0.64	-0.09	0.33	-0.45	-0.18	-0.04	-0.08	0.15	-0.44	-0.01	-0.22	1.00
Lund a	0.01	0.05	-0.15	-0.06	-0.25	0.25	0.17	0.42	0.14	0.33	-0.62	1.00	-0.22
$E_{cut\ off}$	-0.16	-0.16	-0.02	-0.14	0.34	-0.37	-0.09	0.06	-0.13	-0.23	1.00	-0.62	-0.01
$P_{<p_T^2>}$	-0.33	-0.42	0.30	0.11	0.21	-0.02	-0.20	-0.25	0.13	1.00	-0.23	0.33	-0.44
P_{ss}	-0.10	0.29	0.58	-0.21	-0.00	-0.28	-0.61	-0.14	1.00	0.13	-0.13	0.14	0.15
$\beta_{\bar{\nu} n}$	0.43	0.42	-0.14	-0.48	-0.36	-0.24	0.16	1.00	-0.14	-0.25	0.06	0.42	-0.08
$\beta_{\bar{\nu} p}$	-0.25	-0.31	-0.92	0.36	0.04	0.54	1.00	0.16	-0.61	-0.20	-0.09	0.17	-0.04
$\beta_{\nu n}$	-0.53	-0.44	-0.48	0.55	0.46	1.00	0.54	-0.24	-0.28	-0.02	-0.37	0.25	-0.18
$\beta_{\nu p}$	-0.87	-0.70	0.01	0.08	1.00	0.46	0.04	-0.36	-0.00	0.21	0.34	-0.25	-0.45
$\alpha_{\bar{\nu} n}$	-0.25	-0.27	-0.48	1.00	0.08	0.55	0.36	-0.48	-0.21	0.11	-0.14	-0.06	0.33
$\alpha_{\bar{\nu} p}$	0.20	0.27	1.00	-0.48	0.01	-0.48	-0.92	-0.14	0.58	0.30	-0.02	-0.15	-0.09
$\alpha_{\nu n}$	0.72	1.00	0.27	-0.27	-0.70	-0.44	-0.31	0.42	0.29	-0.42	-0.16	0.05	0.64
$\alpha_{\nu p}$	1.00	0.72	0.20	-0.25	-0.87	-0.53	-0.25	0.43	-0.10	-0.33	-0.16	0.01	0.34

i.e. hadronization uncertainties

YOUR

ANALYSIS

Reweighting low-AGKY and PYTHIA

- Test sample $^1/_x$ ν_μ flux on H
- Parameters of interest
  - KNO-Alpha-νp, unweighted (0.8), reweighted (1.8)
  - PYTHIA-Lunda, unweighted (1.9), reweighted (1.0)

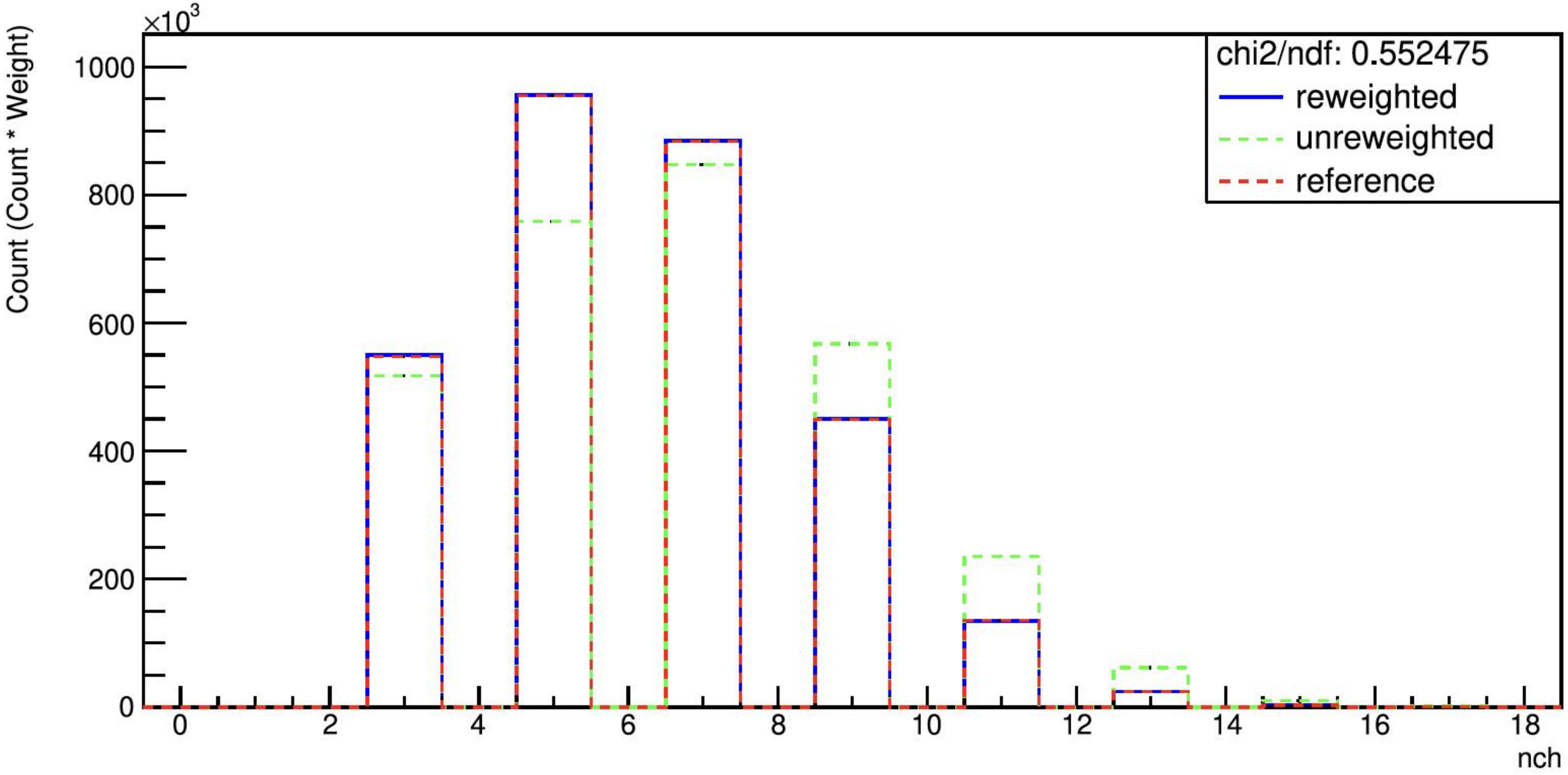
The current observable is W plus another variable:

- p_T of hadronic system
- p_T, p of leading hadron in the final state
- y value
- p of leading π

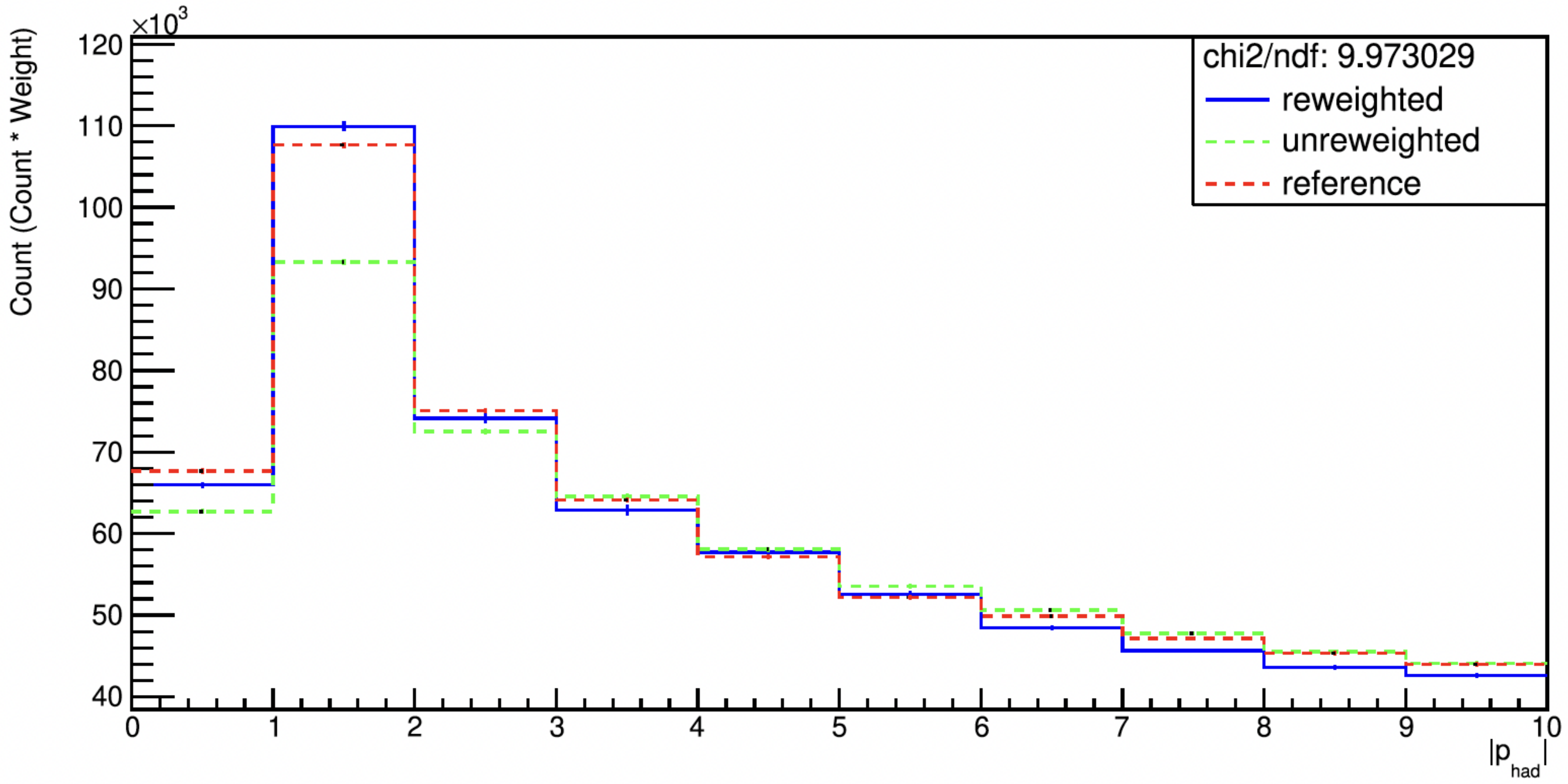
and channel information

- n_{ch} (chapped at 17)
- $n_{neutral}$ (chapped at 3)

Reweighting low-AGKY and PYTHIA



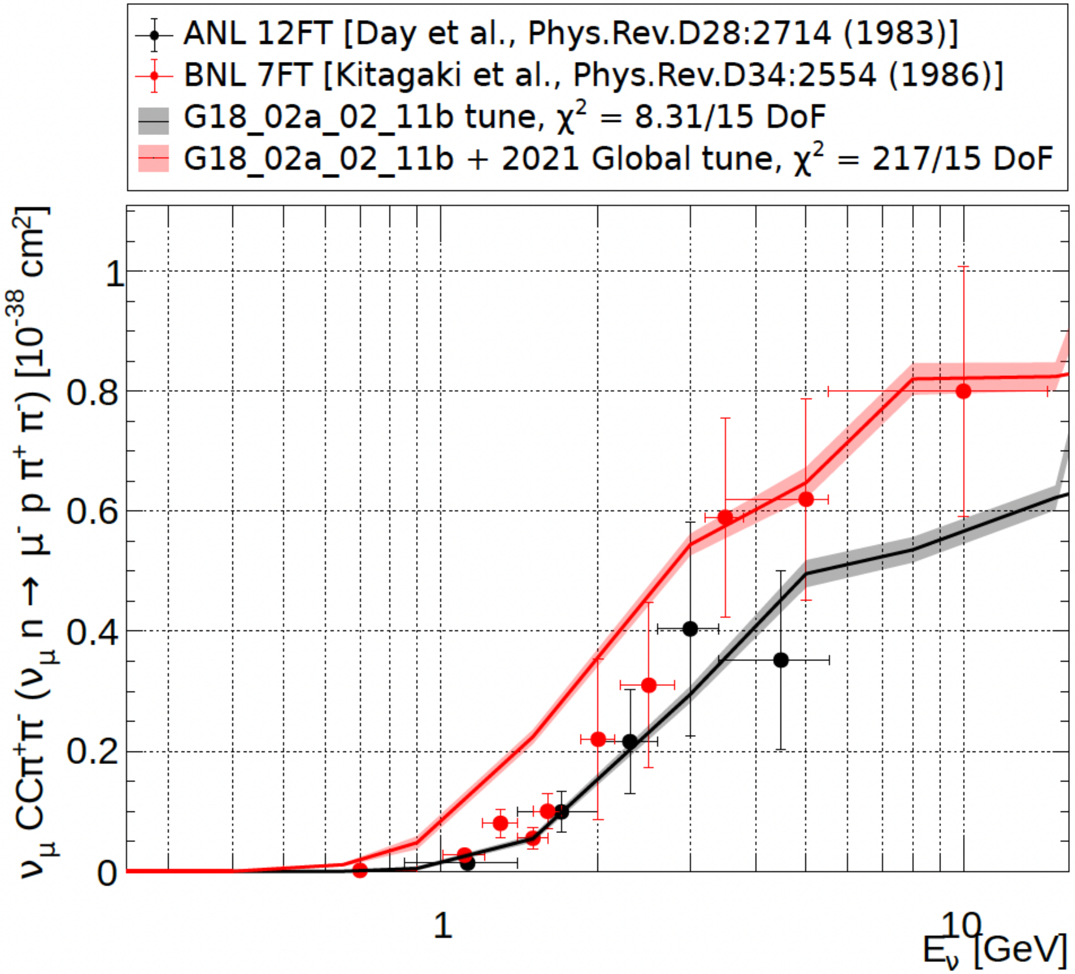
Reweighting low-AGKY and PYTHIA



Not possible to tune without hadronization

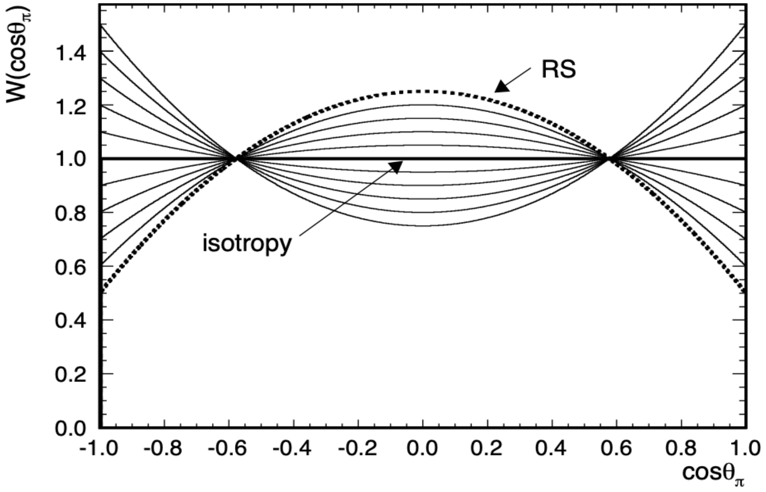
We must start with the hadronization tune for a consistent description of the SIS region

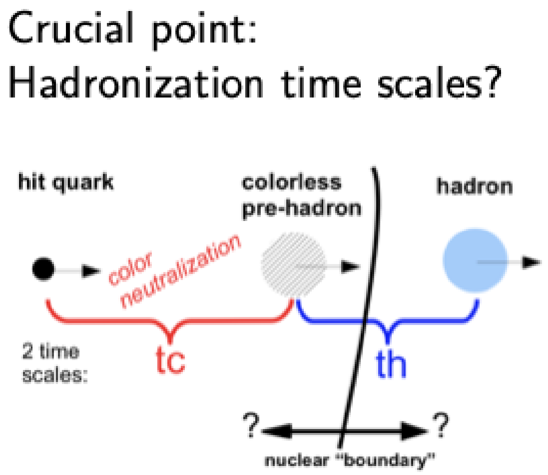
- The SIS region in GENIE is affected by low-W AGKY parameters
    - We simplified the problem into two separate tunes
- When the hadronization tune results are applied on the SIS region, we observe an increase of two-pion production
- The tunings are not fully independent in this configuration
    - This difference is absorbed as an increase of $R_{\nu p}^{CC2\pi}$ and $R_{\nu n}^{CC2\pi}$ in the free nucleon tune



Other parameters of interest

Angular decay in Resonance rest frame



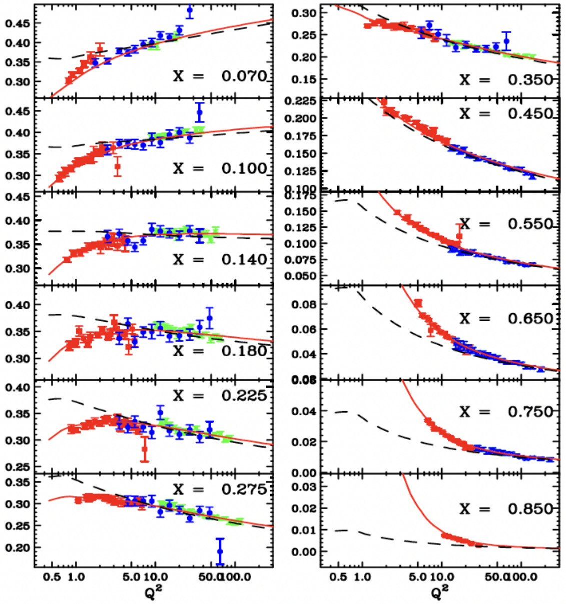


Crucial point:
Hadronization time scales?

In GENIE it is a single formation time

Additional Low-W AGKY

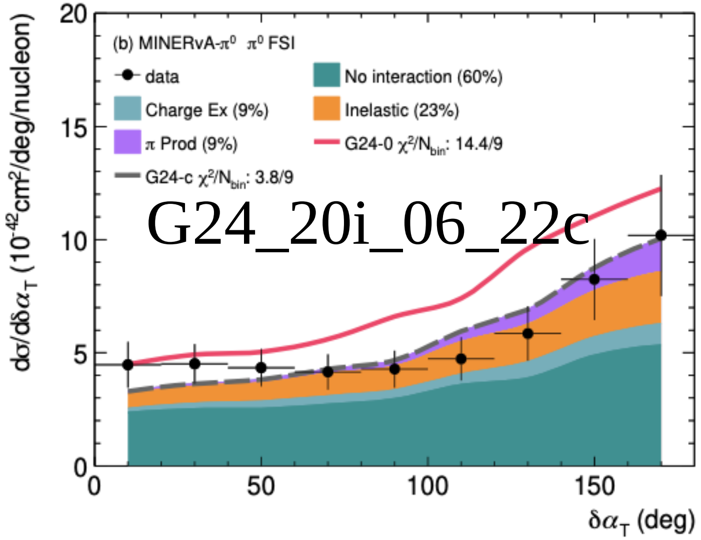
- Parameters determine which type of baryon is produced – always one, n or p
  - Hyperon production
- Shower kinematics …



BY model parameters

And many more!

Nuclear effects
I.e. FSI



G24_20i_06_22c

Conclusions

RES, SIS and DIS region must be tuned altogether

- For νN, complications arise due to the lack of data
  - Electron-scattering is a clear alternative
- Large set of parameters to tune SIS contribution in GENIE, but need to tune it consistently with RES model
- Hadronization plays a non-crucial role and should be first tuned model of SIS region
- Little emphasis on systematic quantification, especially for hadronization
- DIS models tuned to electron data have well defined parameters to tune

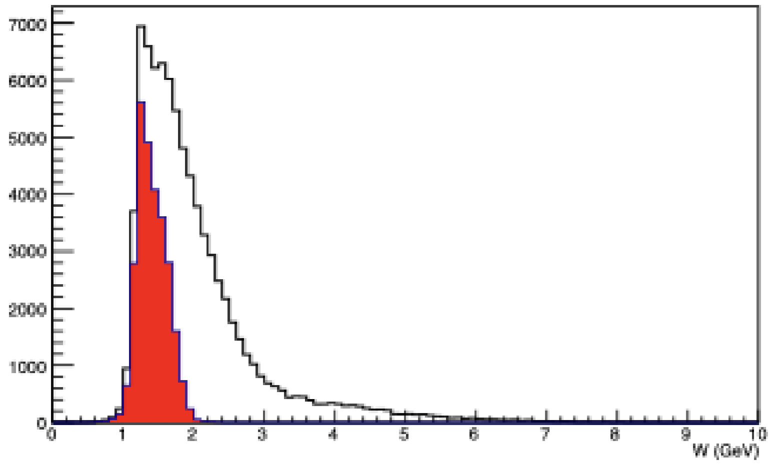
More effort needed – revise hadronization models at energies relevant for neutrinos and quantify uncertainties with electron data

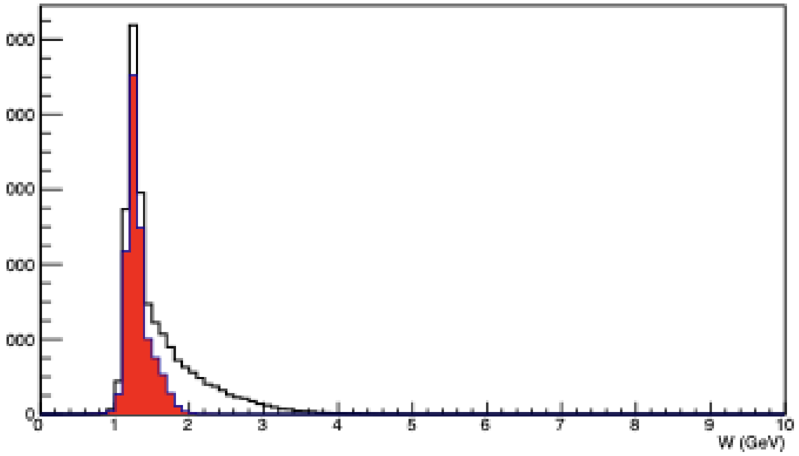
PYTHIA

- LUND string fragmentation model
- Uses the assumption of linear confinement as a starting point.
- As partons move apart, their colour flux tube gets stretched.
- Stored potential energy increases linearly with distance of colour charges.
- You can think of the ”string” as the axis of the flux tube.
- The string constant is $\sim$ 1 GeV/fm.
- As the potential energy increases, the string may break producing a q$\bar{q}$ pair.
- String breaks causally disconnected; simulated in a convenient order. A break typically creates a meson.
- Baryons also produced; A string can break by antidiquark-diquark production, or baryons can be produced using a ‘popcorn’ model. With every break, a produced hadron takes away a fraction of the available energy/momentum.
- Continuing till some cut-off point.

Extending the validity of GENIE model to lower W

- On the right, the invariant mass distributions for inelastic events. (Distribution is smeared due to Fermi momentum.)
- Up plot: DUNE, Down: HK Red component: Resonances
- Kinematic area below 2.5 - 3.0 GeV in invariant mass is critically important.
- Augment PYTHIA with an empirical GENIE model, anchored to data and valid in the area below 3 GeV. Install handles to express uncertainty.

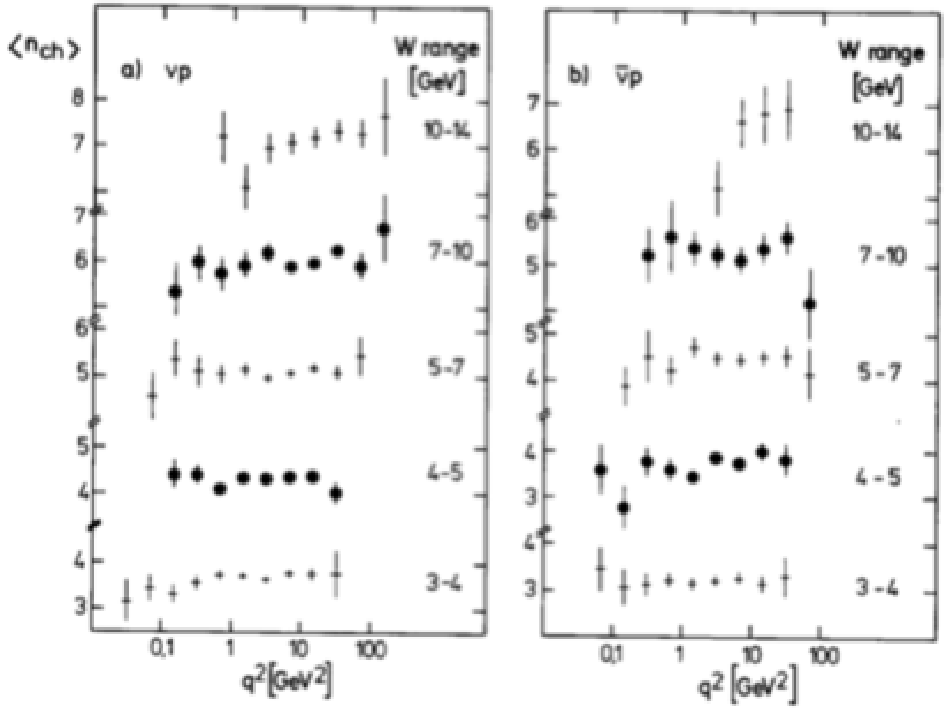




Empirical low-W model: How many hadrons are produced?

Average charged hadron multiplicities $< n_{ch} >$ could, more generally, have an additional Q^2 dependence:

$$< n_{ch} >= a + b \cdot ln(W^2/GeV^2) + b' ln(Q^2/GeV^2)$$

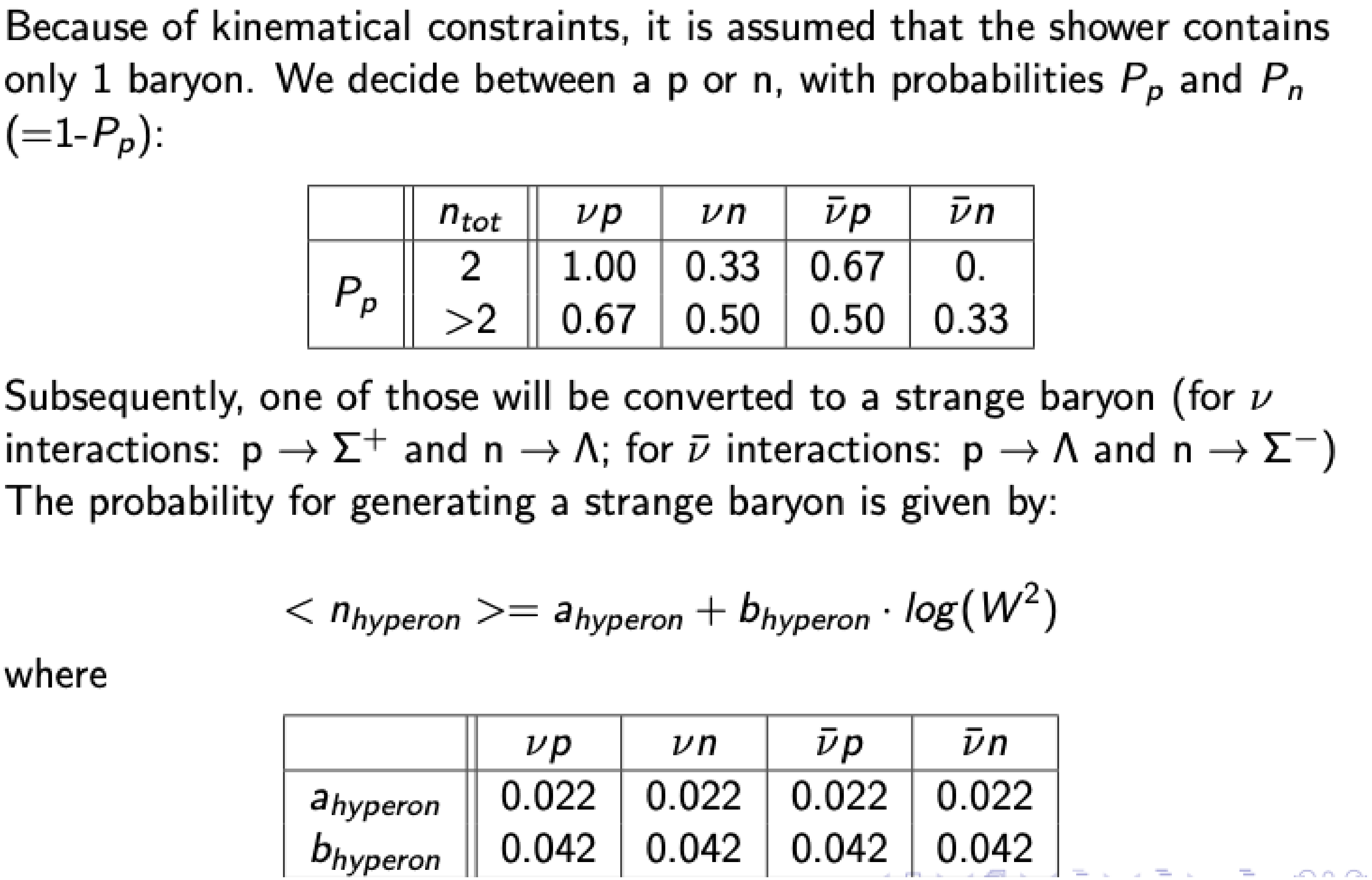


No Q^2 dependence has been observed in $\nu/\bar{\nu}$ scattering [H. Grassler et al., Nucl. Phys., **B223**, 269 (1983)].

Values of b' are 0.04 ± 0.02 for νp and 0.05 ± 0.04 for $\bar{\nu}$p

In GENIE, $b' = 0$ for all channels.

Empirical low-W model: Generating the particle spectrum



Because of kinematical constraints, it is assumed that the shower contains only 1 baryon. We decide between a p or n, with probabilities P_p and P_n (=1-P_p):

	n_{tot}	νp	νn	$\bar{\nu} p$	$\bar{\nu} n$
P_p	2	1.00	0.33	0.67	0.
	>2	0.67	0.50	0.50	0.33

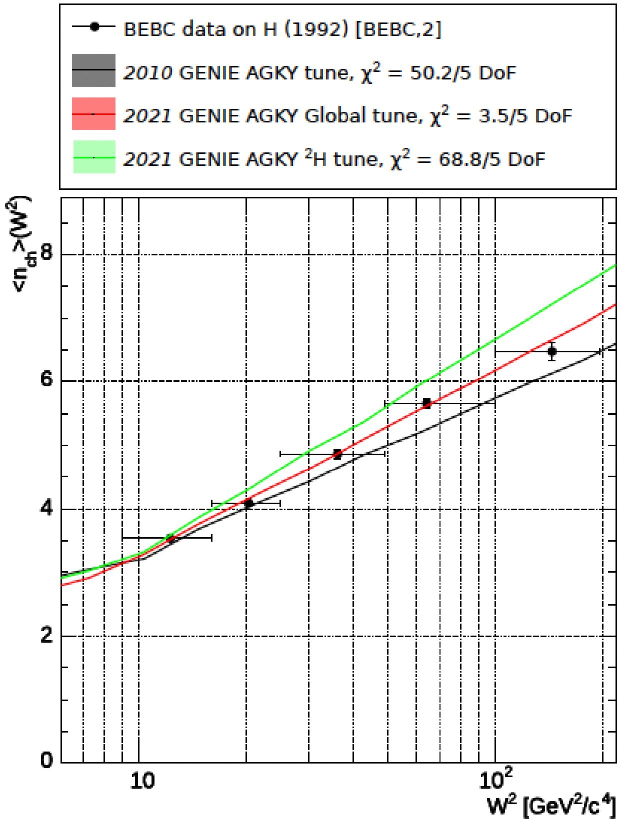
Subsequently, one of those will be converted to a strange baryon (for ν interactions: p $\rightarrow \Sigma^+$ and n $\rightarrow \Lambda$; for $\bar{\nu}$ interactions: p $\rightarrow \Lambda$ and n $\rightarrow \Sigma^-$)
The probability for generating a strange baryon is given by:

$$< n_{hyperon} >= a_{hyperon} + b_{hyperon} \cdot log(W^2)$$

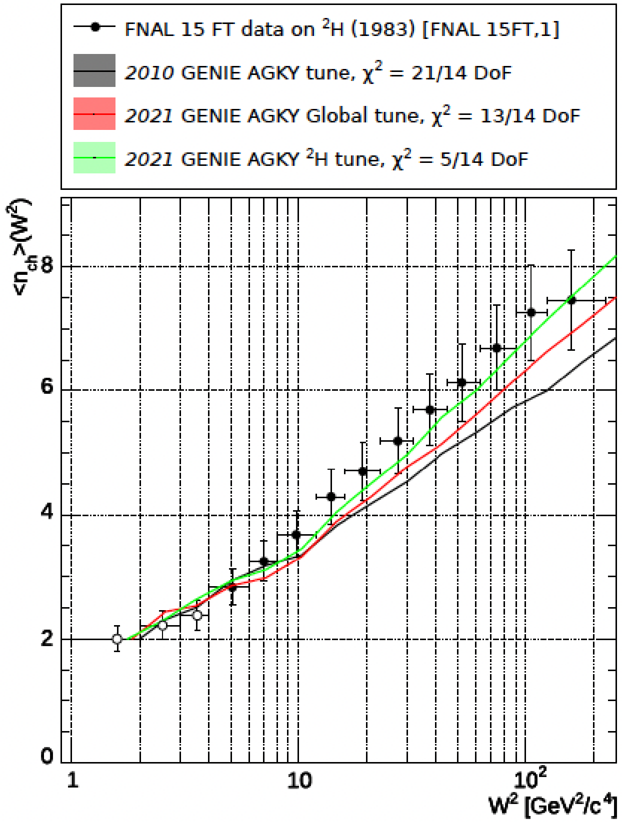
where

	νp	νn	$\bar{\nu} p$	$\bar{\nu} n$
$a_{hyperon}$	0.022	0.022	0.022	0.022
$b_{hyperon}$	0.042	0.042	0.042	0.042

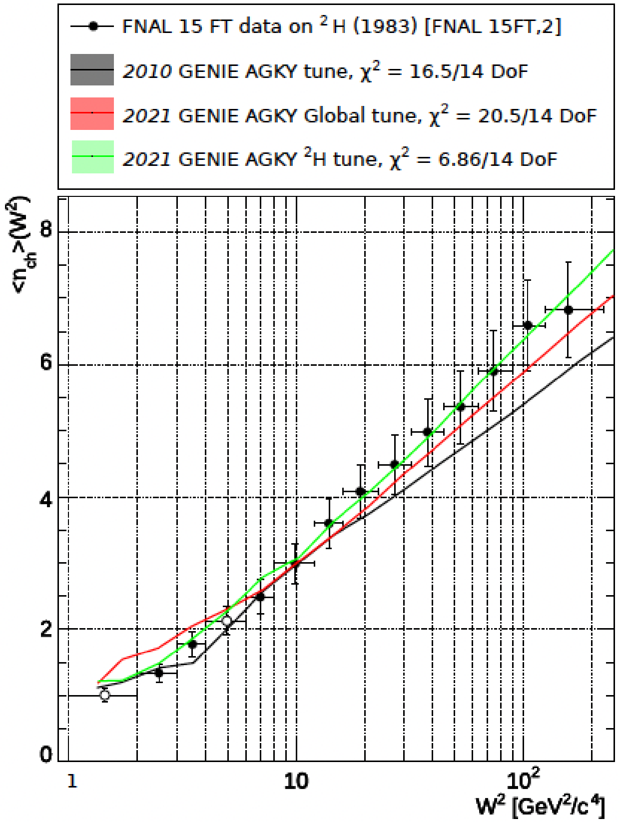
GENIE AGKY tune results



$\nu_\mu + p$ on hydrogen.



$\nu_\mu + p$ on deuterium.



$\nu_\mu + n$ on deuterium.

- The *2021* GENIE global tune (red) underpredicts ^{2}H data
- The *2021* GENIE ^{2}H tune (green) overpredicts H data

Additional Low-W AGKY parameters of interest

- KNO scaling
- Draw actual multiplicities from a Poisson distribution with given average
- $< n > P(n) = f\,(n/ < n >)$ is independent of W
- The function is parameterized using the Levi function

$$L(z; c) = \frac{2e^{-c}c^{cz+1}}{\Gamma(cz+1)}$$

	νp	νn	$\bar{\nu} p$	$\bar{\nu} n$
c	7.93	5.22	as in νn	as in νp