Combining current and historical cross section data

Callum Wilkinson for the T2K NIWG group

University of Bern

31/05/16

UNIVERSITÄT BERN

AEC
ALBERT EINSTEIN CENTER
FOR FUNDAMENTAL PHYSICS

- Cross section uncertainties are/will be a limiting factor for current/future neutrino oscillation experiments.
- No theoretical consensus on neutrino cross section models at relevant energies, and limited amount of data.
- The Neutrino Interactions Working Group (NIWG) provides cross section model parameter errors for T2K.
- Focus on recent work on one cross section channel[1], but the same problems apply to all channels of interest.

[1] CW, *et al.*, *Phys. Rev.* **D93**, 072010 (2016)

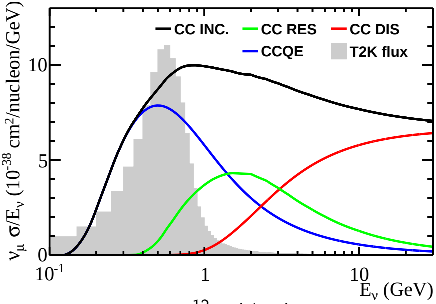
$$R(\vec{\mathbf{x}}) = \overbrace{\underbrace{\Phi(E_\nu) \times \sigma(E_\nu, \vec{\mathbf{x}}) \times \epsilon(\vec{\mathbf{x}})}^{\text{Near}} \times P(\nu_A \rightarrow \nu_B)}_{\text{Far}}$$

- **Event rate**; **Neutrino flux**; **Cross section**; **Detector smearing**; **Oscillation parameters**.
- Different near/far detector acceptances.
- Different targets for some near/far detectors.
- E_ν spectrum different at near/far (no cross section cancellation)

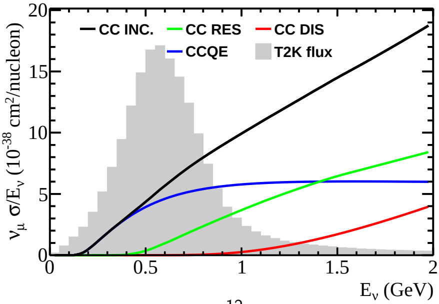
1. How should we calculate goodness of fit and select a model with limited information about the data?
2. How should we define the parameter errors for that model?
3. Should we exclude any datasets from our fits and how should we decide which?

I will reiterate these and plead for help at the end.
But please interrupt at any point!

Charged-current quasi-elastic (CCQE) interactions



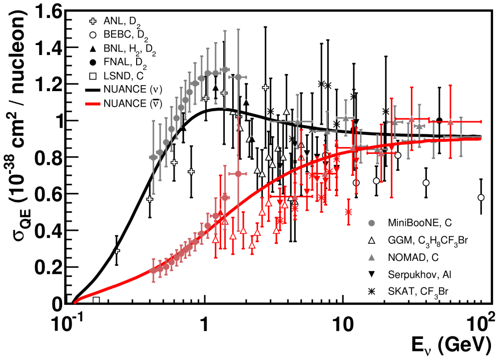
ν_μ–^{12}C $(/E_\nu)$



ν_μ–^{12}C

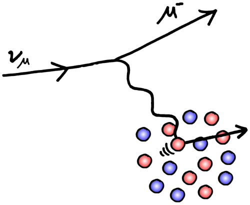
- "Simple" interaction, which is dominant for T2K and other oscillation experiments:

$$\nu_l + n \rightarrow l^- + p,$$
$$\bar{\nu}_l + p \rightarrow l^+ + n.$$

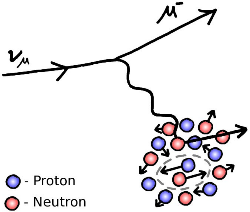


- This is (most) of the available data for CCQE. **Difficult to draw a line on this plot!**
- Recent (low E_ν, heavy target) results show 30-40% discrepancy with old data (high E_ν or light target)

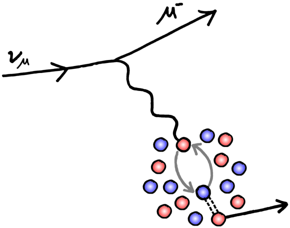
- Many new models proposed to explain the low E_ν, higher mass discrepancies.
- Broadly two categories of model development:
  1. Better descriptions of the ground state of the nucleus
  2. Additional interaction effects due to many nucleons being close together
- Non-trivial to mix and match models, or to introduce fit parameters which vary between them $\rightarrow$ **binary choices**.



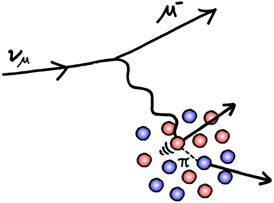
RFG



SF



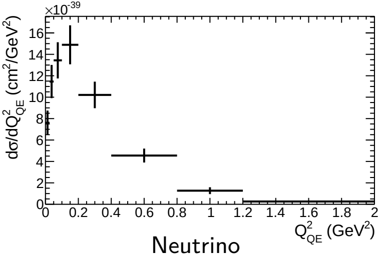
RPA



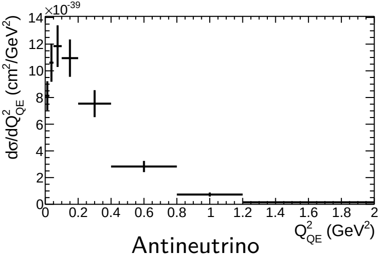
2p2h

Fitting CCQE data

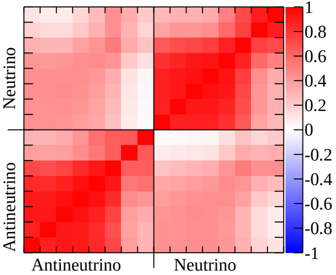
- **Aim:** constrain model parameters using modern heavy target CCQE data.
- Two candidate model combinations:
  - **Model A**: SF+2p2h
  - **Model B**: RFG+RPA+2p2h
- Available fit parameters:
  - $\mathbf{p}_1$: axial mass, M_{A}
  - $\mathbf{p}_2$: normalization of the 2p2h model component
  - $\mathbf{p}_3$: Fermi momentum, p_{F}



Neutrino



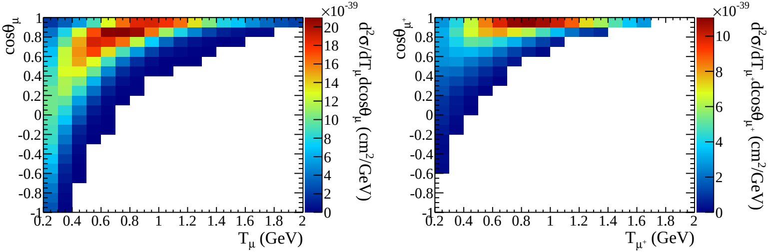
Antineutrino



Correlation matrix

- **Unfolded distributions** differential in Q^2_{QE} (derived from the muon kinematics), on $\sim$CH target
- **Background subtracted**.
- MINERνA provide the full covariance.

MiniBooNE CCQE data



Neutrino　　Antineutrino

- **Unfolded distributions** differential in muon angle, $\cos\theta_\mu$, and kinetic energy, T_μ on a $\sim CH_2$ target.
- **Background subtracted**.
- MiniBooNE only released the diagonals of the shape-only covariance matrix and the normalization uncertainty for each dataset:
  - ν: $\pm 10.7\%$
  - $\bar{\nu}$: $\pm 13.0\%$
- **No information is available about correlations.**

Fit details

- Build prediction for any given parameter set and each experiment. Fit to data using χ^2 minimization and steepest gradient descent:

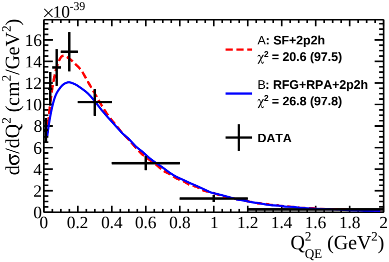
$$
\begin{aligned}
\chi^2(\vec{\mathbf{x}}) = & \left[\sum_{k=0}^{N} \left(\frac{\nu_k^{DATA} - \lambda_{\nu}^{-1} \nu_k^{MC}(\vec{\mathbf{x}})}{\sigma_k} \right)^2 + \left(\frac{\lambda_{\nu} - 1}{\varepsilon_{\nu}} \right)^2 \right] \rightarrow \text{MiniBooNE } \nu \\
& + \left[\sum_{l=0}^{M} \left(\frac{\nu_l^{DATA} - \lambda_{\bar{\nu}}^{-1} \nu_l^{MC}(\vec{\mathbf{x}})}{\sigma_l} \right)^2 + \left(\frac{\lambda_{\bar{\nu}} - 1}{\varepsilon_{\bar{\nu}}} \right)^2 \right] \rightarrow \text{MiniBooNE } \bar{\nu} \\
& + \left[\sum_{i=0}^{16} \sum_{j=0}^{16} \left(\nu_i^{DATA} - \nu_i^{MC}(\vec{\mathbf{x}}) \right) V_{ij}^{-1} \left(\nu_j^{DATA} - \nu_j^{MC}(\vec{\mathbf{x}}) \right) \right] \rightarrow \text{MINER}\nu\text{A}
\end{aligned}
$$

- $\lambda_{\alpha,\beta}$ are normalization parameters, $\varepsilon_{\alpha,\beta}$ are the normalization uncertainties, and $\vec{\mathbf{x}} = (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ are the fit parameters.
- We are given Gaussian errors, so we treat them as such. Very likely that at the edges of phase space ($\sigma \rightarrow 0$), this is incorrect. **Can we mitigate for this?**

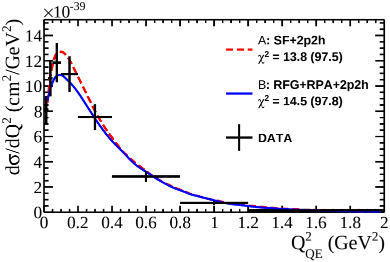
- $\mathbf{p}_1$ (M_A), $\mathbf{p}_2$ (p_F), $\mathbf{p}_3$ (2p2h norm.) and $\lambda_{\alpha,\beta}$ (MiniBooNE norm.) are allowed to vary in the fits:

Model	χ^2/DOF	$\mathbf{p}_1$: M_A (GeV)	$\mathbf{p}_2$: 2p2h (%)	$\mathbf{p}_3$: p_F (MeV)
Theory expectation		$\sim$1.00	$\sim$100	$\sim$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

- Confusing situation, essentially no difference between the models.
  - Neither fit with expectations from old data ($M_\mathrm{A} \approx 1$ GeV), or with the new model for nuclear effects (2p2h $\approx$ 100%).
  - Fake data studies suggested that **A** could not mimic model **B** or **B** mimic **A** (large χ^2 penalties), so this is surprising.
- Suggests that neither model is perfect, but it is unclear which is the better model $\rightarrow$ **model choice is not obvious**

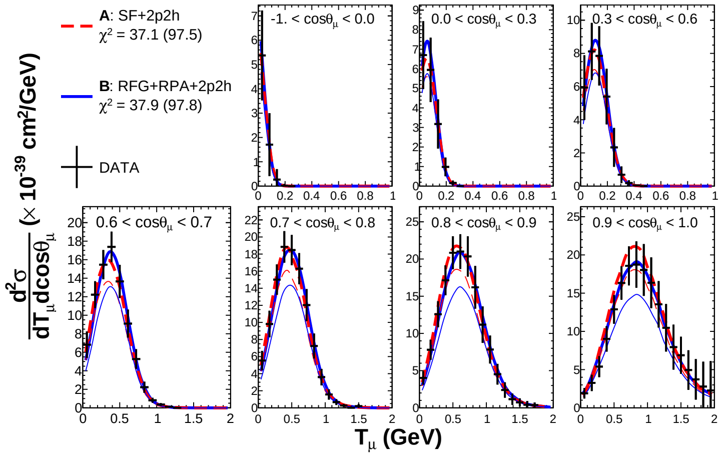


Neutrino



Antineutrino

- The χ^2 value in parentheses gives the **contribution** from each dataset by itself (ignoring correlations).
- **Total** χ^2 is the same for both models, but there are clear differences between them.



- Thick lines have $\lambda_{\alpha,\beta}$ parameters applied, thin lines do not. At the best fit points, there are significant pulls applied to these terms.
- The MiniBooNE antineutrino equivalent plot is in the backups.

- The best fit $\chi^2_{\min}$ values appear good for both fits:
  - **A:** SF+2p2h, $\chi^2_{\min} = 97.5/228$
  - **B:** RFG+RPA+2p2h, $\chi^2_{\min} = 97.8/228$
- **However**, correlations are not always available, so Gaussian statistics no longer work:
  - Standard goodness of fit tests are unreliable
  - $\Delta\chi^2 = 1$ is no longer appropriate for calculating parameter errors
- We believe we have no way to correctly calculate a p-value because we cannot throw fake datasets according to the true covariance.
- Need a better test statistic to properly assess agreement.

- The PGoF test statistic[2] is defined:

$$\chi^2_{\mathrm{PGoF}}(\vec{\mathbf{x}}) = \chi^2_{tot}(\vec{\mathbf{x}}) - \sum_{r=1}^{D} \chi^2_{r,min}(\vec{\mathbf{x}}),$$

for D datasets, where χ^2_{tot} is the minimum χ^2 in a fit to all D datasets.

$$\mathrm{DOF} = \sum_{r=1}^{D} P_r - P_{tot}$$

where P_r and P_{tot} are the number of free parameters varied in each fit.

- Compares the best fit parameters for the complete dataset and for subsets of the data. Assesses the compatibility between datasets.
- **Still assumes Gaussian statistics**, but allows us to roughly compare the tension between different models.

[2] M. Maltoni, T. Schwetz, *Phys. Rev.* **D68**, 033020 (2003)

- In each fit, $\mathbf{p}_1$ (M_A), $\mathbf{p}_2$ (p_F), $\mathbf{p}_3$ (2p2h norm.) and $\lambda_{\alpha,\beta}$ (MiniBooNE norm.) are allowed to vary
- **Model A:** SF+2p2h (see slide 34 for a more detailed breakdown)

	χ^2_{min}/DOF	χ^2_{PGoF}/DOF	PGoF (%)
All		41.1/6	0.00
MINERνA vs MiniBooNE	97.5/228	34.6/3	0.00
ν vs $\bar{\nu}$		8.5/3	3.59

- **Model B:** RFG+RPA+2p2h (see slide 35 for a more detailed breakdown)

	χ^2_{min}/DOF	χ^2_{PGoF}/DOF	PGoF (%)
All		17.9/6	0.66
MINERνA vs MiniBooNE	97.8/228	15.9/3	0.12
ν vs $\bar{\nu}$		-3.3/3	100.00

- There is much less tension in the PGoF results for **B** (RFG+RPA+2p2h) than **A** (SF+2p2h).

- $\Delta\chi^2 \neq 1$ for 1σ errors.
- Parton density distribution fitters had similar problems, many papers discuss these issues[3]. **Solution:** inflate the $\Delta\chi^2$ value that defines the error.
- PGoF gives a value for the incompatibility between the datasets: how much the χ^2 increases between the best fit point of each experiment, and the best fit for the combined dataset.
- Scaled errors such that $\Delta\chi^2 = \sqrt{\chi^2_{\mathrm{PGoF}}/\mathrm{DOF}} \approx 2.3$
  - This scaling does make the errors cover the various datasets, so we adopted it.
  - Basically a guess, but also makes $\chi^2/\mathrm{DOF} \approx 1$, so seems reasonable.

[3] e.g. J. Pumplin, D. Stump, W. Tung, *Phys. Rev.* **D65**, 014011 (2001)

- Table of parameter errors, with and without scaling:

Fit type	χ^2/DOF	$\mathbf{p}_1$: M_A (GeV)	$\mathbf{p}_2$: 2p2h (%)	$\mathbf{p}_3$: p_F (MeV)
Unscaled	97.84/228	1.15±0.03	27±12	223±5
PGoF scaling		1.15±0.06	27±27	223±11

- Recommend model **B**, RFG+RPA+2p2h, with PGoF scaled parameters for T2K analyses[4].
- Note that we preserve the correlations from the fit with this procedure. Not all parameters will contribute equally to the problems we see in the fit.

[4] K. Abe, *et al.*, *Phys. Rev. Lett.* **116**, 181801 (2016)

1. **How should we calculate goodness of fit and select a model with limited information about the data?**
    - Both models have serious flaws, but no strong physics motivation for choice (both have advantages).
    - Best fit predictions are different for our models, but similar χ^2_{min} and we cannot calculate a p-value correctly without the correlations. **Or can we?**
    - We used the PGoF test to pick the model which had more consistent parameters between datasets. Not quite the same as goodness of fit test. **Can we do better?**

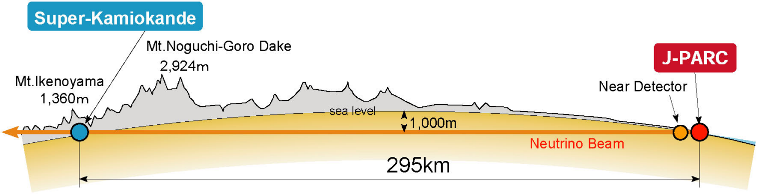
2. **How should we define the parameter errors for that model?**

   - No correlations for some datasets → increase $\Delta\chi^2$ used to define errors (but retain parameter correlations) with a rough guess. **Can we formalize/verify this procedure?**

   - **Alternative:** do not apply external constraints when fitting near detector data.
     - Arguably, we understand this channel better than others.
     - Will limit future sensitivity.

3. **Should we exclude any of the datasets from our fits?**
   - We have simply taken data at face value and used all information available. Unlikely to get more on older datasets.
   - Currently have no method for excluding datasets, but may want to with future fits. **What methods can we use?**
   - In some bins of the data near phase space limits, Gaussian assumptions may not hold. Remove bins for which **our** predicted event rate is low? What prediction? **All advice welcome!**

Backup



- T2K primary generator: **NEUT**[5].
- **ND280 targets:** *Carbon*, *Oxygen*, *Hydrogen*, Iron, Lead, Brass, and Argon.
- **SK targets:** *Oxygen*, *Hydrogen*.

Italicised targets are used for the T2K oscillation analyses.

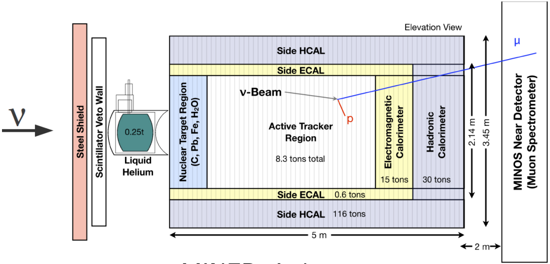
[5] Y. Hayato, *Acta Physica Polonica B* **40**, 2477 (2009)

- Reconstructed neutrino energy E_ν^{QE} and reconstructed four-momentum transfer Q^2_{QE} are given by the equations:

$$E_\nu^{QE} = \frac{2M'_n E_\mu - (M'^2_n + m^2_\mu - M^2_p)}{2(M'_n - E_\mu + \sqrt{E^2_\mu - m^2_\mu}) \cos\theta_\mu}$$

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

- Where $E_\mu = T_\mu + m_\mu$ and $M'_n = M_n - E_B$
- $E_B = 34$ MeV for all datasets except MINERνA antineutrino, where $E_B = 30$ MeV is assumed.

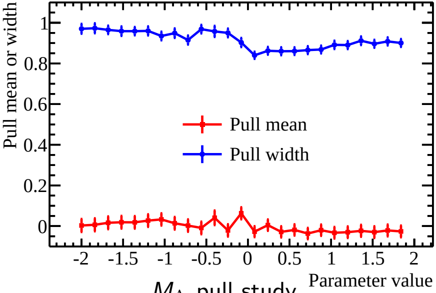


MINERνA detector

- Two presentations of MINERνA CCQE results are currently available:
  - $\theta_\mu < 20^\circ$
  - FULL (previous slide)
- The $\theta_\mu < 20^\circ$ sample excludes the unsampled regions of phase space (where MINOS can't tag muons), so is less model dependent
- Both samples are official MINERνA results.

[6] M. Kordosky, Fermilab Wine and Cheese, June 2012

1. Run one iteration of the fitting code to produce nominal distributions for any datasets for any parameter set.
2. Pass the fake data to a fit, the fit reads in the nominal distributions from the first step and treats them as data.
3. The errors are scaled bin by bin to make errors for the MiniBooNE dataset. $\sigma_P \times \frac{N_F}{N_P}$ where σ_P is the published error, and N_F and N_P are the fake data and published bin contents.
4. Produce a covariance matrix for MINERvA using a similar technique (and the published correlation matrix).
$M_{ij} = \left(\sigma_P \times \frac{N_F}{N_P}\right)_i \mathrm{Cov}(i,j) \left(\sigma_P \times \frac{N_F}{N_P}\right)_j$



M_A pull study

- The pull is calculated for each model/parameter by throwing the errors and other fit parameters:

$$\text{Pull} = \frac{\text{fitted} - \text{input}}{\text{error}}$$

- For each value of the parameter, 1000 throws are made, and a fit is performed on the resulting pull distribution.
- Expect a pull distribution centred on 0 with width 1 for every parameter value in the fit range

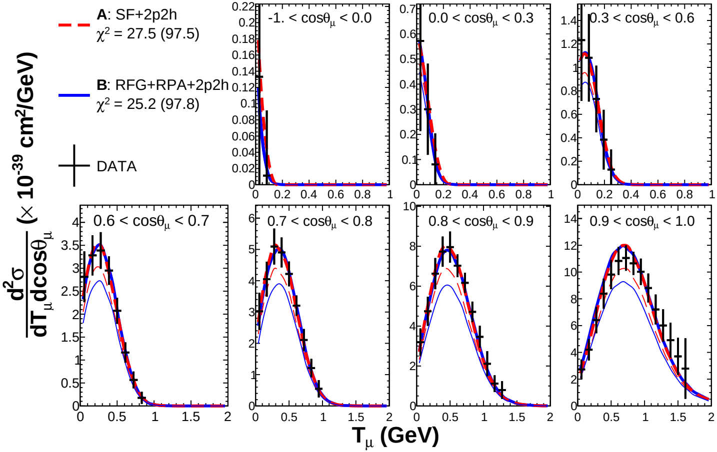
- Asimov fake data fits[7]: fit fake data without throwing errors.

Fake data	Fit model	χ^2/DOF	$\mathbf{p}_1$: M_{A} (GeV)	$\mathbf{p}_2$: 2p2h (%)	$\mathbf{p}_3$: p_{F} (MeV)
A	**A**	0.00/228	1.01±0.03	100±21	217±10
	B	83.63/228	0.94±0.02	181±17	212±7
B	**A**	66.54/228	1.21±0.02	44±16	209 (at limit)
	B	0.00/228	1.01±0.03	100±18	217±3

$$\mathbf{A} = \text{SF+2p2h};\ \mathbf{B} = \text{RFG+RPA+2p2h}$$

- Show the expected parameter error sizes for real data fits.
- Get an idea about model sensitivity by fitting **A** $\rightarrow$ **B** and **B** $\rightarrow$ **A**:
  - Large distortions in the parameters are observed.
  - Poor χ^2 values

[7] G. Cowan, K. Cranmer, E. Gross, O. Vitells, *Eur. Phys. J.* **C71**, 1554 (2011)


A: SF+2p2h
χ^2 = 27.5 (97.5)
B: RFG+RPA+2p2h
χ^2 = 25.2 (97.8)
DATA
$\frac{d^2\sigma}{dT_\mu d\cos\theta_\mu}$ ($\times 10^{-39}$ cm^2/GeV)
-1. < $\cos\theta_\mu$ < 0.0
0.0 < $\cos\theta_\mu$ < 0.3
0.3 < $\cos\theta_\mu$ < 0.6
0.6 < $\cos\theta_\mu$ < 0.7
0.7 < $\cos\theta_\mu$ < 0.8
0.8 < $\cos\theta_\mu$ < 0.9
0.9 < $\cos\theta_\mu$ < 1.0
T_μ (GeV)

- In each fit, $\mathbf{p}_1$ (M_{A}), $\mathbf{p}_2$ (p_{F}), $\mathbf{p}_3$ (2p2h norm.) and $\lambda_{\alpha,\beta}$ (MiniBooNE norm.) are allowed to vary

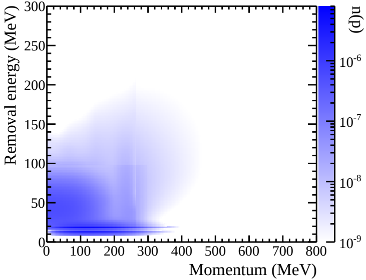
	χ^2_{min}/DOF	SGoF (%)	χ^2_{PGoF}/DOF	PGoF (%)
All	97.5/228	100.00	41.1/6	0.00
MINERνA	12.6/13	47.75	1.0/3	79.49
MiniBooNE	50.2/212	100.00	6.5/3	8.92
ν	54.8/142	100.00	25.1/3	0.00
$\bar{\nu}$	34.1/83	100.00	8.5/3	3.61
MINERνA vs MiniBooNE	97.5/228	100.00	34.6/3	0.00
ν vs $\bar{\nu}$	97.5/228	100.00	8.5/3	3.59

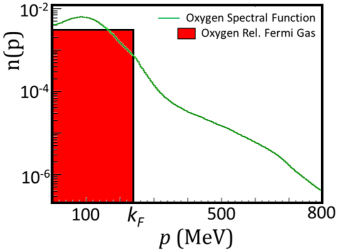
- Generally poor agreement between the datasets. In particular, nothing agrees well with MiniBooNE neutrino.
- Because of the lack of consistency, SF+2p2h is a poor choice for the default T2K MC model.

- In each fit, $\mathbf{p}_1$ (M_{A}), $\mathbf{p}_2$ (p_{F}), $\mathbf{p}_3$ (2p2h norm.) and $\lambda_{\alpha,\beta}$ (MiniBooNE norm.) are allowed to vary

	χ^2_{min}/DOF	SGoF (%)	χ^2_{PGoF}/DOF	PGoF (%)
All	97.8/228	100.00	17.9/6	0.66
MINERνA	23.4/13	3.74	1.0/3	79.03
MiniBooNE	58.6/212	100.00	2.0/3	57.69
ν	62.6/142	100.00	16.1/3	0.11
$\bar{\nu}$	38.5/83	100.00	6.1/3	10.75
MINERνA vs MiniBooNE	97.8/228	100.00	15.9/3	0.12
ν vs $\bar{\nu}$	97.8/228	100.00	-3.3/3	100.00

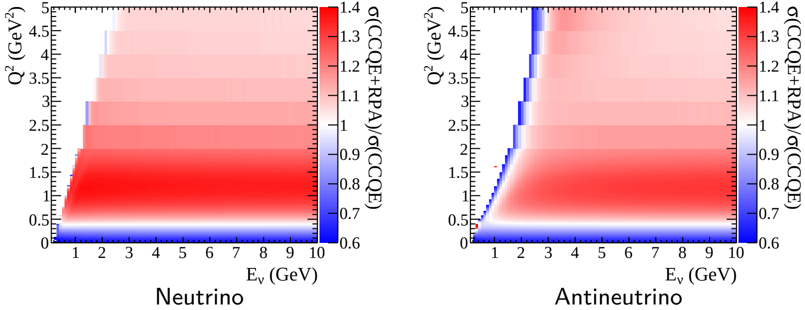
- There is a reasonable amount of tension, but in general the agreement is not bad.
- In particular, the agreement between ν and $\bar{\nu}$ is good, this is very important as we want to use the same systematics for antineutrino running.





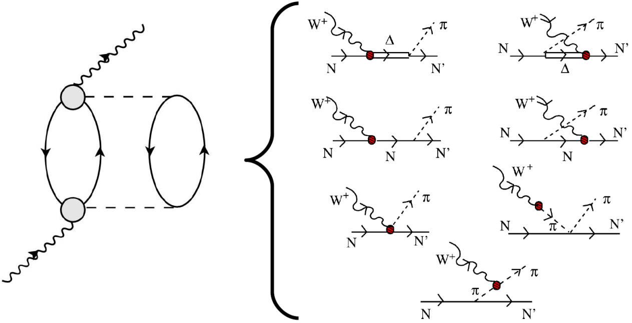
- Relativistic Fermi Gas (RFG), used for a long time in generators due to its simplicity (NEUT $<$v5.3.1).
- Omar Benhar's 2D Spectral Function[8] in momentum and removal energy has been implemented in NEUT (v5.3.1).

[8] O. Benhar, A. Fabrocini, *Phys. Rev.* **C62**, 034304 (2000)



- Random Phase Approximation (RPA), nuclear screening effect due to long range nucleon-nucleon correlations[9].
- NEUT implementation is dependent on Q^2 and E_ν.

[9] J. Nieves, I. R. Simo, M. J. V. Vacas, *Phys. Rev. C* **83**, 045501 (2011)



- Multi-nucleon interactions (2p2h) from Nieves *et al.*[10], see Peter Sinclair's NuInt2014 talk for full implementation details (NEUT v5.3.2).
- Includes the high E_ν extension[11]. The low q3 part of the cross-section is accurate up to high energies.

[10] J. Nieves, I. R. Simo, M. J. V. Vacas, *Phys. Rev. C* **83**, 045501 (2011)

[11] R. Gran, J. Nieves, F. Sanchez, M. Vicente Vacas, *Phys. Rev.* **D88**, 113007 (2013)