Fermilab | U.S. DEPARTMENT OF ENERGY | Office of Science

Experimental data from a theoretical needs POV

Alexis Nikolakopoulos

Towards on consensus in neutrino cross sections, NuXTract 2023

4th October 2023, CERN

Preliminary

Experimental data from a theoretical needs POV

The assignment:

“A talk about your experiences using experimental data releases and to give a theory-perspective on what experiments should provide for the future. This should focus mostly on "technical" needs (e.g. what should be in the data releases)”

What happened:

Examples of things that we ideally shouldn’t do

Uses for neutrino data

Oscillation/BSM analysis → not discussed

Fitting & parameter searches

Large scale tunes on global sets of data:

MicroBooNE tune [Phys. Rev. D 105, 072001]

GENIE tunes : [J.T. Vidal PhysRevD.105.012009], [J.T. V.PhysRevD.104.072009]

	MaCCQE fitted value	CC2p2h Norm. fitted value	CCQE RPA Strength fitted value	CC2p2h Shape fitted value	T2K $\chi^2_{\text{diag}}/N_{\text{bins}}$	T2K $\chi^2_{\text{Koch}}/N_{\text{bins}}$	T2K $\chi^2_{\text{full}}/N_{\text{bins}}$
Nominal (untuned)	0.961242 GeV	1	100%	0	106.7/58	149.83/58	97.56/58
"MicroBooNE tune"	1.10 ± 0.07 GeV	1.66 ± 0.19	$(85 \pm 20)\%$	$1^{+0}_{-0.74}$	52.5/58	110.58/58	103.84/58
"Alternate fit"	1.04 ± 0.10 GeV	1.44 ± 0.42	$(67 \pm 16)\%$	$0.91^{+0.09}_{-0.18}$	55.51/58	100.59/58	91.68/58

Need reliable ways to calculate goodness-of-fit

Will be important for the future as data amasses

TABLE X. Total χ^2 calculated with the datasets included in each fit: 2010 GENIE, χ^2_{2010}, 2021 GENIE, $\chi^2_{2021(\text{Global})}$, and 2021 GENIE, $\chi^2_{2021(^2\text{H})}$.

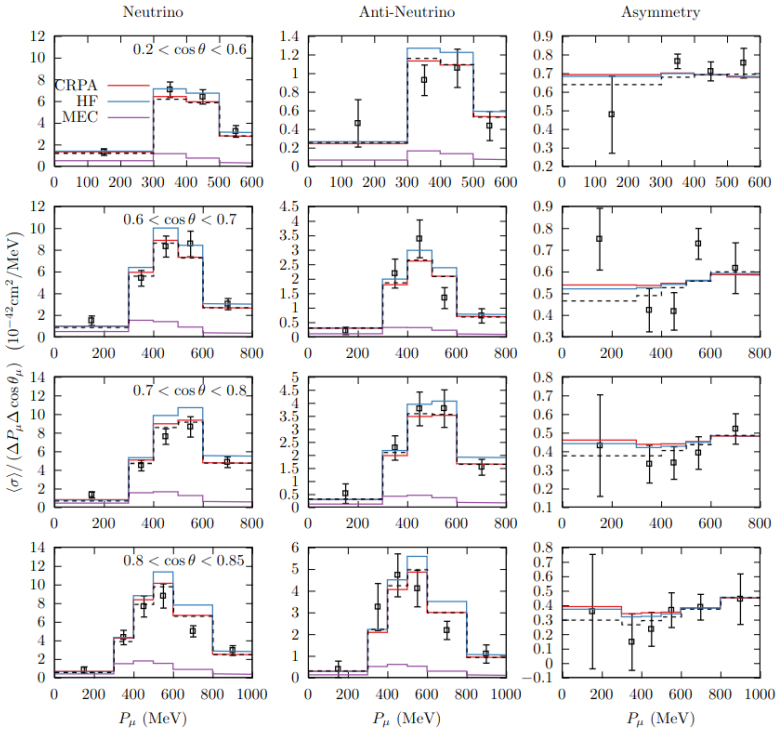
Datasets	χ^2_{2010}	$\chi^2_{2021(\text{Global})}$	$\chi^2_{2021(^2\text{H})}$	Degrees of freedom
All Data in tune	486	242	410	109
^{2}H Data in tune	230	105	37	52
H Data in tune	256	138	374	57

Uses for neutrino data

Oscillation/BSM analysis → not discussed

Fitting & parameter searches

What I use data for:



[Arxiv:2110.11321]

[S. Dolan e.a. 2110.14601]

	Carbon and oxygen	ν_μ and $\bar{\nu}_\mu$
Number of bins	58	116
HF-CRPA	135	740
HF	143	683
SuSAv2	140	741
LFG-RPA	59	446
LFG (no RPA)	184	1028

Data-model comparisons

It's nice to have a number…

But doesn't teach me much

Fermilab

Uses for neutrino data

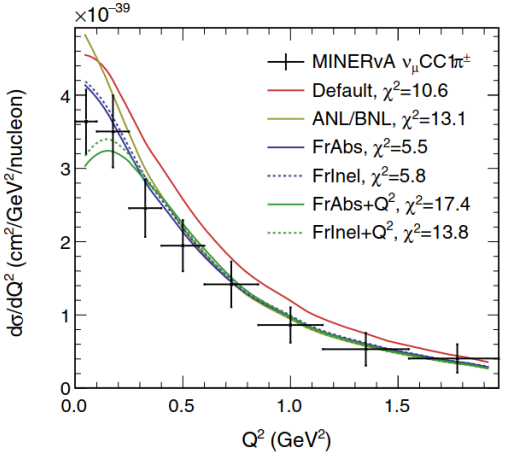
Oscillation/BSM analysis → not discussed

Fitting & parameter searches → Need goodness of fit metric

Validation of methods/models → nice to have a number

Interpretation of results!

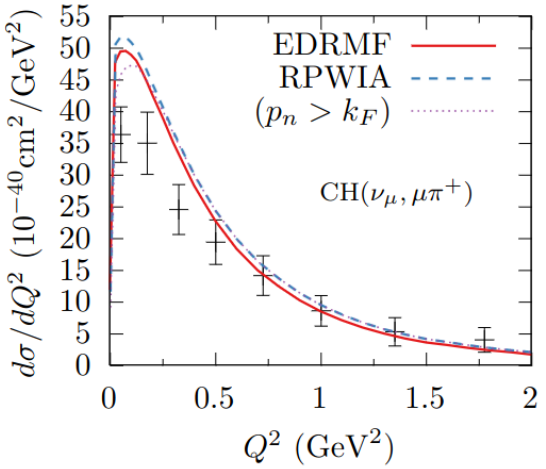
MINERvA 1pi$^+$ data



[Stowell et al. PhysRevD.100.072005]

Combined fit to MINERvA & ANL/BNL points to discrepancies

Calculations!



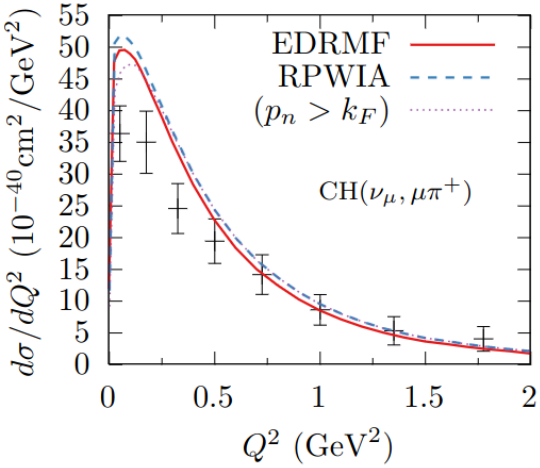
[A.N. et al. PhysRevD.107.053007]

Uses for neutrino data

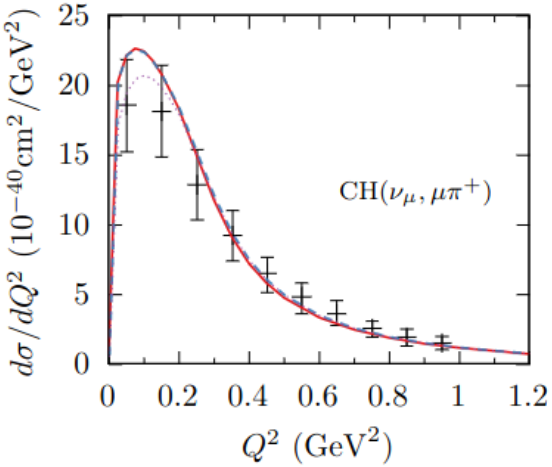
[A.N. et al. PhysRevD.107.053007]

Interpretation of results!

MINERvA 1pi⁺ data



T2K 1pi⁺ data



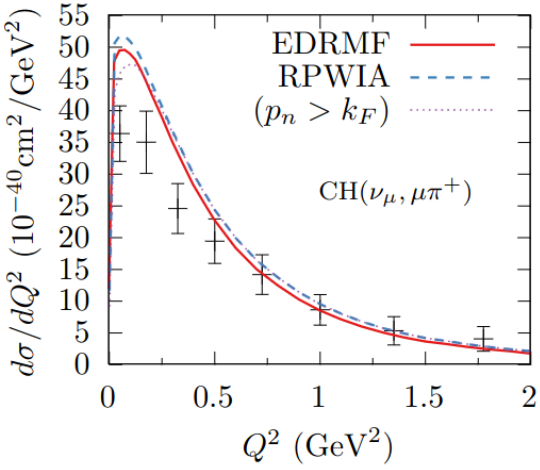
We see some possible energy-dependence

Uses for neutrino data

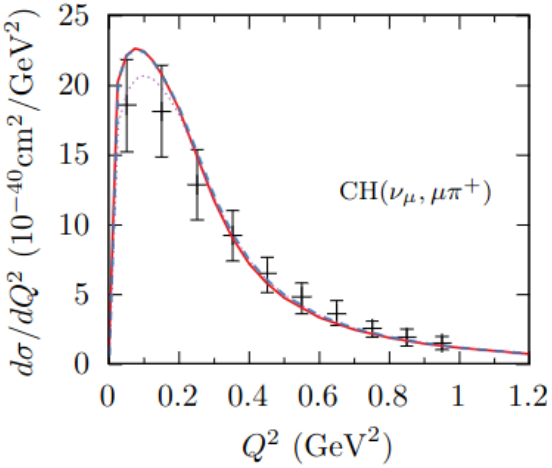
[A.N. et al. PhysRevD.107.053007]

Interpretation of results!

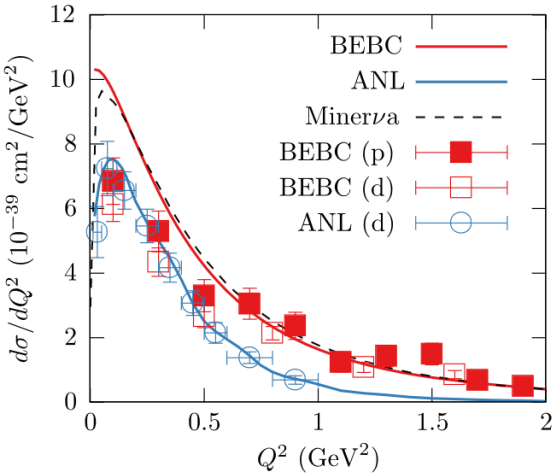
MINERvA 1pi⁺ data



T2K 1pi⁺ data



nucleon data



We conclude/interpret/assert/assume/propose

The delta/BG coupling is not properly constrained by ANL data

→ Phd project M. Hooft (Ugent) on improvements of nucleon-level couplings

How we calculate cross sections

$$\langle\sigma\rangle = \frac{\int_V dV \int dE\ \Phi(E_\nu)\frac{d\sigma(E)}{dV}}{\int dE\Phi(E)}$$

V: phase space in measured kinematics

ϕ : Neutrino flux
→ We need these from experiment
→ Turns out its trickier:
[Koch,Dolan, PRD 102.113012]
But I can deal with errors on the flux

σ : ‘Nature’
→ Might be wrong, but supposed to be nature

In a ‘perfect’ experiment this is exactly the event rate

$$\langle\sigma\rangle = \frac{N}{\int dE\Phi(E)}$$

How we calculate cross sections

$$\langle\sigma\rangle = \frac{\int_V dV \int dE\ \Phi(E_\nu)\frac{d\sigma(E)}{dV}}{\int dE\Phi(E)}$$

V: phase space in measured kinematics

ϕ : Neutrino flux

→ We need these from experiment

σ : ‘Nature’
→ Might be wrong, but supposed to be nature

In a ‘perfect’ experiment this is exactly the event rate

$$\langle\sigma\rangle = \frac{N}{\int dE\Phi(E)}$$

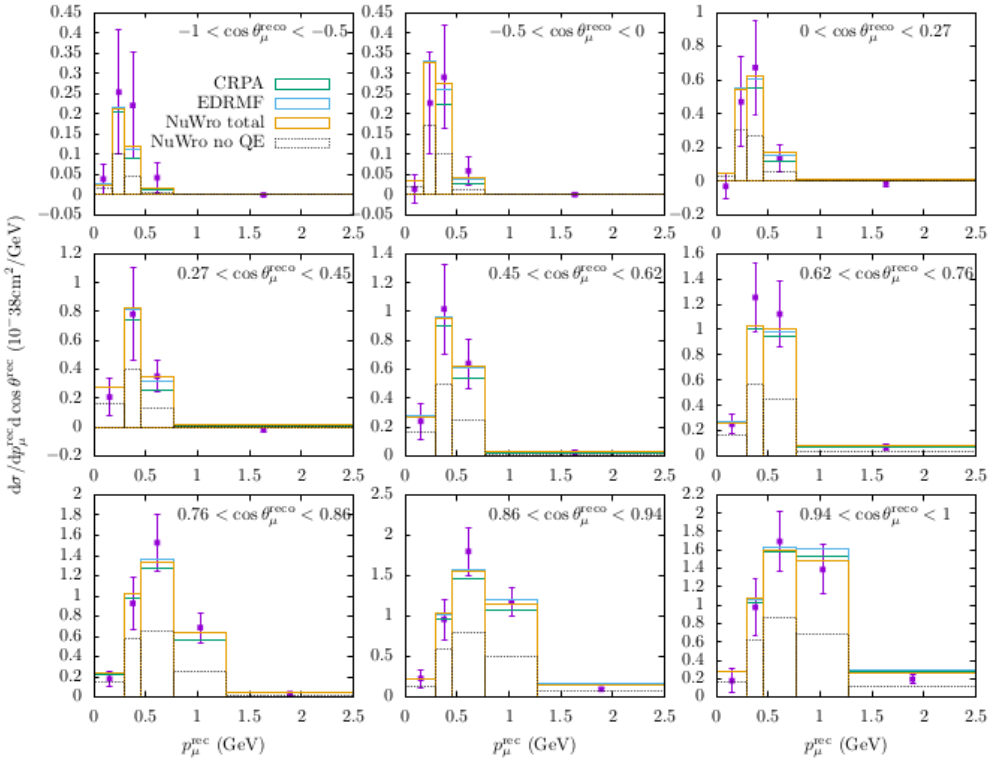
Turns out the perfect experiment doesn’t exist
Efficiencies, smearing, error propagation, flux uncertainties, …
Should be documented
But we must / can only assume that this is taken into account in error

How we calculate cross sections

Turns out the perfect experiment doesn't exist
Efficiencies, smearing, error propagation, flux uncertainties, …
Should be documented
But we must / can only assume that this is taken into account in error

Or… Forward folding ?

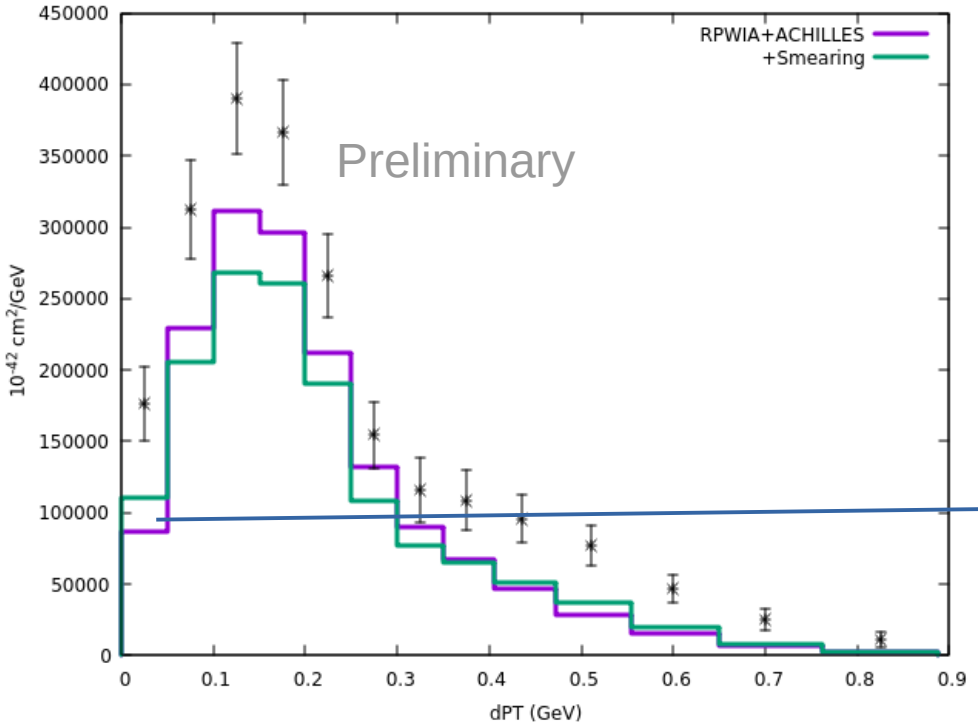
μbooNE data [PRL 123, 131801]



How we calculate cross sections

Turns out the perfect experiment doesn't exist
Efficiencies, smearing, error propagation, flux uncertainties, …
Should be documented
But we must / can only assume that this is taken into account in error

Or… Forward folding ? Wiener SVD + smearing ? μbooNE data [Arx:2301.03700]



Straightforward to compute

But…

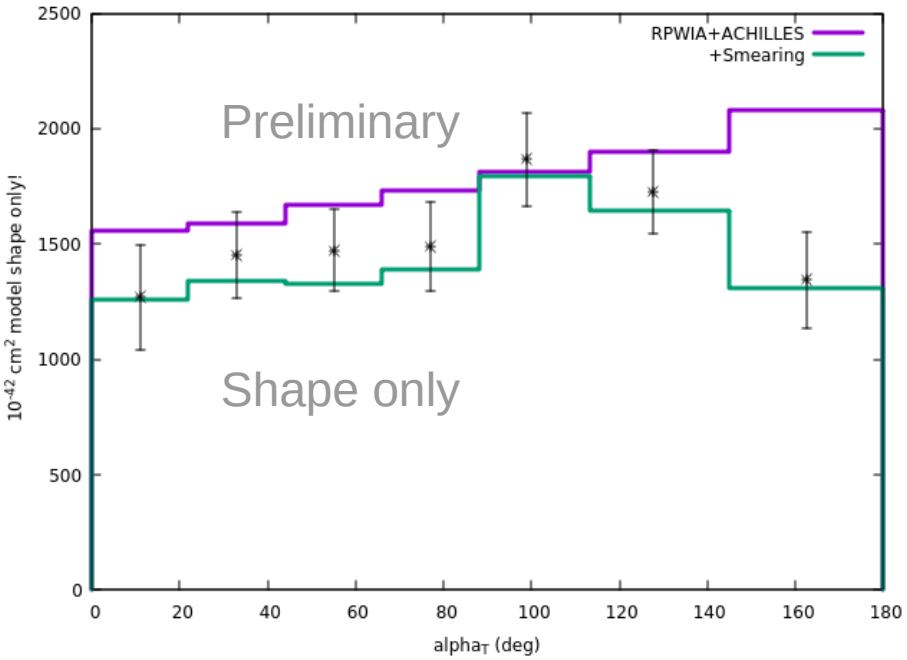
Interpretation is sometimes lost

Not really low-dP_T
Smearing non-negligible up to 0.3 GeV

How we calculate cross sections

Turns out the perfect experiment doesn't exist
Efficiencies, smearing, error propagation, flux uncertainties, …
Should be documented
But we must / can only assume that this is taken into account in error

Or… Forward folding ? Wiener SVD + smearing ? μbooNE data [Arx:2301.03700]



Interpretation is sometimes lost

I can compare true vs smeared
Can check if smearing is diagonal

Does smearing become diagonal with better statistics ?

If smearing is diagonal, there is no point to it

What we need from a data-release that is ‘nature’

~~‘Signal definition’~~
Moved to definition of topologies, instead of interaction mechanisms.
‘CC0pi’ is **not** clear: This is an analysis that ‘includes’ pions
→ Need to know how pions are (not) detected, thresholds

1.) Need full kinematic **phase space**: example CC1p0piX (MicroBooNE)

”1 single proton with 300 < k_p < 1000 MeV.

No charged π with $k_\pi > 70$ MeV

No neutral π at all”

Clear! but complete ?!
+A description of the methods on how particles are distinguished/detected

2.) Flux: Need the flux as function of energy.
Need the total flux assumed in experiment
Ideally provide errorbars on the flux

What we need from a data-release that is ‘nature’

~~‘Signal definition’~~
Moved to definition of topologies, instead of interaction mechanisms.
‘CC0pi’ is **not** clear: This is an analysis that ‘includes’ pions
→ Need to know how pions are (not) detected, thresholds

1.) Need full kinematic **phase space**: example CC1p0piX (MicroBooNE)

”1 single proton with 300 < k_p < 1000 MeV.

No charged π with $k_\pi > 70$ MeV

No neutral π at all”

Clear! but complete ?!
+A description of the methods on how particles are distinguished/detected

What to avoid (from worst to least bad)

1. Unreproducable model-dependent assumptions
2. reproducable model-dependent assumptions
3. model-dependent assumptions

1. Unreproducable model-dependent assumptions
2. reproducable model-dependent assumptions
3. model-dependent assumptions

These often come from:

- **Cuts on or measurements of unmeasurable variables**
- **Background subtraction**
- **Sideband fitting**
- **Reconstructing ‘true’ variables from a MC simulation**

Example: Unfolding into unmeasurable variables: MiniBooNE

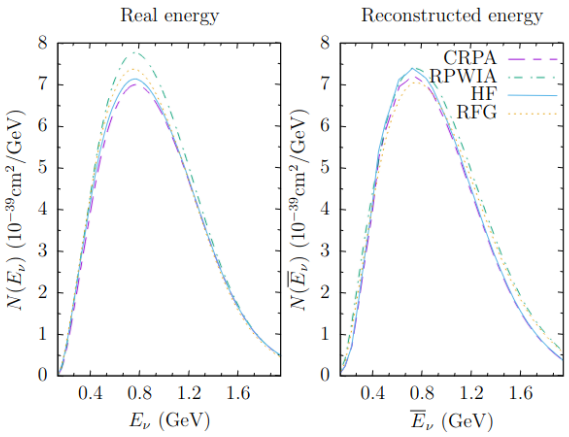
$$E_\nu^{\mathrm{QE}} = \frac{2(M_n')E_\mu - ((M_n')^2 + m_\mu^2 - M_p^2)}{2\cdot[(M_n') - E_\mu + \sqrt{E_\mu^2 - m_\mu^2}\cos\theta_\mu]}, \qquad (1)$$

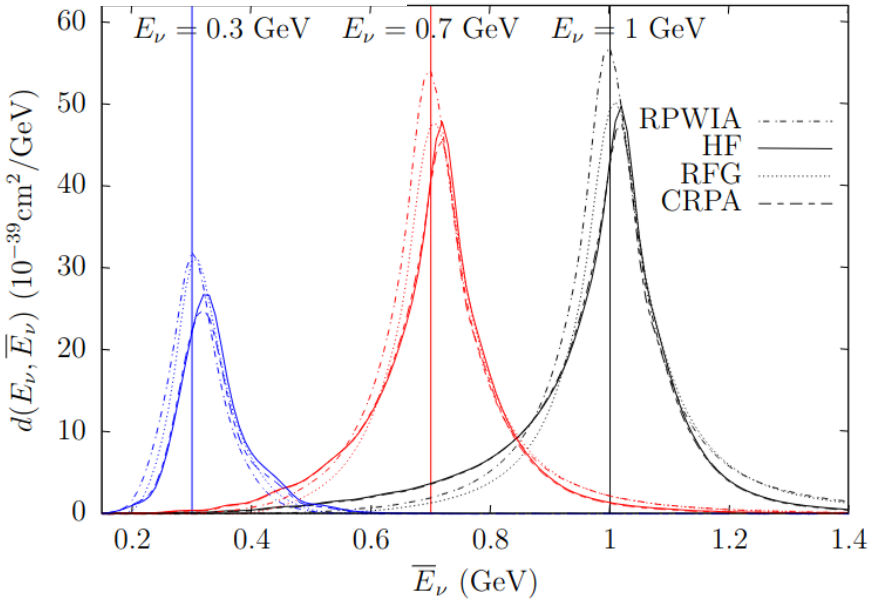
E^{QE} is a model independent kinematic variable, function of measurable kinematics
Very straightforward to compute:

$$\frac{d\sigma}{dE_l\, d\overline{E}_\nu} = \left|\frac{\partial\cos\theta}{\partial\overline{E}_\nu}\right| \frac{d\sigma}{d\omega\, d\cos\theta}\Bigg|_{\omega=E_\nu-E_l,\cos\theta(E_l,\overline{E}_\nu)}$$

[Phys. Rev. C 98, 054603 (2018)]

The cross section as function of E^{QE}
Is of course model-dependent





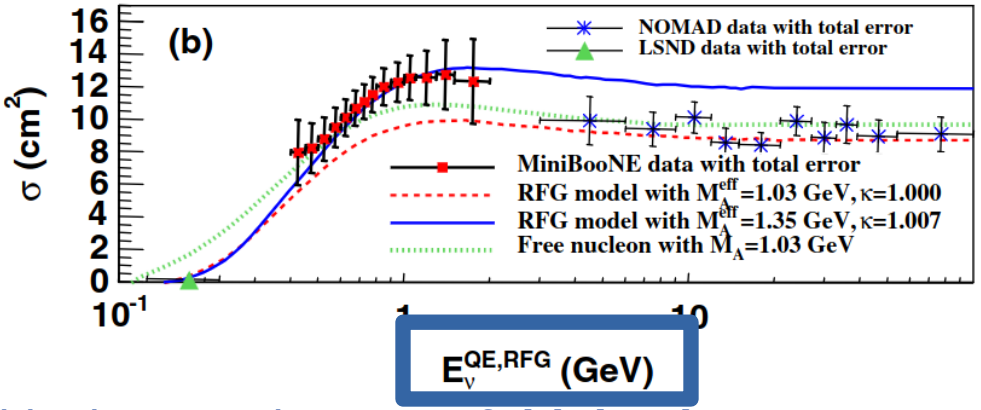
Example: Unfolding into unmeasurable variables: MiniBooNE

$$E_\nu^{QE} = \frac{2(M_n')E_\mu - ((M_n')^2 + m_\mu^2 - M_p^2)}{2 \cdot [(M_n') - E_\mu + \sqrt{E_\mu^2 - m_\mu^2}\cos\theta_\mu]}, \quad (1)$$

E^{QE} is a model independent kinematic variable, function of measurable kinematics

MiniBooNE did **not** publish the CS
As function of E^{QE} !

Instead a **flux-unfolded** cross section
Using a **RFG model** is used



To compare a theoretical model to this data one has to **unfold the theory**

$$\sigma(E_\nu^{uf}) = \int d\overline{E}_\nu \left\{ \int dE_\nu \Phi(E_\nu) d(E_\nu, \overline{E}_\nu) \right\} \times \left[\frac{\frac{d\sigma}{d\overline{E}_\nu}(E_\nu, \overline{E}_\nu)}{\int dE_\nu \Phi(E_\nu) \frac{d\sigma}{d\overline{E}_\nu}} \right]^{RFG}$$

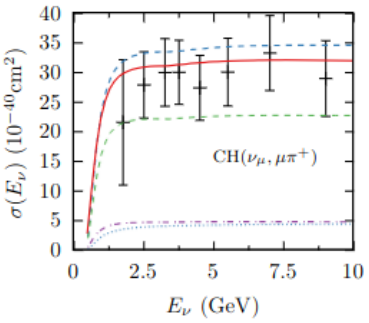
[Nieves et al., PhysRevD.85.113008]

In conclusion: the unfolding is a **barrier**, similar conclusions on 2p2h could have probably been obtained from the model-independent cross section in terms of E^{QE}

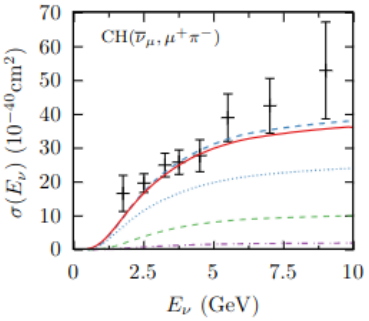
Measurements of unmeasurable variables: E_ν

The most popular unmeasurable variable seems to be $\sigma(E_\nu)$

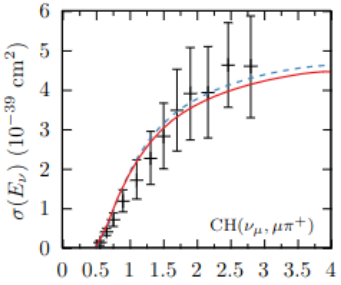
Minerva



Minerva



T2K



[A.N. et al. PhysRevD.107.053007]

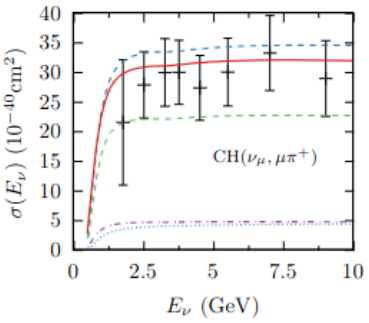
Why ?

"Allows for inter-experiment comparison"
→ Untrue, all these data have different kinematic cuts

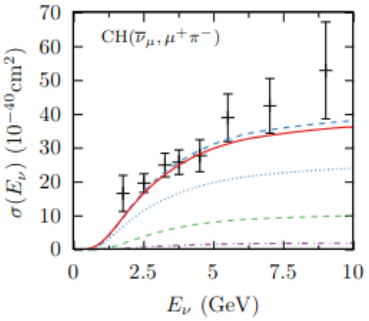
Measurements of unmeasurable variables: E_ν

The most popular unmeasurable variable seems to be $\sigma(E_\nu)$

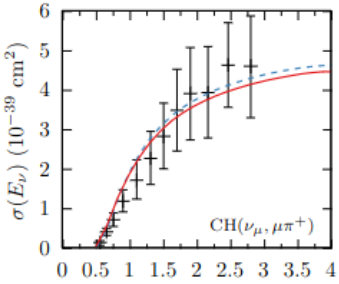
Minerva



Minerva



T2K



[A.N. et al. PhysRevD.107.053007]

Why ?

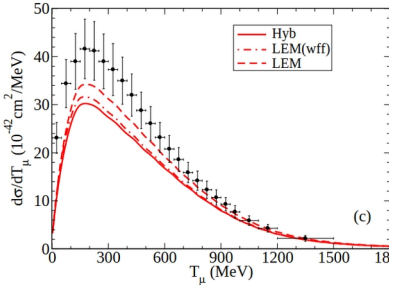
"Allows for inter-experiment comparison"

→ Untrue, all these data have different kinematic cuts

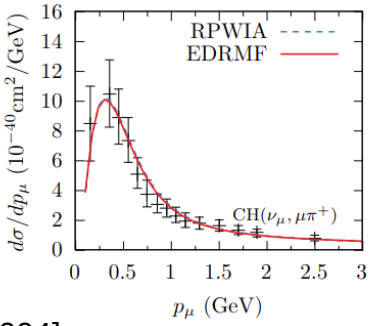
→ Might as well look at flux-folded results for p_μ

for inter-experiment comparison: **the model-data comparison** will show inconsistency

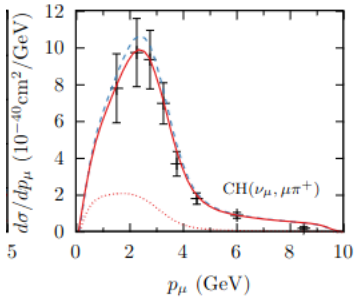
MB



T2K



Minerva



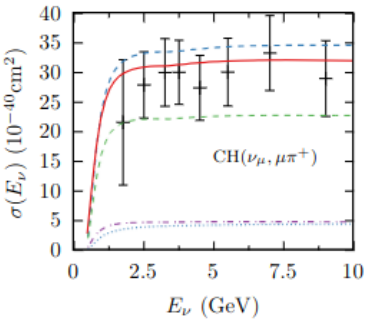
[RGJ et al. PhysRevD.97.013004]

Fermilab

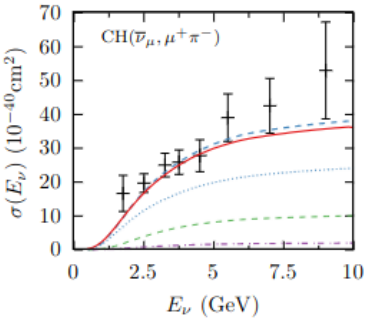
Measurements of unmeasurable variables: E_ν

The most popular unmeasurable variable seems to be $\sigma(E_\nu)$

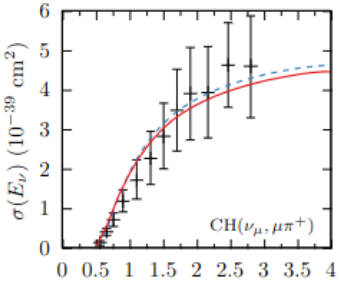
Minerva



Minerva



T2K



[A.N. et al. PhysRevD.107.053007]

σ(E) 'measurements' are:

At-best: Indication that your model has a similar energy-dependence as mine
→ Might as well compare the models

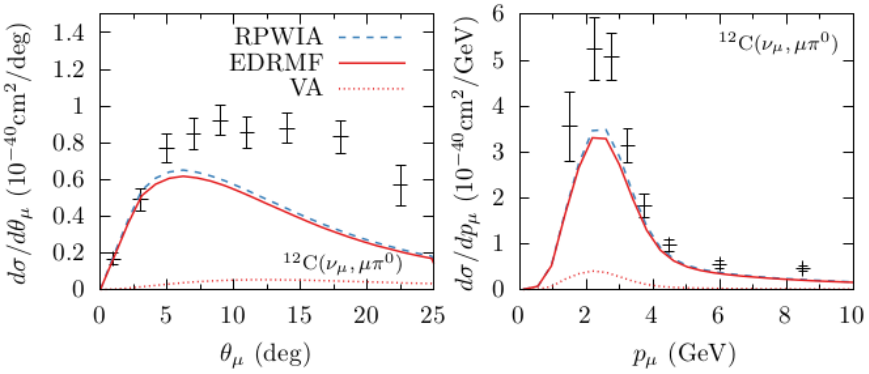
Usually: not taken seriously

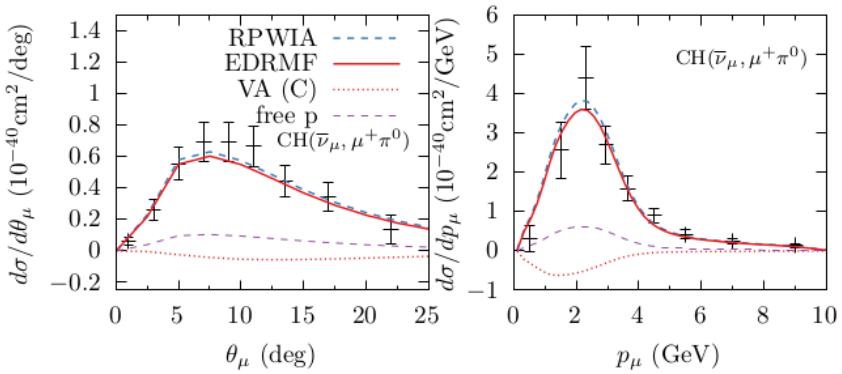
At-worst: misleading and confusing

σ(E) 'measurements' are not data but analysis
And we should start treating them as such

1. Don't put it in a data-release
2. Release the σ(E) of the used model
3. Release the full model + analysis used
4. Write a separate paper and see if it passes review

Sidebands and background subtraction





Neutral π^0 on carbon

Anti-neutrino & neutrino responses are very directly connected by isospin
→ same d.o.f

Underpredict neutrino ~factor 2
Agree well with anti-neutrino

→ Interesting puzzle

But…

Fermilab

Sidebands and background subtraction

lected signal events. The data events that populate these "sideband samples" do not appear in the selected signal sample as they do not satisfy one or more of the event selections. The individual sidebands have different mixtures of the background categories so that a combined fit has sensitivity to the normalizations of all categories.

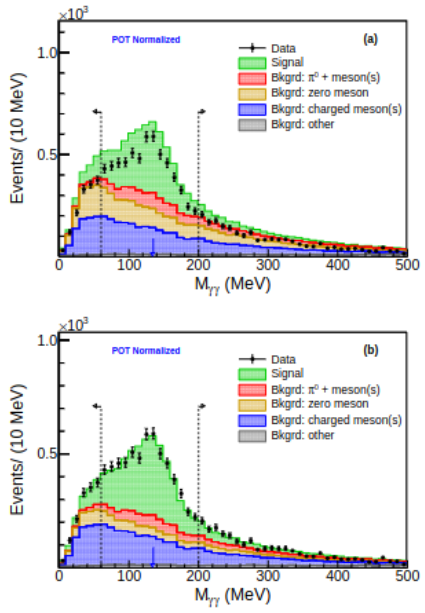


FIG. 6. Data versus MC $\gamma\gamma$ invariant mass distributions before (a) and after (b) revision of the background normalizations to match the values obtained by the overall fit to the four data sideband samples. The low-side and high-side $M_{\gamma\gamma}$ sidebands, denoted by arrows in the plot, are fitted together with the Michel-tag and low-proton-score sidebands, while the data of the signal region between 60 MeV and 200 MeV are excluded. The MC agreement with the data in the signal region improves dramatically as the result of constraining the backgrounds in the sideband regions.

The analysis uses four separate sideband samples to achieve good constraints on the normalizations of the

A third sideband contains events tagged as having Michel EM showers from endpoint π^+ decays. This Michel sideband (1803 events) has abundant charged-meson(s) background but also contains a sizable π^0+meson(s) contribution. There are very few zero-meson events in the Michel sideband. In the fourth sideband, the muon is accompanied by a second reconstructed track which has a low likelihood score for the proton hypothesis (3933 events). This low-proton-score sideband is also predicted to be mostly composed of charged meson(s) plus π^0+meson(s) backgrounds but with their apportionment differing somewhat from that in the Michel-tag sideband. The estimated compositions of the latter sidebands after background tuning by the fit-to-sidebands described below, are displayed as component histograms in Figs. 7a and 7b.

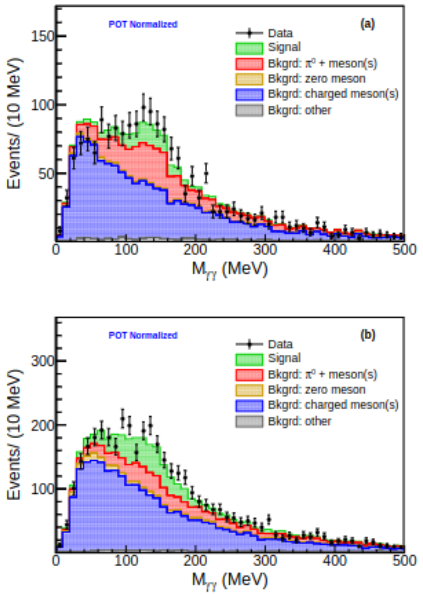


FIG. 7. Comparisons of $M_{\gamma\gamma}$ distributions of data (solid points) to predictions of the reference MC (histograms), for

BG subtraction is often neccesary, but ...
Measurement of single-π^0

Subtractions made of
- multi-π^0
- multi-π
- …

Should I trust the GENIE MC to Correctly describe **more complicated processes** in order to get the data which we still cannot reproduce ?

Then why would I worry ?
Can just compare to GENIE

Sidebands and background subtraction

lected signal events. The data events that populate these "sideband samples" do not appear in the selected signal sample as they do not satisfy one or more of the event selections. The individual sidebands have different mixtures of the background categories so that a combined fit has sensitivity to the normalizations of all categories.

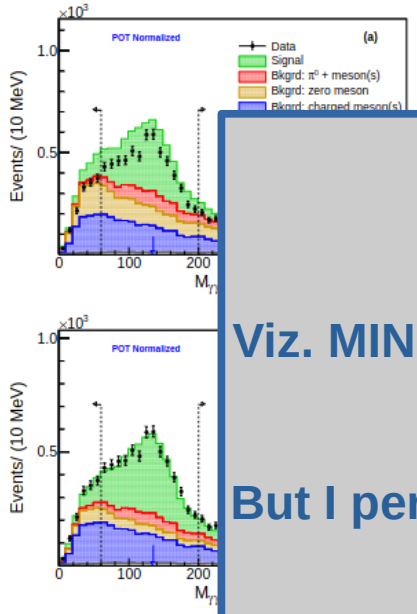


FIG. 6. Data versus MC $\gamma\gamma$ in
fore (a) and after (b) revision
tions to match the values obta
four data sideband samples. Th
sidebands, denoted by arrows i
with the Michel-tag and low-
the data of the signal region b
are excluded. The MC agreeme
region improves dramatically as
backgrounds in the sideband regions.

The analysis uses four separate sideband samples to achieve good constraints on the normalizations of the

A third sideband contains events tagged as having Michel EM showers from endpoint π^+ decays. This Michel sideband (1803 events) has abundant charged-meson(s) background but also contains a sizable π^0+meson(s) contribution. There are very few zero-meson events in the Michel sideband. In the fourth sideband, the muon is accompanied by a second reconstructed track which has a low likelihood score for the proton hypothesis (3933 events). This low-proton-score sideband is also predicted to be mostly composed of charged meson(s) plus π^0+meson(s) backgrounds but

FIG. 7. Comparisons of $M_{\gamma\gamma}$ distributions of data (solid points) to predictions of the reference MC (histograms), for

BG subtraction is often neccesary, but ...
Measurement of single-π^0

Subtractions made of

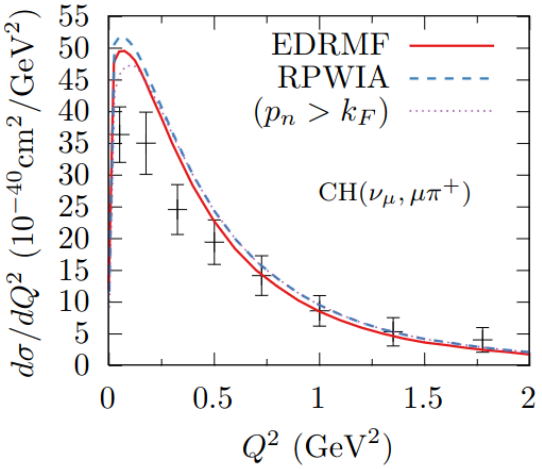
The data is still useful
Viz. MINERvA fits in [Stowell et al. PRD 100 072005]

But I personally cannot trust it to draw conclusions
(I still do it, but shouldn't)

GENIE MC to
more
cesses in
ata which we
luce ?

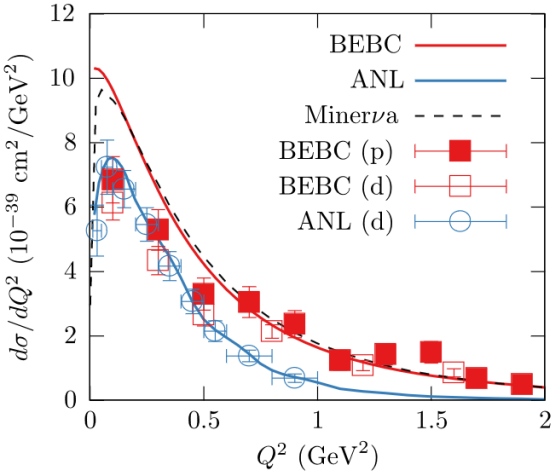
worry ?
to GENIE

Model-dependent cuts MINERvA and ANL pi+



1) Cuts on E_ν
2) Reconstruction of W^{free} from MC
3) Sideband subtractions
4) BG subtractions model-dependent W

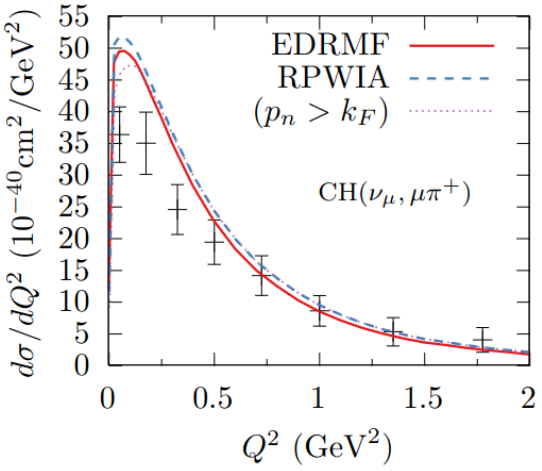
But I don't have anything else...

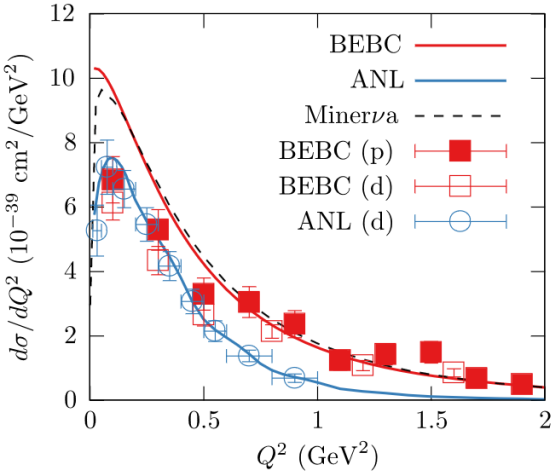


1) Flux uncertainties
2) Unknown deuteron corrections
3) No idea how they actually did this

But I don't have anything else...

Model-dependent cuts MINERvA and ANL pi+





Thinking about the future

If

(do this please!!)

new bubble chamber data comes
→ We will dismiss the old data

What about MINERvA ?
If we understand modeling better
When new data comes out

1) Either it was correct to begin with
2) Or we will dismiss the old data

→ Need a clear 'raw data' preservation plan

(MINERvA does this, but everyone should)

Conclusions

- Calculations for CS (if ‘nature’) are hard, but in principle straightforward
  Need **all** phase-space cuts & the flux + total flux
  Do not refrain from complicated signals for which models don’t work, we should figure it out in the future

- Forward folding/smearing is straightforward, but we lose interpretability

- Model dependence leads to ‘analysis’ instead of ‘data’
  It is a thin line oftentime

- **Why** are the model-dependent analysis used ?
  → Trying to get a ‘interaction mechanism’ (e.g. MB analysis)
  → Trying to get a ‘clean’ kinematic range (e.g. W cuts in Minerva)

  This is **fine** and can be useful, but what is done should be clear
  This is most likely **not future proof**
  → **How to preserve data so that it can be used in the future ?**