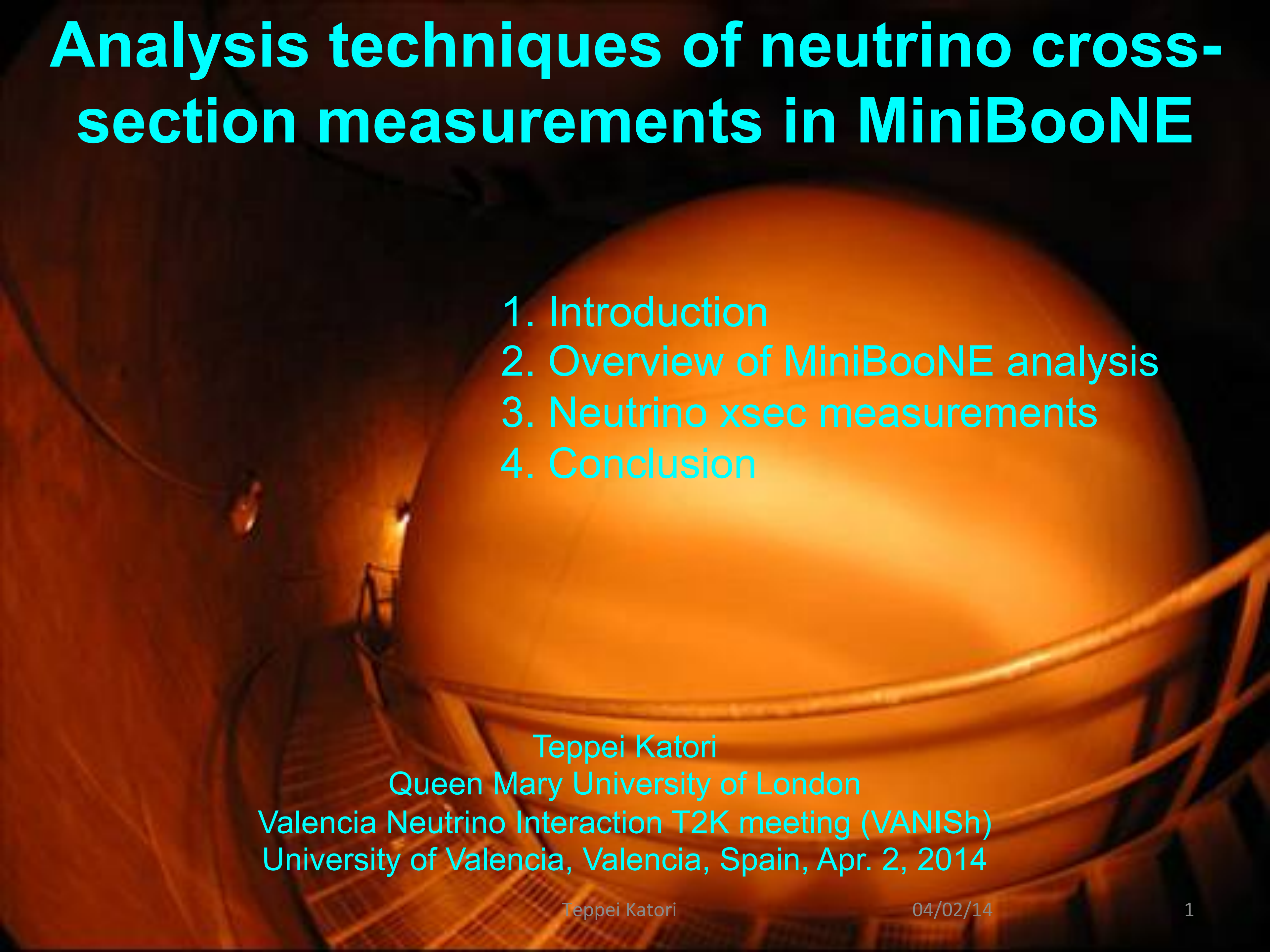
Analysis techniques of neutrino cross-section measurements in MiniBooNE



1. Introduction
2. Overview of MiniBooNE analysis
3. Neutrino xsec measurements
4. Conclusion

Teppei Katori
Queen Mary University of London
Valencia Neutrino Interaction T2K meeting (VANISh)
University of Valencia, Valencia, Spain, Apr. 2, 2014

1. Introduction
2. Overview of MiniBooNE analysis
3. Neutrino cross section measurements
4. Conclusion

PPD neutrino talk (2010), http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=5571

1. Introduction

Purpose of this talk

- This talk is prepared for students who want to measure neutrino cross-sections for their PhD theses.

- It is especially focusing on absolute flux-integrated topological differential cross section measurement [1]. The most important model-independent result to provide to the community.

absolute: normalization is specified
flux-integrated: neutrino flux shape is not unfolded
topological: interaction is defined from final state particles
differential: cross section is function of measured kinematic variables

- MiniBooNE developed number of techniques necessary to measure these, and this talk covers technical aspects from the CCQE, NCEL, NC1π^o, CC1π^+, CC1π^o, antiCCQE, antiNCEL. These are good reference but not the best, T2K should perform even better analyses!

[1] Can anybody invent a better name for this?

1. Introduction

Goal of neutrino cross section measurement

- Goal is to measure model-independent cross section as much as possible. This is what theorists want to study their models.

Model-independent cross-section is

- absolute (flux is not tuned from own measurement)
- the dependence of signal channel MC is minimum
- detector efficiency must be corrected (so it is detector model-dependent)
- no assumption on kinematics (cf neutrino energy reconstruction with lepton kinematics assume CCQE interaction and neutron at rest), which often means cross-section is function of measured variables (differential cross-section)

Formula of flux-integrated differential cross-section

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i (\Phi T) \Delta x_i}$$

1. Introduction

2. Overview of MiniBooNE analysis

3. Neutrino cross section measurements

4. Conclusion

2. Overview of MiniBooNE xs measurements

CCQE (CC0π)
PRD81(2010)092005
FERIMILAB-THESIS-2008-64

$$\nu_\mu + n \rightarrow p + \mu^-$$

$$(\nu_\mu + {}^{12}C \rightarrow X + \mu^-)$$

Signal definition: 1 μ + 0 π + N protons

Why we measure
- Test CCQE models

Why MiniBooNE measure
- Largest sample (~40%) to test detector efficiency, veto efficiency, event uniformity, timing, etc
- Best sample to study ν_eCCQE kinematics (=oscillation signal)
- ν_μCCQE to constraint ν_e from μ-decay in oscillation sample

It is important to measure CCQE!

CCQE

by Denis Perevalov

2. Overview of MiniBooNE xs measurements

MiniBooNE collaboration,
PRD82(2011)092005

NCEL
PRD82(2011)092005
FERIMILAB-THESIS-2009-47

$$\nu_\mu + p \rightarrow \nu_\mu + p$$

$$\nu_\mu + n \rightarrow \nu_\mu + n$$

Signal definition: 0 μ + 0 π + N protons

Why we measure
- Additional test of CCQE models
- Measurement of Δs
    - value is still controversial
    - connection of form factor (elastic) and PDF (inelastic)

$$\int_0^1 dx \Delta s(x) \equiv \Delta s \equiv G_A^s(Q^2 = 0)$$

Why MiniBooNE measure
- NCEL to constrain oil optical property

  oil optical property is the largest detector systematics. NCEL was used to assign variation.

It is important to measure NCEL!

CCQE NCEL

by Colin Anderson

2. Overview of MiniBooNE xs measurements

MiniBooNE collaboration,
PRD81(2010)013005

NC1π^o
PRD81(2010)013005
FERIMILAB-THESIS-2010-49

$$\nu_\mu + N \rightarrow \nu_\mu + \Delta^o \rightarrow \nu_\mu + N + \pi^o$$

$$\nu_\mu + A \rightarrow \nu_\mu + A + \pi^o$$

Signal definition: 0 μ + 1 π^o + N protons

Why we measure

- The biggest misID background for ν_e appearance experiments
  all oscillation experiments perform internal measurement to constrain

Why MiniBooNE measure

- π^o mass peak for energy calibration
- measured π^o kinematics is used to correct simulation

It is important to measure NC1π^o!

CCQE NCEL NC1π^o

by Mike Wilking

2. Overview of MiniBooNE xs measurements

MiniBooNE collaboration,
PRD83(2011)052007

CC1π^+
PRD83(2011)052007
FERIMILAB-THESIS-2009-27

$$\nu_\mu + p(n) \rightarrow \mu + \Delta^{+(+)} \rightarrow \mu + p(n) + \pi^+$$

$$\nu_\mu + A \rightarrow \mu + A + \pi^+$$

Signal definition: 1 μ + 1 π^+ + N protons

Why we measure
- The biggest misID background for ν_μ disappearance experiments
- kinematic distortion by this background must be understood

Why MiniBooNE measure
- highest purity channel (~90%)
  Michel electron tagging achieve extremely pure sample
- Constrain wrong sign background in anti-neutrino mode
  CC1π^+ in anti-ν mode is definitely from ν-contamination

It is important to measure CC1π^+!

CCQE NCEL NC1π^o CC1π^+

Queen Mary
University of London

by Bob Nelson

2. Overview of MiniBooNE xs measurements

MiniBooNE collaboration,
PRD83(2011)052009

CC1π^o
PRD83(2011)052009
FERIMILAB-THESIS-2010-9

$$\nu_\mu + n \rightarrow \mu + \Delta^+ \rightarrow \mu + p + \pi^o$$

Signal definition: 1 μ + 1 π^o + N protons

Why we measure

- There is no coherent channel, so it is useful to understand coherent/resonance Δ production

Why MiniBooNE measure

- The last possibly measurable channels, by this, MiniBooNE measure 90% of interaction cross sections in ν-mode

It is important to measure CC1π^o!

CCQE | NCEL | NC1π^o | CC1π^+ | CC1π^o

Queen Mary
University of London

by Joe Grange

2. Overview of MiniBooNE xs measurements

MiniBooNE collaboration,
PRD88(2013)032001

anti-CCQE
PRD88(2013)032001
FERIMILAB-THESIS-2013-14

$$\bar{\nu}_\mu + p \rightarrow n + \mu^+$$

$$(\bar{\nu}_\mu + {}^{12}C \rightarrow X + \mu^+)$$

$$(\bar{\nu}_\mu + {}^{1}H \rightarrow n + \mu^+)$$

Signal definition: 1 μ + 0 π + N protons

Why we measure
- Additional test of CCQE models
- necessary measurement for CPV measurement

Why MiniBooNE measure
- To understand anti-ν_eCCQE interaction kinematics (oscillation signal)
- anti-ν_μCCQE constraint for anti-ν_e from μ-decay
- Tune wrong sign components in anti-ν mode beam

It is important to measure antiCCQE!

CCQE | NCEL | NC1π^o | CC1π^+ | CC1π^o | anti CCQE by Ranjan Dharmapalan

2. Overview of MiniBooNE xs measurements

MiniBooNE collaboration,
arXiv:1309.7257

anti-NCEL
arXiv:1309.7257
FERIMILAB-THESIS-2012-29

$$\bar{\nu}_\mu + p \rightarrow \bar{\nu}_\mu + p$$

$$\bar{\nu}_\mu + n \rightarrow \bar{\nu}_\mu + n$$

Signal definition: 0 μ + 0 π + N protons

Why we measure

- Complete all 4 QE measurements (CCQE, NCEL, antiCCQE, antiNCEL)

Why MiniBooNE measure

- test beam-dump mode run (dark matter search)
  anti-ν mode beam was used to test MiniBooNE dark matter sensitivity. antiNCEL was the biggest background.

It is important to measure antiNCEL!

CCQE NCEL NC1π^o CC1π^+ CC1π^o anti CCQE anti NCEL

Queen Mary
University of London

2. Overview of MiniBooNE xs measurements

To be published

- CC inclusive cross section
- 1 μ + 0 π +1 proton (2 track CCQE)

Possibly measured in T2K, but not in MiniBooNE

NC1π^+ (0 μ + 1 π^+ + N protons)

- another resonance only channel

1 μ + 1 π^+ + 1 proton

- 3 track CC1π^+, high pure Δ^{++} measurement

1 μ + 0 π + 2 protons

- high pure MEC?

CCQE NCEL NC1π^o CC1π^+ CC1π^o anti CCQE anti NCEL

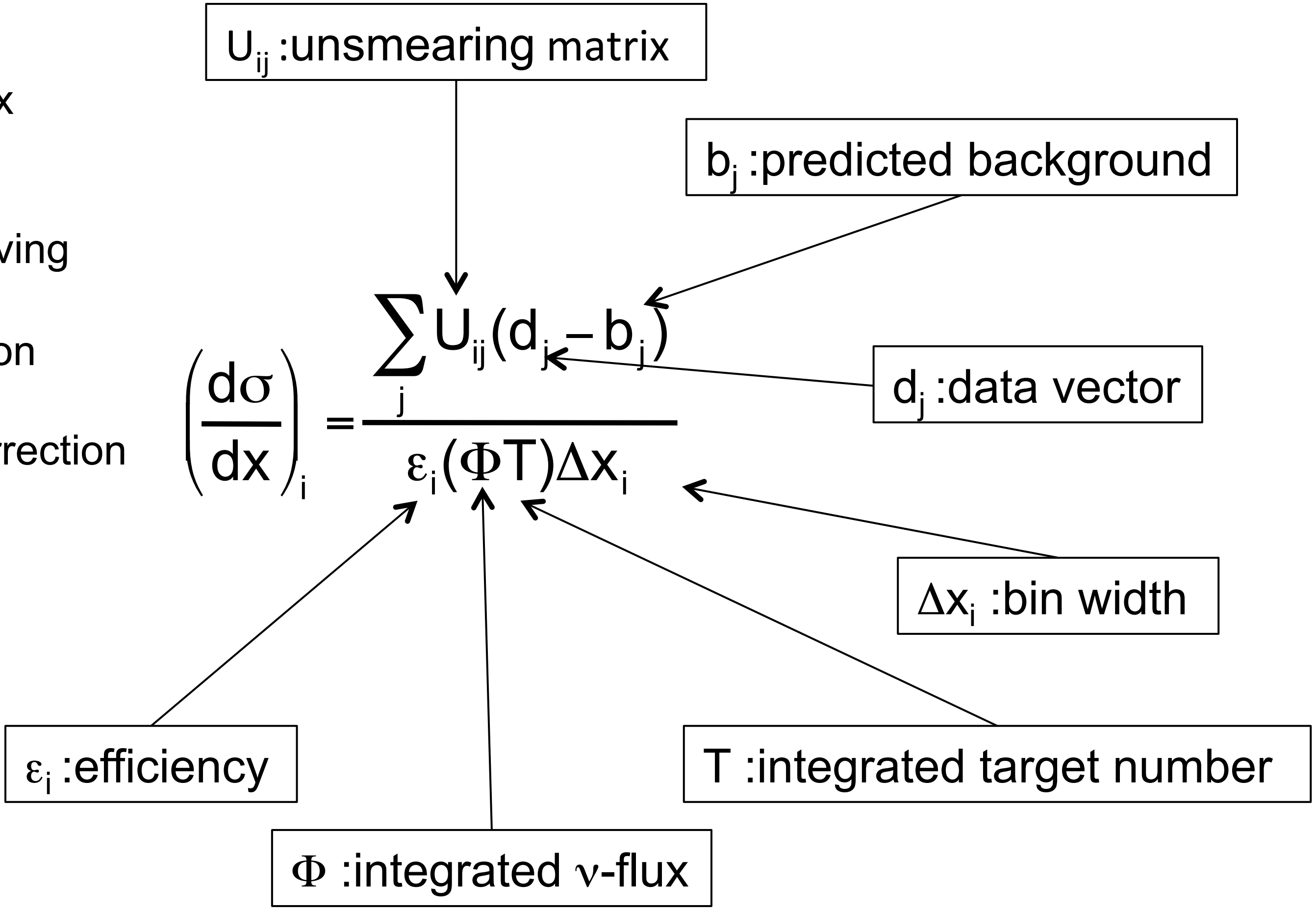
1. Introduction
2. Overview of MiniBooNE analysis
3. Neutrino cross section measurements
4. Conclusion

3. Neutrino cross-section measurements

Absolute flux-integrated topological differential cross section formula

i :true index
j : reconstructed index

3.1 Signal definition
3.2 Background removing
3.3 Unsmearing
3.4 Efficiency correction
3.5 Flux correction
3.6 Target number correction
3.7 Binning
3.8 Systematic errors
3.9 Data format



$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.0 Before you start cross section analysis...

Reconstructions and cuts

- You have good reconstruction and all cuts to select your data sample, congratulations, you are ready to measure cross sections!

(you can spend another year in grad school from there!)

3.0 Event selection

Reconstructions and cuts

- You have good reconstruction and all cuts to select your data sample, congratulations, you are ready to measure cross sections!

(you can spend another year in grad school from there!)

Good sample

- depends
- reasonable statistics to make few bins
- measure distributions as many as possible
- Purity ~ 50% or more (improving purity → improving systematics)

Before you start systematics analysis, you should know rough total final error.

- lose/tight cuts
- change bin size
- sideband constraint
- which differential cross section to measure

3.0 Event selection

Reconstructions and cuts

- You have good reconstruction and all cuts to select your data sample, congratulations, you are ready to measure cross sections!

(you can spend another year in grad school from there!)

Good sample

- depends
- reasonable statistics to make few bins
- measure distributions as many as possible
- Purity ~ 50% or more (improving purity → improving systematics)

Before you start systematics analysis, you should know rough total final error.

- lose/tight cuts
- change bin size
- sideband constraint
- which differential cross section to measure

We don't expect more statistics in near future, and most of analyses will be ~ 10-20% error.
Any cross section data from T2K are unique and precious. WE SHOULD PUBLISH ASAP!

3.1 Signal definition

$\boxed{CC1\pi^o}$

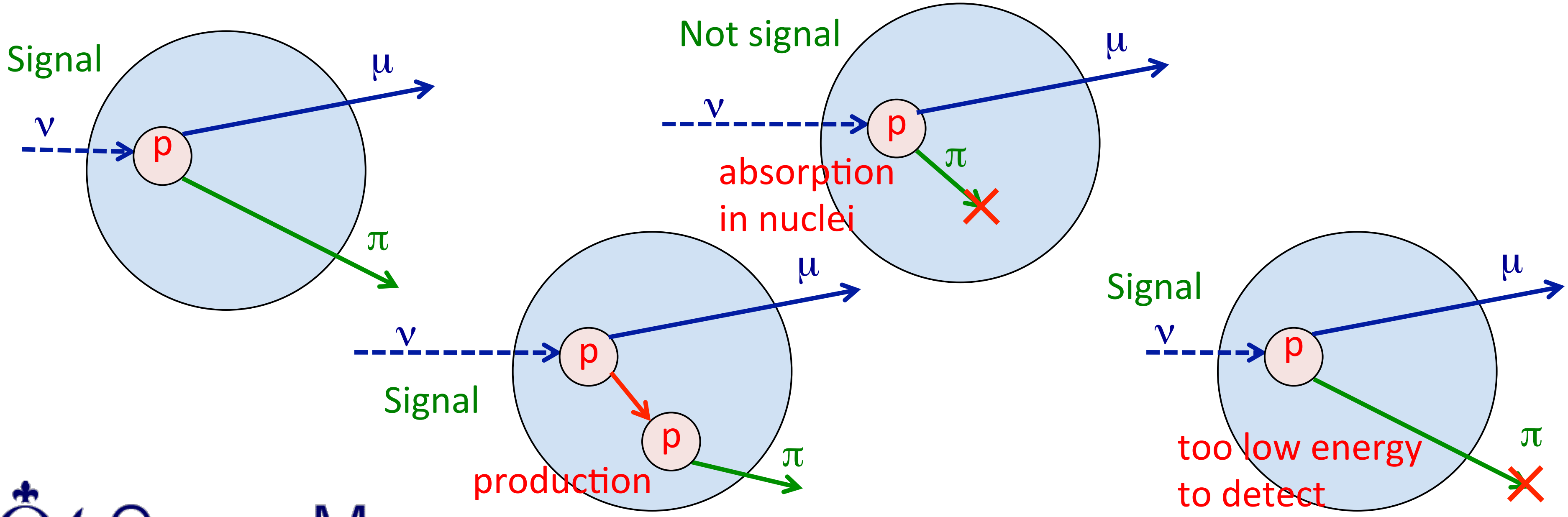
Topological cross section

ex) CC1π^o = " 1 μ + 1 π^o + N protons"

i. This definition includes π^o production by final state interactions (FSIs).

ii. This definition excludes CC1π^o interaction when π^o is lost by FSIs.

This is the necessary definition for the theorists to understand final state interactions (FSIs) without biases. Don't rely on the definition given by your interaction generator. "Signal" needs to be added to signal MC, and "Not signal" needs to be removed from signal MC. By this definition, FSI error of pion shouldn't be big, but detector pion absorption is part of final error.



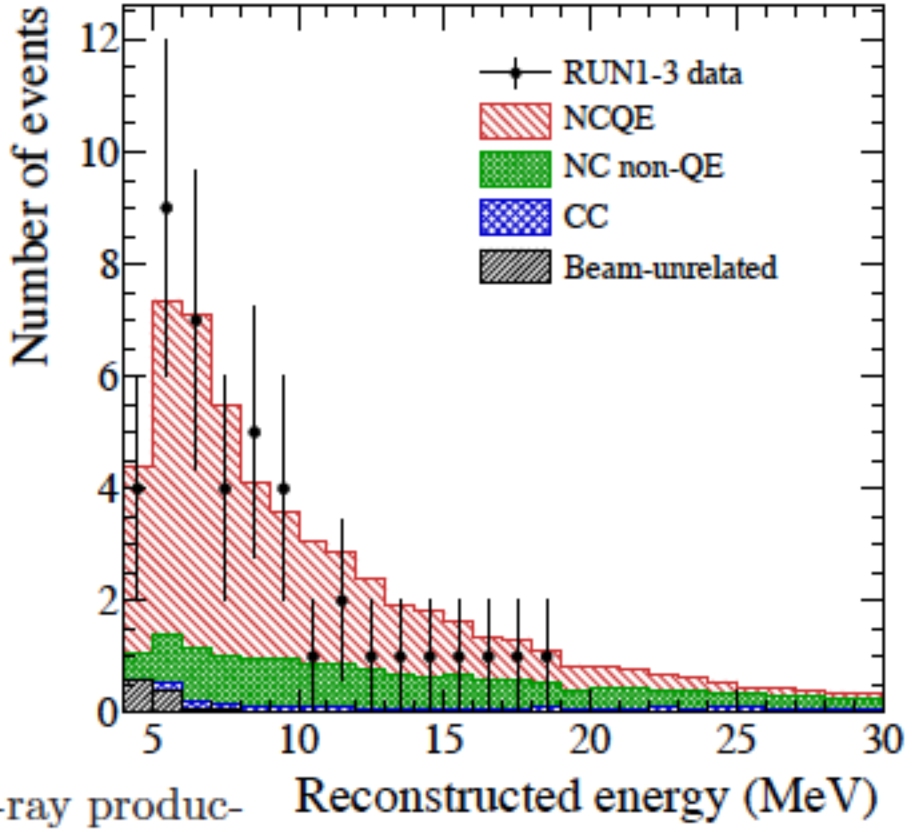
T2K, ArXiv:1403.3140

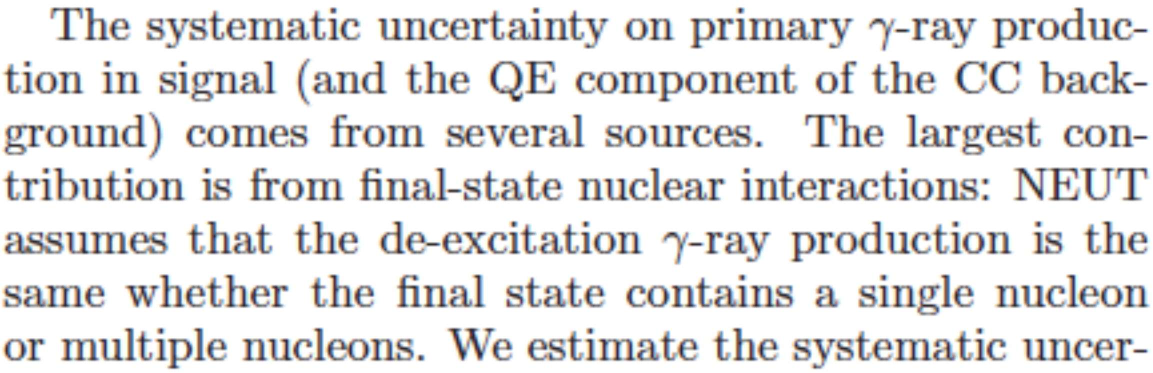
3.1 Signal definition

NCQE gamma measurement

- signal is defined to be gamma from NCQE interactions.
- gamma ray from FSI is not signal.

Alternatively, any gammas from any NC interaction can be defined "signal" (topological cross section), in this way, most of FSI error is gone (smaller systematics), signal statistics is higher, and data is less biased. Drawback is now theorist need to calculate FSI by themselves.




The systematic uncertainty on primary γ-ray production in signal (and the QE component of the CC background) comes from several sources. The largest contribution is from final-state nuclear interactions: NEUT assumes that the de-excitation γ-ray production is the same whether the final state contains a single nucleon or multiple nucleons. We estimate the systematic uncer-

3.1 Signal definition

Topological signal definition

1. Statistics is higher

You have more signal, so statistics is higher

2. Systematic error is lower

Nuclear effect is now signal, not error. Likewise, signal channel model error should be large.

3. Less biased

There is no cross-section model dependent selection nor correction. Data is less biased and preferred by theorists.

Theorists want to find how much MEC from our data using their state-of-the-art nuclear models. We are responsible to provide unbiased data containing these information.

If we use MEC model in our simulation for the selection of events, this mean we try to find how much MEC in our data based on our knowledge. This has 2 bad consequences,

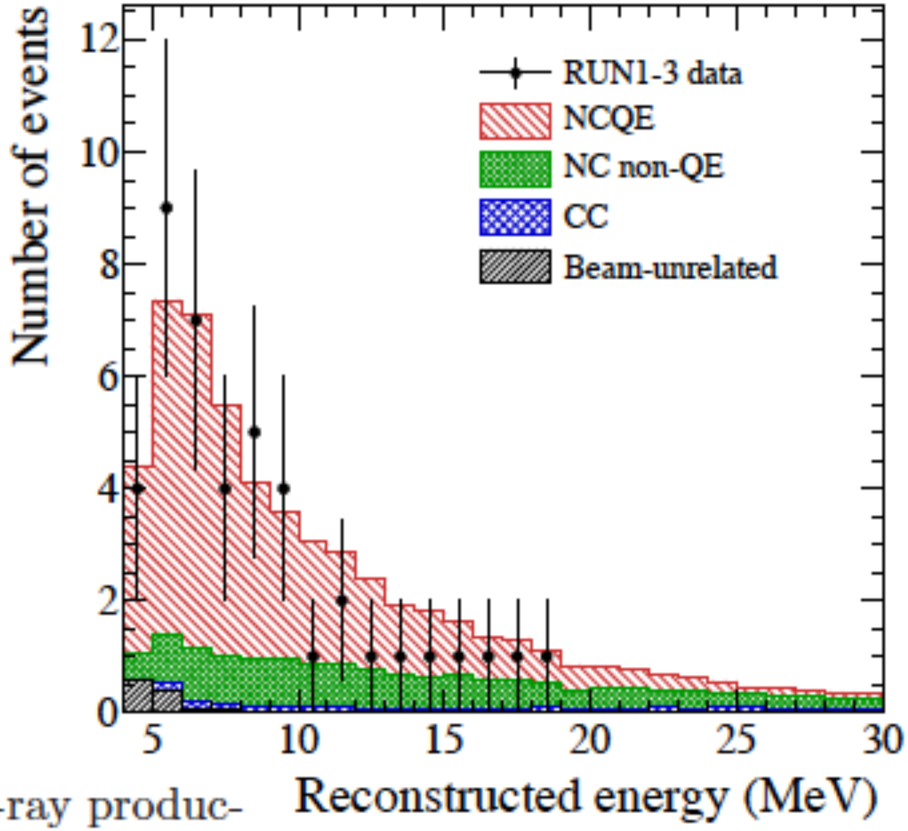
1. theorists lose jobs
2. result is wrong

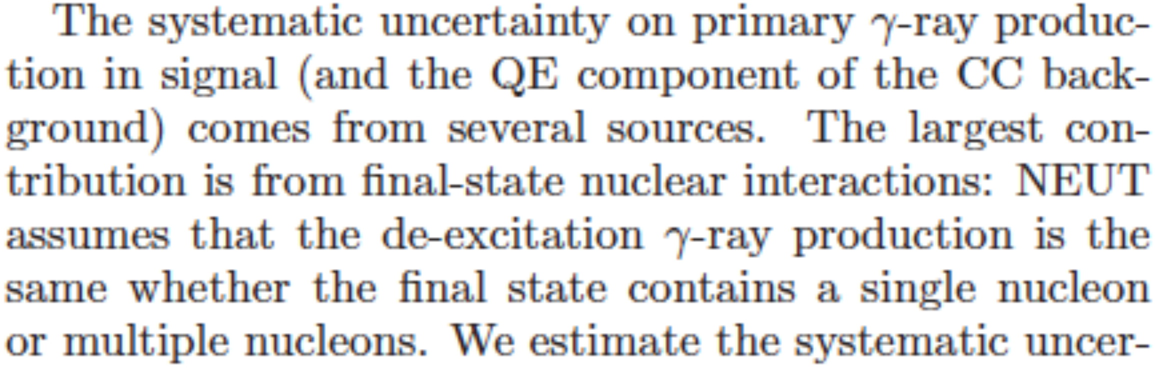
T2K, ArXiv:1403.3140

3.1 Signal definition

NCQE gamma measurement
- signal is defined to be gamma from NCQE interactions.
- gamma ray from FSI is not signal.

Alternatively, anay gamma from NC interaction can be defined "signal" (topological cross section), in this way, most of FSI error is gone (smaller systematics), signal statistics is higher, and data is less biased. Drawback is now theorist need to calculate FSI by themselves.





The systematic uncertainty on primary γ-ray production in signal (and the QE component of the CC background) comes from several sources. The largest contribution is from final-state nuclear interactions: NEUT assumes that the de-excitation γ-ray production is the same whether the final state contains a single nucleon or multiple nucleons. We estimate the systematic uncer-

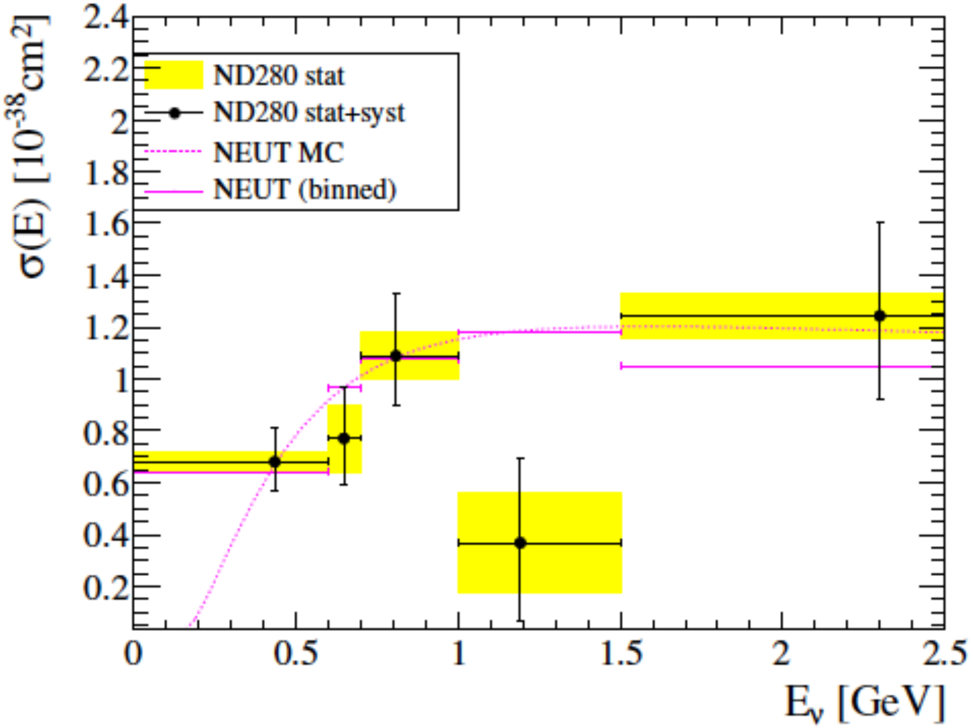
T2K, TN147

3.1 Signal definition

Tracker CCQE analysis

- signal is defined from NEUT channel number.

Alternatively, signal can be defined “1 muon-like track”. You can measure total cross section by this way, too. Error should be smaller because any processes making “1 muon-like track” becomes signal. Total cross section has reconstruction bias so it’s still harder for theorists to get info of MEC from this shape of this.



parameter	nominal value	fitted value
CC1π norm E1	1.63 ± 0.43	$1.34^{+0.34}_{-0.32}$
CC1π norm E2	1.00 ± 0.40	$1.06^{+0.40}_{-0.40}$
M_A^{QE}	1.21 ± 0.20	$1.38^{+0.16}_{-0.16}$
M_A^{RES}	1.11 ± 0.11	$1.16^{+0.10}_{-0.10}$
CC Other Shape	0.00 ± 0.40	$0.02^{+0.40}_{-0.40}$
NC1π^0 norm	1.19 ± 0.43	$1.03^{+0.40}_{-0.40}$
p_F	217.00 ± 30.38	$234.90^{+14.87}_{-15.05}$

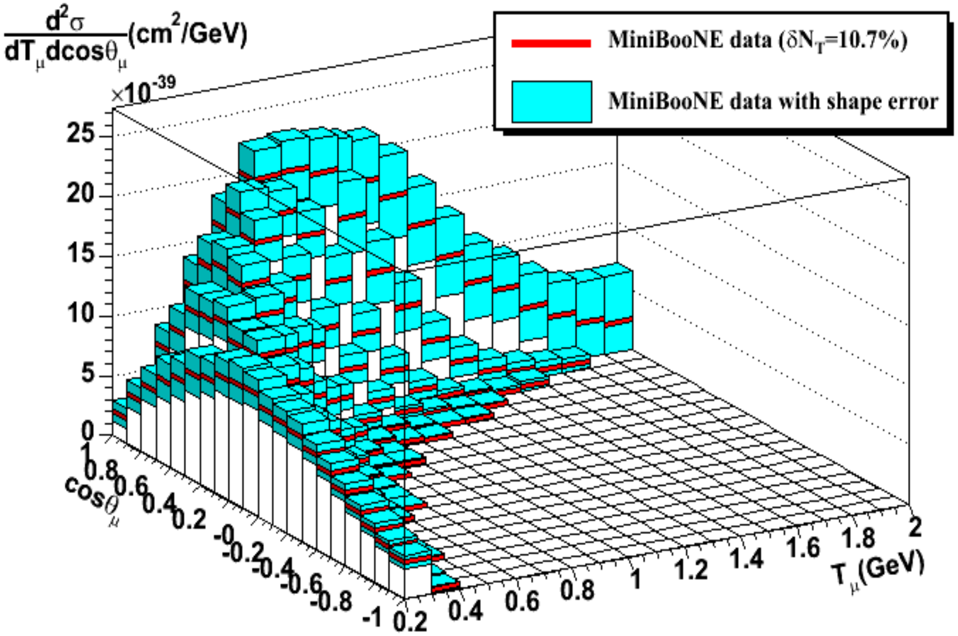
$$-2\ln\lambda(\boldsymbol{\theta}) = 2\sum_{i=1}^{p_\mu-\cos(\theta_\mu)\text{bins}} \left[N_i^{\text{predicted}}(\boldsymbol{\theta}) - N_i^{\text{observed}} + N_i^{\text{observed}}\ln\frac{N_i^{\text{observed}}}{N_i^{\text{predicted}}(\boldsymbol{\theta})}\right] + \ln\frac{\pi_d(d)}{\pi_d(d_{\text{nominal}})} + \ln\frac{\pi_f(f)}{\pi_f(f_{\text{nominal}})} + \ln\frac{\pi_x(x)}{\pi_x(x_{\text{nominal}})}$$

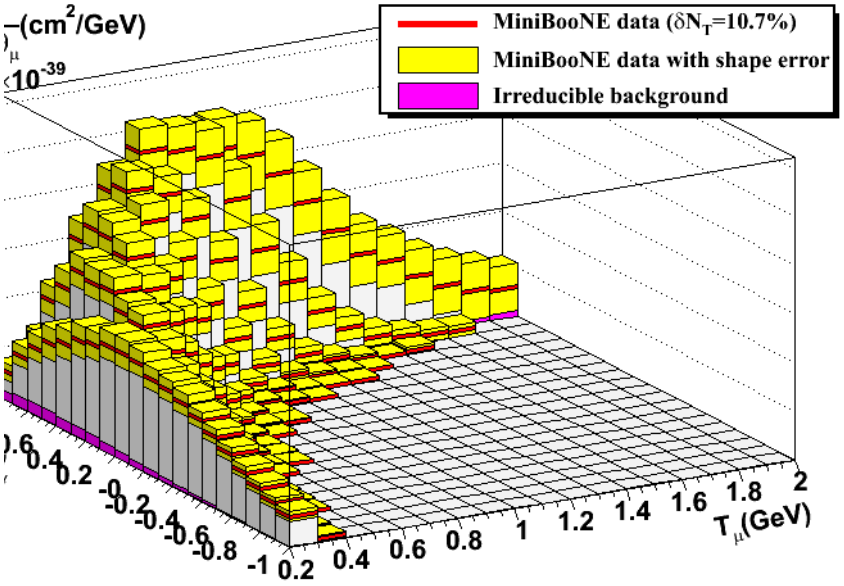
3.1 Signal definition

CCQE

Primary channel cross section

By subtracting CC1π^+ without pion background distribution (from MC with sideband correction), "pure" CCQE cross section is published (which is not, due to 2p-2h). Pure channel cross section is helpful for some theorists who cannot simulate whole set of cross sections.





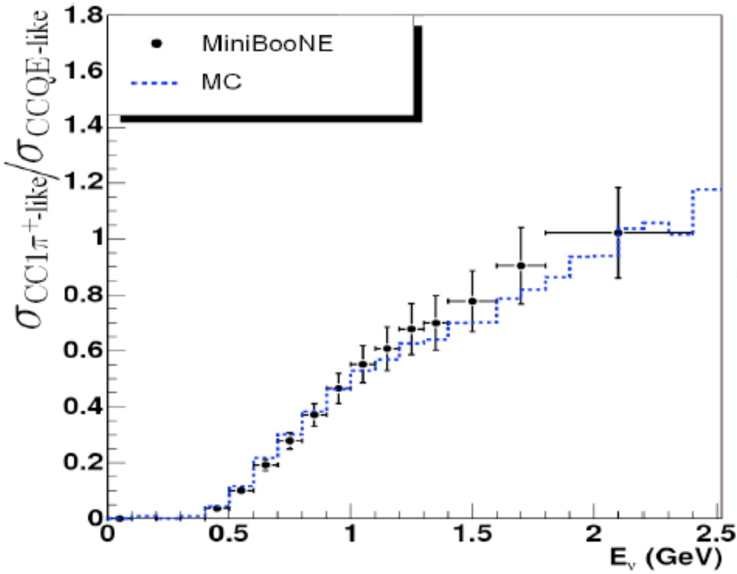
3.1 Signal definition

CC1π^+/CCQE

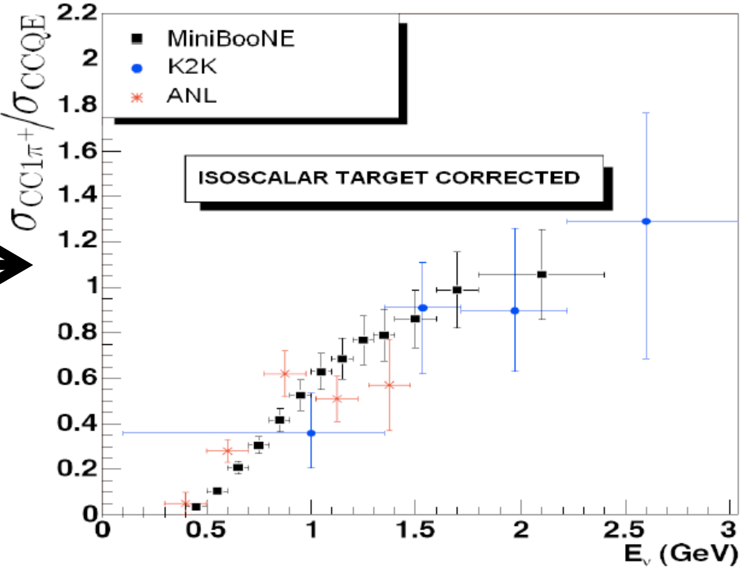
Primary channel cross section

Pure channel cross section is also useful to compare with other data (especially most of old data are published in pure channel cross section).

CC1π^+/CCQE topological cross section ratio



CC1π^+/CCQE primary chnnel cross section ratio



3.1 Signal definition

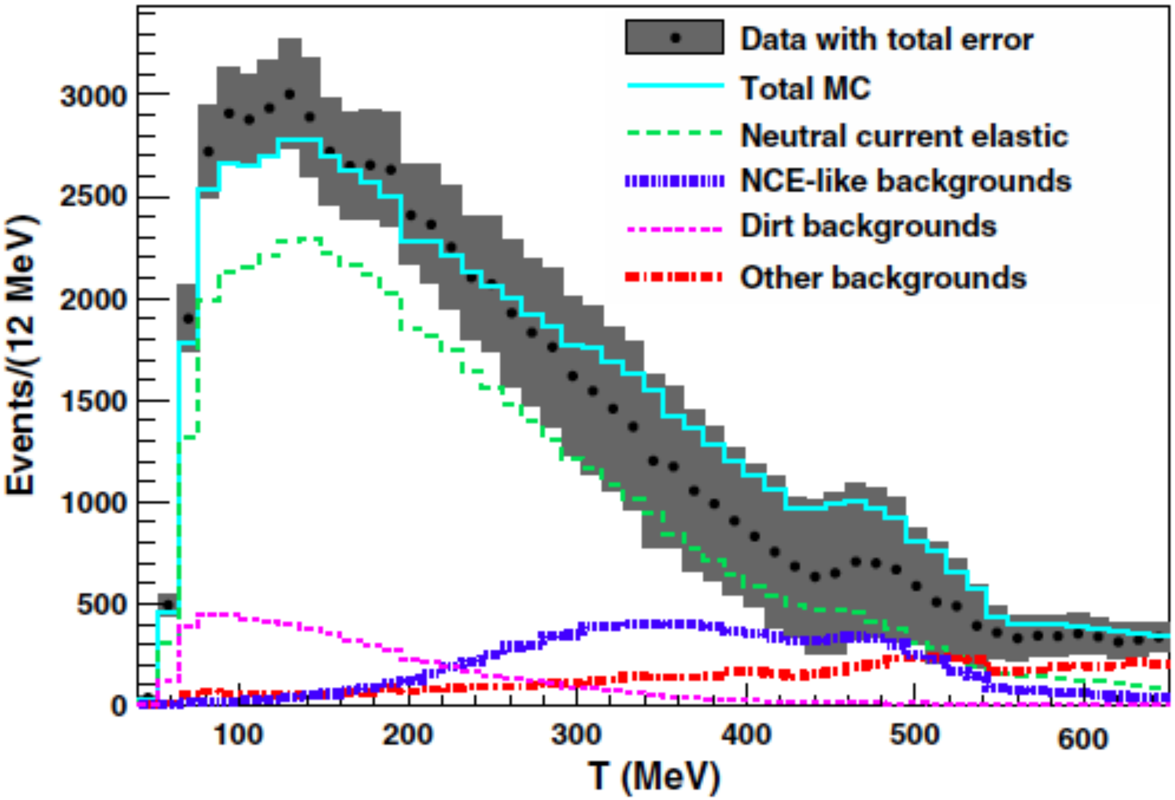
NCEL

Mixed target

Cross section result from mixed target is complicated. In NCEL, differential cross section is interpreted ν+N, which means sum of ν+p (carbon), ν+p (hydrogen), and ν+n (carbon).

Forward folding

Migration matrices (T_p^{true} and T_p^{recon}) are provided for ν+p (carbon), ν+p (hydrogen), and ν+n (carbon), with and without FSI, so that theorists can fold their xsec models to compare with measured Tp.



3.1 Signal definition

NCEL

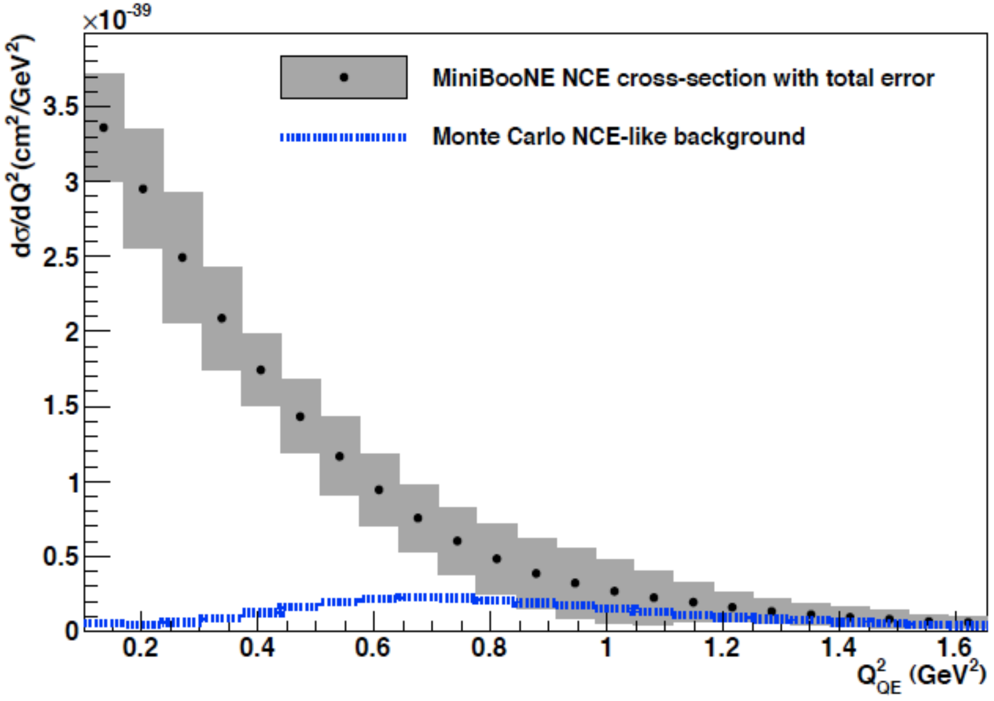
Mixed target

Cross section result from mixed target is complicated. In NCEL, differential cross section is interpreted ν+N, which means sum of ν+p (carbon), ν+p (hydrogen), and ν+n (carbon).

Efficiency difference

The efficiency difference of each interaction is provided to reproduce differential cross section from microscopic modes.

$$\frac{d\sigma_{\nu N\to\nu N}}{dQ^2} = \frac{1}{7}C_{\nu p,H}(Q^2_{\rm QE})\frac{d\sigma_{\nu p\to\nu p,H}}{dQ^2} + \frac{3}{7}C_{\nu p,C}(Q^2_{\rm QE}) \times \frac{d\sigma_{\nu p\to\nu p,C}}{dQ^2} + \frac{3}{7}C_{\nu n,C}(Q^2_{\rm QE})\frac{d\sigma_{\nu n\to\nu n,C}}{dQ^2}, \quad \text{(B6)}$$



3.1 Measured variables

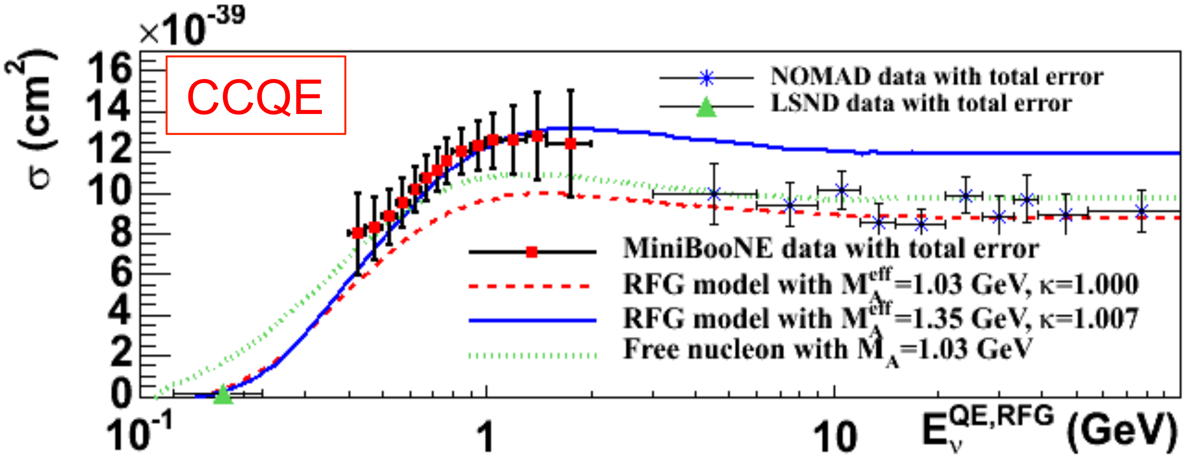
CCQE NCEL NC1π^o CC1π^+ CC1π^o anti CCQE anti NCEL

Function of measured variables

Differential cross section results with function of measured variables (momentum, direction, etc) has no reconstruction bias and preferred.

Function of reconstructed variables

Cross section results of reconstructed variables (especially Q^2 and Eν) are model-dependent. However, flux-unfolded total cross section function (function of true neutrino energy) is the only way to compare results with other experiments.



3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

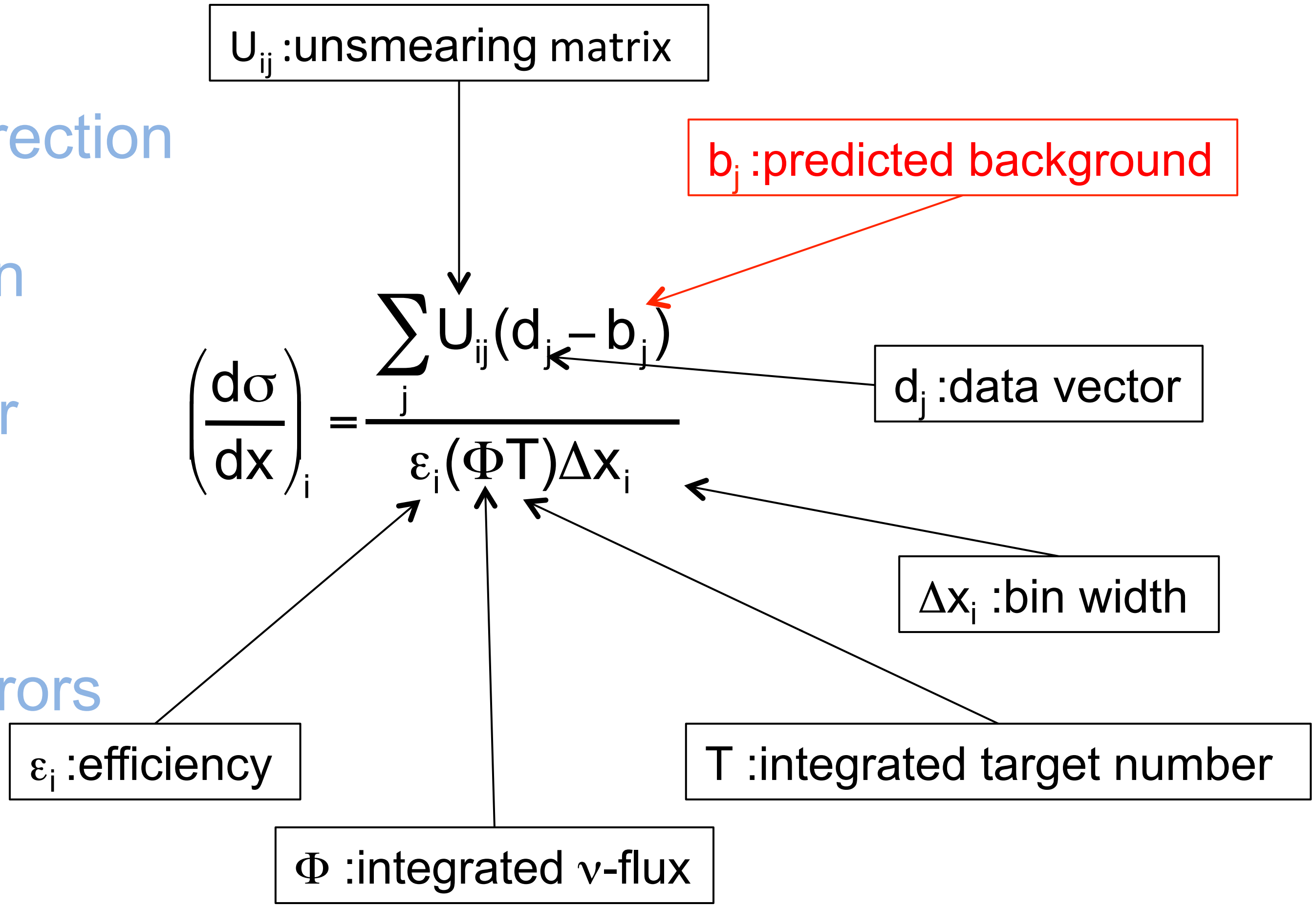
3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format



$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

T :integrated target number

Φ :integrated ν-flux

3.2 Background removing

CCQE NCEL $NC1\pi^o$ $CC1\pi^+$ $CC1\pi^o$ anti CCQE anti NCEL

Background subtraction method

- It is preferred because signal doesn't depend on signal MC explicitly.
- Normalization of background must be known.

Background subtraction

$$d_i - b_i$$

Purity correction

$$d_i \times \frac{s_i}{s_i + b_i}$$

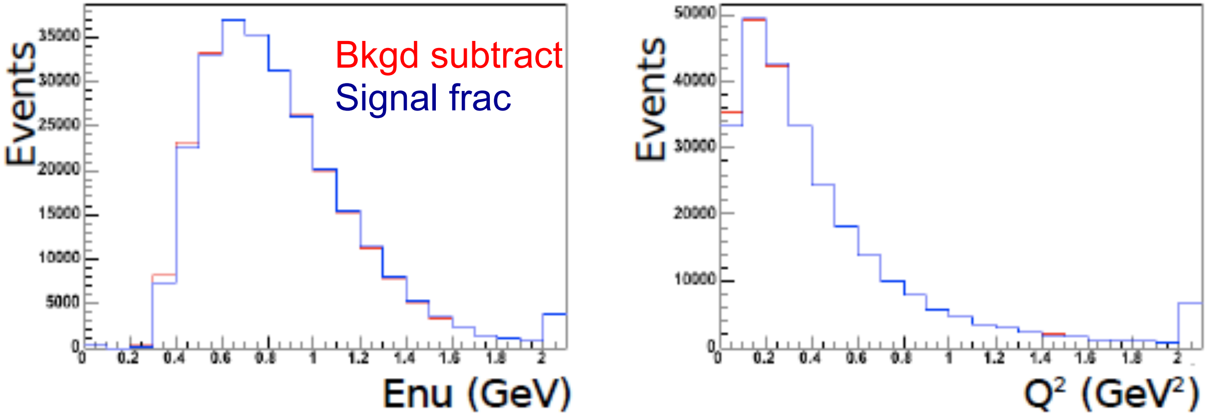
Purity correction method

- Signal explicitly depends on signal MC (=bad), potentially shape is distorted by signal to be measured (depending on size of error).
- However, if you don't know the normalization of background, this may be justified?

For high purity CCQE sample (77%), difference is only large at lowQ^2, where background is ~30%

Background subtraction vs Purity correction

CCQE



3.2 Background removing

CCQE | NCEL | $NC1\pi^o$ | $CC1\pi^+$ | $CC1\pi^o$ | anti CCQE | anti NCEL

Background subtraction method

- It is preferred because signal doesn't depend on signal MC explicitly.
- Normalization of background must be known.

Background subtraction

$$d_i - b_i$$

Purity correction

$$d_i \times \frac{s_i}{s_i + b_i}$$

Purity correction method

- Signal explicitly depends on signal MC (=bad), potentially shape is distorted by signal to be measured (depending on size of error).
- However, if you don't know the normalization of background, this may be justified?

Hybrid

NCEL | anti NCEL

- External background (=cosmic rays) are measured from sideband (=known normalization), and subtracted.
- Internal background (=$NC1\pi^+$ where pion is below detection threshold) is removed by purity correction.

$$(d_i - e_i) \times \frac{s'_i}{s'_i + b_i}$$

External background (measured)

Internal background
(no constraint)

3.2 Background removing

CCQE NCEL $CC1\pi^{+}$ $CC1\pi^{o}$ anti CCQE anti NCEL

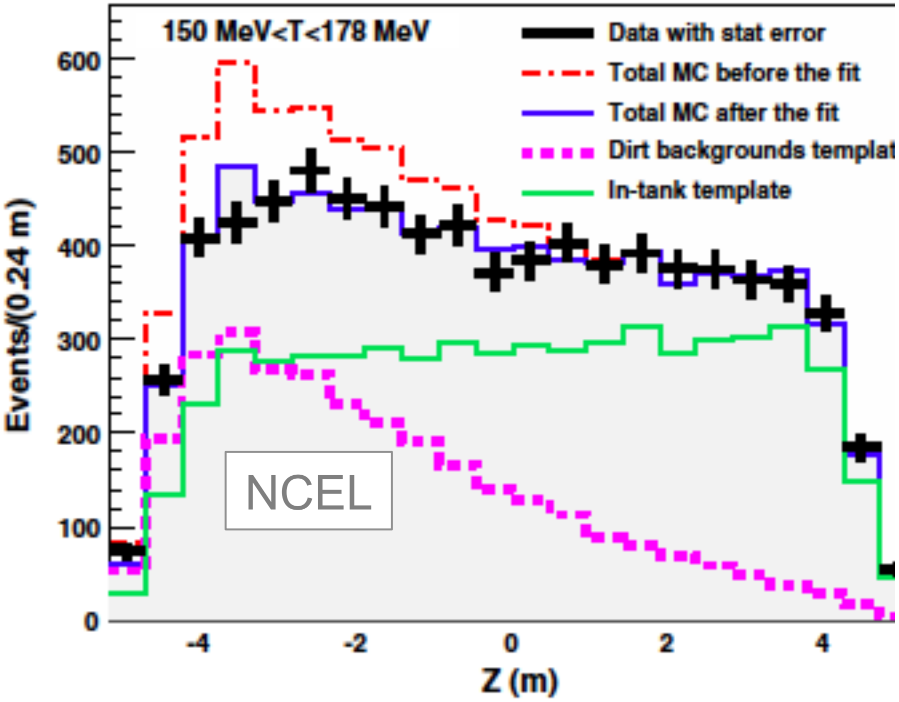
Data driven correction

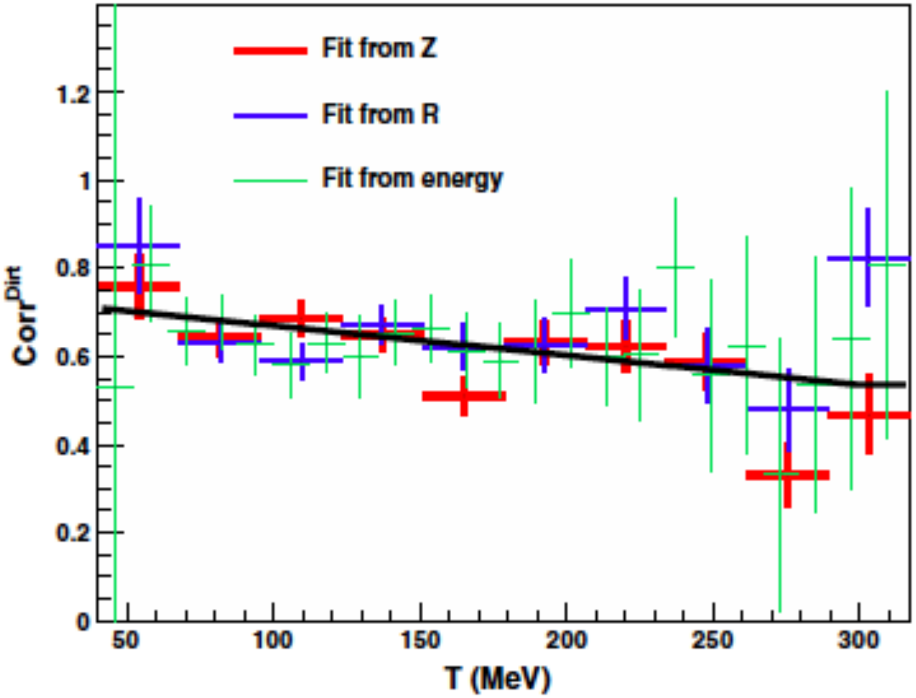
- Almost all analyses use sideband data to correct background distribution in signal boxes

External background measurement NCEL

- NCEL analysis measure amount of background coming from outside, with function of R and Z, to extrapolate background in fiducial volume

External background enhanced sample





3.2 Background removing

CCQE
NCEL
CC1π^+
CC1π^o
anti CCQE
anti NCEL

Data driven correction

- Almost all analyses use sideband data to correct background distribution in signal boxes

Flux error double counting?

- For background dominant sample, background subtraction makes flux error larger.
- Way to avoid is to define background excursion after removing normalization correlated with signal (shape-only background subtraction)

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - \boxed{b_j})}{\varepsilon_i(\Phi T)\Delta x_i}$$

3.2 Background removing

CCQE NCEL $CC1\pi^+$ $CC1\pi^o$ anti CCQE anti NCEL

Data driven correction

- Almost all analyses use sideband data to correct background distribution in signal boxes

Cross section error

- Main xsec error is the error assigned on background models
- Xsec models, such as FSIs?, change true-recon relationship for hadrons (smearing)
- Xsec models kinematics, i.e., cuts (=efficiency)

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - \boxed{b_j})}{\varepsilon_i(\Phi T)\Delta x_i}$$

3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

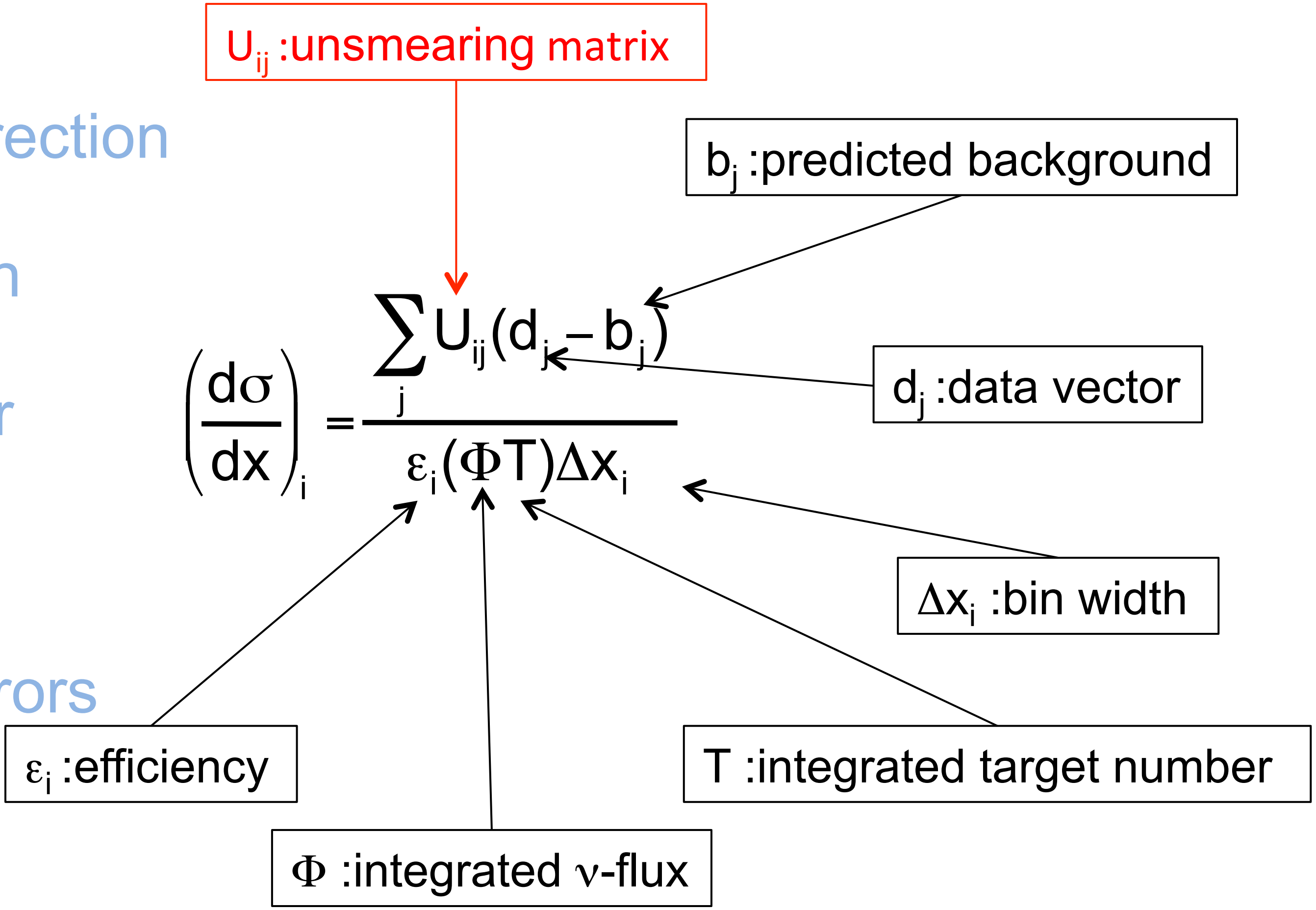
3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format



$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.3 Unsmearing

Unfolding

- The process removing the detector effects, mainly smearing and detector cut, is called unfolding. It is often easier to think by separating unfolding process to 2 parts, unsmearing and efficiency correction. We focus on unsmearing here.

Detector error

- Detector model affect smearing.

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j \boxed{U_{ij}}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

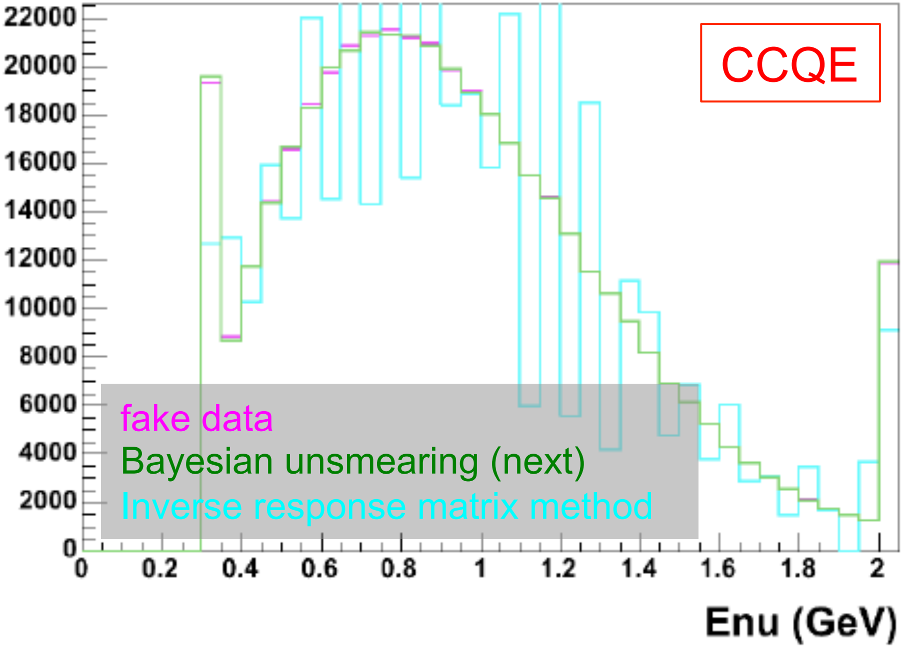
3.3 Unsmearing

Inverse response matrix method None

- Inverse response matrix method is the bias-free unfolding method, but this method doesn't work for anybody. Typically, it makes rapid oscillated solution (Gibb's phenomenon). Say, response matrix R gives the smearing and detector cut of true distribution a to measured distribution b in MC, it's inverse can be used to unfold data b to true distribution a

$$\beta_i = R_{ij}\alpha_j \rightarrow a_j = (R)^{-1}_{ji} b_i$$

Inverse response matrix method is very sensitive with MC statistics. It doesn't work for sparse matrix, it cannot handle large number of bins, it cannot deal histogram with zero-event bins. But all these are features for differential cross section!



Höcker, Kartvelishvili
NIM.A372(1996)469

3.3 Unsmearing

Tikhonov regularization method **NC1π^o**

- The regularization term from the prior knowledge of distribution (e.g., how smooth is) can stabilize inverse response matrix. The bias is introduced through the linear operator L and τ.

$$\beta_i = R_{ij}\alpha_j \to a_j = (R)^{-1}_{ji} b_i \to (Ra - b)^T V(b)^{-1}(Ra - b) + \tau(La)^T(La) \sim 0$$

Regularization parameter τ should be chosen with care.
- too small τ doesn't regulate matrix inversion
- too large τ too much smooth out response matrix R

Solution is,

$$a = U' \cdot b + \left[\sum_i [(I - U') \cdot b]_i\right] \cdot s$$

$$U' = [R + \tau V \cdot R^{T-1}(L^T L)]$$

$$s_i = \frac{\sum_j [U'VR^{-1}]_{ij}}{\sum_{jk} [U'VR^{-1}]_{jk}}$$

D'Agostini,
NIM.A362(1995)487

3.3 Unsmearing

CCQE | NCEL | NC1π^{o} | CC1π^{+} | CC1π^{o} | anti CCQE | anti NCEL

Iterative Bayesian method

- Unsmearing is based on the Bayesian statistics, so bias is introduced from MC knowledge

- Efficiency ε is defined by true distribution after cut μ to true distribution before cut α.

M-matrix gives transformation from measured distribution to true distribution after cut. It give the true distribution after cut μ on projection on one axis.

$$\varepsilon_i = \frac{\mu_i}{\alpha_i} \qquad \sum_{j=1}^{n} M_{ij} = \mu_i = \varepsilon_i \alpha_i$$

Now, define U-matrix by normalizing M-matrix with other axis,

$$U_{ij}^{0th} = \frac{M_{ij}}{\sum_{k=1}^{n} M_{kj}}$$

So, background subtracted data d^{0th} can be unsmeared and efficiency corrected to obtain unfolded cross section d^{1th}

$$d_i^{1st} = \frac{1}{\varepsilon_i} \sum_{j=1}^{n} U_{ij}^{0th} d_j^{0th}$$

Queen Mary
University of London

D'Agostini,
NIM.A362(1995)487

3.3 Unsmearing

CCQE NCEL $NC1\pi^o$ $CC1\pi^+$ $CC1\pi^o$ anti CCQE anti NCEL

Iterative Bayesian method

- Unsmearing is based on the Bayesian statistics, so bias is introduced from MC knowledge

- It is based on Bayes' theorem

$$M_{ij} = P(recon_j \mid true_i) P(true_i)$$

$$P(true_i \mid recon_j) = \frac{P(recon_j \mid true_i) P(true_i)}{\sum_k P(recon_j \mid true_k) P(true_k)} = \frac{M_{ij}}{\sum_k M_{kj}} = U_{ij}^{0th}$$

$$d_i^{1st} = \frac{1}{\varepsilon_i} \sum_{j=1}^{n} U_{ij}^{0th} d_j^{0th}$$

D'Agostini,
NIM.A362(1995)487

3.3 Unsmearing

CCQE | NCEL | NC1π^o | CC1π^+ | CC1π^o | anti CCQE | anti NCEL

Iterative Bayesian method

- Unsmearing is based on the Bayesian statistics, so bias is introduced from MC knowledge

If initial guess μ (=prior probability of Bayesian statistics) is not so close to nature, we can improve U-matrix by assuming d^{1th} is close to nature

$$\omega_i^{1st} = \frac{d_i^{1st}}{\alpha_i} \qquad U_{ij}^{1st} = \frac{\omega_i^{1st} M_{ij}}{\sum_{k=1}^{n} (\omega_k^{1st} M_{kj})} \qquad d_i^{2nd} = \frac{1}{\varepsilon_i} \sum_{j=1}^{n} U_{ij}^{1st} d_j^{0th}$$

- This iteration process usually converge <5 times. 0th iteration is not bad at all.
- Signal model dependence will become systematic error, this is done by varying M-matrix by changing systematics. So signal cross error is part of final error, but shape only.
- This method also fails if M-matrix is highly non-diagonal.
- 0th and 1st iteration difference of data is also included as systematic error (?).

Iterative Bayesian method works for, any number of bins, including zeros, sparse matrix, MxN matrix, background non-subtracted sample etc...

D'Agostini,
NIM.A362(1995)487

3.3 Unsmearing

How to construct M-matrix? Measured variables

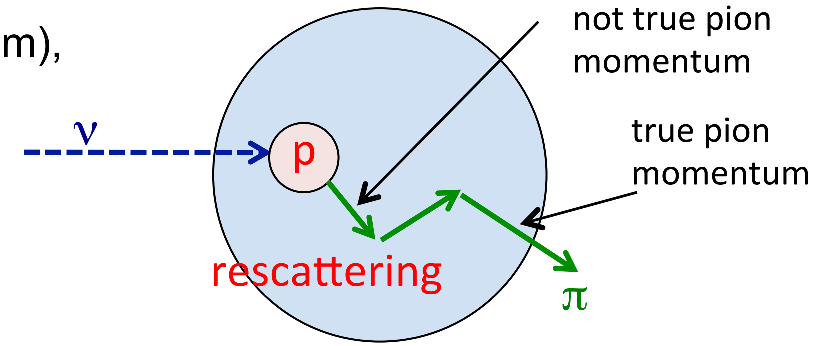
- It is desired to present differential cross section with function of measured quantities, such as muon energy, pion angle, etc, because they are not biased by reconstruction.

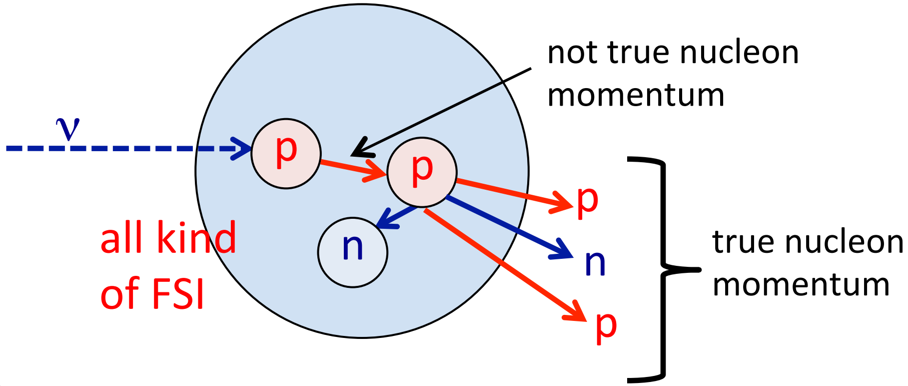
i. It is straightforward if you measure lepton kinematics.
True lepton kinematics are the true information for M-matrix.

CCQE | anti CCQE

NC1π^o | CC1π^+ | CC1π^o

ii. If you measure hadronic events (e.g., π^o momentum),
your "true" kinematics is after FSI,
i.e., particle exiting the nuclei.





NCEL | anti NCEL

iii. "true" momentum is defined by sum of all outgoing nucleons, because that is the observables.

3.3 Unsmearing

How to construct M-matrix? Reconstructed variables

- The definition of true kinematics is tricky, because you have choices.

CCQE NCEL anti CCQE anti NCEL

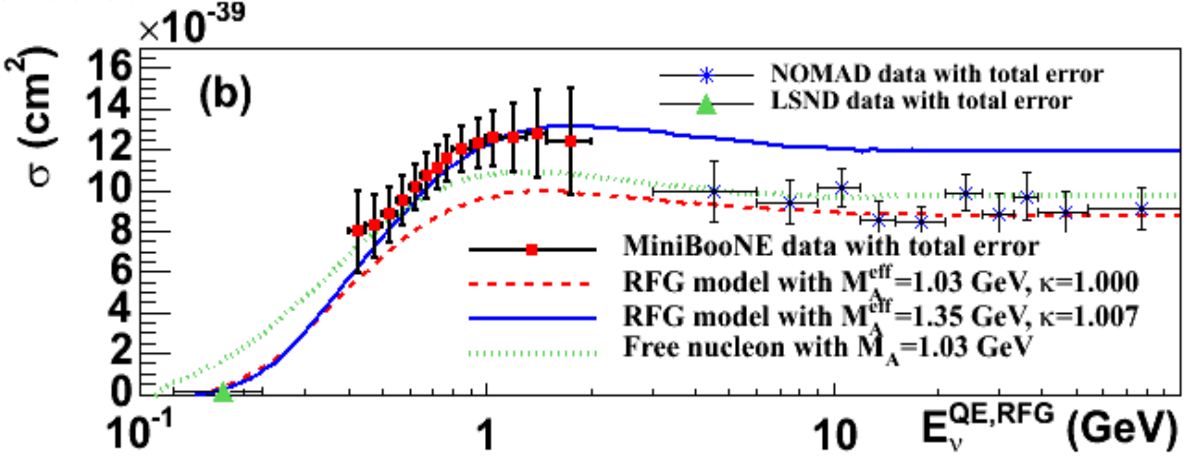
i. True Q^2 is defined by reconstructed Q^2 from true kinematics
For example, CCQE, true Q^2 is defined "reconstructed Q^2 from true muon energy and angle", and we call it "Q^2_{QE}" to remind people this is reconstructed under QE assumption.

CC1π^+ CC1π^o

ii True Q^2 is defined by true Q^2 in MC
This may be useful to compare with old data, only presented by this way

CCQE NC1π^o CC1π^+ CC1π^o anti CCQE

iii True $E\nu$ is defined by true $E\nu$ in MC
For example, CCQE, $E\nu$ is called "$E\nu^{QE,RFG}$" to remind people this is reconstructed under QE assumption then unfolded by assuming RFG model.



3.3 Unsmearing

No unsmearing NC1π^o

- If you know smearing is weak and statistics is low, no unsmearing may be the best option.

So what is the criteria?

Cowan,
"Statistical Data Analysis"

3.3 Unsmearing

No unsmearing NC1π^o

- If you know smearing is weak and statistics is low, no unsmearing may be the best option.

So what is the criteria?

Bias of unsmearing

- There is no perfect unfolding, unfolding method can be different depending on your distribution. Biases may be one of criteria.

i. Inverse response matrix method:
No bias, but it only works for very few bin histogram

$$B^{\mathrm{Inverse}} = 0$$

ii. Tikhonov regularization
Bias is introduced from linear function and regularization parameter. It also requires a fair amount of events.

$$B^{\mathrm{Tikhonov}} = \sum_i \left[U' \cdot (Ra - b)\right]_i$$

iii. Iterative Bayesian method
Bias is introduced from prior knowledge of MC. It works for any distribution.

$$B^{\mathrm{Bayesian}} = \sum_i \left[M \cdot (Ra - b)\right]_i$$

iv. No unsmearing

$$B^{\mathrm{NoUnsmear}} = \sum_i \left[(U - I) \cdot b\right]_i$$

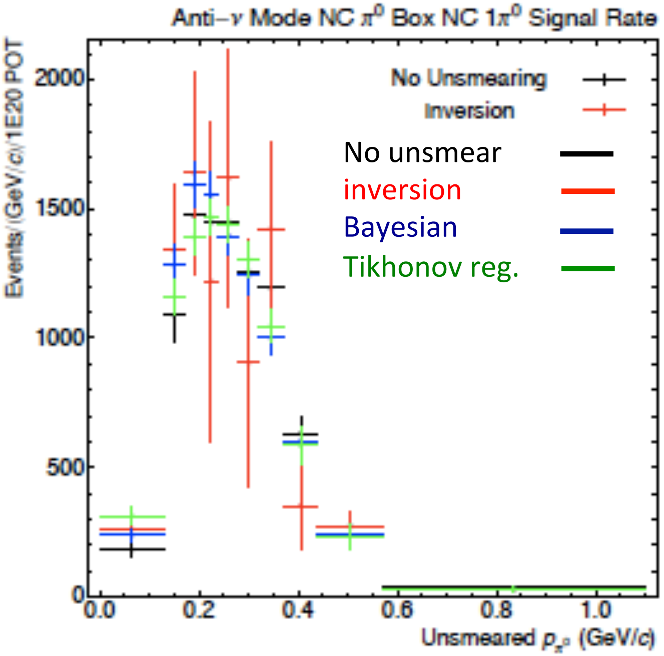
3.3 Unsmearing

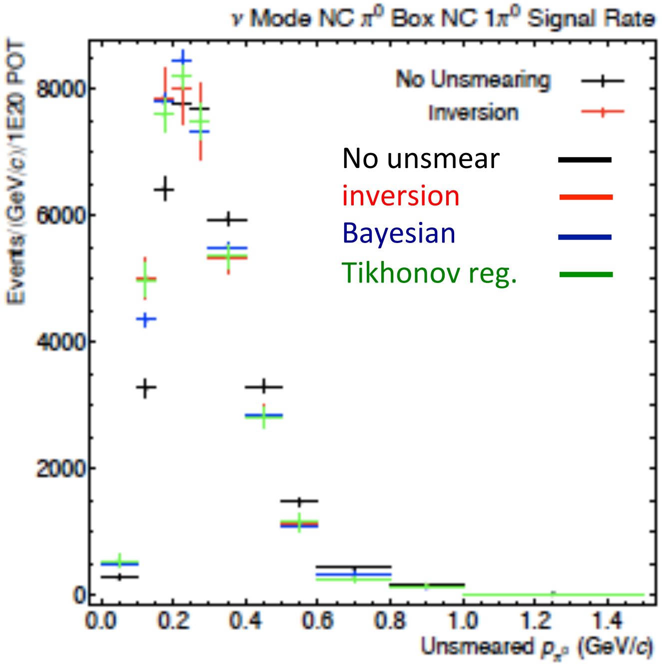
π^o kinematics

NC1π^o

- Comparing biases and histograms by eyes, π^o-kinematics are unfolded by 3 different methods.

$P_\pi(\nu)$: Tikhonov regularization
$\cos\theta_\pi(\nu)$: Iterative Bayesian
$P_p(\text{anti-}\nu)$: Iterative Bayesian
$\cos\theta_\pi(\text{anti-}\nu)$: No unsmearing





3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

T :integrated target number

Φ :integrated ν-flux

ε_i :efficiency

3.4 Efficiency correction

CCQE | NCEL | NC1π^{o} | CC1π^{+} | CC1π^{o} | anti CCQE | anti NCEL

Looks straightforward, no?

- The efficiency is defined as true distribution after cut μ divided by that before cut α. Because of the nature of ratio, ε is insensitive with many systematic variations common for numerator and denominator such as flux error and cross section error. The detector error is important.

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\boxed{\varepsilon_i}(\Phi T)\Delta x_i}$$

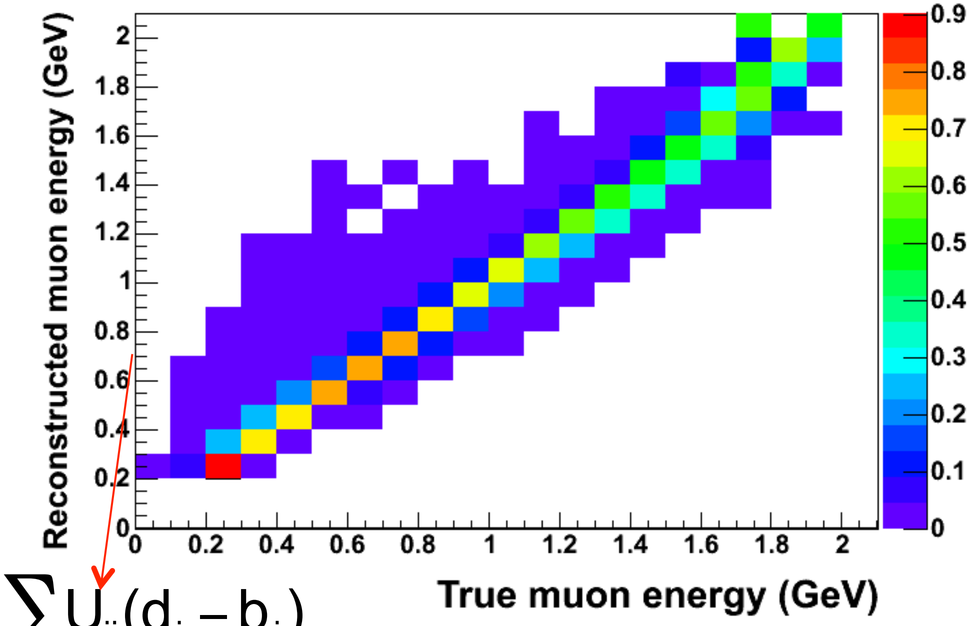
$$\varepsilon_i = \frac{\mu_i}{\alpha_i} = \frac{N_i(\mathrm{AfterCut})}{N_i(\mathrm{BeforeCut})}$$

3.4 Efficiency correction

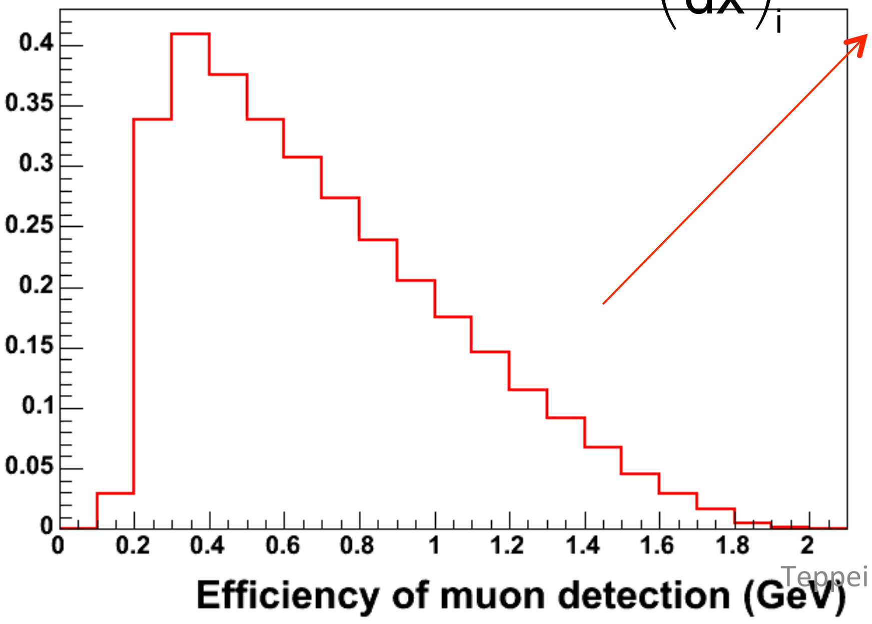
CCQE

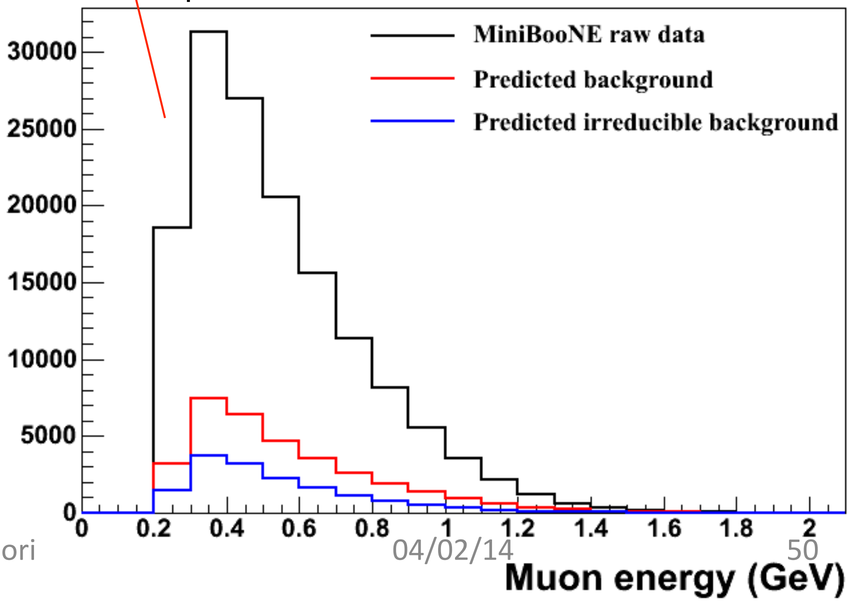
Muon energy unfolding

- Because of the resolution of muon energy, it is possible to recover events outside of kinematic cut by unfolding process (detector model dependent)



$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$



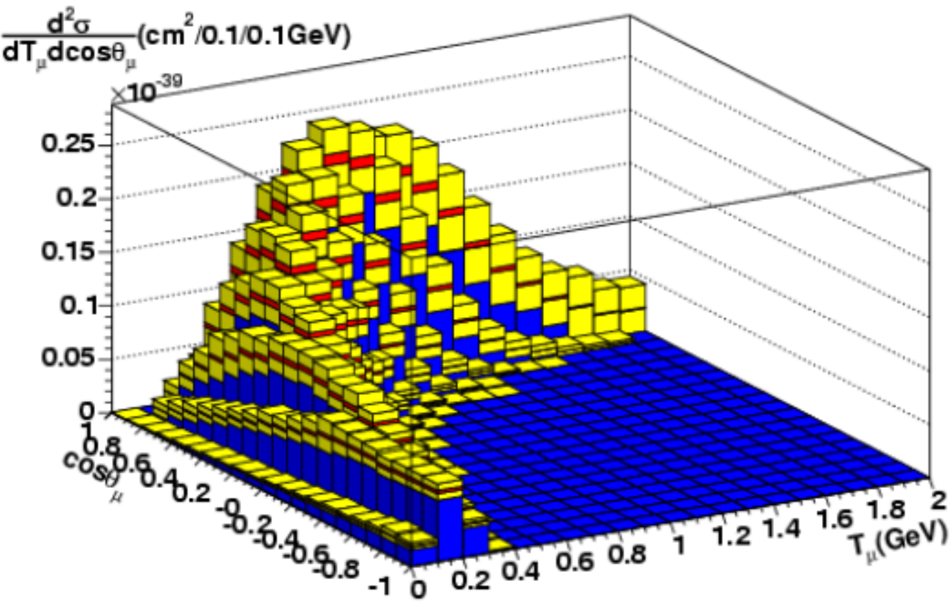


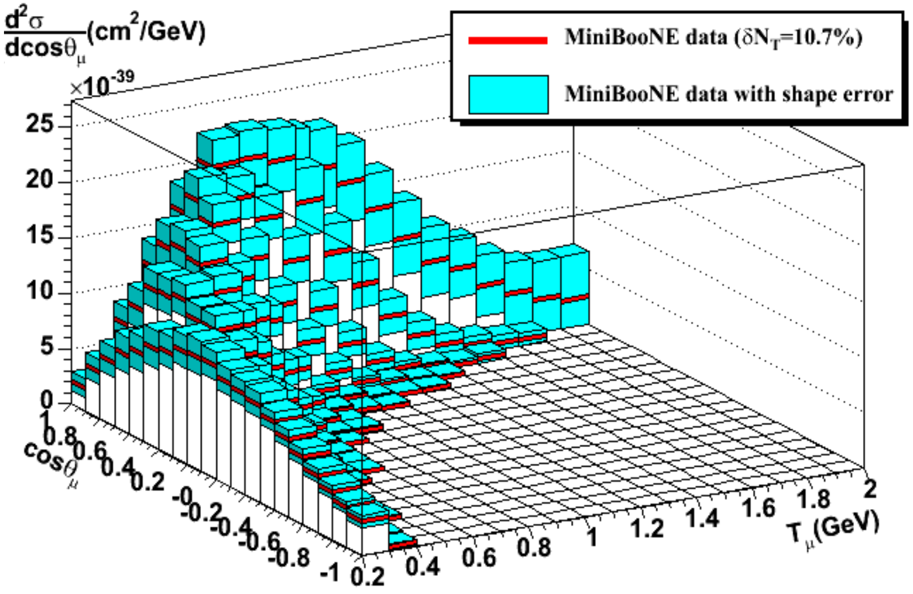
3.4 Efficiency correction

CCQE

Muon energy unfolding

- Because of the resolution of muon energy, it is possible to recover events outside of kinematic cut by unfolding process (detector model dependent)





3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

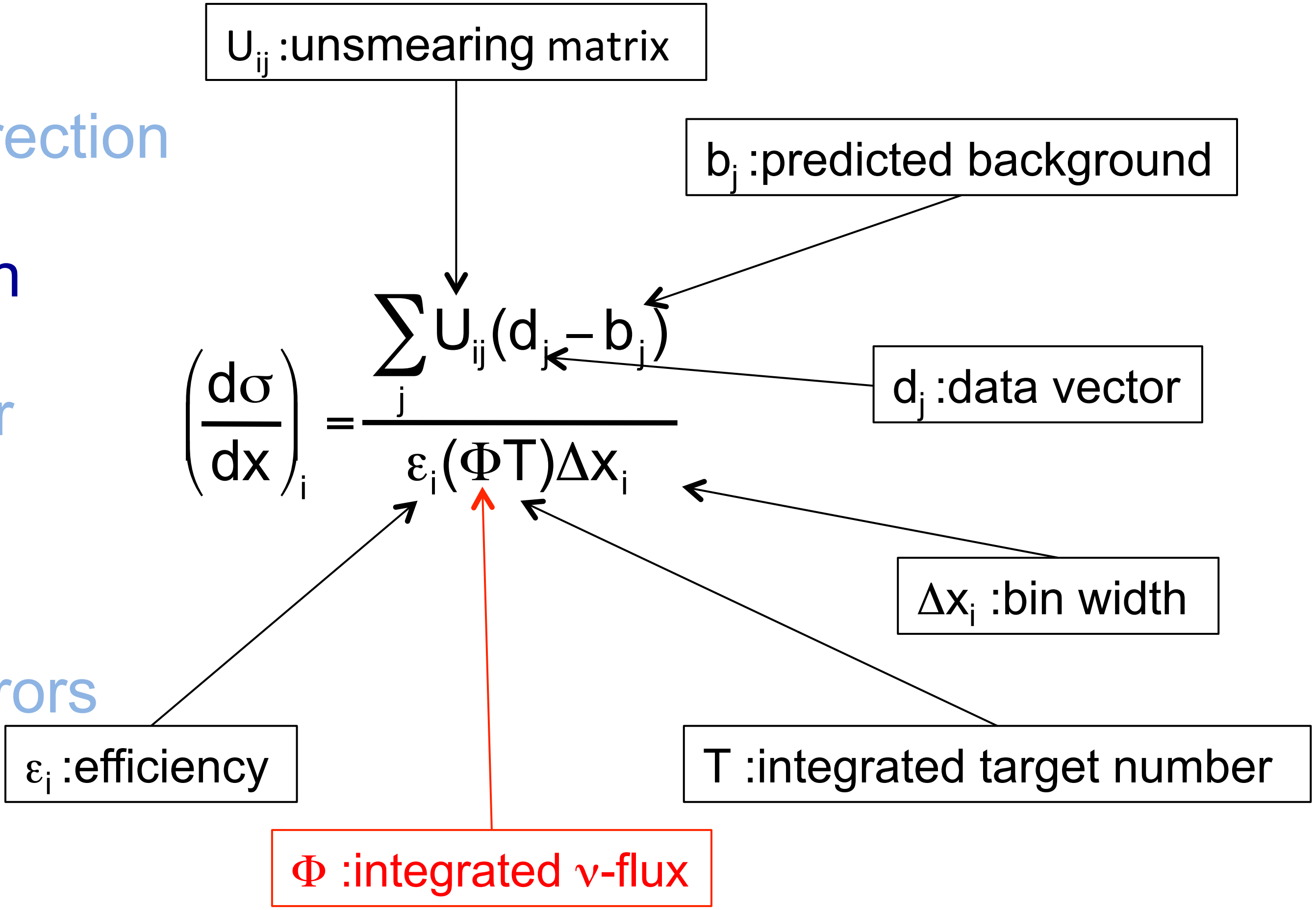
3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format



U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

Δx_i :bin width

ε_i :efficiency

T :integrated target number

Φ :integrated ν-flux

3.5 Flux correction

Integral region of flux

- Flux is integrated and removed. There are many ways how to introduce flux error.

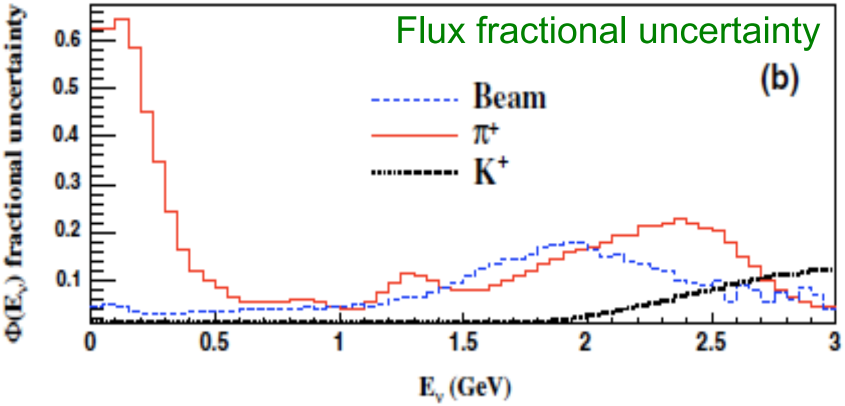
i. Flux is integrated in all spectrum region and it's variation is the flux error. This choice gives rather large flux error (e.g., ~12% for NC1π^o).

NC1π^o

ii. Cutoff for flux integration

CC1π^o

Flux is integrated in [0.5-2.0] GeV, and error is variation of that. In this way, you can avoid flux variation at low energy which don't contribute to the channel. Error is smaller, ~7%.



CCQE

NCEL

anti CCQE

anti NCEL

iii. Flux is integrated all region, but flux error is calculated separately
Flux variation is calculated by variation of numerator of efficiency term (rate).
In this way, flux variation is automatically limited within the region relevant to cross section measurement. Both normalization and shape flux error are taken into account.
Error is smaller, ~8%.

3.5 Flux correction

Integral region of flux

- Flux is integrated and removed. There are many ways how to introduce flux error.

$$\left(\frac{d\sigma}{dx}\right)_i^F = \frac{\sum_j U_{ij}^F (d_j - b_j^F)}{\left(\frac{N_{after}}{N_{before}}\right)_i^F (\Phi^F T)\Delta x_i} \longrightarrow \left(\frac{d\sigma}{dx}\right)_i^F = \frac{\sum_j U_{ij}^F (d_j - b_j^F)}{\left(\frac{N^F_{after}}{N_{before}}\right)_i (\Phi T)\Delta x_i}$$

CCQE | NCEL | anti CCQE | anti NCEL

iii. Flux is integrated all region, but flux error is calculated separately
Flux variation is calculated by variation of numerator of efficiency term (rate).
In this way, flux variation is automatically limited within the region relevant to cross section measurement. Both normalization and shape flux error are taken into account.
Error is smaller, ~8%.

3.5 Flux correction

Integral region of flux

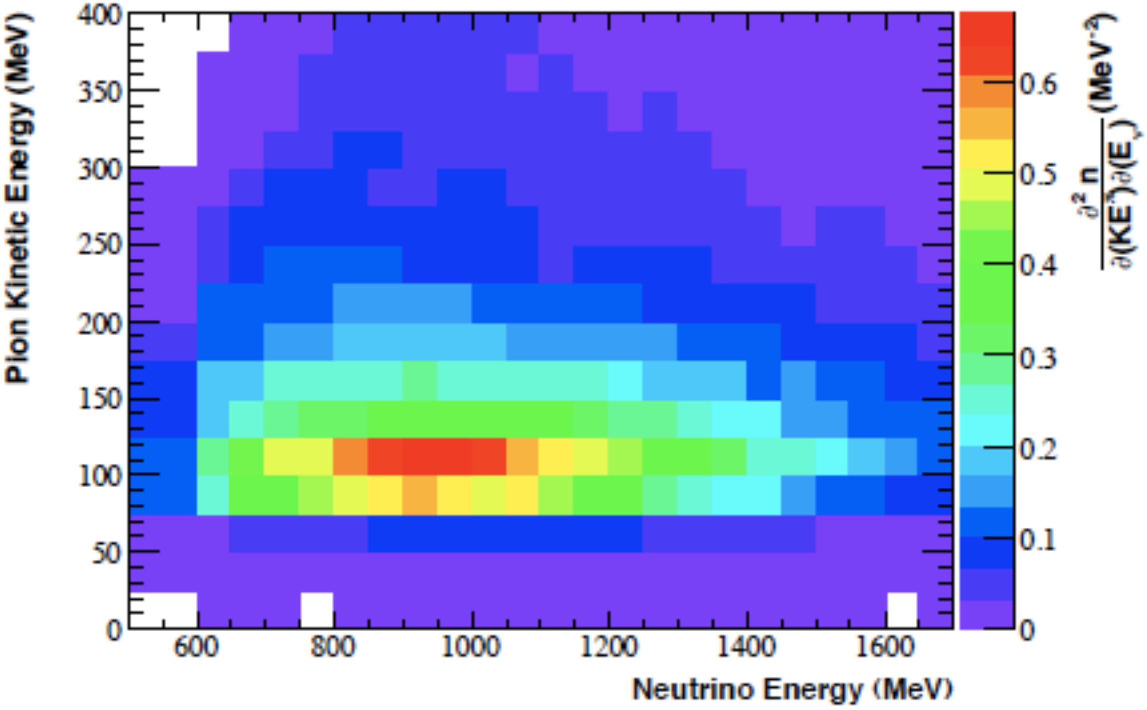
- Flux is integrated and removed. There are many ways how to introduce flux error.

iv. cross sections are function of neutrino energy

CC1π^+

In this way, integrated flux in Eν bin is unfolded in each bin of measured variables (e.g., pion kinetic energy), then flux error only relevant Eν region apply to measured variables. This minimizes flux error at many region.

Pion kinetic energy- neutrino energy 2-dimentional cross section



3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

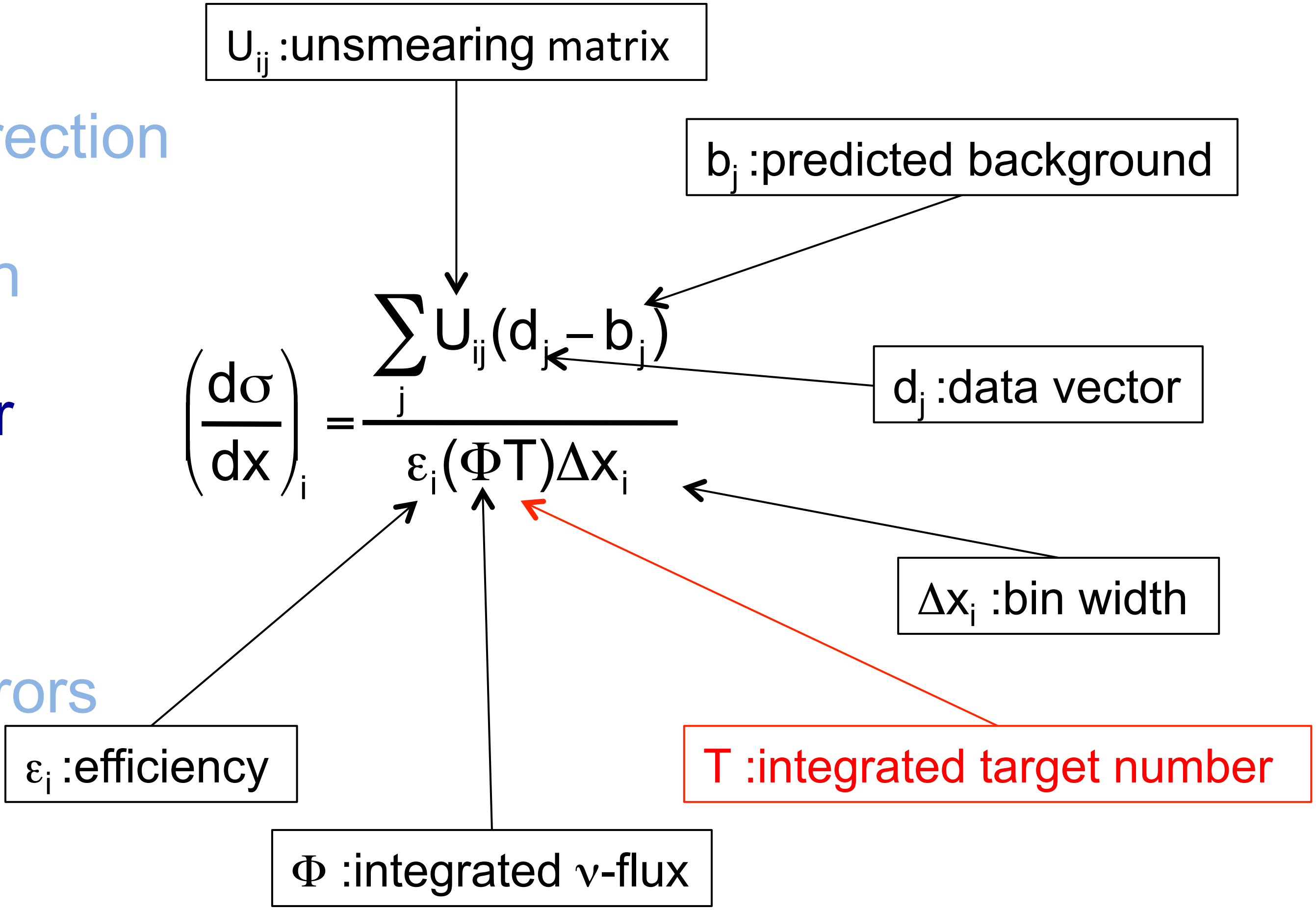
3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format



$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.6 Target number correction

What is the real fiducial volume?

- Fiducial cut is made based on reconstructed vertices.
- Fiducial volume is based on true dimension.

In MiniBooNE, fiducial cut is smaller than fiducial volume, to take account possible vertex mis-reconstruction.

i. MiniBooNE is ~600cm radius sphere.
ii. MC is generated within 550cm sphere.
iii. The fiducial cut is 500cm sphere.

In this way, we can guarantee cross section is calculated in the region where we believe uniform.

- In general, data-MC agreement is not enough for absolute cross section measurement. Even data and MC perfectly agree in reconstructed spectrum, you need to worry the absolute calibration of vertex, target volume, and density.

3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.7 Bin width

CCQE NCEL $NC1\pi^o$ $CC1\pi^+$ $CC1\pi^o$ anti CCQE anti NCEL

Statistics

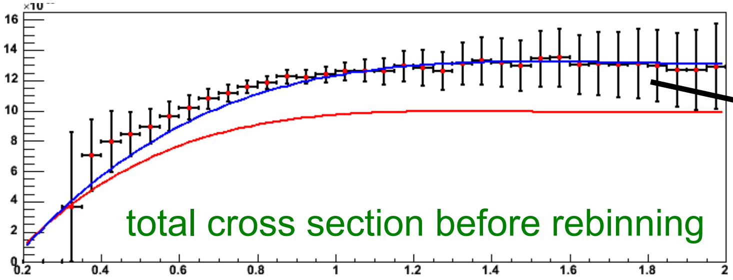
Bin width is finer at high statistics region, and coarser at low statistics region.

Systematics

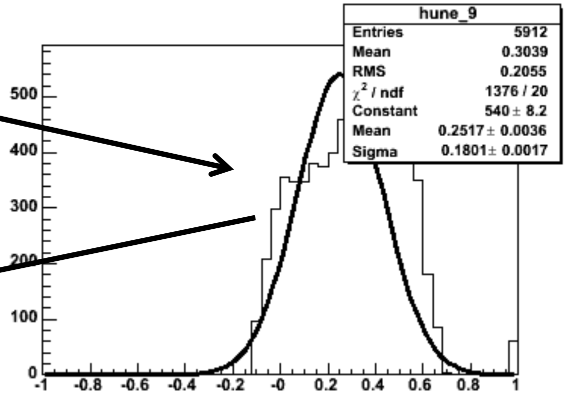
Too fine bins with large shape systematic make no sense?

Reconstruction bias CCQE

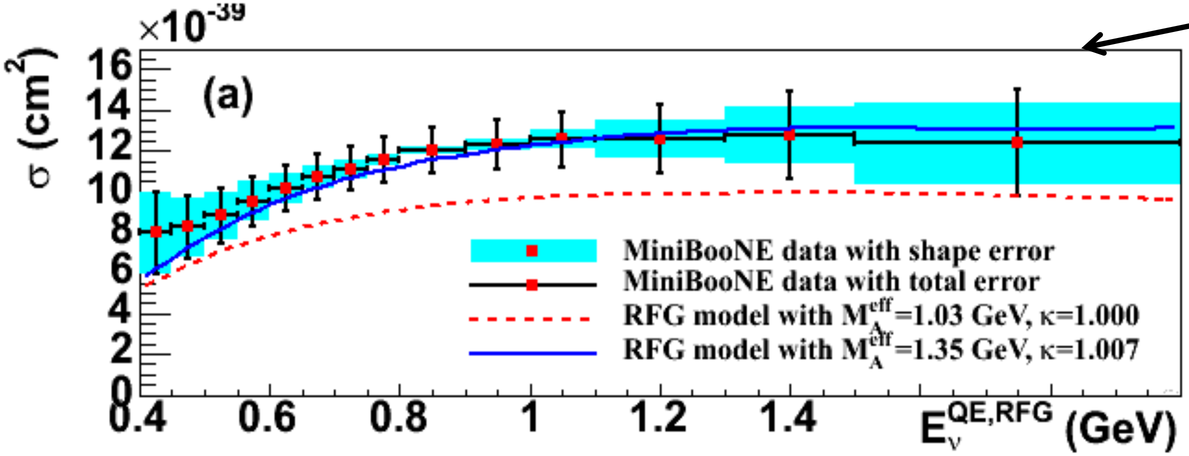
For CCQE analysis, reconstruction bias was added to bin resolution.



total cross section before rebinning



resolution + reconstruction bias



3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i (\Phi T) \Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

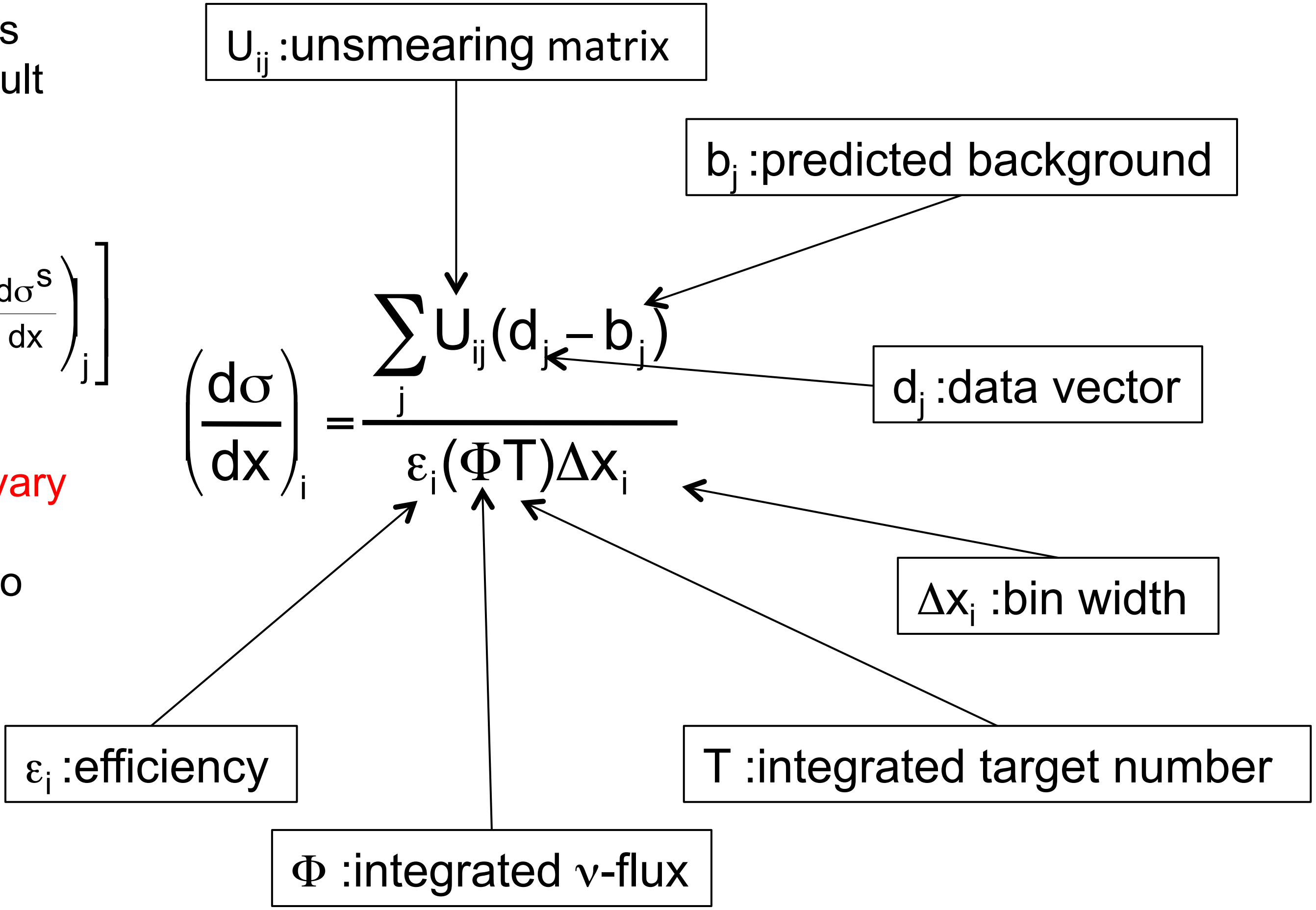
3.8 Systematic errors

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij}=\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^{s}}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^{s}}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don't vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background
3.8.2 U-matrix
3.8.3 Efficiency
3.8.4 Flux term
3.8.5 Target number



$$\left(\frac{d\sigma}{dx}\right)_i=\frac{\sum_j U_{ij}(d_j-b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.8 Systematic errors

$$\left(\frac{d\sigma}{dx}\right)_i^s \neq \frac{\sum_j U_{ij}^s (d_j - b_j^s)}{\varepsilon_i^s (\Phi^s T^s) \Delta x_i}$$

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij} = \left[\left(\frac{d\sigma}{dx}\right)_i - \left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j - \left(\frac{d\sigma^s}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don't vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background
3.8.2 U-matrix
3.8.3 Efficiency
3.8.4 Flux term
3.8.5 Target number

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j \boxed{U_{ij}} (d_j - \boxed{b_j})}{\boxed{\varepsilon_i} (\boxed{\Phi} \boxed{T}) \Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.8 Systematic errors

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij}=\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^s}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don’t vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background
3.8.2 U-matrix
3.8.3 Efficiency
3.8.4 Flux term
3.8.5 Target number

b_j :predicted background

$$\left(\frac{d\sigma}{dx}\right)_i=\frac{\sum_j \boxed{U_{ij}}(d_j-\boxed{b_j})}{\boxed{\varepsilon_i}(\boxed{\Phi}\boxed{T})\Delta x_i}$$

To reduce cross section error
- higher purity
- sideband (then error is measurement)

3.8 Systematic errors

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij} = \left[\left(\frac{d\sigma}{dx}\right)_i - \left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j - \left(\frac{d\sigma^s}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don't vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background

3.8.2 U-matrix

3.8.3 Efficiency

3.8.4 Flux term

3.8.5 Target number

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j \boxed{U_{ij}}(d_j - \boxed{b_j})}{\boxed{\varepsilon_i}(\boxed{\Phi}\boxed{T})\Delta x_i}$$

b_j :predicted background

To reduce flux error
- use cancellation with signal MC
  - purity correction method
  - shape-only background subtraction method

3.8 Systematic errors

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij}=\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^s}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don’t vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background

3.8.2 U-matrix

3.8.3 Efficiency

3.8.4 Flux term

3.8.5 Target number

U_{ij} :unsmearing matrix

$$\left(\frac{d\sigma}{dx}\right)_i=\frac{\sum_j \boxed{U_{ij}}(d_j-\boxed{b_j})}{\boxed{\varepsilon_i}(\boxed{\Phi}\boxed{T})\Delta x_i}$$

The error is dominated by detector error (flux error cancels). For Bayesian unfolding, signal MC gives error here, too.

3.8 Systematic errors

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij}=\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^s}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don't vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background

3.8.2 U-matrix

3.8.3 Efficiency

3.8.4 Flux term

3.8.5 Target number

Detector error goes here (flux and xsec error cancel).

$$\left(\frac{d\sigma}{dx}\right)_i=\frac{\sum_j \boxed{U_{ij}}(d_j-\boxed{b_j})}{\boxed{\varepsilon_i}(\boxed{\Phi}\boxed{T})\Delta x_i}$$

Efficiency variation

3.8 Systematic errors

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij}=\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^s}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don’t vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background
3.8.2 U-matrix
3.8.3 Efficiency
3.8.4 Flux term
3.8.5 Target number

Flux normalization error is here. You may need to apply cutoff to remove flux variation irrelevant for cross section measurement, to avoid overestimation of error.

$$\left(\frac{d\sigma}{dx}\right)_i=\frac{\sum_j \boxed{U_{ij}}(d_j-\boxed{b_j})}{\boxed{\varepsilon_i(\boxed{\Phi}T)}\Delta x_i}$$

Integrated flux variation

3.8 Systematic errors

Systematic error is calculated from the difference of systematics varied cross section result and central value cross section result.

$$E_{ij}=\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^s}{dx}\right)_j\right]$$

4 parts are related with systematic error. Don’t vary all of them with all systematics! You need to think about the effect of each term

3.8.1 background
3.8.2 U-matrix
3.8.3 Efficiency
3.8.4 Flux term
3.8.5 Target number

Precise definition of active volume may remove the bias of target number.
(this is not simulated effect, i.e., incorrect fiducial volume just give wrong answer)

$$\left(\frac{d\sigma}{dx}\right)_i=\frac{\sum_j \boxed{U_{ij}}(d_j-\boxed{b_j})}{\boxed{\varepsilon_i}(\boxed{\Phi}\boxed{T})\Delta x_i}$$

Total target number

3.8.6 Systematics error matrix production

Unisim

The error matrix can be made by changing one of systematics and calculate differential cross section ($d\sigma^s/dx$), then take a difference with differential cross section calculated with central value MC ($d\sigma/dx$).

$$E_{ij}=\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^s}{dx}\right)_j\right]$$

Multisim

If there is a correlation between systematics (input error matrix), it should propagate correctly. In this case, number of $d\sigma^s/dx$ with different set of systematics drawn from input error matrix make many error matrices. Then, we take average of them to construct output error matrix.

$$E_{ij}=\frac{1}{M}\sum_s\left[\left(\frac{d\sigma}{dx}\right)_i-\left(\frac{d\sigma^s}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_j-\left(\frac{d\sigma^s}{dx}\right)_j\right]$$

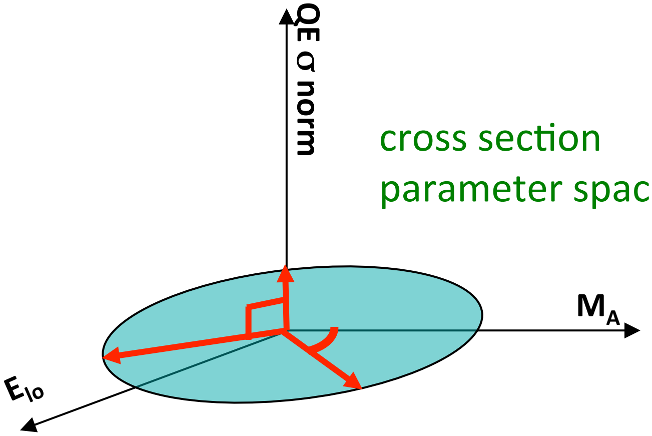
3.8.6 Multisim

ex) cross section uncertainties

M_AQE	6%
E_{lo}sf	2%
QE σ norm	10%

correlated

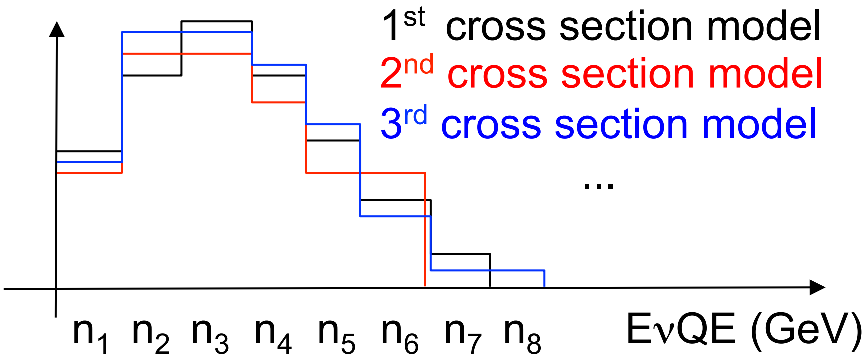
uncorrelated



cross section parameter space

$$M_{input}(xs) = \begin{pmatrix} var(M_A) & cov(M_A, E_{lo}) & 0 \\ cov(M_A, E_{lo}) & var(E_{lo}) & 0 \\ 0 & 0 & var(\sigma - norm) \end{pmatrix}$$

cross section error for EνQE



repeat this exercise many times to create smooth error matrix for EνQE

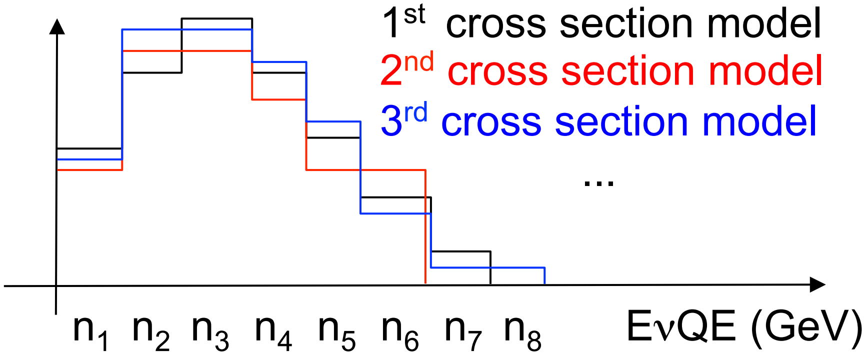
3.8.6 Multisim

Output cross section error matrix for EνQE

$$\left[M_{output}(xs)\right]_{ij} \approx \frac{1}{S}\sum_{k}^{S}\left(N_i^k(xs)-N_i^{MC}\right)\left(N_j^k(xs)-N_j^{MC}\right)$$

$$M_{output}(xs)=\begin{pmatrix} var(n_1) & cov(n_1,n_2) & cov(n_1,n_3) & \cdots \\ cov(n_1,n_2) & var(n_2) & cov(n_2,n_3) & \cdots \\ cov(n_1,n_3) & cov(n_2,n_3) & var(n_3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

cross section error for EνQE



repeat this exercise many times to create smooth error matrix for EνQE

3.8.7 Statistics error

NC1π^o CC1π^o

Statistical error propagation

Due to unfolding, there is data statistical error on off-diagonal term of error matrix. The diagonal statistical error can be propagated through Jacobian. It is weakened, and smoothly migrate to off diagonal. MC statistics can be transferred by similar way if it is large.

$$V_{ij}\left[\frac{d\sigma}{dx}\right] = V_{ij}\left[\frac{\sum U(d-b)}{\varepsilon(\Phi T)\Delta x}\right] = \left(\frac{\partial\left[\frac{\sum U(d-b)}{\varepsilon(\Phi T)\Delta x}\right]}{\partial d}\right)_{ki} V_{km}[d]\left(\frac{\partial\left[\frac{\sum U(d-b)}{\varepsilon(\Phi T)\Delta x}\right]}{\partial d}\right)_{mj}$$

Statistical error through Multisim

NCEL

Fake data set is made by applying fluctuation on data within data statistics. Then statistics multisim output error matrix is made from fake data set.

CCQE CC1π^+ anti CCQE anti NCEL

Statisical error through detector error matrix

Detector error multisim MC set is made with data statistics (MiniBooNE historic reason), so the multisim output error matrix has ~data statistical error, too.

3.8.8 Correlated systematic errors between samples

Correlated errors between T2K cross section results

T2K used many MiniBooNE cross section results for the global fit, however, MiniBooNE data errors should be correlated because all of them are measured by the same beamine and the detector.

To avoid same mistake, we should provide correlated errors? For example

1. flux normalization
2. π/K ratio (Mark H.)

Providing these 2 numbers could reduce over all errors dramatically for the global fit.

3.1 Signal definition

3.2 Background removing

3.3 Unsmearing

3.4 Efficiency correction

3.5 Flux correction

3.6 Target number

3.7 Binning

3.8 Systematic errors

3.9 Data format

$$\left(\frac{d\sigma}{dx}\right)_i = \frac{\sum_j U_{ij}(d_j - b_j)}{\varepsilon_i(\Phi T)\Delta x_i}$$

U_{ij} :unsmearing matrix

b_j :predicted background

d_j :data vector

Δx_i :bin width

ε_i :efficiency

Φ :integrated ν-flux

T :integrated target number

3.9 Data format

CCQE NCEL NC1π^o CC1π^+ CC1π^o anti CCQE anti NCEL

Tables on MiniBooNE data release website

- In MiniBooNE, all cross section tables, as well as flux table, are released in website http://www-boone.fnal.gov/for_physicists/data_release/

Cross secti

- Flux-integ

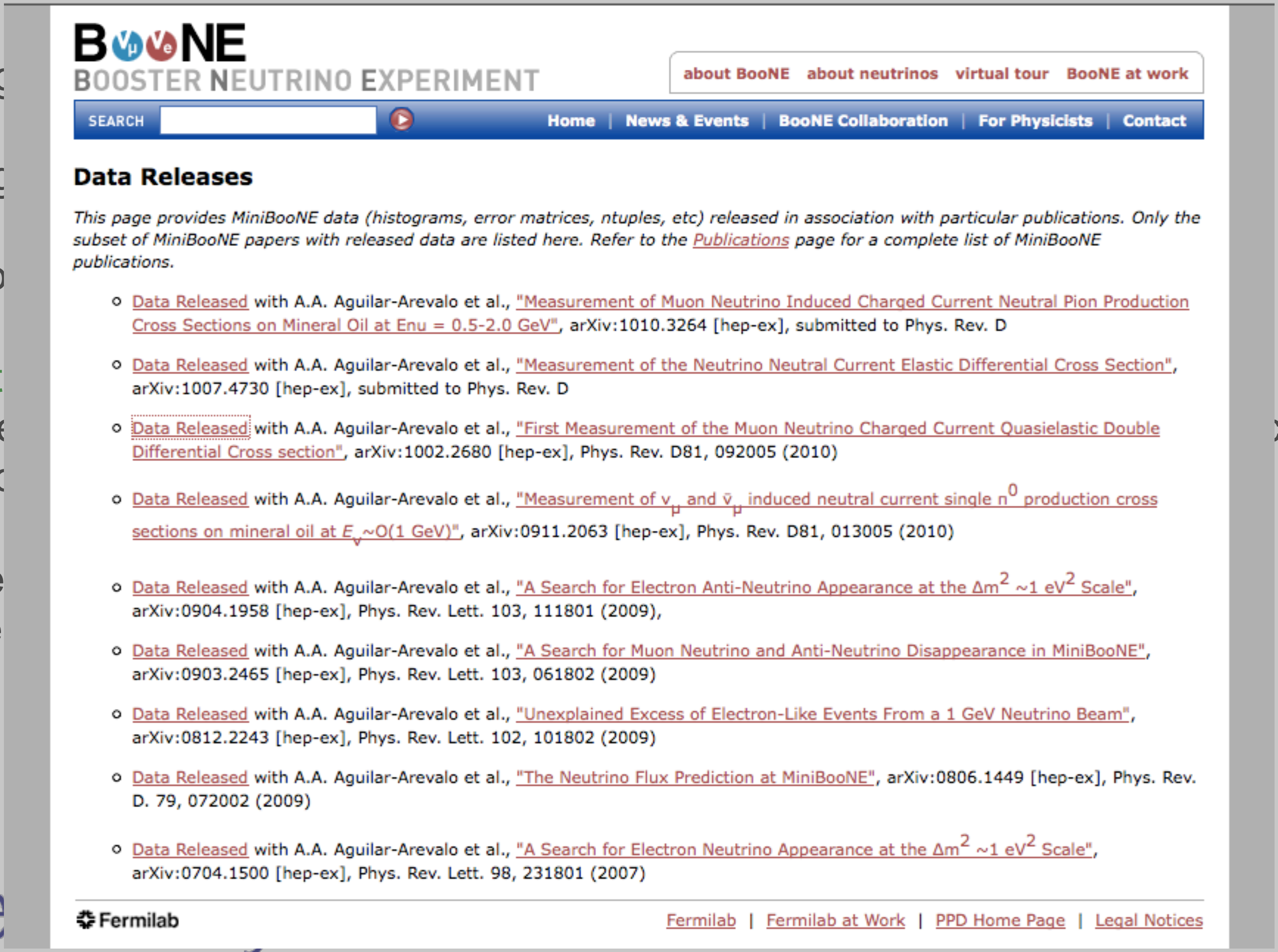
- Flux integ

- Flux-unfo

Additional t

- signal-like … clusive cross sectio

- response observed e



Data Releases

This page provides MiniBooNE data (histograms, error matrices, ntuples, etc) released in association with particular publications. Only the subset of MiniBooNE papers with released data are listed here. Refer to the Publications page for a complete list of MiniBooNE publications.

- Data Released with A.A. Aguilar-Arevalo et al., "Measurement of Muon Neutrino Induced Charged Current Neutral Pion Production Cross Sections on Mineral Oil at Enu = 0.5-2.0 GeV", arXiv:1010.3264 [hep-ex], submitted to Phys. Rev. D
- Data Released with A.A. Aguilar-Arevalo et al., "Measurement of the Neutrino Neutral Current Elastic Differential Cross Section", arXiv:1007.4730 [hep-ex], submitted to Phys. Rev. D
- Data Released with A.A. Aguilar-Arevalo et al., "First Measurement of the Muon Neutrino Charged Current Quasielastic Double Differential Cross section", arXiv:1002.2680 [hep-ex], Phys. Rev. D81, 092005 (2010)
- Data Released with A.A. Aguilar-Arevalo et al., "Measurement of ν_μ and $\bar{\nu}_\mu$ induced neutral current single n^0 production cross sections on mineral oil at E_ν~O(1 GeV)", arXiv:0911.2063 [hep-ex], Phys. Rev. D81, 013005 (2010)
- Data Released with A.A. Aguilar-Arevalo et al., "A Search for Electron Anti-Neutrino Appearance at the Δm^2 ~1 eV^2 Scale", arXiv:0904.1958 [hep-ex], Phys. Rev. Lett. 103, 111801 (2009),
- Data Released with A.A. Aguilar-Arevalo et al., "A Search for Muon Neutrino and Anti-Neutrino Disappearance in MiniBooNE", arXiv:0903.2465 [hep-ex], Phys. Rev. Lett. 103, 061802 (2009)
- Data Released with A.A. Aguilar-Arevalo et al., "Unexplained Excess of Electron-Like Events From a 1 GeV Neutrino Beam", arXiv:0812.2243 [hep-ex], Phys. Rev. Lett. 102, 101802 (2009)
- Data Released with A.A. Aguilar-Arevalo et al., "The Neutrino Flux Prediction at MiniBooNE", arXiv:0806.1449 [hep-ex], Phys. Rev. D. 79, 072002 (2009)
- Data Released with A.A. Aguilar-Arevalo et al., "A Search for Electron Neutrino Appearance at the Δm^2 ~1 eV^2 Scale", arXiv:0704.1500 [hep-ex], Phys. Rev. Lett. 98, 231801 (2007)

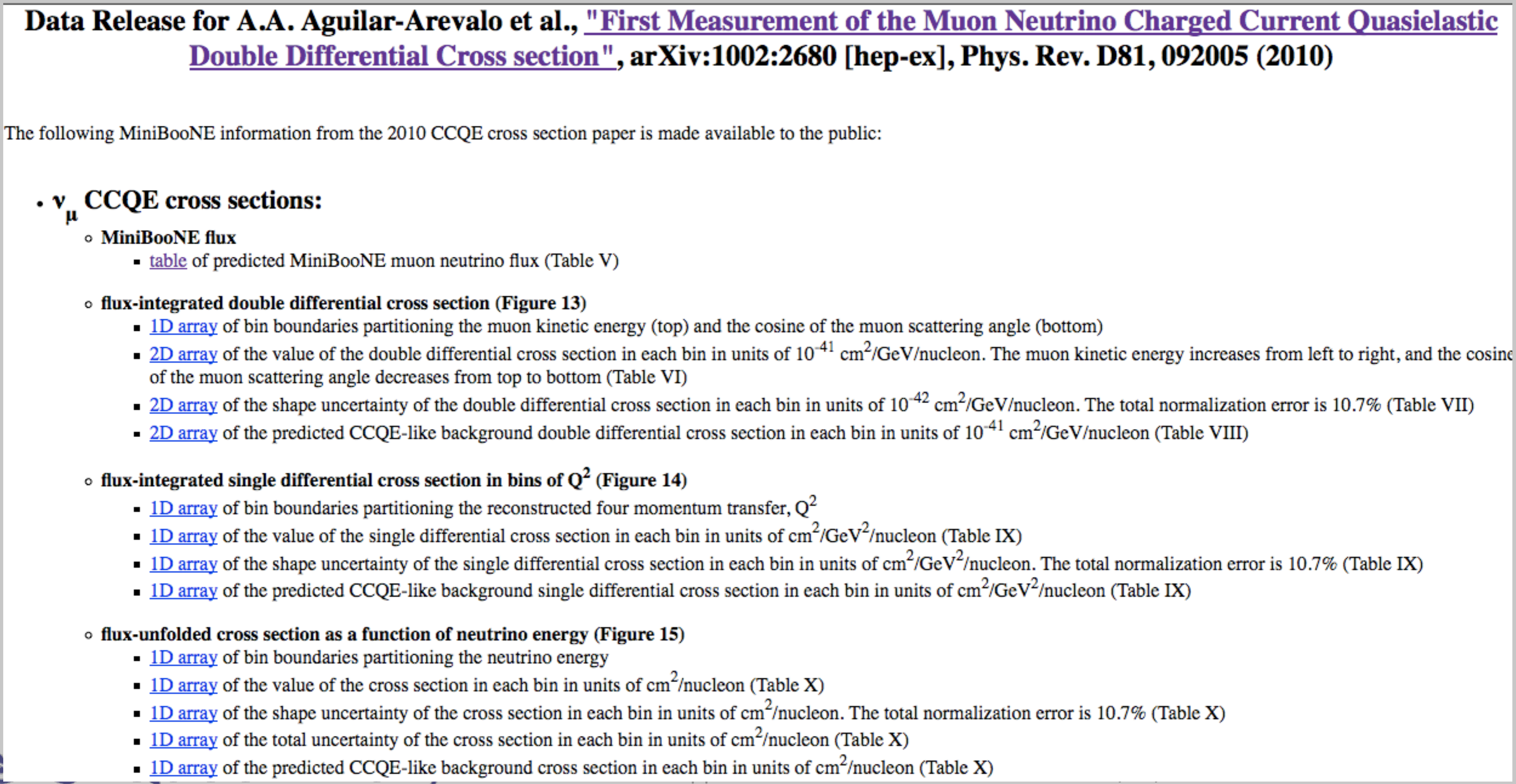
3.9 Data format

CCQE | NCEL | NC1π^o | CC1π^+ | CC1π^o | anti CCQE | anti NCEL

Tables on MiniBooNE data release website

- In MiniBooNE, all cross section tables, as well as flux table, are released in website

http://www-boone.fnal.gov/for_physicists/data_release/



Data Release for A.A. Aguilar-Arevalo et al., "First Measurement of the Muon Neutrino Charged Current Quasielastic Double Differential Cross section", arXiv:1002:2680 [hep-ex], Phys. Rev. D81, 092005 (2010)

The following MiniBooNE information from the 2010 CCQE cross section paper is made available to the public:

- **ν_μ CCQE cross sections:**
  - **MiniBooNE flux**
    - table of predicted MiniBooNE muon neutrino flux (Table V)
  - **flux-integrated double differential cross section (Figure 13)**
    - 1D array of bin boundaries partitioning the muon kinetic energy (top) and the cosine of the muon scattering angle (bottom)
    - 2D array of the value of the double differential cross section in each bin in units of 10^{-41} cm^2/GeV/nucleon. The muon kinetic energy increases from left to right, and the cosine of the muon scattering angle decreases from top to bottom (Table VI)
    - 2D array of the shape uncertainty of the double differential cross section in each bin in units of 10^{-42} cm^2/GeV/nucleon. The total normalization error is 10.7% (Table VII)
    - 2D array of the predicted CCQE-like background double differential cross section in each bin in units of 10^{-41} cm^2/GeV/nucleon (Table VIII)
  - **flux-integrated single differential cross section in bins of Q^2 (Figure 14)**
    - 1D array of bin boundaries partitioning the reconstructed four momentum transfer, Q^2
    - 1D array of the value of the single differential cross section in each bin in units of cm^2/GeV^2/nucleon (Table IX)
    - 1D array of the shape uncertainty of the single differential cross section in each bin in units of cm^2/GeV^2/nucleon. The total normalization error is 10.7% (Table IX)
    - 1D array of the predicted CCQE-like background single differential cross section in each bin in units of cm^2/GeV^2/nucleon (Table IX)
  - **flux-unfolded cross section as a function of neutrino energy (Figure 15)**
    - 1D array of bin boundaries partitioning the neutrino energy
    - 1D array of the value of the cross section in each bin in units of cm^2/nucleon (Table X)
    - 1D array of the shape uncertainty of the cross section in each bin in units of cm^2/nucleon. The total normalization error is 10.7% (Table X)
    - 1D array of the total uncertainty of the cross section in each bin in units of cm^2/nucleon (Table X)
    - 1D array of the predicted CCQE-like background cross section in each bin in units of cm^2/nucleon (Table X)

3.9 Data format

CCQE NCEL NC1π^o CC1π^+ CC1π^o anti CCQE anti NCEL

Tables on MiniBooNE data release website

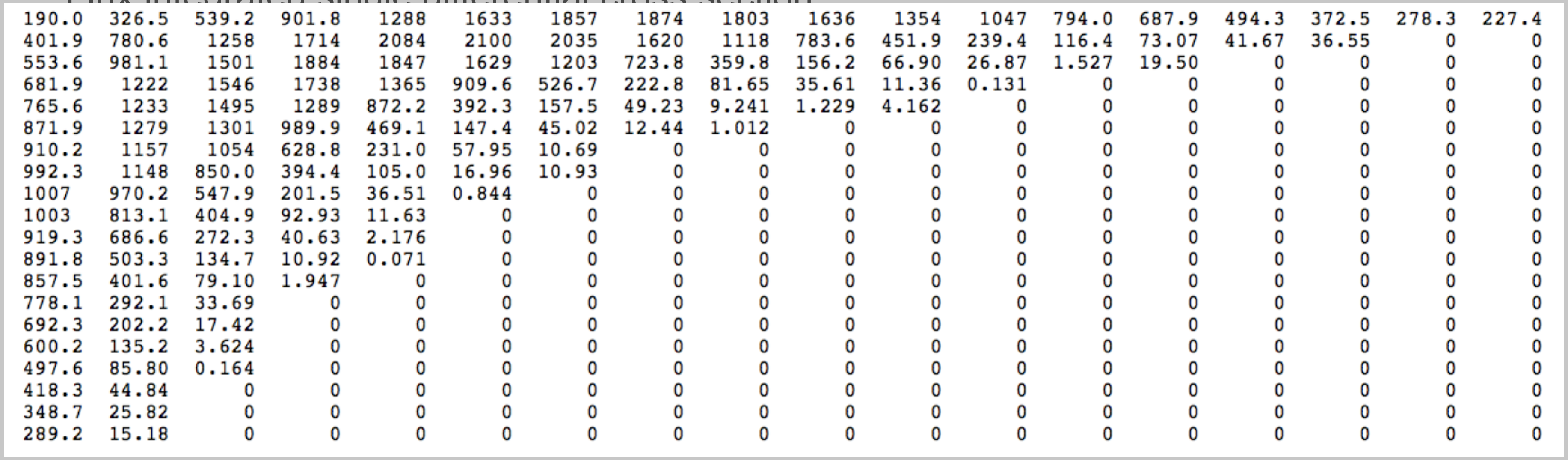
- In MiniBooNE, all cross section tables, as well as flux table, are released in website

http://www-boone.fnal.gov/for_physicists/data_release/

Cross section format

- Flux-integrated double differential cross section

- Flux integrated single differential cross section



```
190.0  326.5  539.2  901.8   1288   1633   1857   1874   1803   1636   1354   1047  794.0  687.9  494.3  372.5  278.3  227.4
401.9  780.6   1258   1714   2084   2100   2035   1620   1118  783.6  451.9  239.4  116.4  73.07  41.67  36.55      0      0
553.6  981.1   1501   1884   1847   1629   1203  723.8  359.8  156.2  66.90  26.87  1.527  19.50      0      0      0      0
681.9   1222   1546   1738   1365  909.6  526.7  222.8  81.65  35.61  11.36  0.131      0      0      0      0      0      0
765.6   1233   1495   1289  872.2  392.3  157.5  49.23  9.241  1.229  4.162      0      0      0      0      0      0      0
871.9   1279   1301  989.9  469.1  147.4  45.02  12.44  1.012      0      0      0      0      0      0      0      0      0
910.2   1157   1054  628.8  231.0  57.95  10.69      0      0      0      0      0      0      0      0      0      0      0
992.3   1148  850.0  394.4  105.0  16.96  10.93      0      0      0      0      0      0      0      0      0      0      0
1007   970.2  547.9  201.5  36.51  0.844      0      0      0      0      0      0      0      0      0      0      0      0
1003   813.1  404.9  92.93  11.63      0      0      0      0      0      0      0      0      0      0      0      0      0
919.3  686.6  272.3  40.63  2.176      0      0      0      0      0      0      0      0      0      0      0      0      0
891.8  503.3  134.7  10.92  0.071      0      0      0      0      0      0      0      0      0      0      0      0      0
857.5  401.6  79.10  1.947      0      0      0      0      0      0      0      0      0      0      0      0      0      0
778.1  292.1  33.69      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
692.3  202.2  17.42      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
600.2  135.2  3.624      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
497.6  85.80  0.164      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
418.3  44.84      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
348.7  25.82      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
289.2  15.18      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
```

3.9 Data format

CCQE NCEL NC1π^o CC1π^+ CC1π^o anti CCQE anti NCEL

Tables on MiniBooNE data release website

- In MiniBooNE, all cross section tables, as well as flux table, are released in website
http://www-boone.fnal.gov/for_physicists/data_release/

Cross section format

CCQE CC1π^+ anti CCQE

- Flux-integrated double differential cross section

CCQE NCEL NC1π^o CC1π^+ CC1π^o anti CCQE anti NCEL

- Flux integrated single differential cross section
- Flux-unfolded total cross section

CCQE NC1π^o CC1π^+ CC1π^o anti CCQE

Additional tables

- signal-like background tables are presented, so that people can use either exclusive cross section or effective cross section.

CCQE NCEL

- response matrix R is presented so that people can calculate MiniBooNE observed energy spectrum

NCEL

3.9 Data format

MINERvA style table?

- In MINERvA, data tables are all in “supplemental material” and the papers only have cross section plots, to fit in PRL page limit. I think this is clever.

3.9 Cross section errors

NC1π^o CC1π^+ CC1π^o

Cross section error format

- Complete error matrix of differential cross sections

NCEL

- Complete error matrix for reconstructed energy spectrum

CCQE

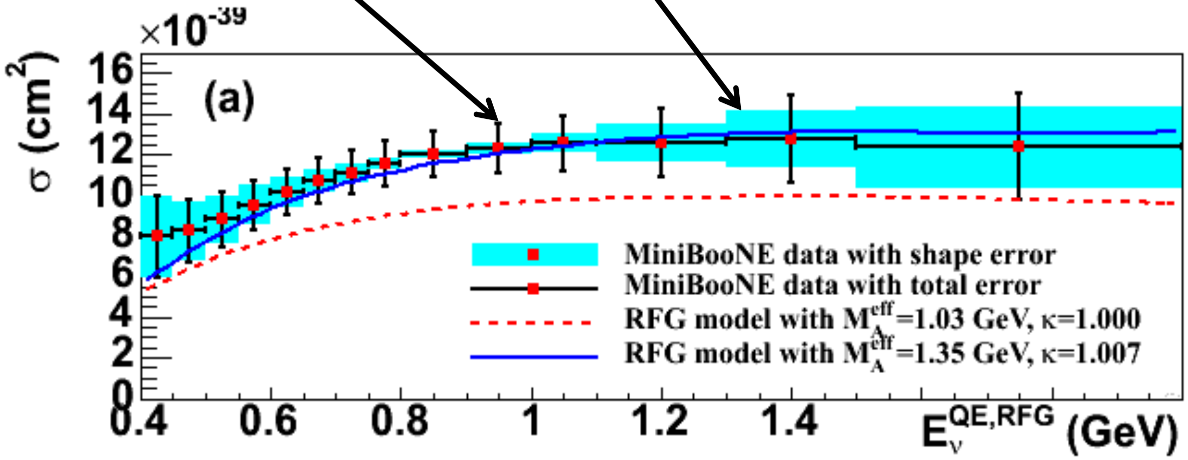
anti CCQE

- Diagonal term of shape only error matrix and total normalization error.

Diagonal term of shape only error matrix has information of covariance of total error matrix, so this is a convenient way to show bin-bin correlation in 1-dimention

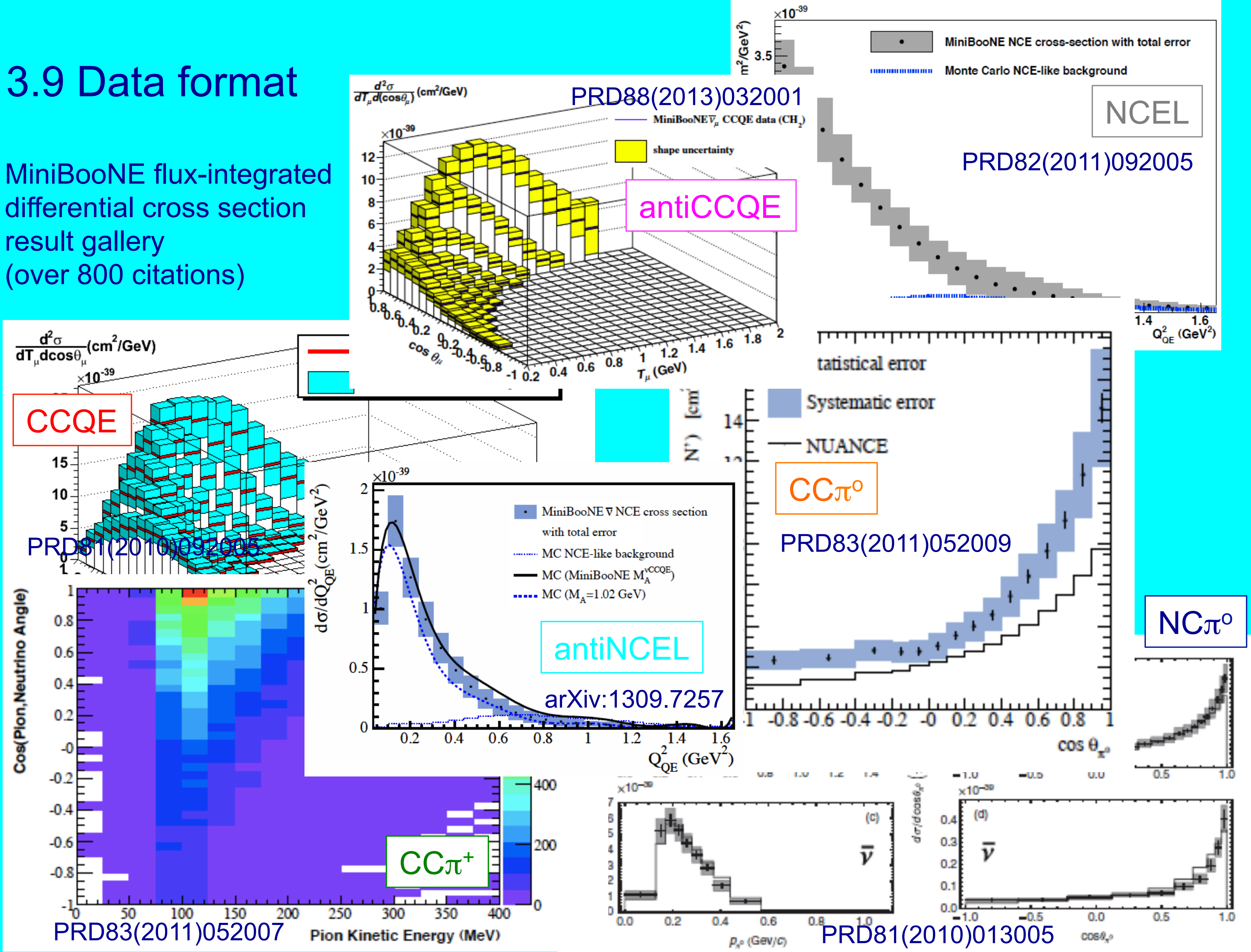
$$\delta V_i^{total} = \sqrt{E_{ii}}$$

$$\delta V_i^{shape} = \sqrt{E_{ii} - 2\frac{V_i}{N}\sum_k^n E_{ki} + \frac{V_i^2}{N^2}\sum_{k,m}^n E_{km}}$$



3.9 Data format

MiniBooNE flux-integrated differential cross section result gallery
(over 800 citations)



CCQE

PRD81(2010)092005

antiCCQE

PRD88(2013)032001

NCEL

PRD82(2011)092005

CCπ^o

PRD83(2011)052009

antiNCEL

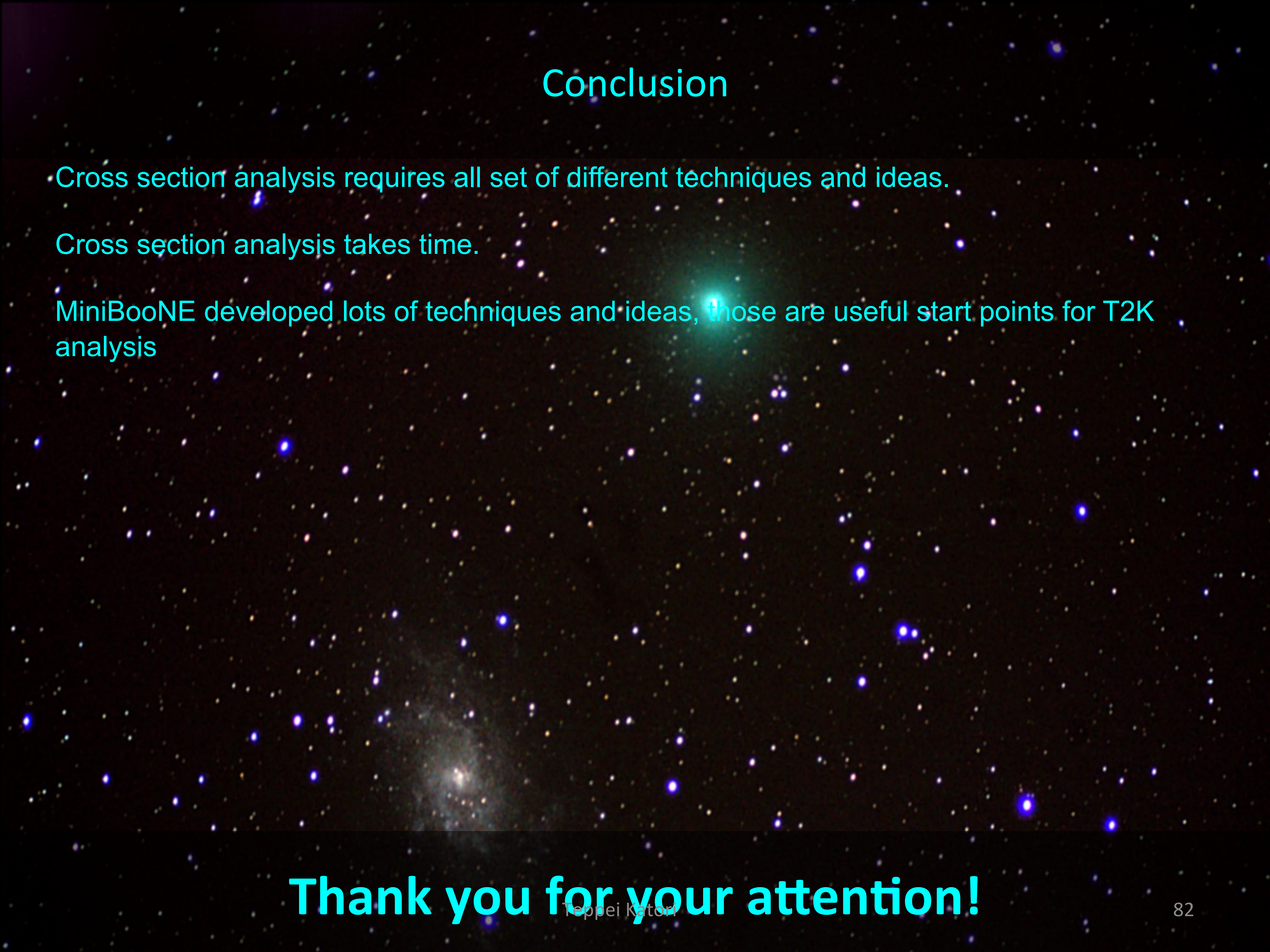
arXiv:1309.7257

NCπ^o

PRD81(2010)013005

CCπ$^+$

PRD83(2011)052007



Conclusion

Cross section analysis requires all set of different techniques and ideas.

Cross section analysis takes time.

MiniBooNE developed lots of techniques and ideas, those are useful start points for T2K analysis

Thank you for your attention!

3.2 Background removing

CCQE NCEL CC1π^+ CC1π^o anti CCQE anti NCEL

Data driven correction

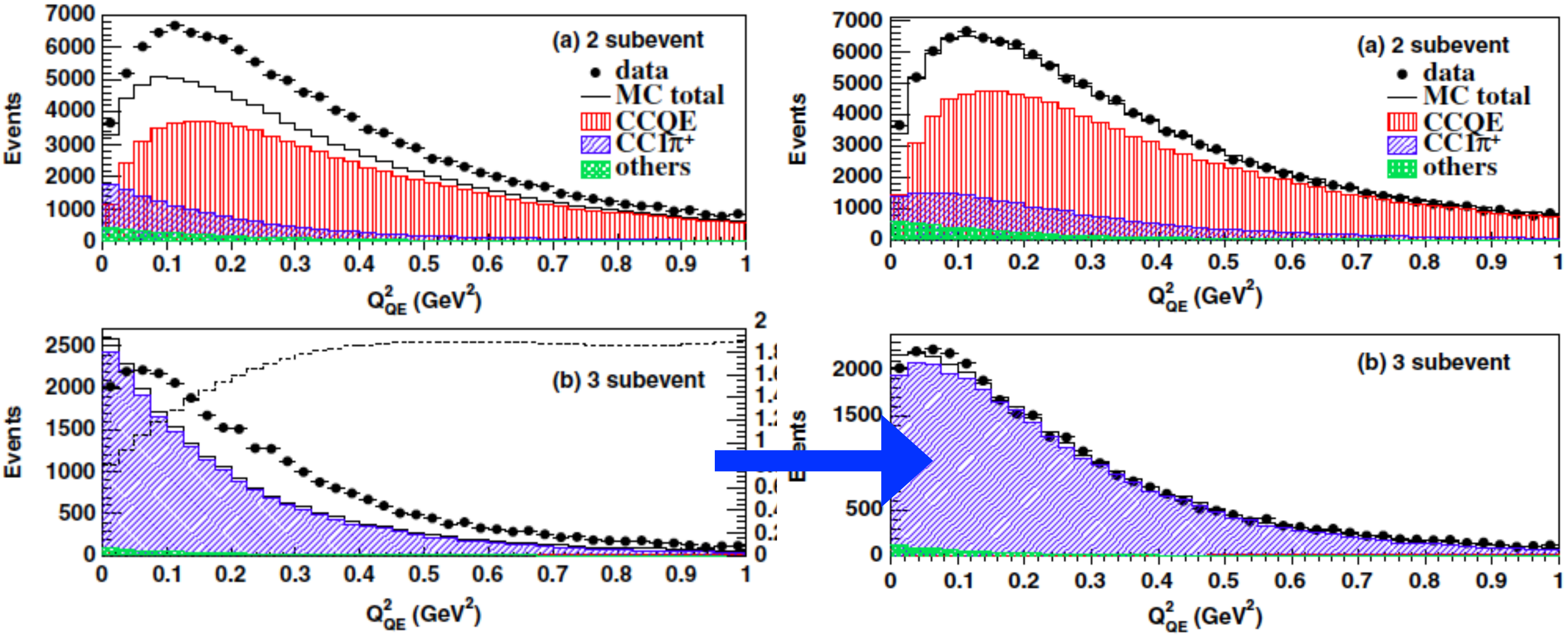
- Almost all analyses use sideband data to correct background distribution in signal boxes

CCQE-CC1π^+ box simultaneous fit

CCQE

- CC1π^+ box is >90% purity. Using this sample, background distribution in CCQE box is modified.

External background enhanced sample



3.2 Background removing

Cross section error of pion kinematics CC1π^+

- Purity is extremely high (~90%), so error from background is negligible.
- Signal is defined from pion in final state, so pion absorption is not systematic error.
  → In general, FSI is small error for all analyses
- The pion absorption in the oil is error.
  → Secondary scattering will be the largest error for hadron measurement.

I think you revise all errors to think which is the important error etc. If you run analysis machinery with all errors on, you will over-estimate systematic errors.

3.3 Unsmearing

How to construct M-matrix? Reconstructed variables

- The definition of true kinematics is tricky, because you have choice.

CCQE NCEL anti CCQE anti NCEL

i. True Q^2 is defined by reconstructed Q^2 from true kinematics

For example, CCQE, true Q^2 is defined "reconstructed Q^2 from true muon energy and angle", and we call it "Q^2_{QE}" to remind people this is reconstructed under QE assumption.

$CC1\pi^+$ $CC1\pi^o$

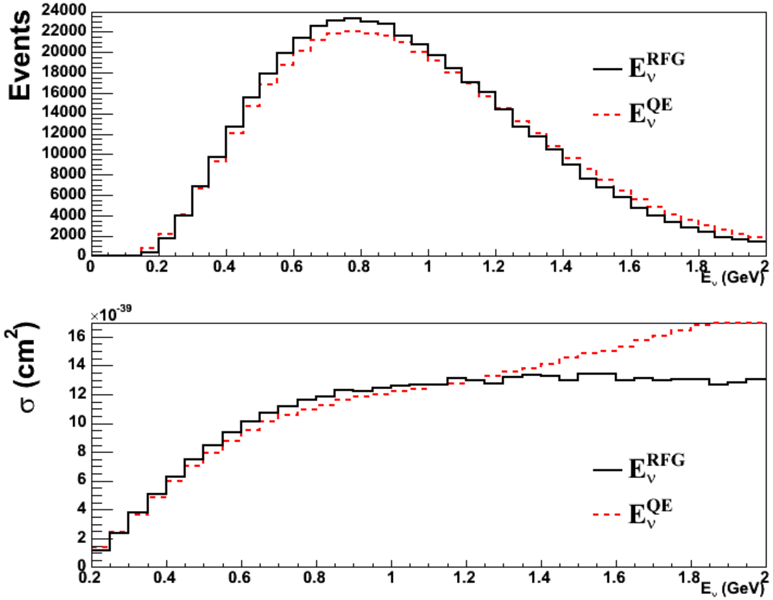
ii True Q^2 is defined by true Q^2 in MC

This may be useful to compare with old data, only presented by this way

CCQE $NC1\pi^o$ $CC1\pi^+$ $CC1\pi^o$ anti CCQE

iii True $E\nu$ is defined by true $E\nu$ in MC

For example, CCQE, $E\nu$ is called "$E\nu^{QE,RFG}$" to remind people this is reconstructed under QE assumption then unfolded by assuming RFG model.



Flux-unfolded total cross section

- it is important to unfold under RFG, otherwise reconstruction bias deviate cross section at the tail significantly.