



UNIVERSITY OF WROCLAW
FACULTY OF PHYSICS & ASTRONOMY
INSTITUTE OF THEORETICAL PHYSICS

**Investigation of MINERVA experiment
data using NuWro Monte Carlo
generator with spectral function model**

Author:

Michał Siemaszko

Supervisor:

prof. Jan T. Sobczyk

12 September 2018

Abstract

Monte Carlo event generators are an essential tool for particle physics research nowadays, which still needs improvements. This paper is dedicated *NuWro* - neutrino generator developed in Wrocław - and more precisely, its discrepancies with experimental data from MINERvA experiment. The procedure for minimization of these differences is shown. It relay on MEC events (meson exchange current) analysis. As result, a discrete reweighting function is proposed. Also, an attempt to find its analytic form was made. After all, χ^2 analysis has been done and the outcome is compared to the results of another group - GENIE.

Streszczenie

Generatory zdarzeń Monte Carlo są obecnie niezbędnymi narzędziami w fizyce cząstek elementarnych, które jednak wciąż potrzebują udoskonalania. Owa praca jest poświęcona *NuWro* - generatorowi poświęconemu neutrinom, który jest rozwijany we Wrocławiu - a dokładniej jego rozbieżnością z danymi doświadczalnymi z eksperymentu MINERvA. Została zaprezentowana procedura mająca na celu zminimalizowanie tych niezgodności. Opiera się ona na analizie zdarzeń MEC (*meson exchange current*). Jako wynik zaprezentowana została dyskretna funkcja przewagowania. Została również przeprowadzona próba znalezienia jej analitycznej formy. Na końcu została przeprowadzona analiza χ^2 , a jej rezultat został porównany do wyników innej grupy badawczej - GENIE.

Contents

1	Introduction	9
2	NuWro	11
2.1	Dynamics description	11
2.1.1	Quasi-elastic neutrino scattering (QEL)	12
2.1.2	Resonant single pion production (RES)	12
2.1.3	Deep inelastic scattering (DIS)	13
2.1.4	Meson exchange current (MEC)	13
2.1.5	Coherent pion production (COH)	14
2.2	Cross-section	14
3	Reweighting	16
3.1	Preparation	16
3.2	Analysis of results	17
3.3	Problem	19
3.4	Investigation	20
3.5	Changing domain	22
3.6	<i>GENIE</i> study	24
4	Results	26
4.1	Introduction to χ^2 test	26
4.2	χ^2 calculation	28
5	Summary	30

List of Figures

2.1	Feynman diagram of a QEL CC event - n and p stand for neutron and proton accordingly	12
2.2	Feynman diagram of a RES event. Symbol Δ means Δ^{++} , Δ^+ , Δ^0 or Δ^- and π stands for π^+ , π^0 or π^- - every combination, which fulfils charge conservation law, is possible	12
2.3	Feynman diagram of a DIS event - N stands for nucleon and three empty lines means produced in this process particles	13
2.4	MEC event scheme. π is particle exchanged between two nucleons, N_1 and N_2	14
2.5	COH event scheme	14
2.6	Φ is neutrino flux and A is area of target. With gray stars interactions is marked. σ is effective area covered by one particle	15
3.1	Energy distribution of beam particles	16
3.2	Three phases of interaction (from left): in, out and post	17
3.3	Used parameters of histograms	19
3.4	All of generated histograms. Each of them represent different range of p_L . On x-axis is p_T and on y-axis is cross-section. With black crosses experimental results of T2K is marked.	20
3.5	$\delta(p_T, p_L)$ matrix. Inside red polygon are bins which will be analyzed	21
3.6	Histograms later used to calculate $\delta(\omega, q)$	22
3.7	$\delta(\omega, q)$ with bins, where events number is less than 10 (a), and without them (b)	23

3.8	All of improved histograms. Each of them represent different range of p_L . On x-axis is p_L and on y-axis is cross-section. With black crosses experi- mental results of T2K is marked.	24
3.9	Graphical form of fitting functions	25
4.1	Example - starting bins (middle) can be improved by one of two rescaling: left or right. Red is data points (with uncertainty) and blue is what was added to the bin	27
4.2	χ^2 for NuWro discrete without covariance with different numbers of bins	29

1 | Introduction

At the beginning of 20th century, scientists had observed that beta decay, as they knew it (i.e. $n \rightarrow p + e^-$), apparently does not conserve energy. This was a really shocking discovery, because physicists believed (and still believe), that no matter what process is happening, energy is always conserved. After that, several theories were proposed to save this one of the most fundamental law of physic - among them was Pauli theory (1930). He suggested, that after decay there is an additional particle in the final state (i.e. $n \rightarrow p + e^- + \bar{\nu}_e$), which he called neutron and now we call it neutrino. This particle is supposed to carry missing energy and momentum. For experimental evidence we had to wait over 25 years, when in 1956 Clyde Cowan and Frederick Reines have proven the existence of neutrinos.

Since then, those particles have become the topic of intense research. Engineers with scientists all around the world have been building detectors like Homestake (USA), Kamiokande (Japan) or IceCube (Antarctica). They help us to understand the properties of neutrinos and their interaction with matter. It is a difficult task, because for those particles regular matter is almost transparent and interactions happen very rarely.

After first detectors had been built, scientists have started to observe neutrinos from the Sun. It was really surprising when it turns out, that they were detecting 2-3 times fewer particles, than it was predicted by theory. It was later explained by new mechanism - neutrino oscillation. It means, that neutrino can change its flavor (family) while traveling. E.g. Sun emits only electron antineutrinos but we can detect all the flavors (electron, muon, and taon) in our detector. For that discovery Nobel Prize has been granted in 2002 to Raymond Davis Jr. and Masatoshi Koshihira.

An integral part of every neutrino experiment is Monte Carlo event generator. It is a theoretical tool, that helps experimentalists in an analysis of gathered data, like distinction

interesting event from the background or calculating uncertainties. Due to complicated and in practice impossible to solve equations, which describe neutrinos interactions, theorists have to use numerical methods - based on Monte Carlo algorithm. It is often said, that the generators serve as a bridge between theory and experiment.

Today there are number of Monte Carlo neutrino event generator and numerical tools, for example GENIE, NEUT, GiBUU or NuWro. They use different models for calculation and thus discrepancies between them are sometimes huge. For their comparison a separate tool has been created, called NUISANCE. But after all, the goal is one - to match the experimental data. In my thesis, I work with the *NuWro* event generator, created by the group of theoretical physicists from Wrocław University under the direction of the prof. Jan Sobczyk. It has been developed since 2005.

2 | NuWro

2.1 | Dynamics description

Neutrino interacts with matter only ¹ by weak force exchanging W^\pm or Z^0 boson. But there is no one universal rule to calculate strength/probability of interactions, because there are many kinds of them and each has to be treated individually. In NuWro 10 dynamics (i.e. types of interaction) are implemented. They can be divided into two main categories related to a type of exchange boson:

- charged current (CC) – due to exchange of W^\pm boson; it means that incoming neutrino changes to charged lepton in the same family e.g.



- neutral current (NC) – due to exchange of Z^0 boson; neutrino does not change during interaction e.g.



In the next five sections, NuWro dynamics are shortly described. Each of them can be either CC or NC type. Here I discuss only first of them, because further I will work only with those.

¹And also gravitationally ever since they have mass, but it will be not included in this paper as it is negligible.

2.1.1 | Quasi-elastic neutrino scattering (QEL)

It is the simplest neutrino interaction and dominant in energy below 1.5 GeV. In this case, a neutrino (ν_l) is converted to a charged lepton - l (in the same family region) and nucleon change charge as it describes diagrams in Figure 2.1



Figure 2.1: Feynman diagram of a QEL CC event - n and p stand for neutron and proton accordingly

Of course, it is hard to find protons or neutrons in an unbounded state (outside nucleus). But if neutrino has enough energy it starts to *see* single nucleons when hitting nucleus and this interaction becomes possible.

2.1.2 | Resonant single pion production (RES)

This process is very important within neutrino energy range between 1.5 GeV and ~10 GeV. Typically it occurs, when neutrino transfer enough energy to struck nucleon (N) that it gets excited. In that case, baryon resonance is produced - usually it is Δ particle. Because it is unstable, it quickly decays - most often into nucleon (N') and pion (π). Example diagrams is shown in Figure 2.2.

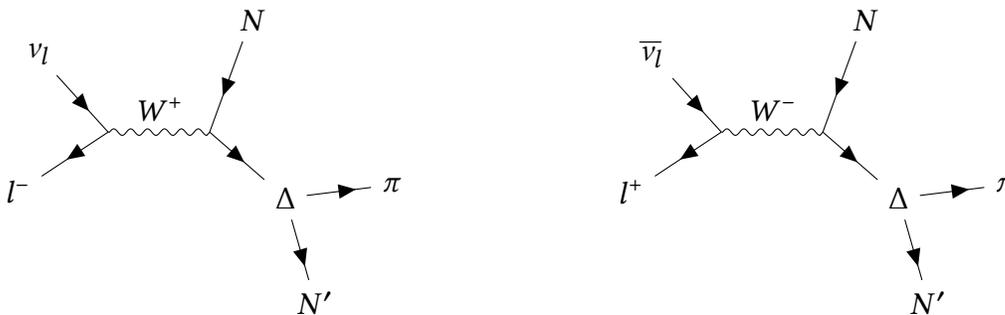


Figure 2.2: Feynman diagram of a RES event. Symbol Δ means Δ^{++} , Δ^+ , Δ^0 or Δ^- and π stands for π^+ , π^0 or π^- - every combination, which fulfils charge conservation law, is possible

Produced pion does not necessary leave nucleus. Because interaction occurs inside nucleus, pion have to travel through it. During its journey it can be absorbed by nucleon. So even though pion is produced, it can remain undetected.

2.1.3 | Deep inelastic scattering (DIS)

When a neutrino has a significantly high energy (over 20 GeV) it starts to scatter on single quark. It causes, like earlier, change lepton to charged one but also hadron shower i.e. lots of particles in the final state. The basic scheme is shown in Figure 2.3.

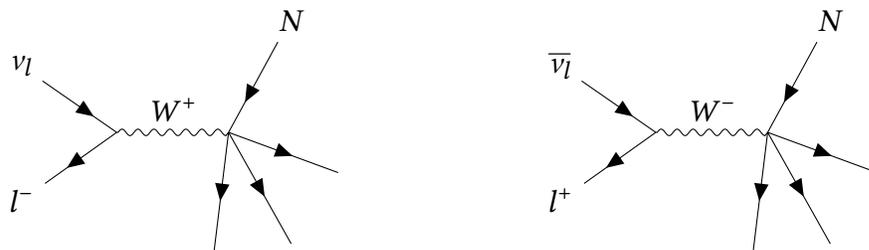


Figure 2.3: Feynman diagram of a DIS event - N stands for nucleon and three empty lines means produced in this process particles

2.1.4 | Meson exchange current (MEC)

It is yet the most inaccurate described interaction. Here, exchange boson is interacting not with only one nucleon, but with two of them at the same time. In result, those two particles are liberated from a nucleus. Other name of this dynamic is "2 particle - 2 holes" ("2p-2h") meaning 2 particles are emitted from a nucleus and 2 holes are left behind in it. The basic scheme of this process is shown in Figure 2.4.

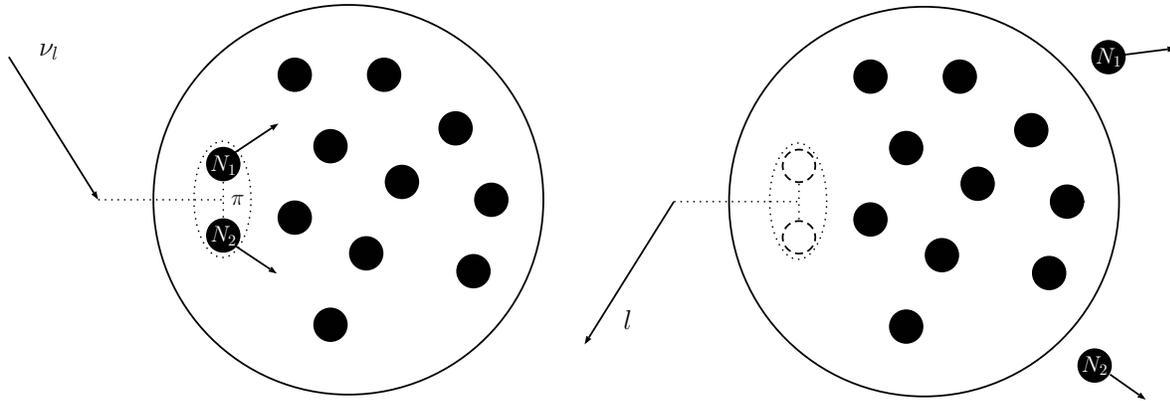


Figure 2.4: MEC event scheme. π is particle exchanged between two nucleons, N_1 and N_2

2.1.5 | Coherent pion production (COH)

In this interaction an incoming neutrino does not see individual particles which made up nucleus, but perceive it as a unity. It causes only momentum and energy transfers. If a neutrino has larger energy, there is a possibility to also produce pion (see Figure 2.5). It can happen, when transfer of four-momentum in this interaction is small.

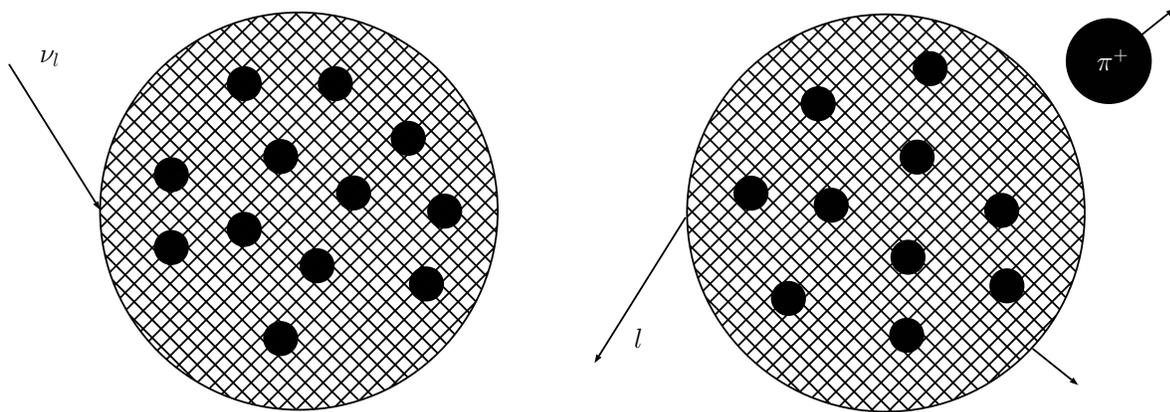


Figure 2.5: COH event scheme

2.2 | Cross-section

A cross-section is a measure of the probability of process to happen. It does not have classical properties like value in the range $[0,1]$. Even more - it has its dimension which

is area. It is defined in the following way. Let a flux of particles Φ hit a target with a small thickness dx and area A , containing N particles per unit volume. That implies, that a number of particles in target equals $\mathcal{N} = NAdx$. If for every of target particle the cross-section (which is effective area covered by it) is equal to σ then probability of interaction is equal

$$P = \frac{n\sigma}{A} = \sigma Ndx \quad (2.3)$$

which is a fraction of area covered by target particles.

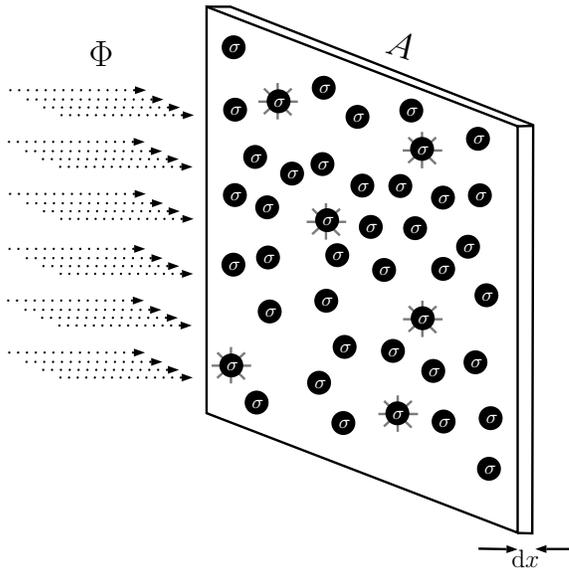


Figure 2.6: Φ is neutrino flux and A is area of target. With gray stars interactions is marked. σ is effective area covered by one particle

The number of neutrinos passing through the target in unit time is equal to ΦA . Thus, the number of interactions in unit time equals to

$$\Phi AP \stackrel{2.3}{=} \Phi A \sigma Ndx = \Phi \sigma \mathcal{N} \quad (2.4)$$

which is also a frequency of interactions happening.

Described above model requires a thickness of the target to be very small (microscopic). Only then data analysis and theoretical calculation are simple. With neutrinos is different - for them it does not matter. Thickness of the target can be even macroscopic. It is caused by an incredibly small probability of neutrino interaction.

This parameter is commonly used to compare experimental data with theoretical calculations - also in this thesis.

3 | Reweighting

3.1 | Preparation

A first thing I need is to make NuWro simulation with main parameters presented in Table 3.1.

Number of events	Beam particle	Energy	Target	Active dynamics	Nucleus model
5×10^6	ν_μ	0 – 100 GeV non-uniform	${}^6_6\text{C}$	QEL_CC, RES_CC, DIS_CC, MEC_CC	spectral function (SF)

Table 3.1: Most important parameters of NuWro run

The distribution of beam particles energy is shown in Figure 3.1.

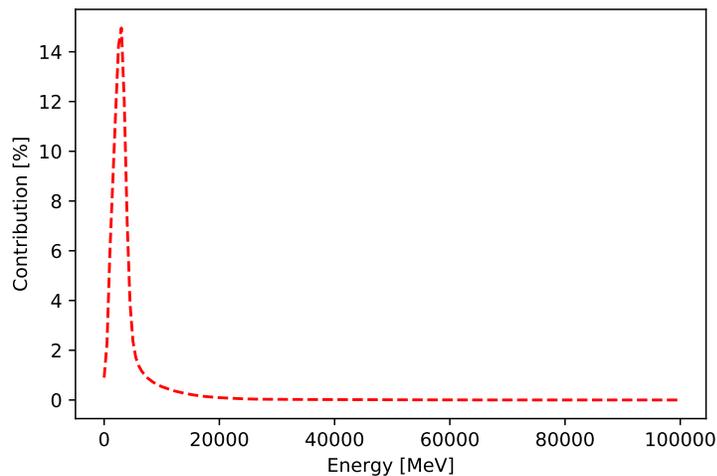


Figure 3.1: Energy distribution of beam particles

Nucleus model is necessary for proper description of interactions. It provides information about removal energy and momentum distribution of nucleons inside a nucleus. Most common and easy to use is the Fermi gas model. It means, that neutrons and protons are treated like free particles bounded by a binding potential of nucleus - no interactions between them are allowed. More accurate results can be obtained by using spectral function model. It includes interaction provided by the nuclear shell model as well as short-range correlations between nucleons.

To fully recreate MINERvA experiment, I have to make proper *cuts* i.e. reject events, which do not fulfill appropriate criteria. In my case, this will be cut on an angle. I do not count events, where angle between outgoing lepton (in my case it is muon) momentum and neutrinos beam directions is below 20° (using notation of Figure 3.3a): $\alpha = \angle(\mathbf{p}, \hat{\mathbf{z}}) < 20^\circ$

3.2 | Analysis of results

As an output I get a file, where information about every event is stored. In it, among other things, three vectors can be found - in, out and post - each one of them contains information about particles in different phases of interaction:

- in - just before interaction - in[0] is a incoming neutrino(anti-neutrino) and the rest are nucleons in nucleus
- out - just after interaction - out[0] is a outgoing neutrino or charged lepton and the remaining are products of interaction and nucleons (still in nucleus)
- post - particle which escape the nucleus - only those are detectable

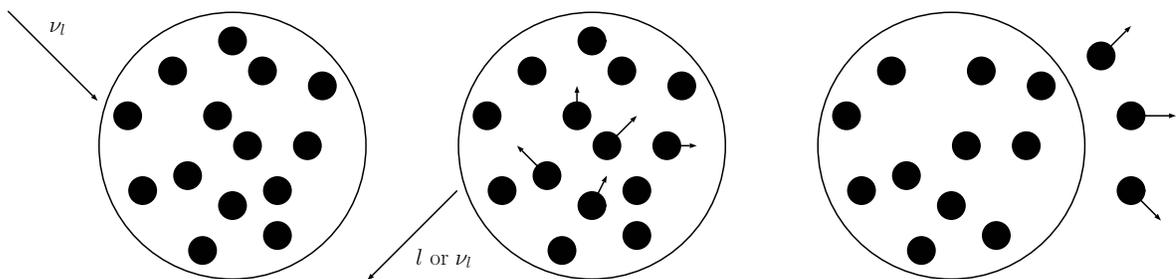


Figure 3.2: Three phases of interaction (from left): in, out and post

Between out and post works algorithm (called *cascade*) which calculates probability of escaping nucleus for every product particle. E.g. there is possibility to have pions in out but not in post. So if we do not see pions in final state, we can not be sure what type of interaction it was (see Section 2.1).

For purpose of this research, I am looking only at events with no pions in the final state (post vector) - so called $CC0\pi$ ¹. Sources of those events can be either CC_QEL or CC_MEC, since they do not produce pions at all. But they are not the only ones - also CC_RES and CC_DIS can produce those (of course, not every event of that type). It is possible, because despite that pion is produced, it can be absorbed by nucleons and thus, it will be not detected (see Section 2.1.2)

$CC0\pi$ made up about 20 – 30% of all generated events. For those, I read two parameters of outgoing muons - longitudinal momenta (\mathbf{p}_L) and transversal momenta (\mathbf{p}_T) (with respect to the beam) - and put them into a histogram (because they are vectors I take their lengths).

Graphical explanation of those parameters is shown in Figure 3.3a. Longitudinal momentum is a projection of whole particle momentum on the chosen axis (beam direction) and transversal is the difference between those two. These parameters have been chosen, because they can be measured in detectors and can be easily compared to data. Nonetheless, there are other pairs of variables to histogramize by - one of them, which we will use later, are a transfer of energy (ω) and transfer of momentum (\mathbf{q}). First of them is the difference between an incoming neutrino and outgoing lepton energy, and the second is similar but with momentum. Those can not be measured experimentally neither compare directly to data. Using Figure 3.3b notation it could be written as

$$\omega = E_\nu - E_l \qquad \mathbf{q} = \mathbf{k}_\nu - \mathbf{k}_l \qquad (3.1)$$

For drawing, a stacked histogram was used. It means, that it is possible to see how big are contributions of each dynamics to every bin. After that normalization has been done. For better look those generated 3D histograms have been sliced to several 2D histograms - it allows visual comparison to experimental data (see Figure 3.4)

¹CC means that only CC types of interaction are included



(a) Particle in detector with momentum \mathbf{p} . \hat{z} is beam directions and α is angle between \mathbf{p} and p_L (b) Incoming neutrino (k_ν) interact with nucleon. One of the product is lepton (k_l)

Figure 3.3: Used parameters of histograms

Used bins are not equal. It is because with larger energy or momentum transfer there are less detected muons. To have better statistic, bins with a small number of detected events are joined. MINERvA divided p_L and p_T range into 12 and 13 pieces accordingly. The values of the first of them extending from 1.5 to 20 GeV and the second - from 0 to 2.5 GeV. For comparison, bin corresponding to the lowest p_L value has a width of 0.5 GeV and for the highest it is 5 GeV. Similarly is with p_T - in this case the values are 0.1 GeV and 1 GeV.

It is worth mentioning, that quick (or even sharp) disappearance of NuWro result as p_T is rising, is caused by cuts on angle described in Section 3.1. Also, as p_L is getting bigger, this boundary shifts in the direction of higher p_T value.

3.3 | Problem

After looking at obtained histograms one say, that NuWro has problems to describe data points. In a few places, NuWro results are above them, but in most cases the opposite situation happening - NuWro "underfit" data points. The complete answer to the question "Why is that?" is unknown/hard and it is not going to be shown. Also, it could be very complicated due to multiple dynamics used - we can not be sure which one of them are causing the discrepancy.

In my work, I am checking if *simple models* can be at least a good approximation of improvement NuWro needs or hint in further NuWro development. By *simple model* I mean, that only one dynamic is broken and the rest are flawless. Because QEL is believed to be

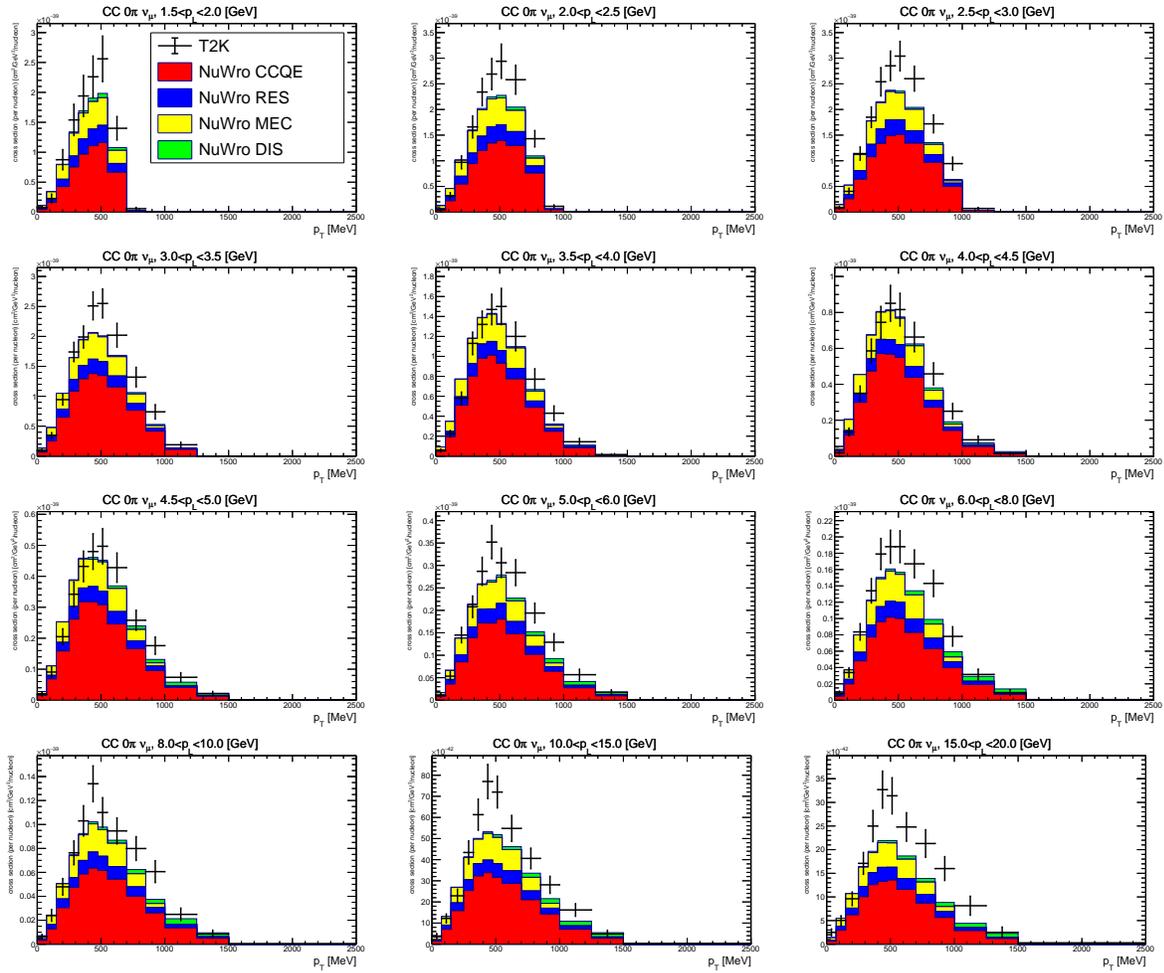


Figure 3.4: All of generated histograms. Each of them represent different range of p_T . On x-axis is p_T and on y-axis is cross-section. With black crosses experimental results of T2K is marked.

well-defined (it is the easiest to describe model) and RES with DIS are less numerous than MEC, the last one mentioned is chosen by me to be fixed.

3.4 | Investigation

First thing I have done, was a calculation of NuWro results disagreement with the data points. For this, I chose the following method. In each bin, I calculate a factor (it will be called δ) by which I have to multiply MEC events contribution to the cross-section to fit perfectly with data points. In other words

$$\overbrace{\sigma_{nuwro} - \sigma_{mec}}^{\text{non-mec}} + \overbrace{\sigma_{mec} \times \delta}^{\text{new mec}} = \sigma_{data} \quad (3.2)$$

After solving this in respect to δ :

$$\delta = \frac{\sigma_{data} - \sigma_{nuwro}}{\sigma_{mec}} \quad (3.3)$$

where

- σ_{data} is cross-section from experiment
- σ_{nuwro} is cross-section from NuWro (for all dynamics)
- σ_{mec} is cross-section from NuWro only for MEC dynamic

After this operation, I obtain the matrix of δ factors. Its graphical form is shown in Figure 3.5. The first thing, which catches the eye, is the right side of the histogram. White color in last two of them (in the first and second row even more) means that there is a lack of MEC events. In my analysis, I reject all these bins, because they are not topic of my interest. Yellow bins are interesting too. Their δ value are above 10 and in one case is even larger than 70. It is because MEC events are in a trace amount in those bins - they made up about 2% of the total cross-section. In further analysis, I will not take into account white neither yellow bins.

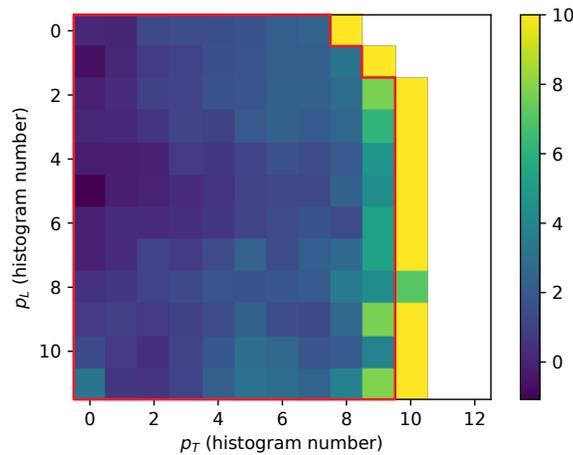


Figure 3.5: $\delta(p_T, p_L)$ matrix. Inside red polygon are bins which will be analyzed

Looking at this histogram in $\{p_T, p_L\}$ domain does not tell me about the NuWro problem because of two things. First, is that those variables are dependent from neutrino beam - different flux will give different results. Second, is that $\{p_T, p_L\}$ are not the basic variables. In NuWro case MEC events description relies on Nieves model, which for calculations uses ω and q . Because of that, I am going to change the domain of rescaling to these variables.

3.5 | Changing domain

To do that, I once again go through all generated MEC events and make two histograms (with respect to ω and q) with bins number equal 48×48 . First is just number of events - it means that every event has the same weight (see Figure 3.6a). In second weights are different - they are equal to $\delta(p_T, p_L)$ (see Figure 3.6b).

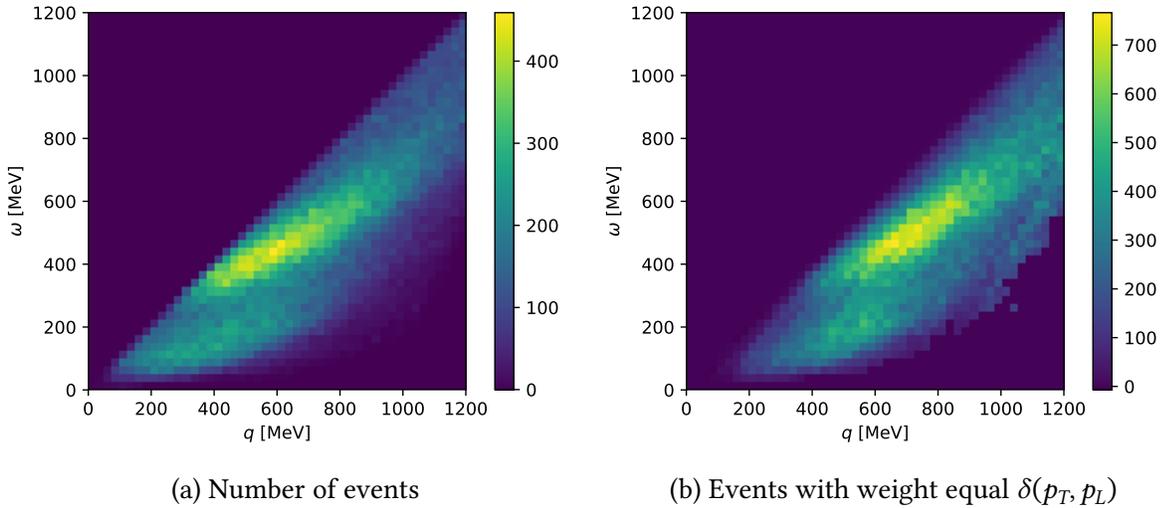


Figure 3.6: Histograms later used to calculate $\delta(\omega, q)$

Obtained histograms seem to be limited by a function $f(x) = x$. Indeed they are, because value of transfer of energy (ω) can not exceed value of momentum transfer (q). Also ranges of axes (less than 1200 MeV) are chosen with purpose - above this value no event has been recorded.

Once I have done it, I divide second histogram by first (it means I divide corresponding bins) to create new one (see Figure 3.7a). When a symbol $0/0$ is spotted, it is converted to 1. It means, that for values, where bins are empty, MEC events are not touched, which is

equivalent to rescaling by a factor of 1. This histogram is what I am looking for - it tells how to rescale MEC events (in $\{\omega, q\}$ domain) to fit data more precisely.

Some bins contain only a few events. It can cause large numerical fluctuation and, in result, wrong value of reweighting factor. Thus, I reject all bins, into which less than 10 events landed. By *reject* I mean to set their δ factor to 1.

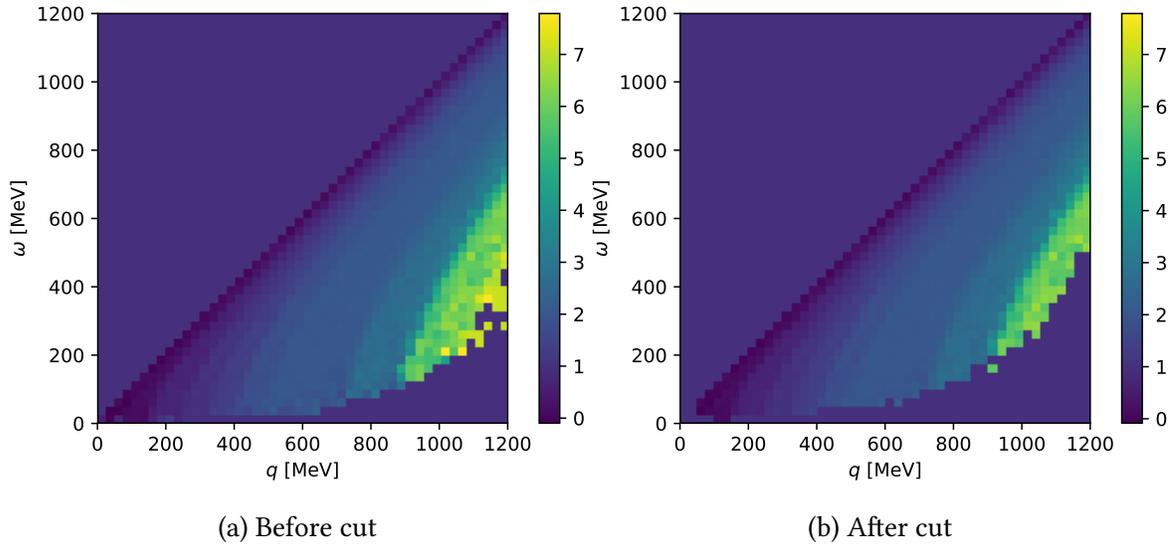
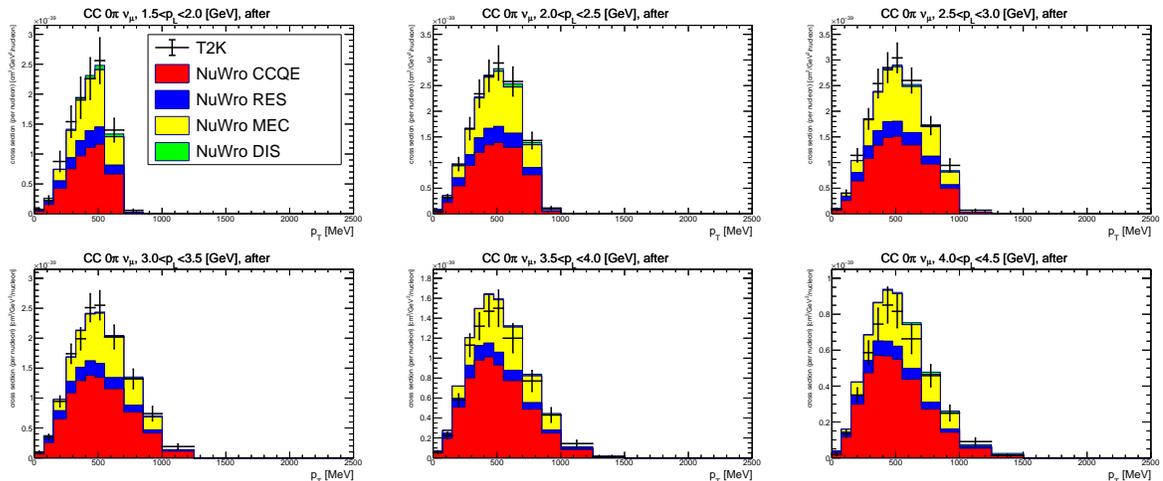


Figure 3.7: $\delta(\omega, q)$ with bins, where events number is less than 10 (a), and without them (b)

Having this histogram, I am going to check how does it change NuWro results. For this I once again histogramize events with respect to $\{p_T, p_L\}$, but this time every MEC event has weight equal $\delta(\omega, q)$. Outcome is shown in Figure 3.8.

Indeed new result differ with old one only at MEC events - rest dynamics are still. This



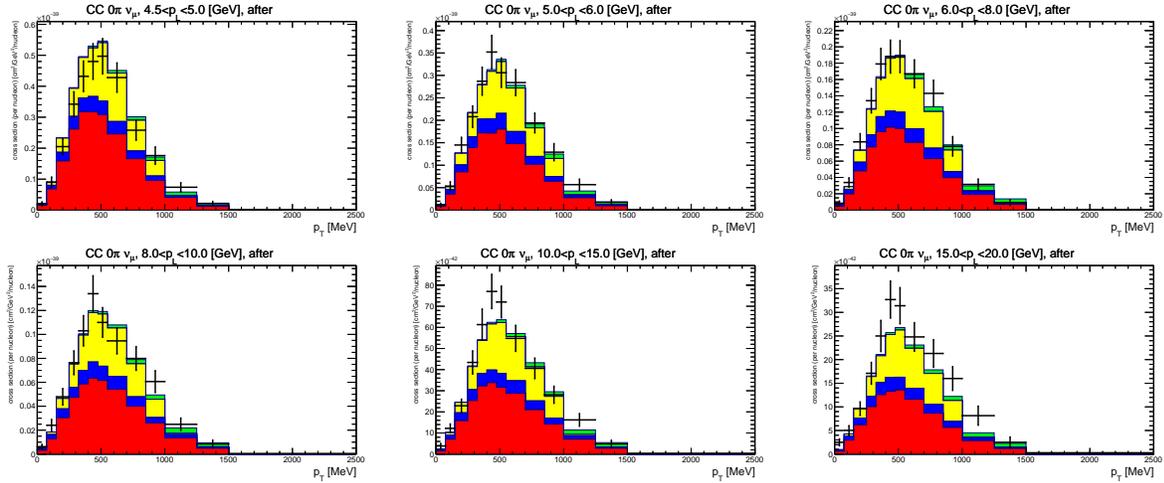


Figure 3.8: All of improved histograms. Each of them represent different range of p_L . On x-axis is p_L and on y-axis is cross-section. With black crosses experimental results of T2K is marked.

is sign of good code working. At the first glance, it seems to agree better to the data.

3.6 | *GENIE* study

Now I have discrete function of reweighting, but it is always better if continuous one will be found. With this can help me GENIE developers. They have done similar work with their generator and eventually came up with following form of fitting function :

$$\delta(q, \omega) = N \times \exp \left(- \frac{\left(\frac{(\omega - \mu_\omega)^2}{\sigma_\omega^2} + \frac{(q - \mu_q)^2}{\sigma_q^2} - 2c \frac{(\omega - \mu_\omega)(q - \mu_q)}{\sigma_\omega \sigma_q} \right)}{2(1 - c^2)} \right) + 1 \quad (3.4)$$

which is 2D Gaussian function with correlation. I also tried to fit this function to my

	N	μ_ω	μ_q	σ_ω	σ_q	c
NuWro	17.27	1976	1018	559	406	0.898
GENIE	10.86	254	508	57	129	0.875

Table 3.2: Parameters of GENIE and NuWro fitting function

results. For that, I used tool provided by *ROOT* framework - the `Fit()` method. I used it with "W" option, which means that every non-empty bin is counted with weight equal 1. This method finds optimal parameters of fitted function by minimizing its χ^2 (for more detail see Chapter 4). Both fitting functions are shown on picture Figure 3.9 and their parameters are shown in Table 3.2.

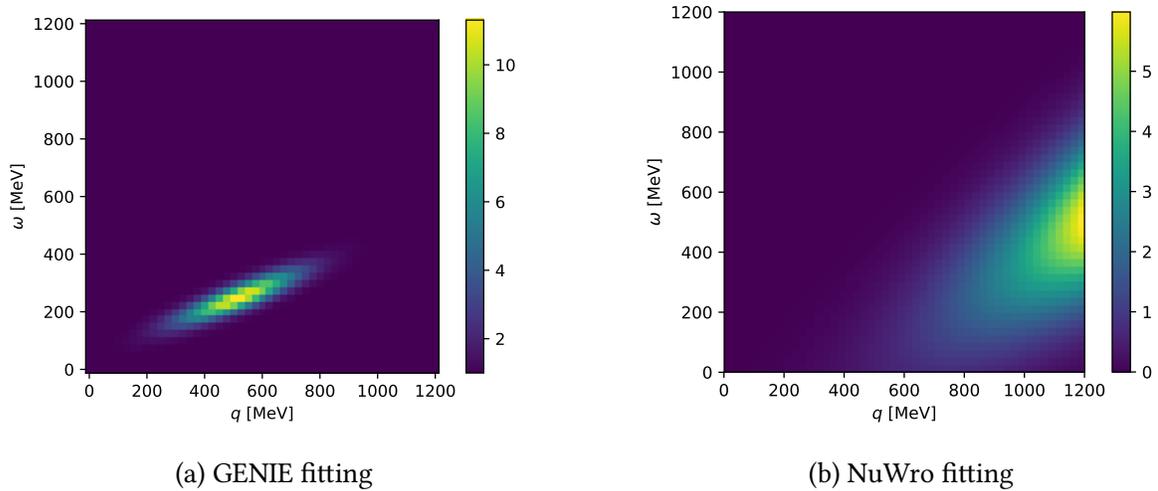


Figure 3.9: Graphical form of fitting functions

Fitting function obtained by me is definitely quite different. It is not very surprising, as GENIE describes $CC0\pi$ in a different way, and also method how scaling function was obtained is very different.

4 | Results

4.1 | Introduction to χ^2 test

χ^2 test is commonly used to quantify data/Monte Carlo comparison. It provides information about goodness of fitting theory to experimental data. Its simplicity and independence from used model are causes of its widely application. That is also why groups of theorist often provide information about χ^2 parameter. It helps intermodular comparison.

This parameter is defined as follow:

$$\chi^2 = \sum_{i=1}^N \sum_{j=1}^N (x_j^t - x_j^e) \mathcal{M}_{ij}^{-1} (x_i^t - x_i^e) \quad (4.1)$$

or in shorter (matrix) form:

$$\chi^2 = \Delta^T \mathcal{M}^{-1} \Delta \quad (4.2)$$

where i and j goes through all bins of historgam and:

- x^t - result predicted by theory
- x^e - result obtained by experiment
- \mathcal{M} - covariance matrix (in my case provided by the experimental group)
- Δ - vector of differences between theory and experiment such that $\Delta_i = x_i^t - x_i^e$

Covariance matrix is matrix in which element in the position i, j is covariance between i -th and j -th element of vector Δ . In the simplest model, where every pair of bins is independent, \mathcal{M} -matrix is diagonal and $\mathcal{M}_{ii} = cov(\Delta_i, \Delta_i) = \sigma_{\Delta_i}^2$ where σ^2 stands for variance. Then χ^2 will be calculated as follow:

$$\chi^2 = \sum_{i=1}^N \left(\frac{x_i^t - x_i^e}{\sigma_i} \right)^2 \quad (4.3)$$

Probability of experimental data set agreement with theoretical model can be obtained by looking at χ^2 distribution. It depends on one parameter - system's degrees of freedom (*d.o.f.*). In my case, this will be number of reweighted bins.

For in-depth understanding of χ^2 with covariance matrix, let me consider simple example - 2 bins and 2 possible reweighting (see Figure 4.1). Left one better fits data (both bins land in uncertainty region) but also change monotonicity (value of bins are rising beside fact that data points are evidently falling). In right one is opposite - worse data fit but monotonicity unchanged.

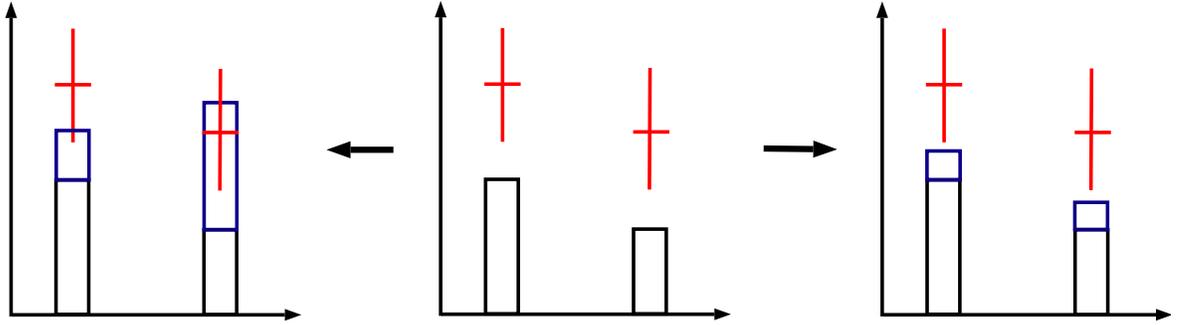


Figure 4.1: Example - starting bins (middle) can be improved by one of two rescaling: left or right. Red is data points (with uncertainty) and blue is what was added to the bin

Now there is a question: which of those rescaling is better (which minimize χ^2)? In case where there is no covariance between bins, better is left one. But if covariance is non-zero, then we do not know - it depends on the matrix values. For example suppose that

$$\mathcal{M}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \Delta_L = \begin{pmatrix} -1/2 \\ 1/3 \end{pmatrix} \quad \Delta_R = \begin{pmatrix} -3/4 \\ -2/3 \end{pmatrix} \quad (4.4)$$

where Δ_L, Δ_R are example vectors of differences between bins value and data points, for left and right histogram accordingly. For this case $\chi_L^2 = 19/18$ and $\chi_R^2 = 73/72$. It means, that right rescaling is better, although it seems to worse fit data points.

Summarizing, χ^2 tells not only about data fit, but also about trend/monotonicity fit.

4.2 | χ^2 calculation

In my case, 118 bins are getting improved ¹, therefore my covariance matrix has dimension equal 118 (in fact it is $12 \times 13 = 156$, but, as I earlier mentioned, I ignore some bins). Unfortunately, it has in some rows and columns only 0's, which causes its singularity. It is huge problem, because what is interesting me, is inverse of this matrix. I am dealing with this problem using following procedure of pseudo-inverting:

change every 0 on diagonal to 1
 \downarrow
 inverse new matrix
 \downarrow
 subtract 1 from the diagonal where it was added

Using this new matrix, I calculate χ^2 for vanilla NuWro and for my improvement.

In Table 4.1 I gather my calculation. I put there result of GENIE rescaling as well as χ^2 without covariance matrix.

$\chi^2/d.o.f.$	NuWro vanilla	NuWro discrete	NuWro continuous²	GENIE
without covariance	344/118	90 (90)/118	227/118	360/118
with covariance	494/118	173 (170)/118	378/118	550/118

Table 4.1: Results of χ^2 calculation

The values in brackets corresponds with rescaling, where every bins are taken into account - even those with less than 10 events in them (see Figure 3.7a)

Surprising is result for NuWro discrete without covariance. The method, which is used to calculate it, is supposed to give much lower result. I suspect, that it could be caused by used binning. For that reason, I have made several more calculations of χ^2 , but this time

¹I do not take into account bins without MEC events and those which δ factor is above 10

²See Section 3.6

for different numbers of bins. Because, as bins' density grows, their sizes get smaller and less events land in them. For that reason, I did not reject bins with less than 10 events in it. Results of this process are shown in Figure 4.2.

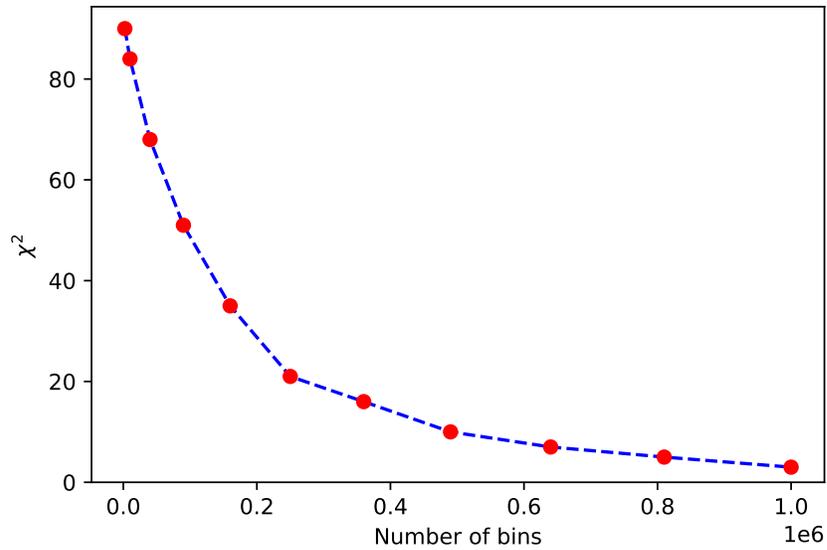


Figure 4.2: χ^2 for NuWro discrete without covariance with different numbers of bins

The results clearly shows, that increasing number of bins decrease χ^2 . Also, in the limit, it should drop to 0.

5 | Summary

As it is shown in Table 4.1 rescaling method presented in this paper significantly reduce χ^2 with covariance, which was expected. For discrete reweighting, it helps to lower this parameter by a factor of almost three, and continuous one reduce it for about 30%. Also, without the covariance matrix improvement is noticeable. It could be seen already, by looking at Figure 3.8, since χ^2 without covariance matrix can be approximate by human eye.

Also, it was shown, that results of reweighting highly depends on a number of bins used - as it grows, χ^2 falls down. But, if continuous function has to be found, a balance of total events number and bins number has to be found as well. It means, that obtained histogram of $\delta(\omega, q)$ can not be too *sharp* - differences between adjacent bins should not be too big. Also, if histogram would not be *smooth* enough, it would seem too *random* - a pattern or a hint, where NuWro has a problem, could not be even visible.

Interesting fact, is that GENIE function make NuWro results worse. It could be seen, that their rescaling does not cover the same region as my rescaling does. The function is also narrower - it raises only a tiny fragment of histogram and even not the important one. Those things adds up and causes the raises of χ^2 . It could be cause be two things. First, is nucleus model used. In my case (spectral function), GENIE function does not work. On the other hand it get along with local Fermi gas model - in this case slightly improvement of χ^2 has been recorded. And second, is the way of neutrino energy reconstruction. It differ in NuWro and in GENIE.

In the future, I would like to still investigate this topic. For now, main goal is to find analytic function of rescaling, which cover up all experiments, and not only MINERvA. It could be done also for anti-neutrinos. If that function is found, it will be good option to implement this improvement to NuWro permanently.

Bibliography

- [1] *NuWro user manual*. <https://nuwro.github.io/user-guide>.
- [2] A. M. ANKOWSKI AND J. T. SOBczyk, *Argon spectral function and neutrino interactions*, Phys.Rev. C 74, 054316, (2006).
- [3] D. PERKINS, *Wstęp do fizyki wysokich energii*, Wydawnictwo Naukowe PWN, Warszawa, 2004.
- [4] J. A. FORMAGGIO AND G. P. ZELLER, *From eV to EeV: Neutrino Cross-Sections Across Energy Scales*, Rev. Mod. Phys. 84, 1307, (2013).
- [5] J. NIEVES, I. RUIZ SIMO, AND M. J. VICENTE VACAS, *Inclusive charged-current neutrino-nucleus reactions*, Phys. Rev. C 83, 045501, (2011).
- [6] K. SARASWAT, P. SHUKLA, V. KUMAR, AND V. SINGH, *Coherent Pion Production in Neutrino Nucleus Scattering*, Phys. Rev. C 93, 035504, (2016).
- [7] T. GOLAN, *Modeling nuclear effects in NuWro Monte Carlo neutrino event generator*, PhD thesis, University of Wrocław, 2014. http://wng.ift.uni.wroc.pl/files/phd_tomasz_golan.pdf.
- [8] T. KATORI, *Meson Exchange Current (MEC) Models in Neutrino Interaction Generators*, in 8th International Workshop on Neutrino-Nucleus Interactions in the Few-GeV Region (NuInt12), 2012.

Michał Siemaszko
PESEL: 96051910294
✉ ul. Jedności Narodowej 88/6, 50-261 Wrocław
✉ michal.siemaszko@protonmail.com

Wydział Fizyki i Astronomii
Kierunek studiów: fizyka
studia stacjonarne, licenjackie
nr albumu: 281162

OŚWIADCZENIE O PRAWACH AUTORSKICH I DANYCH OSOBOWYCH

Ja niżej podpisany Michał Siemaszko, student Wydziału Fizyki i Astronomii, kierunek fizyka, oświadczam, że przedkładana praca dyplomowa na temat: *Investigation of MINERvA experiment data using NuWro Monte Carlo generator with spectral function model*

- jest mojego autorstwa i nie narusza autorskich praw w rozumieniu ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (tekst jednolity: Dz. U. z 2006 r. Nr 90, poz. 631, z późn. zm.) oraz dóbr osobistych chronionych prawem;
- nie zawiera danych i informacji uzyskanych w sposób niedozwolony;
- nie była wcześniej przedmiotem innej urzędowej procedury związanej z nadaniem dyplomu uczelni wyższej lub tytułu zawodowego;
- treść pracy dyplomowej przedstawionej do obrony, zawarta na przekazanym nośniku elektronicznym, jest identyczna z jej wersją drukowaną.

Oświadczam, iż zostałem poinformowany o prawie dostępu do treści moich danych osobowych oraz ich poprawiania. Udostępnienie przez mnie danych osobowych ma charakter dobrowolny. Wyrażam zgodę, na:

- udostępnienie mojej pracy dla celów naukowych i dydaktycznych;
- przetwarzanie moich danych osobowych w myśl ustawy z dnia 29 sierpnia 1997 r. o ochronie danych osobowych (tekst jednolity: Dz. U. z 2014 r., poz. 1182, z późn. zm.);
- umieszczenie mojej pracy w bazie danych Uczelni i jej przechowywanie przez okres stosowny do potrzeb Uczelni;
- wykorzystanie mojej pracy jako elementu komparatywnej bazy danych Uczelni;
- udostępnienie mojej pracy innym podmiotom celem prowadzenia kontroli antyplagiatowej prac dyplomowych i innych tekstów, które zostaną opracowane w przyszłości;
- porównywanie tekstu mojej pracy z tekstami innych prac znajdujących się w bazie porównawczej systemu antyplagiatowego i zasobach Internetu.