Flux measurements with the Near Detector
Draft 4

M. Kordosky and M. Dorman
University College London

December 20, 2005

Executive Summary

Topic

We make a measurement of the ν_μ flux in the LE-10 beam configuration with the Near detector using a technique that is claimed to minimise systematics associated with cross-sections .

Relevance to CC Measurement

Somewhat ancillary, particularly for 1e20 POT analysis, since analysis strategies will use the ND data to correct the prediction of the FD spectrum. Our result is most useful for constraining simulation of NuMI , discriminating between beam and reconstruction related data/MC discrepancies, and as input to cross-section measurements.

Method

The measurement technique employs a quasi-elastic enriched $\nu_\mu-$CC sample to derive the flux shape in the range $0 < E_\nu < 10\,\text{GeV}$. A separate inclusive ν_μ–CC sample is used, along with knowledge of the total cross-section , to fix the flux normalisation in the range $10 < E_\nu < 20\,\text{GeV}$. There is an emphasis on selecting quasi-elastic events such that acceptance is independent of the energy, which reduces the effect of bin-to-bin migration and may ultimately make acceptance correction unnecessary.

Conclusion

We have made a preliminary measurement of the flux in the range $10 < E_\nu < 20\,\text{GeV}$:

$$\Phi = 1.77 \ \pm 0.06\,(\text{stat}) \ {}^{+0.08}_{-0.09}\,(\text{sys}) \ \ \text{m}^{-2}10^{-5}\ \text{POT}^{-1} \quad 10 < E_\nu < 20\,\text{GeV}$$

Though this number includes a systematic error we consider it especially preliminary as some sources of uncertainty must still be studied.

We have also extracted a spectral prediction of the flux shown in Fig. 32. This result should be considered preliminary. At this time the quasi-elastic selection suffers from an uncomfortable background of NC and DIS events, some of which is known to be reconstruction related. Reducing this source of background is a priority for our future work. Efforts are underway to improve the quasi-elastic selection. In particular we are attempting to use the over constrained quasi-elastic kinematics to calculate the expected momentum 4–vector of the recoil proton, thereby predicting the location of hits in the event. Finally, both the shape and normalisation analyses suffer from the small size of the MC sample (about a factor of 10 less than our dataset). To alleviate the situation, we have begun to generate and reconstruct a $\sim 3e18$ POT sample of single events in the LE-10 configuration using the Rutherford Tier1A farm. Most of the systematic errors associated with the quasi-elastic selection have yet to be evaluated though we do comment on the shape invariance of the cross-section and some work has been done to understand the effect of variations in the strength of intra-nuclear re-scattering.

Abstract

We report the preliminary results of measurement of the ν_μ flux of the NuMI beam using data collected by the MINOS Near Detector during an exposure of 3.4e19 POT in the LE-10 beam configuration. The measurement technique employs a quasi-elastic enriched $\nu_\mu-$CC sample to derive the flux shape in the range $0 < E_\nu < 10$ GeV. A separate inclusive $\nu_\mu-$CC sample is used, along with knowledge of the total cross-section , to fix the flux normalisation in the range $10 < E_\nu < 20$ GeV. The resulting flux may be used to constrain the predictions of the `GNuMI` Monte Carlo, understand the systematics of inclusive $\nu_\mu-$CC event reconstruction below 10 GeV/c, and as a basis for cross-section measurements below 10 GeV.

1 Introduction

In its current version the `GNuMI` Monte Carlo predicts the neutrino energy flux in the Near and Far detectors based on the best knowledge of hadron production in the NuMI target, beam optics, downstream interactions and geometrical acceptance in the two detectors. The predictions are not uncertain and as such the ν_μ disappearance analysis will employ a measurement of the (inclusive) $\nu_\mu-$CC energy spectrum in the Near detector to, in effect, correct the `GNuMI` prediction[1] at the Far detector. The currently planned analyses make that correction either by fitting some elements of the beam and cross-section models to the Near detector data, using the best fit to predict the Far detector spectrum, or by extracting a Near-to-Far transfer function directly from the beam model. Both

[1] These methods also correct or mitigate a number of other uncertainties, such as those in knowledge of neutrino cross-sections .

techniques are expected to reduce the effect that uncertainty in the absolute flux has on the $\nu_\mu-$CC disappearance measurement to acceptable levels, at least for the initial exposure of 1e20 POT.

As such, a direct measurement of the ν_μ flux lies somewhat outside of the analysis path for the initial oscillation study. That said, the flux measurement technique presented here employs cross-sections with relatively small uncertainties and utilises a sample of $\nu_\mu-$CC events with systematically different, and arguably better understood, kinematic and reconstructed properties. As an example of the latter, the quasi-elastic enriched sample used to derive the $0 < E_\nu < 10\,\text{GeV}$ energy spectrum is insensitive to the hadronic energy scale, does not suffer from uncertainties in reconstructing short tracks in large showers, and has a better energy resolution than the inclusive $\nu_\mu-\text{CC}$ sample. The resulting flux may therefore be used to identify discrepancies between the ND data and MC as being likely due to inadequacies in the beam-line simulation or, conversely, to inadequacies in the reconstruction or cross-section model. In principle the reconstructed flux could also be used to directly constrain beam model parameters and will be a necessary input for low energy neutrino cross-section measurements.

This note describes the flux measurement technique and presents initial results along with a projection of the analysis effort into the next few weeks and months.

2 Overview of the technique

The flux measurement technique was suggested to the authors by H. Gallagher and J. Thomas in early 2005. Similar methods have been used before, a recent example being analyses done on data collected from the IHEP-JINR detector [1]. Our method is related to the $y \to 0$ strategy employed at higher energies and described in Ref. [2].

Our technique leverages two important properties of $\nu_\mu-\text{CC}$ cross-sections . The first is that the absolute cross-section has been measured very precisely at high energies using iron targets [5, 6, 7]. The second is that the shape of the quasi-elastic cross-section $\frac{1}{\sigma}\frac{d\sigma}{dE_\nu}$ is relatively well understood over a broad range of energies extending down to 1 GeV and below. These properties motivated us to divide the neutrino energy spectrum measured in the Near detector into two regions (see Fig. 1):

Normalisation $10 < E_\nu < 20\,\text{GeV}$ in which an inclusive $\nu_\mu-\text{CC}$ sample along with knowledge of the total cross-section will be used to set the flux normalisation.

Shape $E_\nu < 10\,\text{GeV}$ in which a quasi-elastic enriched sample along with knowledge of the cross-section shape will be used to extrapolate the flux.

In what follows cross-sections from `NEUGEN` [3] are utilised. We have examined the inclusive $\nu_\mu-\text{CC}$ cross-section in the $10 < E_\nu < 20\,\text{GeV}$ region and

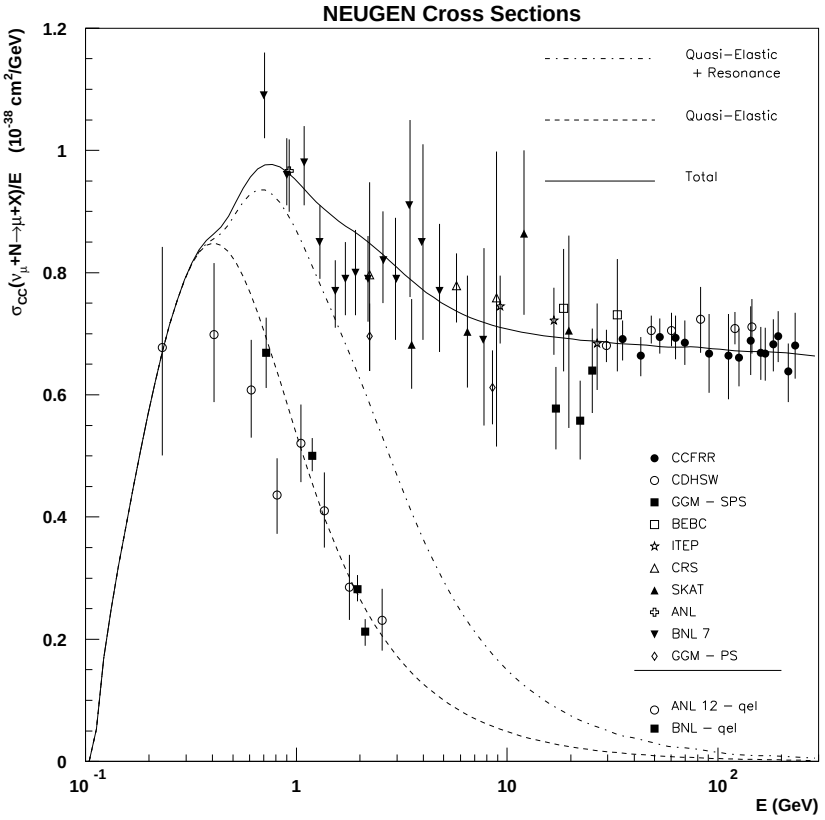


Figure 1: $\nu_\mu - \mathrm{CC}$ cross-section data along with the `NEUGEN` predictions. In the present analysis the energy range is divided into a region $E_\nu < 10\,\mathrm{GeV}$ in which a quasi-elastic enriched sample will be used to predict the flux shape and a second region $10 < E_\nu < 20\,\mathrm{GeV}$ in which an inclusive sample will be used to fix the normalisation. Figure from Ref. [3]

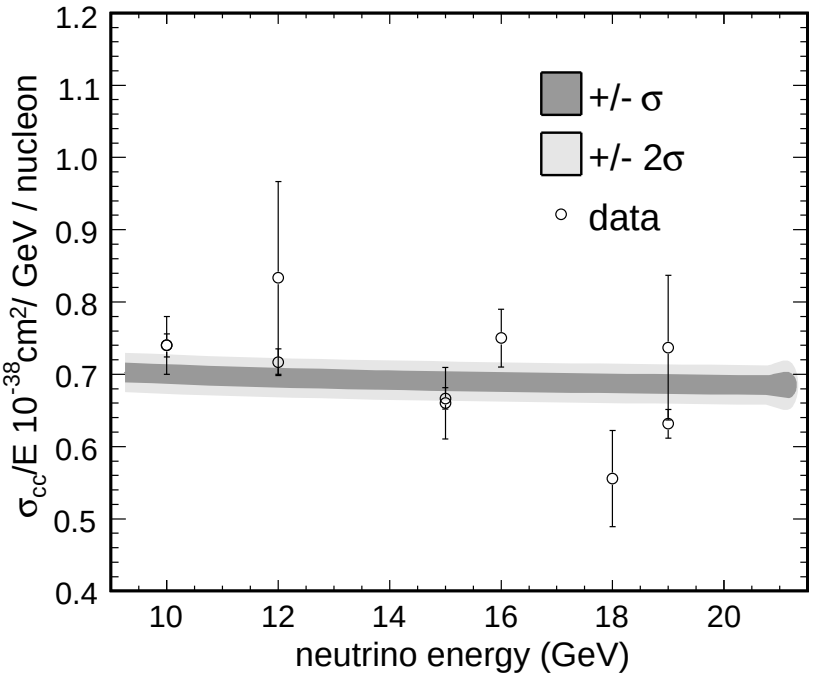


Figure 2: ν_μ–CC inclusive cross-section data compared to the `NEUGEN` prediction. The shaded region indicates the cross-section uncertainty accommodated by the data. Data are from Ref [4].

have found that the NEUGEN prediction is in good agreement with the experimental data, the latter accommodating a $\pm 2\%$ ($1\ \sigma$) variation about the model prediction (see Fig. 2). We have not attempted to quantify the uncertainty in the quasi-elastic cross-section shape but instead refer to Ref. [8, 9].

3 Data and Monte Carlo samples

We utilise data collected in the LE-10 beam configuration between early June and late August 2005. For the normalisation measurement the runs utilised span 7860–8200, 8300–8432 [10], and, after beam quality cuts, correspond to an exposure of 2.34e19 POT as measured by TORTGT [11]. The shape measurement uses the above dataset but also includes suitable runs in the 8200–8300 range, yielding an exposure of 3.4e19 POT. Both analyses utilise a MC dataset consisting of 2.36e18 POT. The dataset was generated and reconstructed by the batch processing group. The shape analysis also uses about 5e4 single events, generated with a $1/E_\nu, 0 < E_\nu < 20$ GeV flux, to construct PDFs.

The data and MC were calibrated and reconstructed with offline release R1.18 during the official batch production process. The resulting long ntuple files (sntp) were used to construct [12] mini-dsts which were used in the event selection and final analysis.

4 Normalisation Measurement

4.1 $\nu_\mu-$CC Event Selection

Events were selected as $\nu_\mu-$CC according to the following criteria:

fiducial volume "Pitt" fiducial volume (molded to ND shape): $0.6 < z < 3.56$ m, $0.3 < u < 1.8$ m, $-1.8 < v < -0.3$ m, $x < 2.4$ m, $r > 0.8$ m (see Fig. 3). Here u, v, x, z refer to the event vertex position and r is the distance from the vertex to the center of the coil hole. This selection was refined to require $z > 1$ m and $r_B < 1$ m during the CC-like selection, where r_B is the radius around the beam axis. In future analyses the refinement will either be dropped or included at this stage.

track quality single track passing the tracker's internal checks, $\chi^2/\text{n.d.f.} < 20$, and $|u - v| < 6$ planes where u, v refer to the track vertex.

fit quality $|\sigma_{q/p}/(q/p)| < 0.3$ for exiting tracks. A track is classified as exiting if (a) the track end point $z < 7$ m and is also outside the fiducial region defined above, or (b) the track endpoint $z > 7$ m and is also outside an ellipse centered at $(x, y) = (0.8, 0.0)$ m with associated major and minor axes $(a, b) = (1.7, 1.4)$ m and $r > 1$ m (see Fig. 4) or (c) $z > 15.6$ m.

CC-like The purity of the sample is enhanced by selecting $p_{cc} > -0.4$ where p_{cc} is calculated for each event $\boldsymbol{x}$ as the product of the output of three

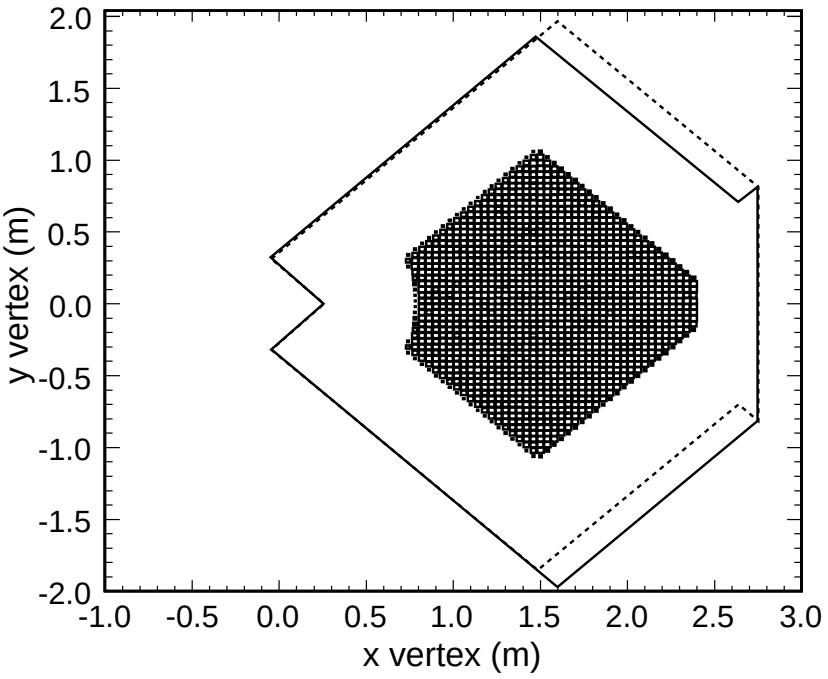


Figure 3: Reconstructed (x,y) vertices of events accepted by the fiducial volume cut. Lines show outlines of the ND partial planes, the coil is centered at (0,0), and the view is looking into the beam.

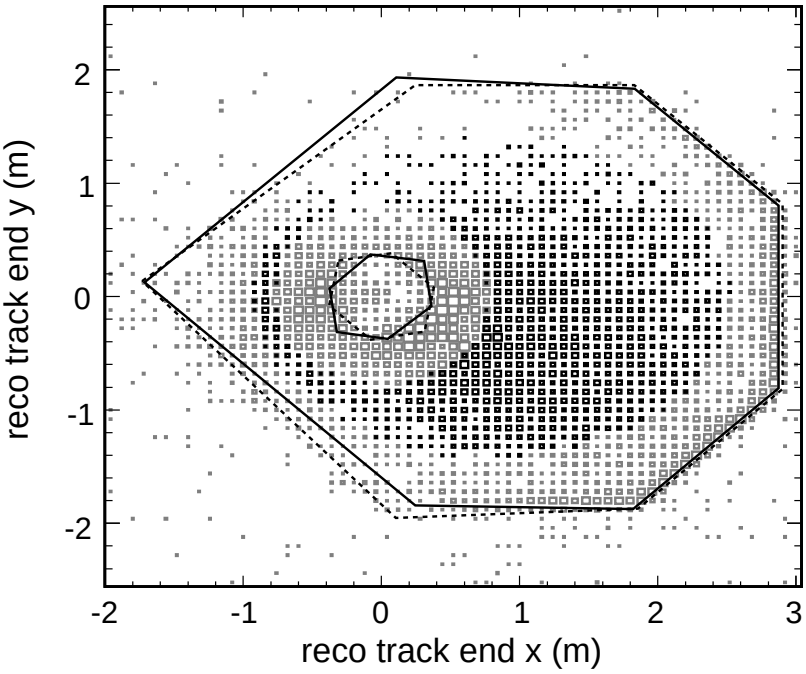


Figure 4: Reconstructed (x,y) track endpoints. The (x,y) region in which a track's momentum may be reconstructed by range is shown with dark boxes. The region in which a track's momentum is reconstructed by curvature is shown with light boxes. This figure was made without a cut on the z endpoint of the track. In practice many of the tracks terminating in the first, darker, region are reconstructed by curvature because they exit through the rear of the spectrometer.

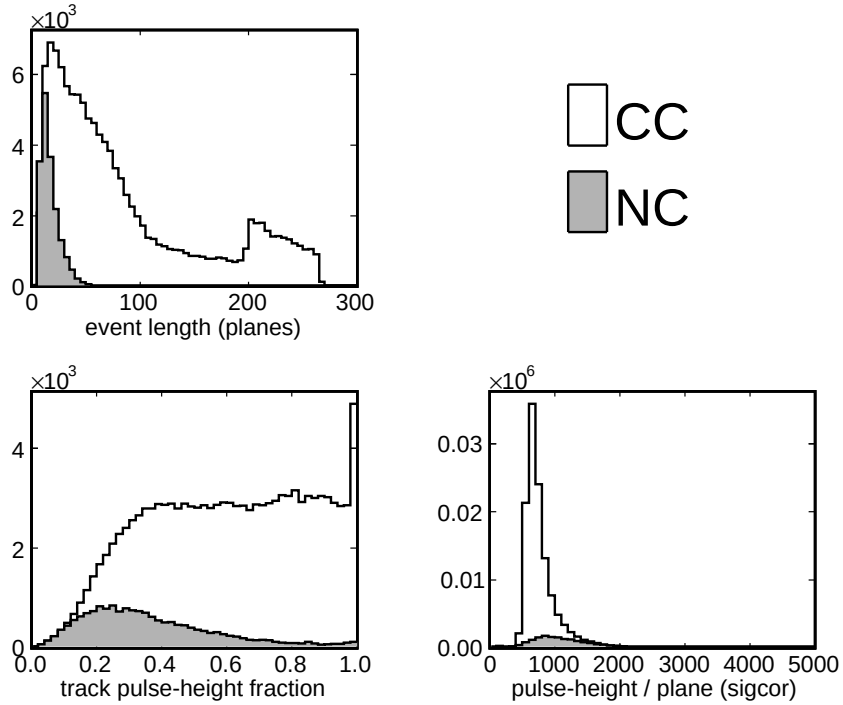
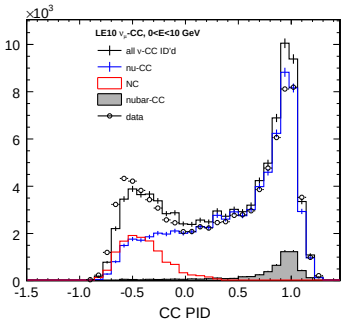
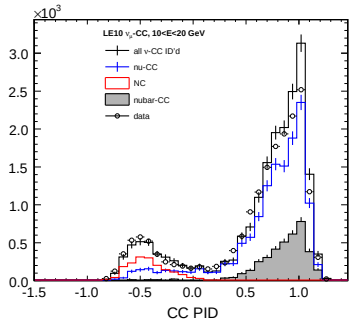


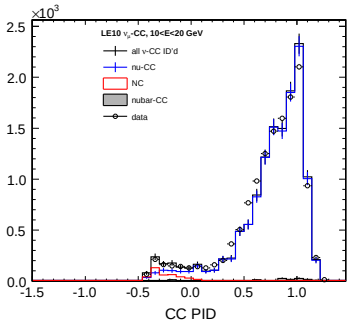
Figure 5: PDFs used to calculate p_{cc} [13].



(a) $E_\nu < 10\,\text{GeV}$: fiducial + track quality + fit quality



(b) $10 < E_\nu < 20\,\text{GeV}$: fiducial + track quality + fit quality



(c) $10 < E_\nu < 20\,\text{GeV}$: fiducial + track quality + fit quality + PID + charge sign

Figure 6: p_{cc} distributions for a few energy ranges and selections. The figures are normalised to POT.

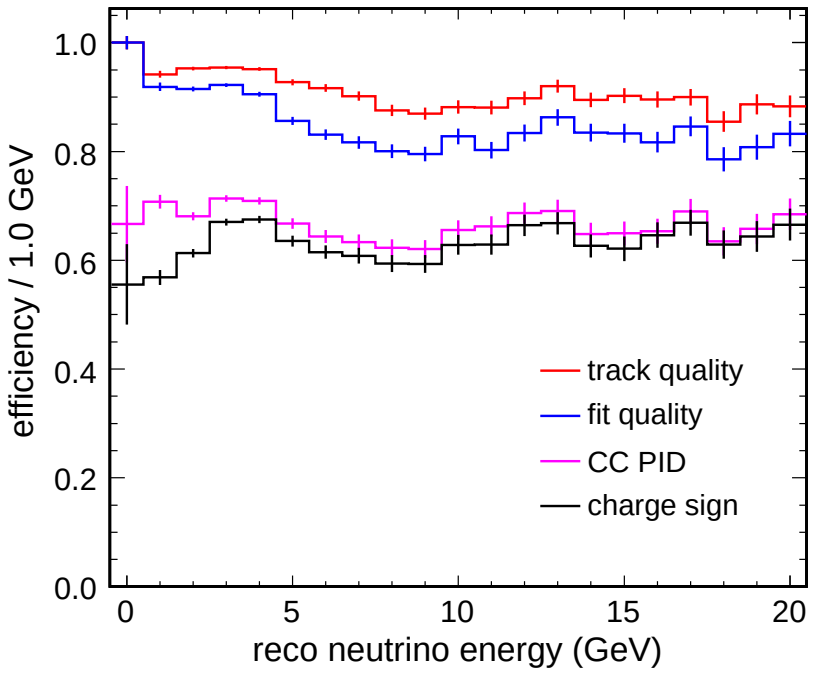


Figure 7: The cumulative efficiency of the cc event selection. The denominator in each bin is the number of $\nu_\mu -$CC events reconstructed inside the fiducial volume.

	Cumulative Efficiency (%)					
	0–20 GeV			10–20 GeV		
Cut	data	MC	$\nu_\mu-$CC	data	MC	$\nu_\mu-$CC
track quality	90.7(0.1)	90.7(0.2)	90.9(0.3)	89.7(0.1)	88.8(0.4)	89.1(0.5)
fit quality	85.2(0.1)	84.6(0.3)	85.0(0.3)	84.0(0.2)	84.6(0.5)	82.6(0.6)
CC-like	60.4(0.1)	62.0(0.4)	66.6(0.5)	62.4(0.2)	62.1(0.6)	66.5(0.7)
charge sign	49.6(0.1)	48.4(0.4)	62.9(0.5)	51.4(0.2)	46.7(0.6)	64.5(0.7)

Table 1: The relative efficiency of the CC selection. The denominator in each case is the number of events reconstructed in the fiducial volume. "MC" indicates the full Monte Carlo dataset, while "$\nu_\mu-$CC " indicates true $\nu_\mu-$CC events. Statistical uncertainties, which are correlated between rows in the same column, are shown in (. . .).

	Composition (%)	
Cut	0–20 GeV	10–20 GeV
QE	18.7(0.3)	06.9(0.5)
RES	27.7(0.4)	13.2(0.6)
DIS	48.4(0.5)	75.7(1.5)
NC	04.6(0.2)	03.0(0.3)
$\bar{\nu}$	0.42(0.05)	01.0(0.2)

Table 2: The expected composition of the selected CC sample. Statistical uncertainties are shown in (. . .).

one dimensional PDFs: $\prod_i \mathcal{L}_i(\boldsymbol{x})$. The $\mathcal{L}_i$ are shown in Fig. 5 [13]. Distributions of p_{cc} are displayed in Fig. 6.

charge sign $(q/p) < 0$ to select ν_μ rather than $\bar{\nu}_\mu$.

The neutrino energy is reconstructed as the sum of the track and shower energies, where the former is determined from the curvature(range) for exiting(contained) tracks and the muon mass is assumed. The shower energy is the "linear CC energy" [14] of the first shower in the event. Under all but exceptional cases this is the vertex shower identified by the reconstruction.

Figure 7 shows the effect of the selection criteria as a function of reconstructed neutrino energy for MC events. The integrated effect of the selection procedure is tabulated in Tab. 1 and the expected composition of the sample is presented in Tab. 2 and shown in Fig. 8.

4.2 Characteristics of the Sample

Table 1 shows fairly good agreement between the data and the MC, with a pair of notable exceptions. The first of these is the 2.6% difference between

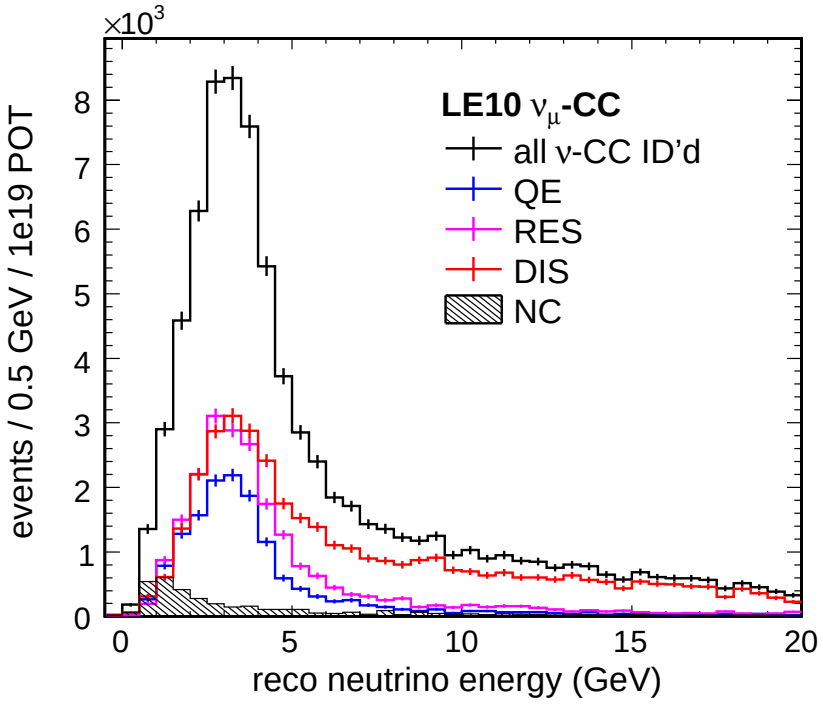


Figure 8: Composition of the of selected $\nu_\mu-$CC sample. See also Tab. 2.

the fraction of events passing the CC-like stage in the 0–20 GeV range which is notably absent in the 10–20 GeV range. There are known instrumental effects in the data that may cause this discrepancy for low reconstructed energies and steps have been taken to reduce their effect in the next round of processing. The second notable difference appears in the charge sign selection for 10–20 GeV. The table indicates that there is an overabundance of positively charged muons reconstructed in the 10–20 GeV MC sample. There are known limitiations in the simulation and reconstruction which may affect this result. These are, namely, the accuracy of the magnetic field maps and of the simulation of tracks passing near or through the coil hole. The next round of processing is expected to address these issues. That said, the relative excess of positive tracks is not substantially reduced by a tighter fit quality cut and the positively charge events appear strongly defocused in the horizontal and are also manifested as an excess at low reconstructed y.

Fig. 9 shows the reconstructed energy spectrum of the selected $\nu_\mu - \mathrm{CC}$ for both data and MC. The energy spectrum displays a (now well worn) disagreement between the data and MC in the region 2–9 GeV. The origin of this disagreement is not known but the region is subject to relatively large errors owing to the neutrino beam optics, and is also the region in which there is some uncertainty from matching the resonant and DIS pieces of the cross-section . The spectral agreement may also be subject to the inadequacies described above in the context of muon charge sign. The lower portion of Fig. 9 shows the range used for the normalisation measurement. The overall agreement between data and MC is remarkably good.

Reconstructed vertex distributions for the normalisation sample are shown in Fig. 10. The data/MC ratio for all three distributions is consistent with a line having zero slope over the entire x, y, z regions shown, though the size of the MC dataset is too small to make this a very rigorous test. It is clear that there is some systematic variation in the z distribution, perhaps caused by an inadequate strip-to-strip calibration but also, at the endpoint, may be subject to an incorrect z offset coupled with the planar structure of the detector.

Reconstructed kinematic distributions are shown in Fig. 11-13. The data/MC agreement in these distributions varies from good (the inelasticity – y) to poor (Bjorken–x). The y distribution is expected *a priori* to have the smallest model uncertainty since it is kinematic in origin. In contrast, the Bjorken–x distribution is known to vary significantly between the different cross-section models available in `NEUGEN` .

4.3 Calculation of the Normalisation

The normalisation is calculated as the ratio data/MC $= \sum_i d_i / \sum_i m_i$ where d_i, m_i are the (POT normalised) number of events reconstructed in energy bin i for data and MC. The calculation is restricted to reconstructed energies $10 < E_\nu < 20$ GeV and the ratio will be interpreted as a correction factor to be applied to the generated flux. This simple method requires no unfolding procedure and utilises the MC coupled to the input flux to account for bin-to-bin migration

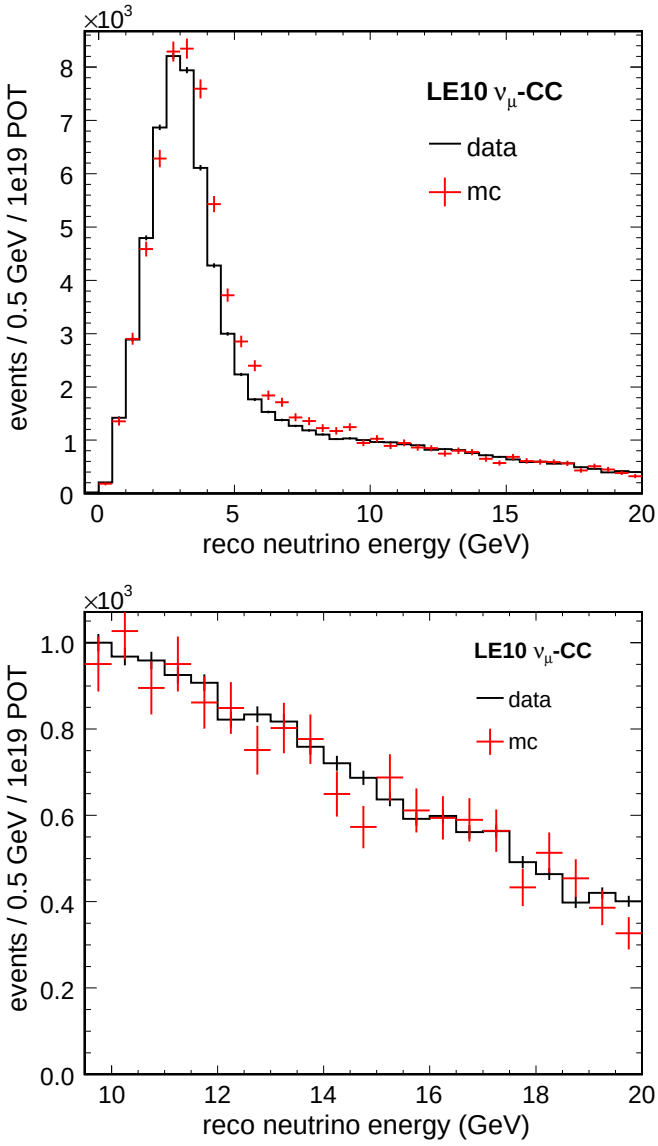


Figure 9: Reconstructed energy distribution of events selected as $\nu_\mu-$CC . The lower figure is zoomed to show the region used in the normalisation calculation.

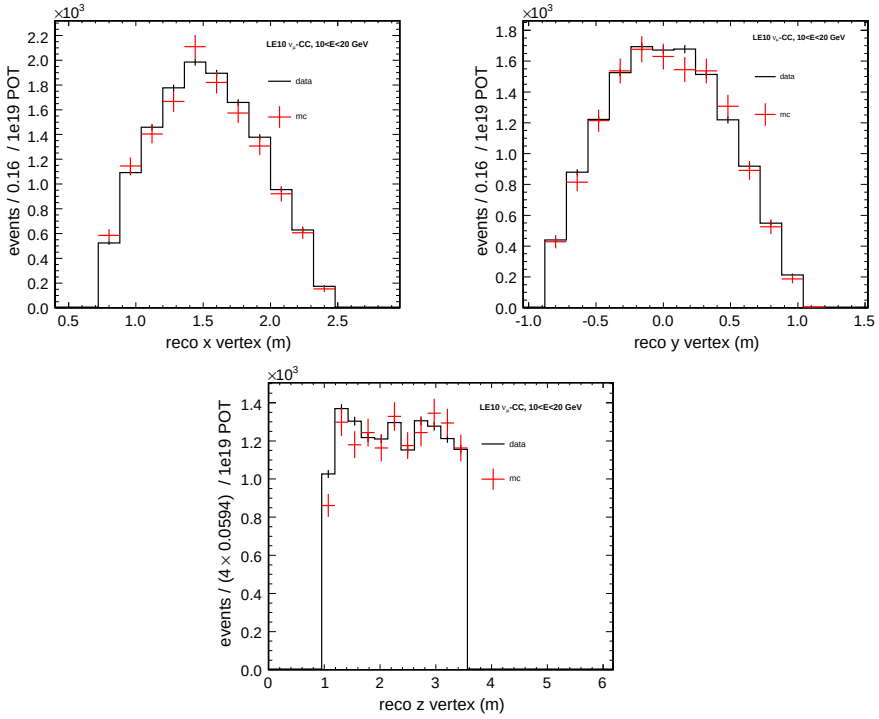


Figure 10: Distributions of the reconstructed vertex positions for the normalisation sample.

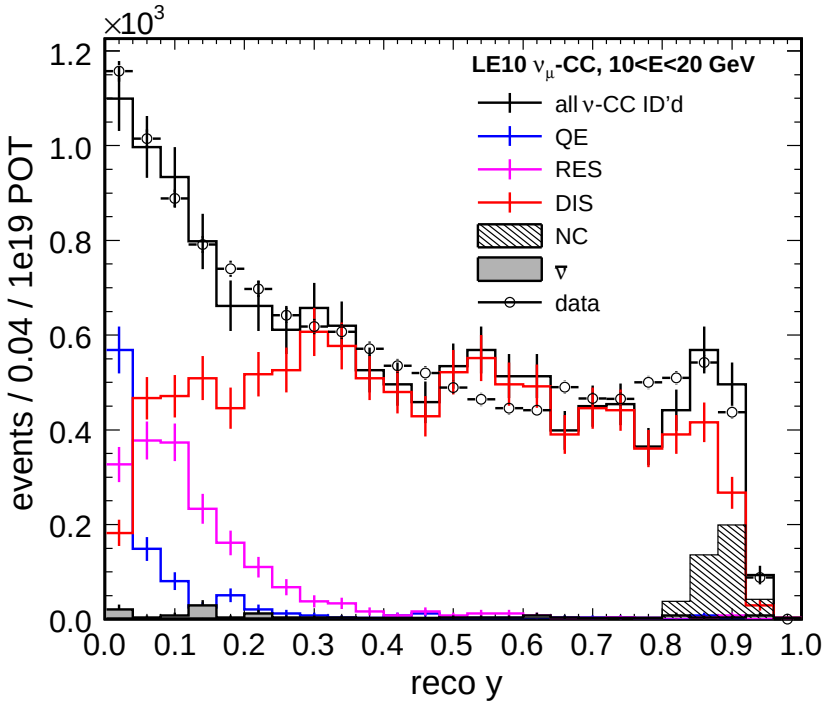


Figure 11: The reconstructed y distribution of the normalisation sample. The χ^2(Kolmogorov-Smirnov) probability for data/MC agreement is 11% (97%).

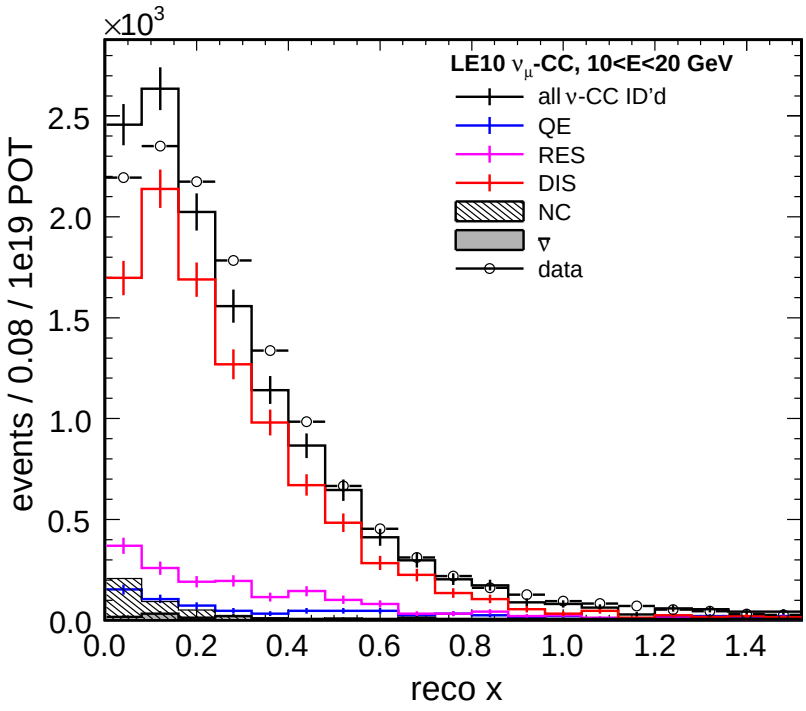


Figure 12: The reconstructed Bjorken-x distribution of the normalisation sample. The χ^2(Kolmogorov-Smirnov) probability for data/MC agreement is 0.04% (9%).

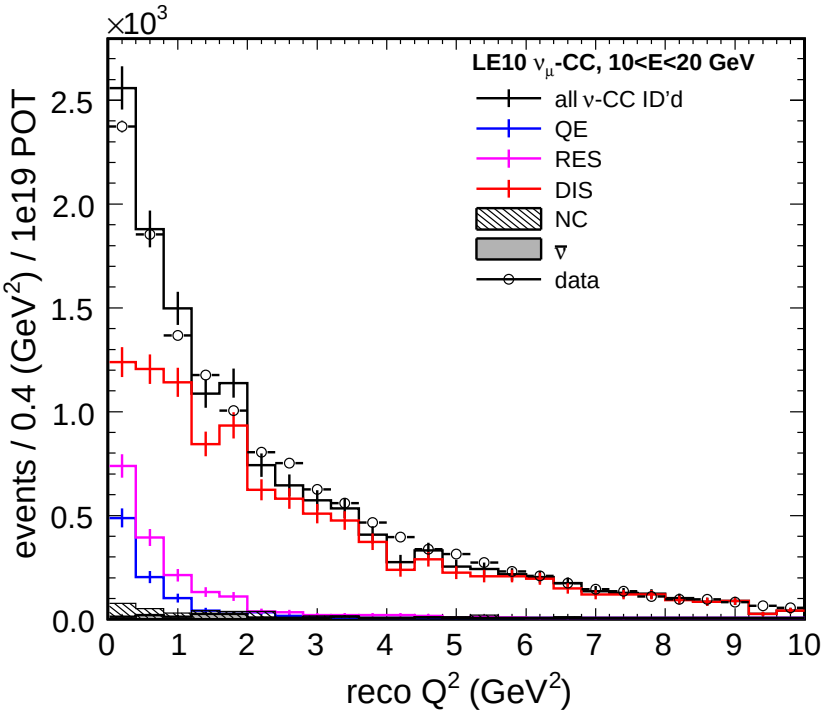


Figure 13: The reconstructed Q^2 distribution of the normalisation sample. The χ^2(Kolmogorov-Smirnov) probability for data/MC agreement is 3% (6%).

data/MC normalisation factor $10 < E_\nu < 20\,\mathrm{GeV}$	
inclusive	0.98(0.2)
inclusive ($0.2 < y < 0.75$)	0.98(0.3)
DIS enriched	1.04(0.3)
DIS enriched ($0.2 < y < 0.75$)	1.02(0.3)

Table 3: The measured data/MC ratio from the normalisation sample. Statistical uncertainties are shown in (...).

and acceptance. This sort of procedure is most valid when applied to energy distributions which change slowly on the scale of the detector resolution (= $14.0 \pm 0.9\,\%$ for this sample) and for which the agreement in spectral shape and vertex distributions is already relatively good. Our normalisation sample satisfies these criteria. An initial calculation, with no additional event selection yields a ratio 0.98 ± 0.02 where the error is statistical and dominated by the size of the MC sample.

The non-$\nu_\mu-$CC background in the normalisation sample is expected to be small (Tab. 2) but not completely negligible. NC events form the only severe component but they appear at high reconstructed y (Fig. 11) and the predicted y-distribution is in good agreement with the data. Nevertheless, NC events must be removed since they feed down into the sample from higher energies. Also, the normalisation ratio will eventually be used to set the scale of the flux distribution derived from quasi-elastic events. It is therefore advisable to limit the overlap between the two samples. With these considerations in mind we recalculate the normalisation for events within $0.2 < y < 0.75$. The resulting ratio is unchanged: 0.98 ± 0.03. Finally, we have selected a 92% DIS enriched sub-sample (53.4% of the original sample) by requiring events to have $Q > 1\,\mathrm{GeV}$ and $W > 2\,\mathrm{GeV}$. The ratios with (without) the y selection are 1.02 ± 0.03 (1.04 ± 0.03). Our results are summarised in Tab. 3.

It is not clear that the cross-section for the DIS selection is as well constrained by the data as the total CC cross-section . We will, therefore, use the ratio 0.98 ± 0.3 for the normalisation of the $E_\nu < 10\,\mathrm{GeV}$ range. To predict an absolute flux we re-normalise the generated flux to obtain:

$$\Phi = 1.77 \pm 0.06\,\mathrm{m}^{-2}10^{-5}\ \mathrm{POT}^{-1} \quad 10 < E_\nu < 20\,\mathrm{GeV}$$

The uncertainty above is statistical. We consider systematic errors in the next section.

4.4 Systematic Uncertainty on the Normalisation

In addition to the uncertainty in the $\nu_\mu-$CC total cross-section discussed above we have examined the following sources of systematic errors:

fiducial mass from assay Based on measurements of the Near detector steel we estimate the plane-to-plane difference in thickness as $\pm 3\%$ [15]. Though

the plane-to-plane differences are known we do not account for them in Monte Carlo and do not attempt to correct our fiducial mass using the information. We estimate the mean thickness as 2.55 ± 0.02 cm. This mean is offset from the plane thickness assumed in MC by an insignificant $+0.3\%$. In addition, the MC assumes a steel density of 7.87 g/cc. Limited measurements suggest that the actual steel density is $\sim 2.5 \pm 0.6\%$ lower than the MC value [16]. We do not correct our data/MC ratio but, based on these considerations, do assign a $\pm 2.5\%$ systematic error to it.

z vertex asymmetry Interaction vertices are distributed uniformly along the z-axis of the detector. The detector acceptance and instrumental non-uniformities may distort the reconstructed distribution but we rely on the MC to accurately model the distortion such that the data/MC ratio is constant along z. We divide our sample into two parts $A = 1.0 < z < 2.3$ m and $B = 2.3 < z < 3.6$ m (see Fig. 10), recalculate the data/MC ratio in each, and use the asymmetry $\frac{|A-B|}{2(A+B)}$ plus it's own statistical error to estimate our uncertainty as $\pm 2.6\%$. We note that if we use the entire $0 - 20$ GeV sample the estimated error shrinks to $\pm 1.2\%$.

x vertex asymmetry As for the z vertex asymmetry but instead we divide the x vertex distribution into $A = 0.5 < x < 1.5$ m and $B = 1.5 < x < 2.5$ m from which we derive an uncertainty $\pm 1.9\%$.

y vertex asymmetry As for the z vertex asymmetry but instead we divide the y vertex distribution into $A = -1.1 < y < 0.0$ m and $B = 0.0 < y < 1.1$ m. We find that the resulting asymmetry is consistent with zero.

energy scale We estimate a 5%(10%) uncertainty in the muon(shower) energy scale and add the two values together assuming $y = 0.5$ to extract a scale uncertainty of 5.6%. We then resample from the measured (data) energy distribution for the entire CC sample, applying the scale uncertainty as a multiplicative factor in each event, and recalculate the data/MC ratio. We do not modify the MC distribution in this process. The ratio varies by $^{+1.1}_{-2.2}\%$ for $\pm 5.6\%$ changes to the energy scale.

z vertex smearing Figure 14 shows the difference between the true and reconstructed z-vertex for our MC sample. The distribution is asymmetric with a negative tail caused by backward going shower particles and spurious hits included onto the recontacted track. We have characterised the tail using a log-normal fit. We then use the distribution to smear z-vertex positions in a stand-alone MC, calculating the smeared/unsmeared ratio while varying the strength of the tail by up to a factor of 3 and the width of the distribution by up to a factor of 10. We are unable to produce a significant change in the smeared/unsmeared ratio. In all cases as many events enter through the front of our fiducial volume as leave through the rear.

Systematic errors on the data/MC ratio $10 < E_\nu < 20\,\mathrm{GeV}$	
source	uncertainty in data/mc
cross-section	$\pm 2\%$
mass assay	$\pm 2.5\%$
z vertex asymmetry	$\pm 2.6\%$
x vertex asymmetry	$\pm 1.9\%$
y vertex asymmetry	–
energy scale	$^{+1.1}_{-2.2}\%$
z vertex smearing	–
hadronisation	–
total	$^{+4.7}_{-5.0}\%$

Table 4: Systematic errors on the data/MC ratio determined from the normalisation sample. See text for details.

hadronisation There is significant uncertainty in the strength of intra-nuclear re-scattering and hadronic final state multiplicities. The former modifies the visible energy while both change event topology. We have accounted for uncertainty in the hadronic energy scale above but here consider the effect that the topological variations may have on event selection. Three separate MC datasets were generated, each with 1e4 ν_μ–CC events between 10–20 GeV. The first had no simulation of intra-nuclear re-scattering, the second had a full simulation in which the energy of absorbed pions disappeared and in the third the energy of absorbed pions was transferred to a *npnp* cluster, the individual nucleons later being passed to `GMINOS` for tracking. The selection efficiency was, with a statistical precision of 1.3%, the same in each of these runs. We generated an additional three datasets: one with nominal hadronic final state multiplicity and the other two with the multiplicity varied by $\pm 20\%$. To within the statistical precision of $\pm 0.6\%$ the selection efficiency was the same in each of these three runs. We assign no systematic error due to hadronisation uncertainties.

The systematic errors are summarised in Tab. 4.

5 Shape Measurement

5.1 Event Selection Technique

5.1.1 Introduction

The selection technique uses several discriminating variables to construct probability distribution function (PDF) based likelihood. All variables used vary with energy and so a separate likelihood analysis is conducted in a number of energy bins in the range $0 < E_\nu < 20\,\mathrm{GeV}$. The bin width varies with energy, accord-

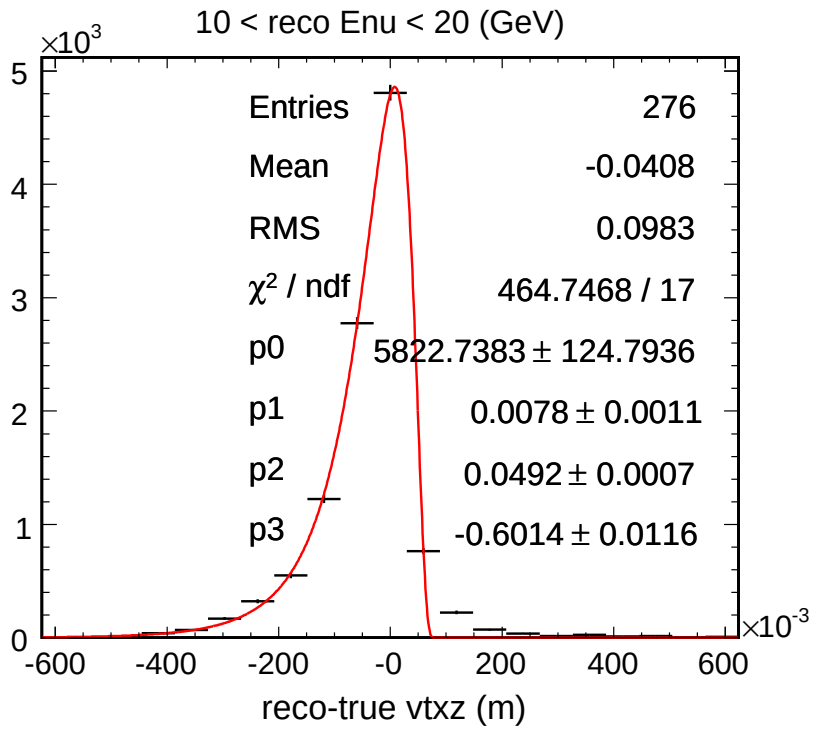


Figure 14: The difference between the reconstructed and true z-vertex for the $10 < E_\nu < 20\,\text{GeV}$ sample. The asymmetry is caused by backward going shower particles as well as a proclivity for inclusion, by the track finder, of spurious hits onto both ends of reconstructed tracks. The fit is to a log-normal distribution and is discussed in the text.

ing to the expected energy resolution for our events, so as to reduce the effect of bin-to-bin event migration. The likelihood analysis takes two PDFs for each variable and in each energy bin; one for true charged current quasi-elastic events (QE) and one for all remaining background events. The background PDFs include contributions from charged current resonant (RES), charged current deep inelastic scattering (DIS) and neutral current (NC) events. The following section describes the discriminating variables used and the construction of an event identification parameter (PID).

5.1.2 Discriminating Variables

The discriminating variables are calculated and tabulated on for events that pass the following pre-selection criteria:

- The event has a reconstructed neutrino energy in the range $0 < E_\nu < 20\,\text{GeV}$.
- The passes the fiducial volume, track quality and fit quality cuts mentioned in section 4.1.

Several of the discriminating variables rely on the identification of hits that correspond to the vertex hadronic shower for a given event. The following steps are applied to a given event and the resulting hits define the vertex hadronic shower for that event in this analysis:

- Any hits that have a pulse height of less than 150 $SigCors$ are removed. This cut is an effort to reduce the number of cross talk-like hits.
- Any hits that are radially further than 2 meters away from the track vertex in the (u, z)-plane or the (v, z)-plane are removed. This cut assumes that protons and pions will not travel further than 2 meters in the detector.
- Any hits that constitute part of the track but not part of any shower are removed.
- Any hits that constitute part of the track and part of a shower are included but have 1 MIPs worth of pulse height subtracted. If this subtraction takes the hit pulse height down to zero or below the hit is removed.

Many of the variables attempt to discriminate quasi-elastic events from background on the basis of the different ionization signatures of protons and pions in the detector:

- Protons will only travel a small distance in the detector and will leave at most a couple of high pulse height hits.
- Pions have a greater range in the detector and will leave more hits of a lower average pulse height than protons.

The following discriminating variables are used in the construction of the event identification parameter:

1. The reconstructed invariant mass squared:

   - True QE events will be strongly peaked about the squared mass of the proton $m_p^2 = (0.938)^2$.
   - True RES events will be more loosely peaked about the squared mass of the $\Delta(1232)$, the main resonance that will be excited in the MINOS detectors.
   - True DIS and NC events will have no definite peaks due to the kinematic natures of these event types but should, on average, have values in excess of that of a true CCQE event.

   Fig. 15 shows the PDFs for the reconstructed invariant mass squared combined into 4 bins of reconstructed E_ν. In all plots of PDFs the true QE events are shown as black histograms and the true background events are shown in red. The figures are based on the MC dataset described in section 4.2.

2. The number of showers:

   - True QE events will have little or no hadronic activity in the detector and so this variable will be mostly zero and one.
   - All other event types are likely to have more showers.

   Fig. 16 shows the PDFs for the number of showers in the event combined into 4 bins of reconstructed E_ν. Most of the power of this variable comes from the zero and one shower bins.

3. The total pulse height of the vertex hadronic shower hits as a fraction of the total event pulse height:

   - True QE events should have small values in this variable as there is little or no hadronic activity associated with the proton.
   - True RES and DIS events will have larger values as there are more final state particles and hence more hadronic activity in the detector.
   - True NC events will have the largest values as there is no PH associated with a muon to be removed from the event.

   Fig. 17 shows the PDFs for the total vertex hadronic shower PH as a fraction of the total event PH combined into 4 bins of reconstructed E_ν.

4. The summed height of the peak in Hough space for a Hough Transform over the hadronic vertex shower hits in the U and V plane views:

- The peak height in Hough space is a measure of the multiplicity of hits and also includes information about the spread of those hits. It could be used to try to quantify how well a series of hits conform to the hypothesis that they are a track or a shower although this has been tested and proved difficult to use given the detector resolution.
- For true QE events the peak should be small as there is little hadronic activity in the detector.
- For the background events this variable should be larger, both because there are more final state particles and because pions from RES and DIS events will tend to travel further in the detector and leave track-like sets of hits that result in a large Hough peak. For NC events the Hough transform will not find a track-like set of points but the peak should still be large because the binning is coarse enough that a number of shower hits can fall into the same bin in Hough space.

Fig. 18 shows the PDFs for the summed Hough peak height from the (u,z) and (v,z) planes combined into 4 bins of reconstructed E_ν.

5. The number of hits in the vertex hadronic shower:

   - For true QE events there should only be a couple high PH hits.
   - For the background events there should be many more hits. For RES and DIS events these will come from the greater number of final state particles and the larger range of pions as compared to protons. For NC events there should be many shower hits.

   Fig. 19 shows the PDFs for the number of hits in the vertex hadronic shower combined into 4 bins of reconstructed E_ν.

5.1.3 Construction of the Event Identification Parameter

PDFs are filled for each of the above discriminating variables for QE and non-QE events in each of the bins of reconstructed $E_\nu$by taking the distributions in each variable and normalising to unity. Then for a given event the probability to be QE or non-QE is formed by looking at the PDFs in the appropriate energy bin and forming compound probabilities according to:

$$P_{QE} = \Pi_{i=1}^{5} p_{QE,i}$$
$$P_{non-QE} = \Pi_{i=1}^{5} p_{non-QE,i}$$

where i indexes the discriminating variables and the $p_{QE,i}$ or $p_{non-QE,i}$ are the individual probabilities obtained from the PDF for the variable in the appropriate bin of reconstructed E_ν. The event identification parameter e_{id} can then be constructed for a given event according to:

$$e_{id} = -\sqrt{-\ln P_{QE}} + \sqrt{-\ln P_{non-QE}}$$

and as such should take values less than zero for non-QE-like events and values greater than zero for QE-like events. Cuts on this parameter are then made in each energy bin to select a sample of QE-like events. Fig. 20 shows the distributions of this event identification parameter for true signal and background events for a $\sim 2.4e18$ POT exposure of R1.18 reconstructed LE-10 MC. A large number of the background events (and some true QE events) are given a default value of -5 where two or more of the probabilities from the PDFs for that event to be QE were equal to zero. The next section shows some example event displays of events passing and failing the selection criteria.

5.1.4 Examples of Events Passing and Failing Cuts

The following are some examples of events that have passed and failed the QE-like selection criteria:

- Fig. 21 shows an example of a signal event that has passed the selection criteria.
- Fig. 22 shows an example of an irreducible background event where, although the event was not true QE, there was no hadronic activity observed in the detector. Events of this type make up a large fraction of the background in the sample.
- Fig. 23 shows an example of an irreducible background event where a RES event looks QE-like as one or other of the proton or pion was not observed in the detector.
- Fig. 24 shows an example of a background event that has been accepted because of a reconstruction failure. Our method treats all the reconstructed events in a snarl independently and utilises only hits that have been associated with a track or shower. In this instance a spurious track has been reconstructed out of a large shower as a separate event which appears to be QE-like. Our event selection criteria will be enhanced to remove such events in the future.

5.2 Efficiency and Background Estimation

This analysis requires a large number of MC events to fill the PDFs and so a series of MC samples were generated with a $\frac{1}{E}$ flux out to 60 GeV in true E_ν. These samples are idealised in that they only contain 1 event per snarl and all events are in the center of the detector. Two samples were generated, the first corresponds to ∼30,000 pre-selected events with reconstructed $E_\nu$less than 20 GeV of all process types and with the appropriate cross-sections applied. A second sample was generated with only true QE events and contains ∼20,000 pre-selected events with reconstructed neutrino energy below 20 GeV. The first MC sample was used along with half of the all-QE MC sample to fill the PDFs for the analysis and the remaining half of the all-QE MC sample was used to train the cut on the event identification parameter in each bin so as to give a

specified flattened acceptance for true QE events as a function of reconstructed $E_\nu$between 0 and 20 GeV. The analysis was then performed over a ∼2.4e18 POTs exposure of LE-10 MC with cut values trained to give a flattened true QE efficiency of 40%. This MC sample gives a prediction of the efficiences of the selection procedure for true QE, RES and DIS events. It also gives an estimate of the number of NC events accepted for a given exposure of POTs. The following plots show some characteristics of the samples from the LE-10 MC and a ∼3.4e19 POTs exposure of data from the Near Detector:

- Fig.25 shows the predicted true QE efficiency and purity of the QE-like sample from MC.
- Fig.26 shows the MC QE-like sample reconstructed $E_\nu$spectrum decomposed by event process type.
- Fig.27 shows the POT normalised data and MC comparison of the QE-like sample reconstructed neutrino energy spectra.
- Fig.28 shows the MC QE-like sample reconstructed q^2 distribution decomposed by event process type.
- Fig.29 shows the POT normalised data and MC comparison of the QE-like sample reconstructed q^2 distributions.
- Fig.30 has been zoomed to highlight the low-q^2 region of the previous figure.
- Fig.31 shows the POT normalised data and MC comparison of the QE-like sample reconstructed y distributions.

5.3 Derivation of the Flux Shape

For a given energy bin i:

$$n\left(E\right) = \Phi_\nu\left(E\right)\left[\sigma_{\mathrm{QE}}\left(E\right)\epsilon_{\mathrm{QE}}\left(E\right) + \sigma_{\mathrm{BG}}\left(E\right)\epsilon_{\mathrm{BG}}\left(E\right)\right] \tag{1}$$

In the above equation $n\left(E\right)$ denotes the number of QE-like events in bin i that contains the true neutrino energy E, $\sigma_{\mathrm{QE/BG}}$ is the cross-section for QE or background events at the true neutrino energy E, and $\epsilon_{\mathrm{QE/BG}}$ is the selection efficiency for QE or background events at the true neutrino energy E.

The background term can be broken down into its components:

$$\sigma_{\mathrm{BG}}\left(E\right)\epsilon_{\mathrm{BG}}\left(E\right) = \sigma_{\mathrm{RES}}\left(E\right)\epsilon_{\mathrm{RES}}\left(E\right) + \sigma_{\mathrm{DIS}}\left(E\right)\epsilon_{\mathrm{DIS}}\left(E\right) + \sigma_{\mathrm{NC}}\left(E\right)\epsilon_{\mathrm{NC}}\left(E\right) \tag{2}$$

Here $\sigma_{\mathrm{RES/DIS/NC}}$ is the cross-section and $\epsilon_{\mathrm{RES/DIS/NC}}$ is the selection efficiency for resonant, DIS and neutral current events.

The resulting equation can then be rearranged to yield the flux for the i^{th} energy bin in terms of variables that are either measured directly or can be predicted from the MC. The cross-sections $\sigma_{\mathrm{QE}}\left(E\right)$,$\sigma_{\mathrm{RES}}\left(E\right)$ and $\sigma_{\mathrm{DIS}}\left(E\right)$ can

be obtained from NEUGEN . NC events do not have an accurately reconstructed $E_\nu$and so the cross section can not be used and the only way to take them into account is to subtract off the expected number of NC events for a given exposure of POTs based on MC predictions. MC also provides the estimates for the selection efficiencies $\epsilon_{\mathrm{QE}}(E)$, $\epsilon_{\mathrm{RES}}(E)$ and $\epsilon_{\mathrm{DIS}}(E)$. The flux can hence be written:

$$\Phi_\nu(E) = \frac{n(E) - n_{NC}(E)}{\sigma_{\mathrm{QE}}(E)\,\epsilon_{\mathrm{QE}}(E) + \sigma_{\mathrm{RES}}(E)\,\epsilon_{\mathrm{RES}}(E) + \sigma_{\mathrm{DIS}}(E)\,\epsilon_{\mathrm{DIS}}(E)} \quad (3)$$

where $n_{NC}(E) = \Phi_\nu(E)\,\sigma_{NC}(E)\,\epsilon_{NC}(E)$ is the number of NC events with true neutrino energy E.

The NC background is particularly problematic in that it is impossible to unfold it from the flux itself as the events feed down from much higher energies. Instead the level of NC contamination in the QE-like sample must be made small enough such that the uncertainty in the number of NC events in a given energy bin will not be the limiting error on the flux measurement. Also, as an alternative to the treatment of the DIS events listed above, we can instead use a similar subtraction method:

$$\Phi_\nu(E) = \frac{n(E) - n_{NC}(E) - n_{DIS}(E)}{\sigma_{\mathrm{QE}}(E)\,\epsilon_{\mathrm{QE}}(E) + \sigma_{\mathrm{RES}}(E)\,\epsilon_{\mathrm{RES}}(E)} \quad (4)$$

where $n_{DIS}(E) = \Phi_\nu(E)\,\sigma_{DIS}(E)\,\epsilon_{DIS}(E)$ is the number of DIS events with true neutrino energy E.

We are motivated by the fact that many "DIS" events are actually small events reconstructed from the detritus of a much larger event. In these cases the reconstructed neutrino energy is an inaccurate estimate of the true neutrino energy, and a correction based on the cross-section is not valid. Below we present a flux which has been derived by subtracting off the DIS background as we feel that this is the superior approach.

The energy E of Eq. 4 formally refers to the true neutrino energy. We, however, utilise Eq. 4 in bins of reconstructed energy. That is, we conduct no unfolding to the true neutrino energy. We do not regard this as a major deficiency in our analysis since the quasi-elastic and resonance cross-sections are weakly dependent on the energy, the efficiency is nearly independent of the energy, and the energy resolution of our quasi-elastic sample ($\sim 10\%$) is significantly smaller than the bin width. It is probably more robust and certainly easier to smear MC flux distributions according to our resolution in order to compare them with our measurement.

6 Conclusions

Fig.32 shows the estimated flux at the Near Detector. The error bars on the data figures reflect the statistical uncertainty in the data, the statistical uncertainty

in the MC backgrounds and efficiencies and the statistical and systematic uncertainties from the normalisation measurement. The statistical errors in the MC sample dominate, and as a consequence we are now generating and processing a large ($\sim 3e19$ POT) single event dataset for use in background and efficiency determination. A sizable portion of this dataset has already been copied to ENSTORE and is available for general use. The MC curve was derived directly from the `GNuMI-v17` [17] flux files and the errors are the uncorrelated errors reported by the beam systematics group [18].

Though the DIS background subtraction relies on the MC accurately modeling the data[2] it is valid in principle even for badly reconstructed events. The new round of processing is expected to reduce the severity of the DIS-detritus problem and in addition we are developing a set of cleaning cuts specific to the QE sample.

Our method relies on knowing the shape of the efficiency with respect to reconstructed neutrino energy but is relatively insensitive to the actual value since the normalisation is derived from the $10 < E_\nu < 20$ GeV region. The important checks then involve studying the shape (with respect to reconstructed neutrino energy) of the efficiency and purity curves under different systematic conditions. Though much of this work remains to be done we have conducted a preliminary study (see Fig. 33) of the effect of intra-nuclear re-scattering on the quasi-elastic sample purity. The sizes of the MC datasets prevent us from drawing any strong conclusions, though the shape distortion caused by intra-nuclear scattering is probably not much larger than 20%.

Finally, efforts are underway to improve the quasi-elastic selection. In particular we using the over constrained quasi-elastic kinematics to calculate the expected momentum 4–vector of the recoil proton, thereby predicting the location of hits in the event. We hope that this new technique will allow us to raise the efficiency of our selection while holding the background from resonant production at about the same level.

[2] Indeed, after-pulsing, whether instrumental or physics in origin, may make the detritus effect worse in the data.

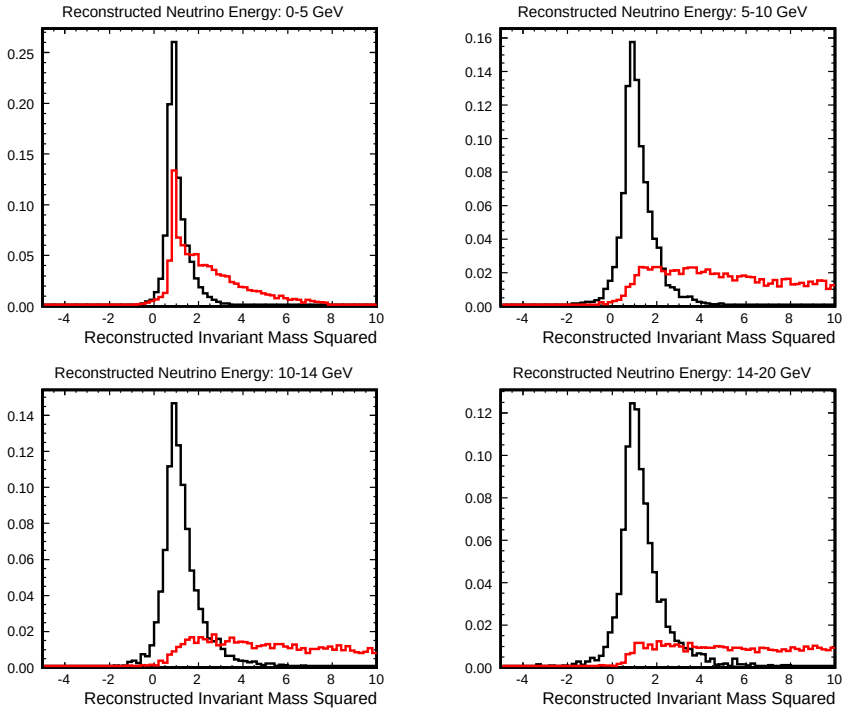


Figure 15: Reconstructed invariant mass squared in 4 bins of E_ν.

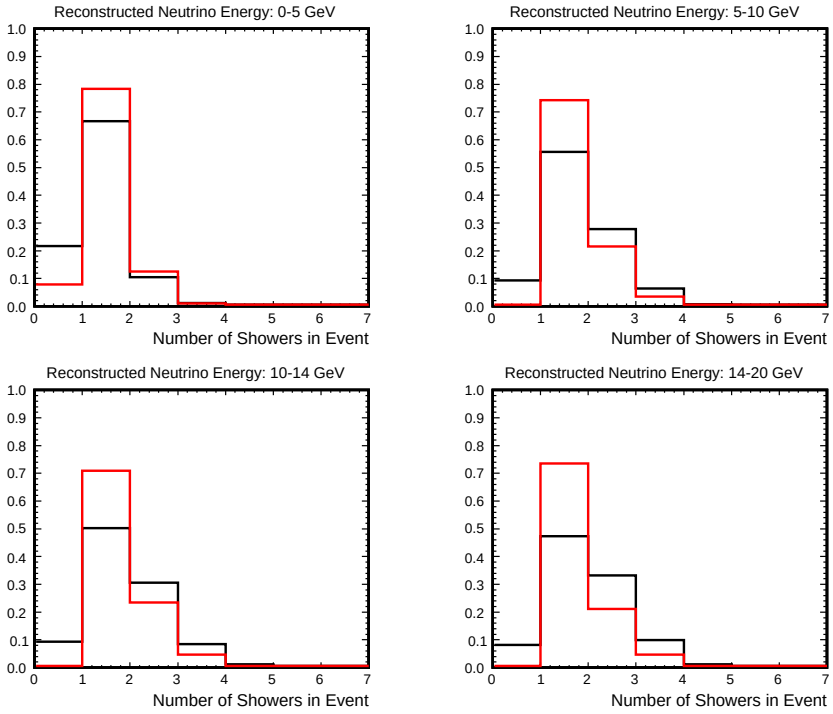


Figure 16: Number of showers in event for 4 bins of E_ν.

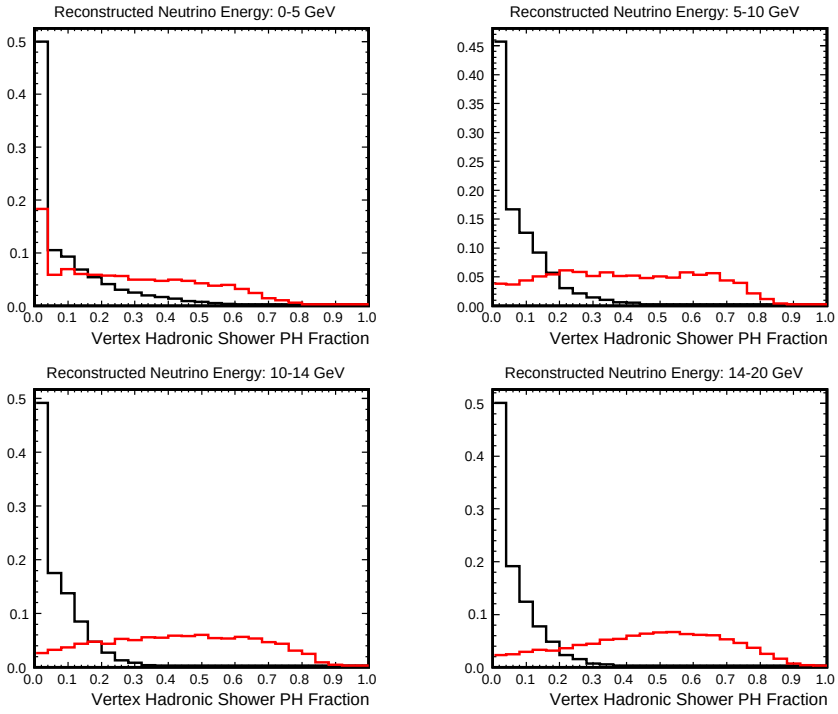


Figure 17: Vertex hadronic shower PH as a fraction of total event PH in 4 bins of E_ν.

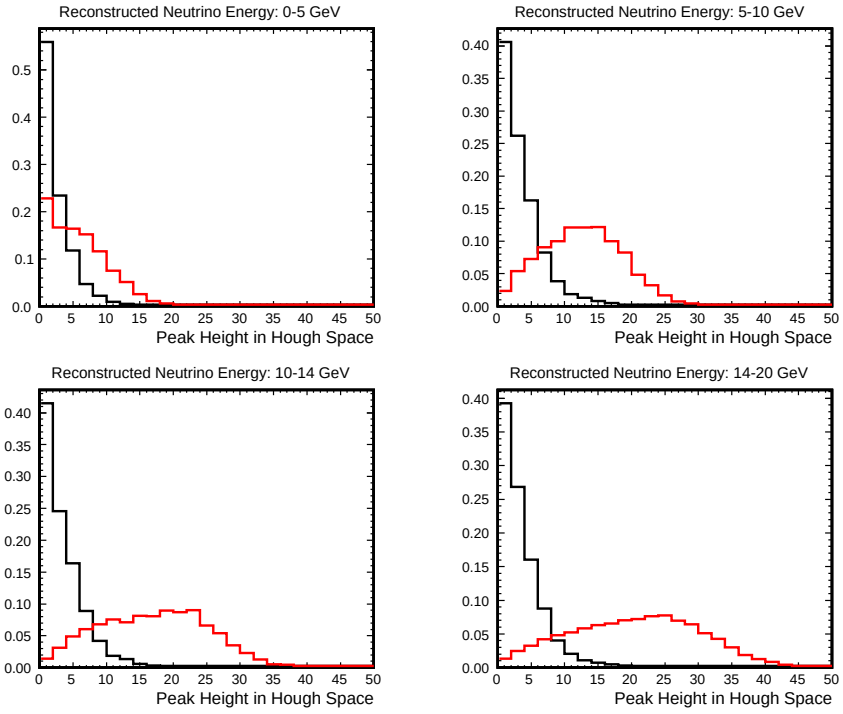


Figure 18: Summed Hough peak height in U and V planes in 4 bins of E_ν.

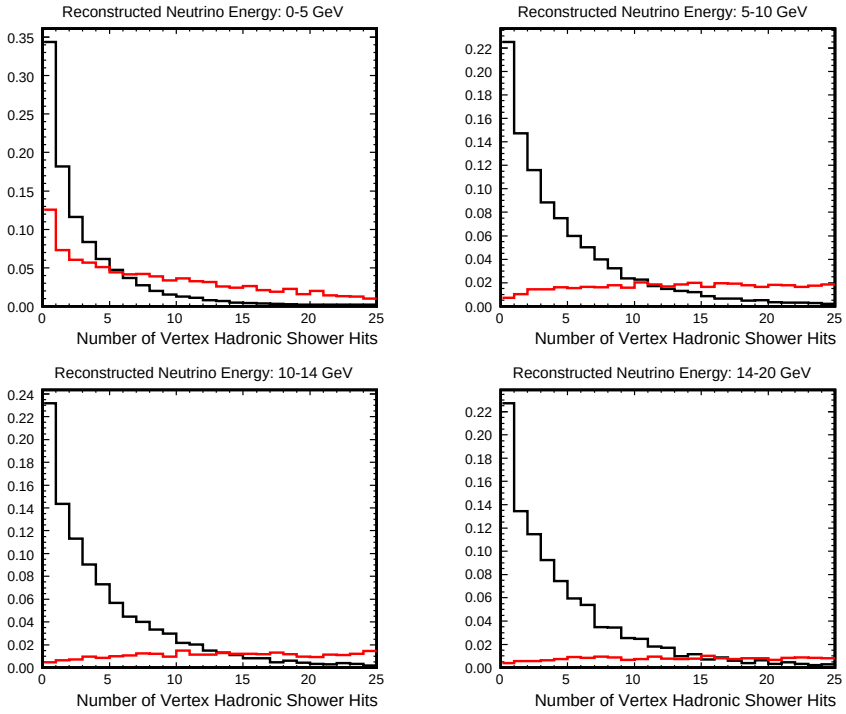


Figure 19: Number of hits in vertex hadronic shower in 4 bins of E_ν.

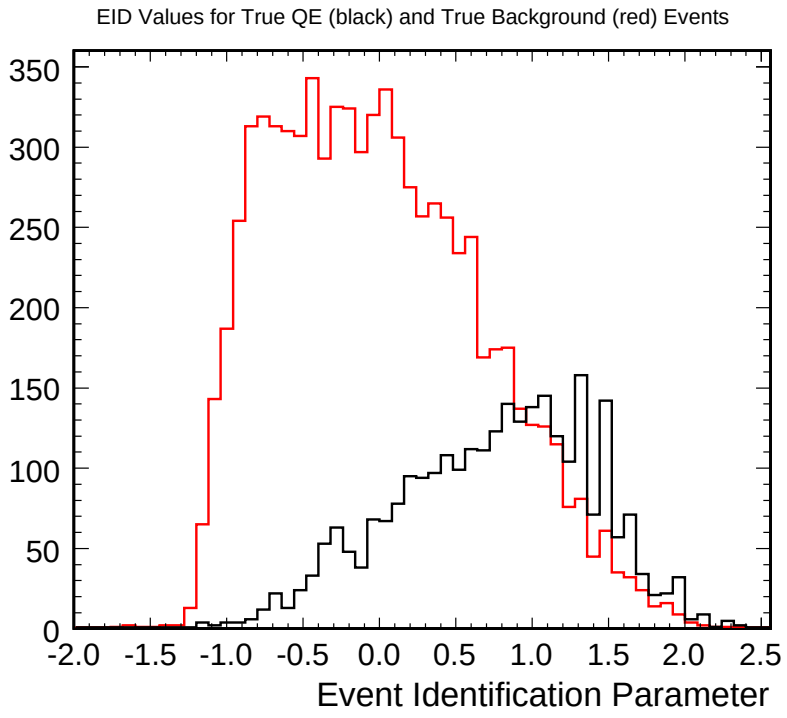


Figure 20: Event identification parameter values for true QE and true background events.

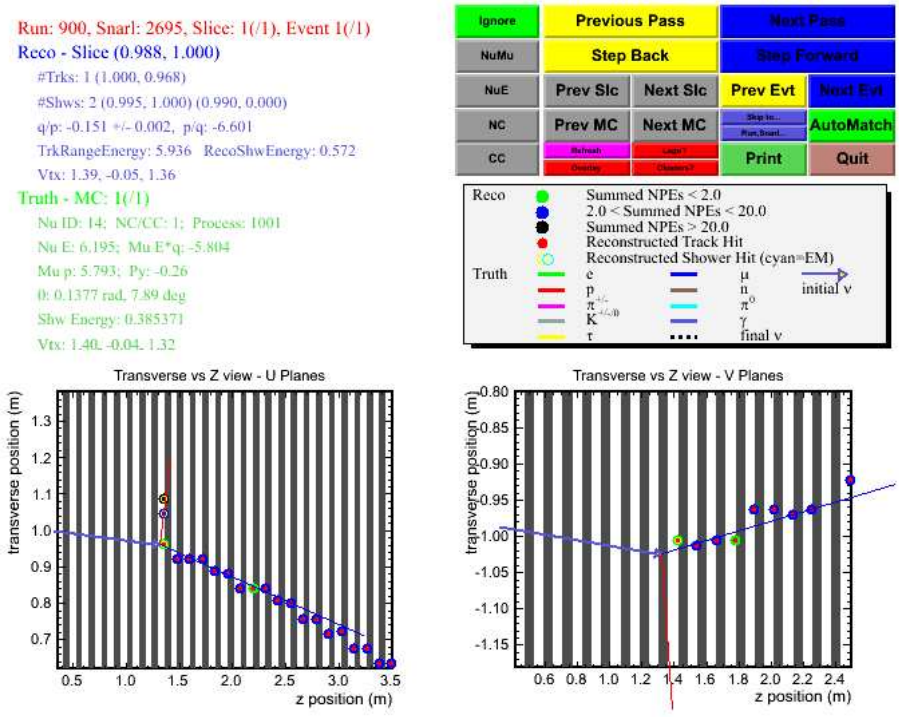


Figure 21: A signal event classified as QE.

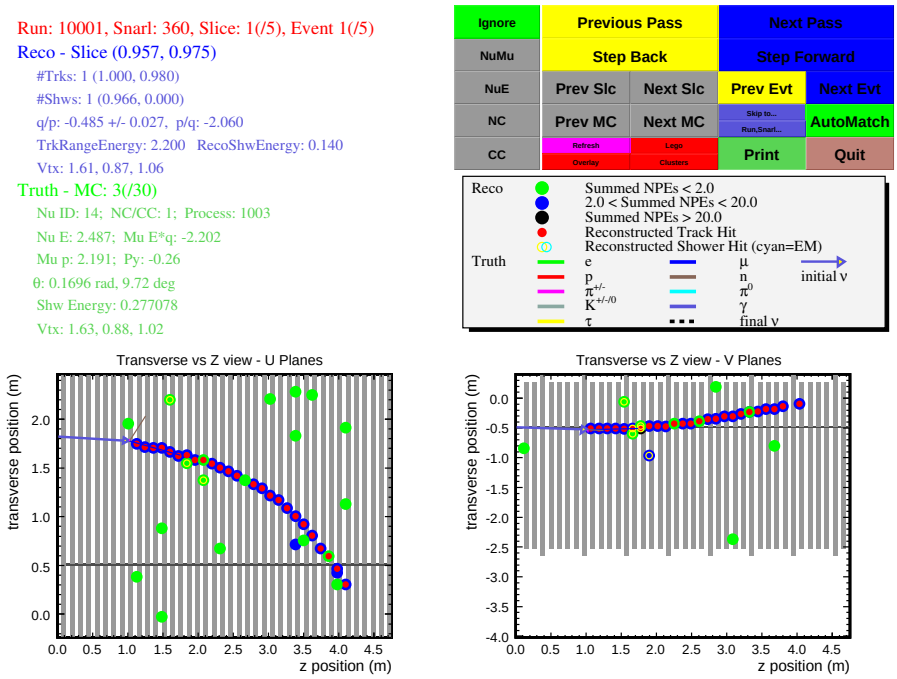


Figure 22: A background event classified as QE.

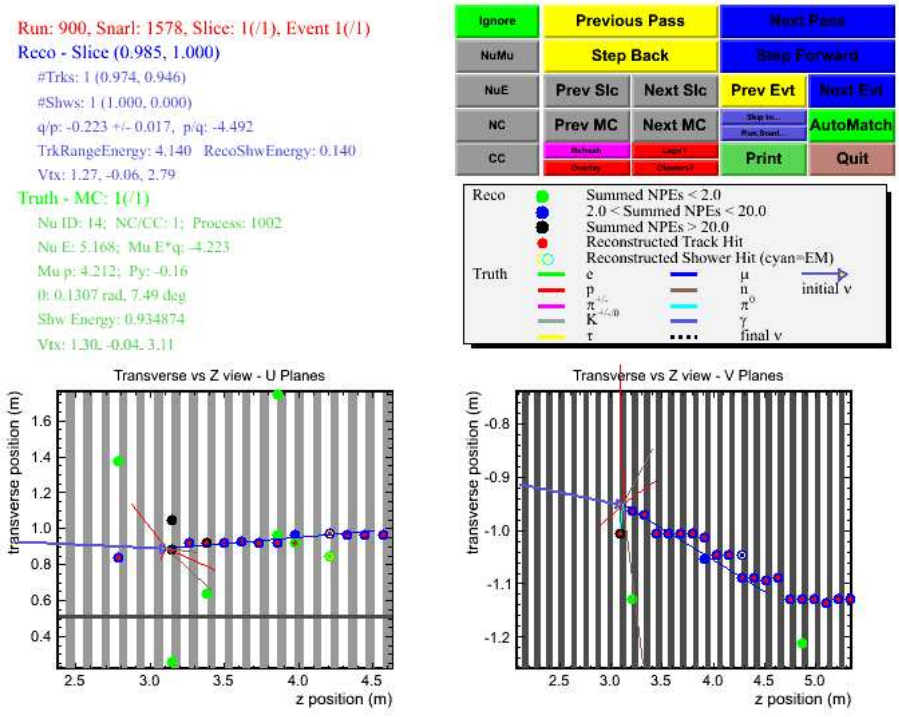


Figure 23: A background event classified as QE.

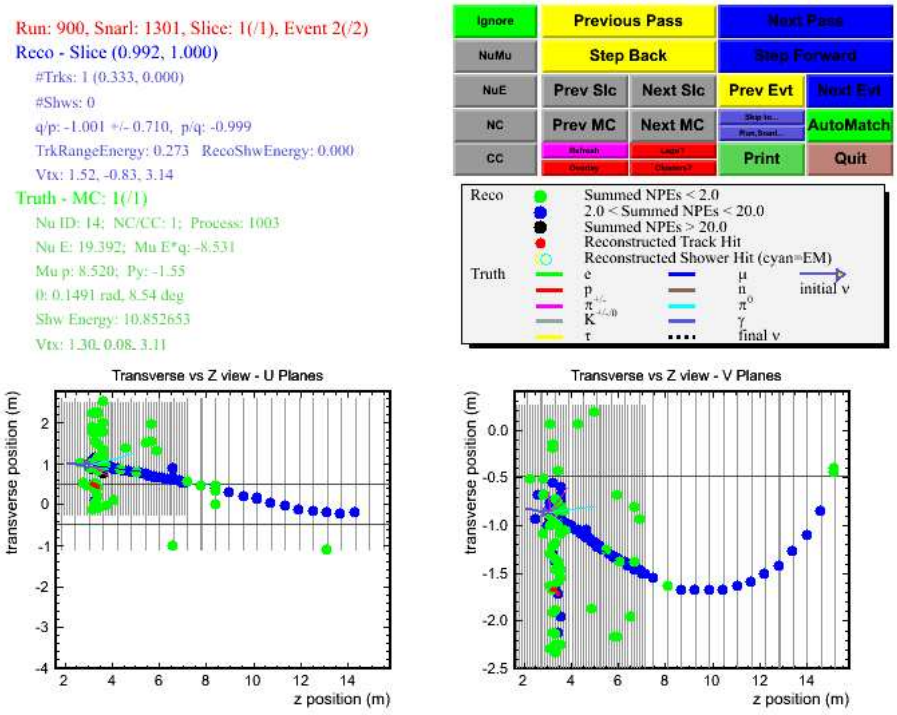


Figure 24: A background event classified as QE.

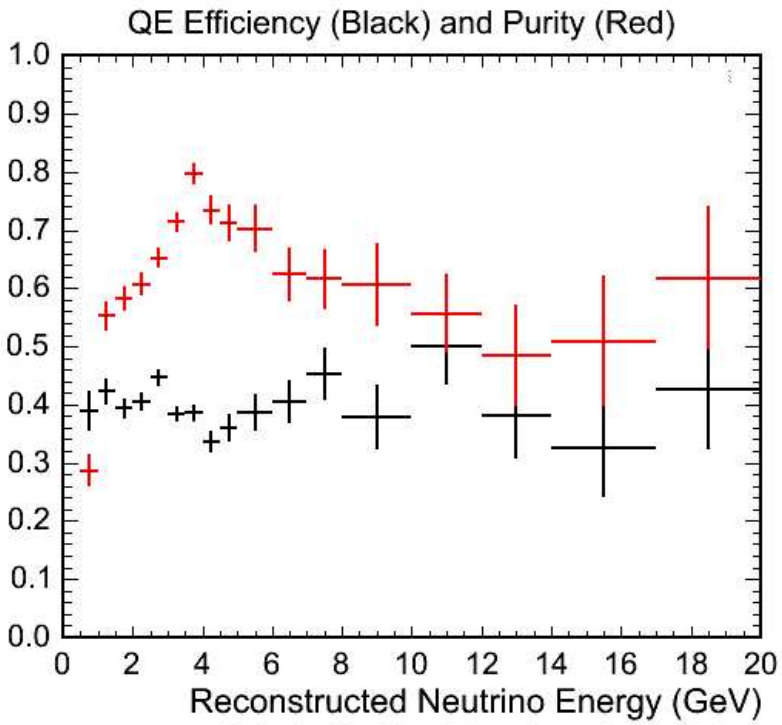


Figure 25: The predicted efficiency (black) and purity (red) of the QE-like sample from MC.

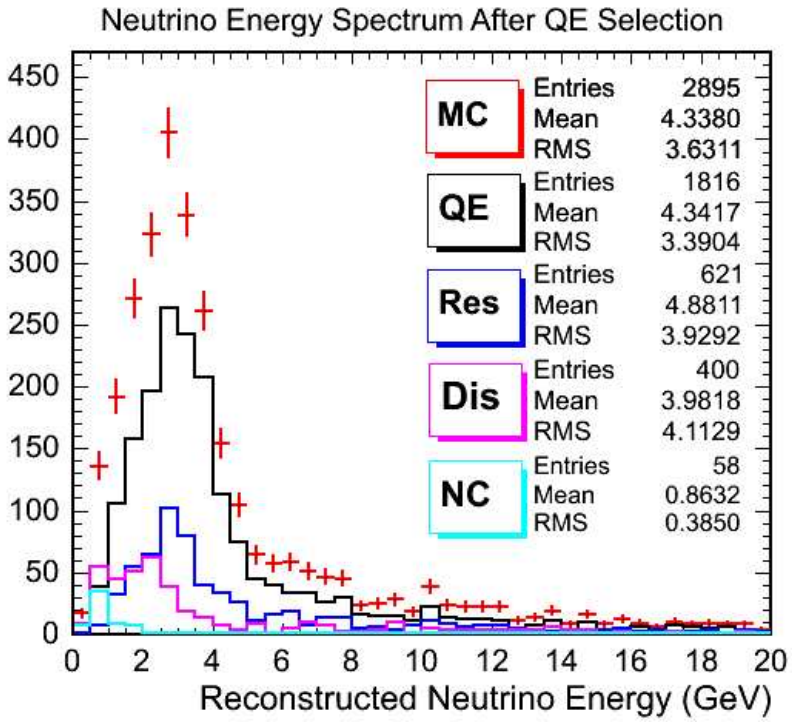


Figure 26: Energy spectrum of the QE-like sample in MC decomposed by event process type.

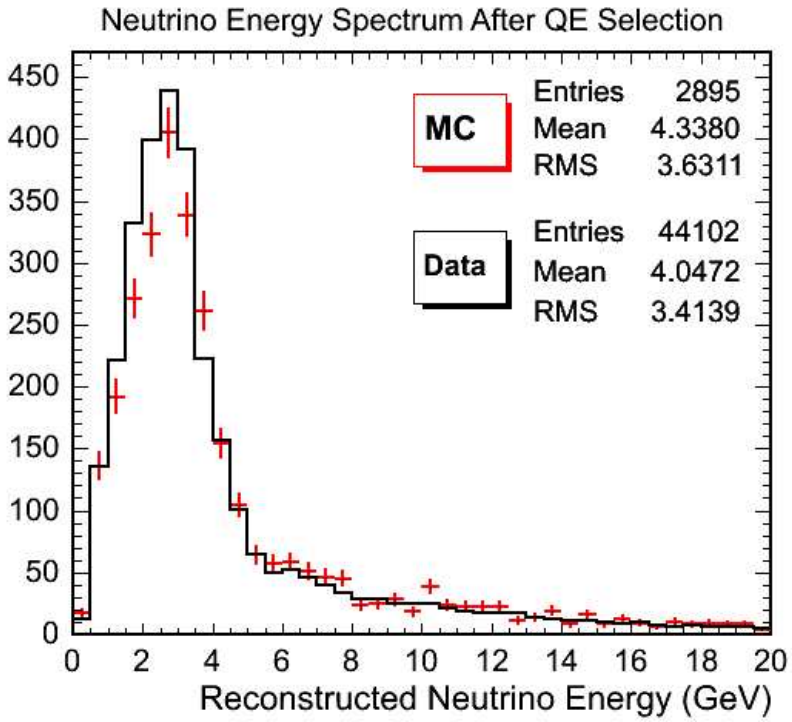


Figure 27: POT normalised data (black) and MC (red) comparison of the QE-like sample reconstructed energy spectra.

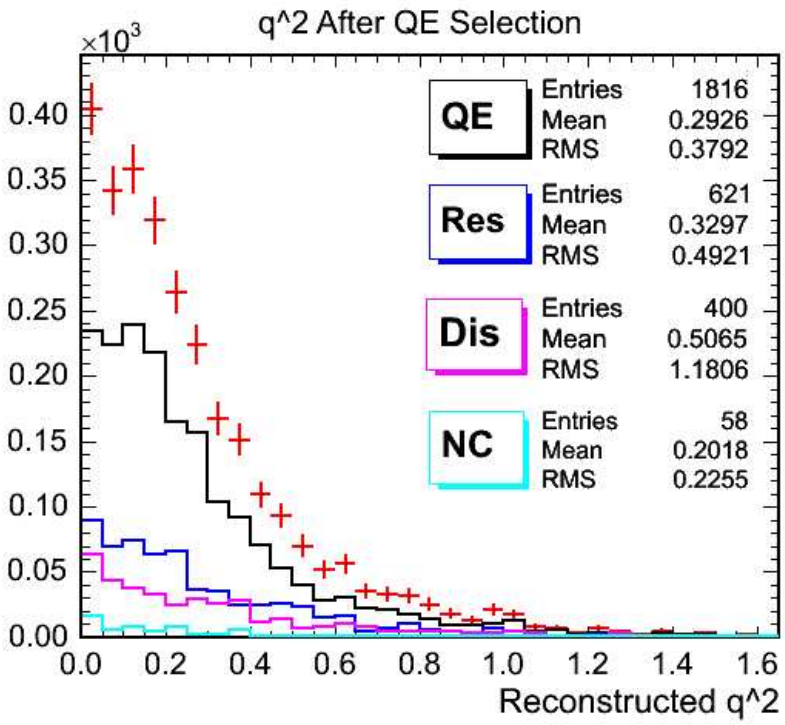


Figure 28: Reconstructed q^2 distribution for the QE-like sample in MC decomposed by event process type.

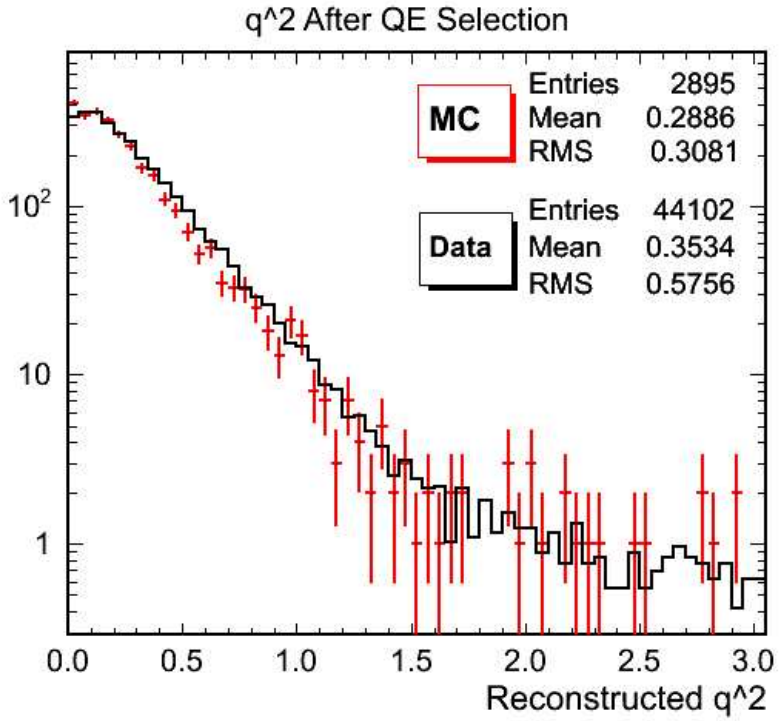


Figure 29: POT normalised data (black) and MC (red) comparison of the QE-like sample q^2 distributions.

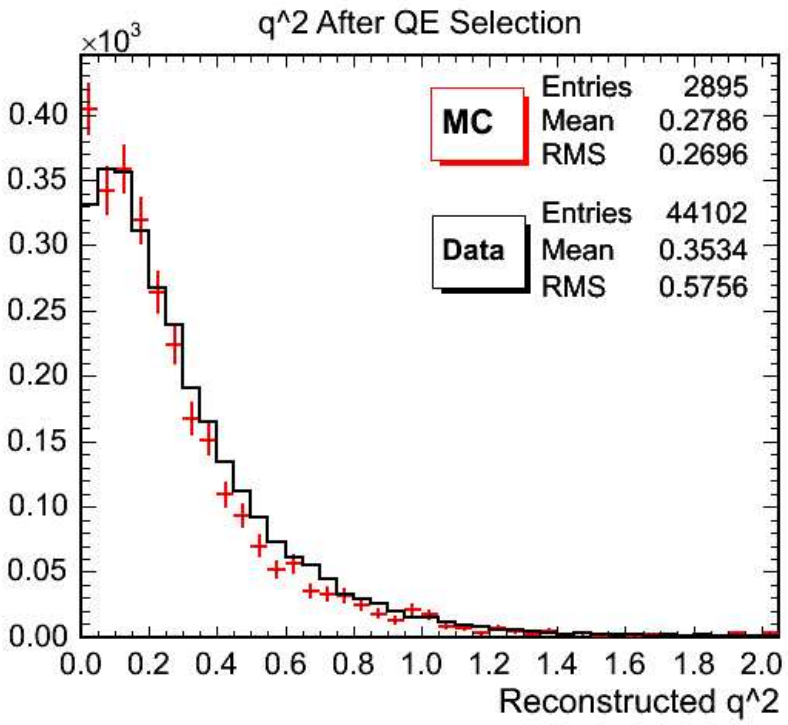


Figure 30: Low q^2 zoom of the POT normalised data (black) and MC (red) comparison of the QE-like sample q^2 distributions.

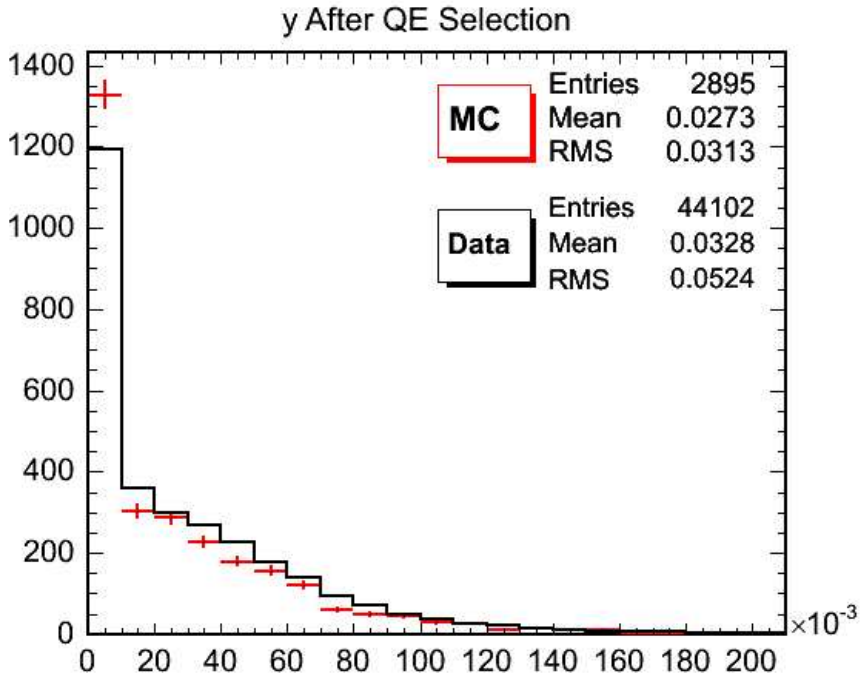


Figure 31: POT normalised data (black) and MC (red) comparison of the QE-like sample y distributions.

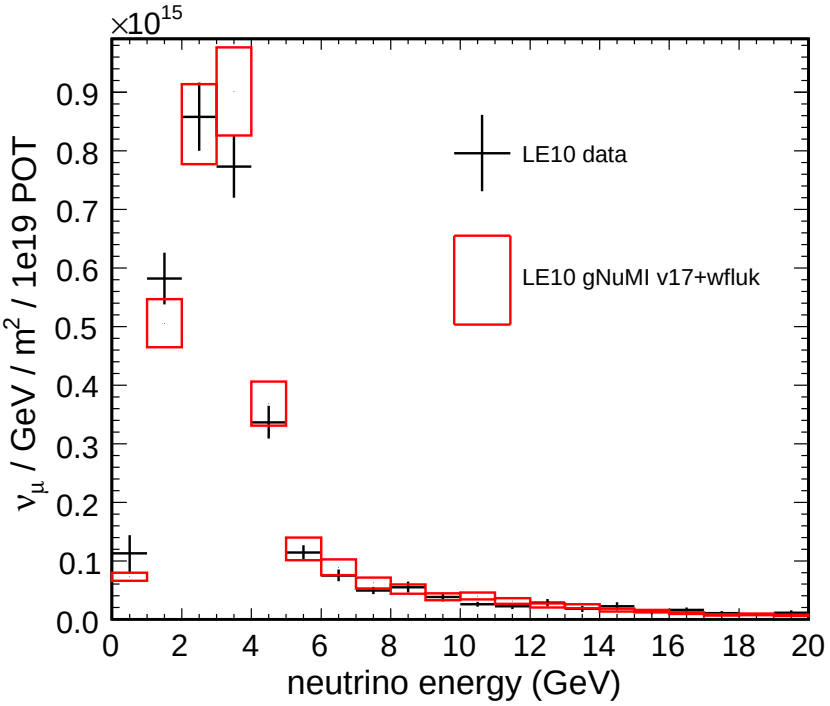


Figure 32: Estimated neutrino flux at the near detector. Both NC and DIS background have been subtracted according to the prescription of Eq. 4. The error bars on the data include statistical uncertainties in the MC efficiencies and backgrounds, statistical uncertainties on the data, and statistical and systematic uncertainties in the scale determined from the normalisation sample. The errors on the MC are those reported by the beam systematics group [18].

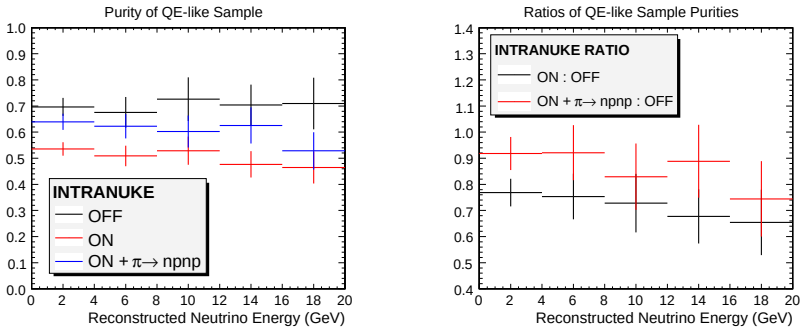
References

[1] V. B. Anikeev, et al., Total cross-section measurements for muon-neutrino, anti- muon-neutrino interactions in 3-gev - 30-gev energy range with ihep-jinr neutrino detector, Z. Phys. C70 (1996) 39–46.

[2] J. M. Conrad, M. H. Shaevitz, T. Bolton, Precision measurements with high energy neutrino beams, Rev. Mod. Phys. 70 (1998) 1341–1392.

[3] H. Gallagher, The neugen neutrino event generator, Nucl. Phys. Proc. Suppl. 112 (2002) 188–194.

[4] Data are from IHEP-JINR, IHEP-ITEP, SKAT, Gargamelle, FNAL-15th BC and BEBC and were taken from the Durham database (`http://`



(a) (b)

Figure 33: (a)Purity of the quasi-elastic sample as a function of energy for different simulations of intra-nuclear re-scattering. The quasi-elastic selection procedure classifies events with low hadronic activity as more likely to be quasi-elastic . The purity of the sample then varies in inverse proportion to the strength of intra-nuclear re-scattering. All curves correspond to an quasi-elastic efficiency of 40%, independent of the reconstructed energy. (b) QE sample purity ratios with "Intranuke Off" in the denominator. In both figures the small size of the MC sample prevents us from drawing strong conclusions.

`durpdg.dur.ac.uk/HEPDATA/`) via the `NuValidator` (`hepunx.rl.ac.uk/~candreop/generators/validator/`) program.

[5] P. S. Auchincloss, et al., Measurement of the inclusive charged current cross-section for neutrino and anti-neutrino scattering on isoscalar nucleons, Z. Phys. C48 (1990) 411–432.

[6] D. MacFarlane, et al., Nucleon structure functions from high-energy neutrino interactions with iron and qcd results, Z. Phys. C26 (1984) 1.

[7] J. P. Berge, et al., Total neutrino and anti-neutrino charged current cross-section measurements in 100-gev, 160-gev and 200-gev narrow band beams, Z. Phys. C35 (1987) 443.

[8] H. Budd, A. Bodek, J. Arrington, Modeling quasi-elastic form factors for electron and neutrino scattering Presented at the Second International Workshop on Neutrino-Nucleus Interactions in the Few GeV Region (NuInt02).

[9] A. Bodek, H. Budd, J. Arrington, Modeling neutrino quasielastic cross sections on nucleons and nuclei, AIP Conf. Proc. 698 (2004) 148–152.

[10] Runs 8200–8300 were removed to simplify the analysis because the range includes a number of runs taken in non-standard beam configurations. In principle the majority of these runs could be recovered but this does not affect the precision of the results since the normalisation method is currently limited by MC statistics.

[11] The exposure has been corrected by -2.8% for variations in the `TORTGT` calibration. An additional overall correction of -0.6% was applied for beam missing the target.

[12] The offline package `Mad` (class `MadMKAnalysis`) was used in the dst construction.

[13] The PDF technique is the one of D. Petyt. PDFs are archived in offline package `Mad/data`.

[14] The `ntsp` ntuple variable is `shw.shwph.linCCgev`. This differs from the often used variable `shw.ph` in that the latter uses the weighted shower energy of A. Culling. Unfortunately a typographical error in the `R1.18` reconstruction code caused the weighed energy be improperly calculated. The authors argue that using the weighted energy is an error for the current reconstruction pass but have high hopes for it in the future.

[15] J. Nelson and D. Boehnlein, private communication. See also `http://www-numi.fnal.gov/numinotes/restricted/html/numi1079/nelson-bfie%ld-mar05.pdf`.

[16] J. Kilmer, private correspondence, March 2004. Steel from three heats was measured and found have a mean 7.755 g/cc with a spread of 0.05 g/cc.

[17] This version of `GNuMI` contains the `wfluk` correction.

[18] Beam related uncertainties are to be described in MINOS-doc-1283. The uncertainties presented here are to be replaced with correlated uncertainties computed with the recently released `zbeam` code of Ž. Pavlović et al.