Analysis of ν_μ Charged Current Inclusive Events in the PØD in Run 1+2+3+4

T2K TN-080 V4

Alex Clifton[1], Rajarshi Das[1], Rob A. Johnson[2], Alysia D. Marino[2], Erez Reinherz-Aronis[1], Walter Toki[1], and Tianlu Yuan[2]

[1] *Colorado State University, Fort Collins, Colorado, USA*
[2] *University of Colorado, Boulder, Colorado, USA*

February 27, 2014

Abstract

We present an analysis of Charged Current inclusive neutrino interactions that originate in the Pi-Zero Detector and have a negative track reconstructed in the Time Projection Chamber. This analysis includes the ND280 Run 1-4 periods and utilized Production 5 samples both for Data and Monte Carlo. The full data sample, which correspond to 56.8×10^{19} Protons On Target, includes both the time periods where the PØD water target bags were filled and empty. The analysis was performed at both the oaAnalysis and oaRecon levels with PØD and Tracker reconstruction objects in conjunction with a Tracker to PØD track matching algorithm. We report the result as a Data to Monte Carlo ratio of the total number of selected Charged Current inclusive interactions for the period where the water bags were fill and empty, which we measure as $0.879 \pm 0.006(stat.)^{+0.024}_{-0.010}(syst.)$. , $0.872 \pm 0.006(stat.)^{+0.024}_{-0.010}(syst.)$. respectively. In addition, detailed systematic studies for the selection have been performed and are reported.

Contents

1 Introduction

This note describes the analysis of Charged Current (CC) inclusive events, $\nu_\mu + N \rightarrow \mu^- + X$, whose vertex is in the Pi Zero Detector (PØD) and its μ^- candidate track is reconstructed by the Tracker Time Projection Chambers (TPC).

The Production 4 samples used in this analysis include several known issues and problems, some of which should be corrected by later productions. The current work have addressed the related analysis issues, which we list below,

- PØD mass in the Monte Carlo.
  The PØD mass in the Monte Carlo (MC) was set slightly different than the constructed one [2] therefore we added correction factors to adjust the MC rates accordingly.

- Displacement between PØD and TPC1.
  Displacement shifts in both X and Y PØD coordinates with respect to TPC1 are present in the Data. To account for these, an alignment correction (7 mm, 26 mm) was applied between the PØD and TPC1.

Section 2 contains a short discussion on the PØD mass followed by Section 3 that describes the Data and MC samples used in the current analysis. Next in Section 4 the Event reconstruction components are laid out. These include energy corrections in the PØD and matching procedure between the PØD and Tracker sub-detectors. In Section 5, the selection cut flow is described. In addition, this section contains the selected distributions for both Data and MC samples and a comparison between them. Section 6 details the different systematic studies that were preformed and finally Section 7 summarizes the final result.

2 Mass of the PØD Detector

The PØD detector is composed of two ECAL SuperPØDules and two Water Target (WT) SuperPØDules [1]. Each ECAL SuperPØDule contains 7 PØDules and 7 lead layers and each WT SuperPØDule contains 13 PØDules, 13 brass layers, and 12 or 13 WT layers. A PØDule is a combination of an X layer and a Y layer of scintillator bars and a radiator later (brass in the water SuperPØDules and lead in the ECaLs). Each WT bag is composed of a frame, water bag and HDPE cover sheet. In this section we summarize the PØD mass by weight as tabulated by Clark McGrew for the detector As Built (AB) for the different MC productions [2]. A more detailed description of the PØD can be found in [1] and [2].

The mass of the PØD detector is estimated inside a fiducial volume that was about 25 cm inside the physical boundary of the PØD detector. This corresponds to about 1.6 m in x, 1.74 m in y and inside the water target in z. These boundaries are taken from Karin Gilje's mass estimates in T2K Tech-note 73, version 3.1.

Table 1: The mass in the PØD fiducial volume for water-in and water-out run periods.

Mass (kg)	Water-in	Water-out	Water-only
As-built (Run 1)	5460.86 ± 37.78	3558.86 ± 34.23	1902 ± 16
As-built (Run 2)	5480.30 ± 37.40	3578.30 ± 33.80	1902 ± 16
Prod. 5	5393.22 ± 0.56	3469.14 ± 0.55	1927.5 ± 0

Table 1 summarizes the mass and the mass uncertainty in the fiducial volume of the PØD. The table is extracted from T2K-TN-73. The differences between the as-built and the MC for each run type will be treated as a correction factor to the Data to MC ratio. The procedure is described in greater detail in Section 6.3.

3 Data and Monte Carlo Samples

The CC inclusive studies reported here have included T2K Run 1-4 periods. In these studies we used production 5 "Spin F"[1] for Data and production 5 "Spin E"[2] for the MC samples.

We note that in the middle of both T2K Run 2 and T2K Run 4 the PØD water bags were drained. The results presented in this report includes both data periods when the PØD water bags were empty and filled with water.

3.1 Data Sample

Data Set	Run Period	Protons On Target	
		Total (Available)	DQ (Used)
Run 1	Jan 2010 - Jun 2010	3.033×10^{19}	2.946×10^{19}
Run 2	Nov 2010 - Feb 2011	6.974×10^{19}	4.286×10^{19}
Run 4	Oct 2012 - Feb 2013	16.45×10^{19}	16.24×10^{19}
Total		26.46×10^{19}	23.47×10^{19}

Table 2: The POT values available and used in the data analysis for the Water-in run periods

The total number of Protons On Target (POT) recorded in the four T2K periods was $\sim 62.18 \times 10^{19}$ POT. The PØD Water-in periods included $\sim 27.95 \times 10^{19}$ POT and the PØD Water-out periods included $\sim 34.23 \times 10^{19}$ POT.

The studies presented here have incorporated both the "ND280" Data Quality flag and the beam "GoodSpill" flag checks. After these checks, the Water-in analyses included 11,237 SubRuns that sum up to around 23.47×10^{19} POT and the Water-in analyses included 12,255 SubRuns that sum up to around 33.54×10^{19} POT. The breakdown for the different run periods is shown in Tables 2 and 3 for the Water-in and Water-out respectively.

Data Set	Run Period	Protons On Target	
		Total (Available)	DQ (Used)
Run 2	Feb 2011 - Mar 2011	3.593×10^{19}	3.552×10^{19}
Run 3	Apr 2012 - Jun 2012	13.58×10^{19}	13.48×10^{19}
Run 4	Feb 2013 - Aug 2013	16.37×10^{19}	15.86×10^{19}
Total		33.54×10^{19}	32.89×10^{19}

Table 3: The POT values available and used in the data analysis for the Water-in run periods

[1] ND280 software version v10r11p21
[2] ND280 software version v10r11p17

3.2 Monte Carlo Sample

Dedicated Monte Carlo samples were produced for the different T2K run periods. These samples incorporated the different detector configurations and the different beam flux profile when weighted. The main change Run 1 and the other run periods in the detector configuration was the addition of the Barrel ECal. In addition, Run 3 and Run 4 MC samples are assume higher beam power which results in more simulated events per POT. The exact amount of MC files used for each run period and their corresponding POT is shown in Table 5 and Table **??** for the Water-in and Water-out respectively.

MC Configuration	Number of files	Protons On Target
Run 1	1,102	55.10×10^{19}
Run 2	1,483	75.15×10^{19}
Run 4	9,863	496.2×10^{19}
Total		625.4×10^{19}

Table 4: The available Water-in MC samples and their corresponding POT used for the Water-in analysis

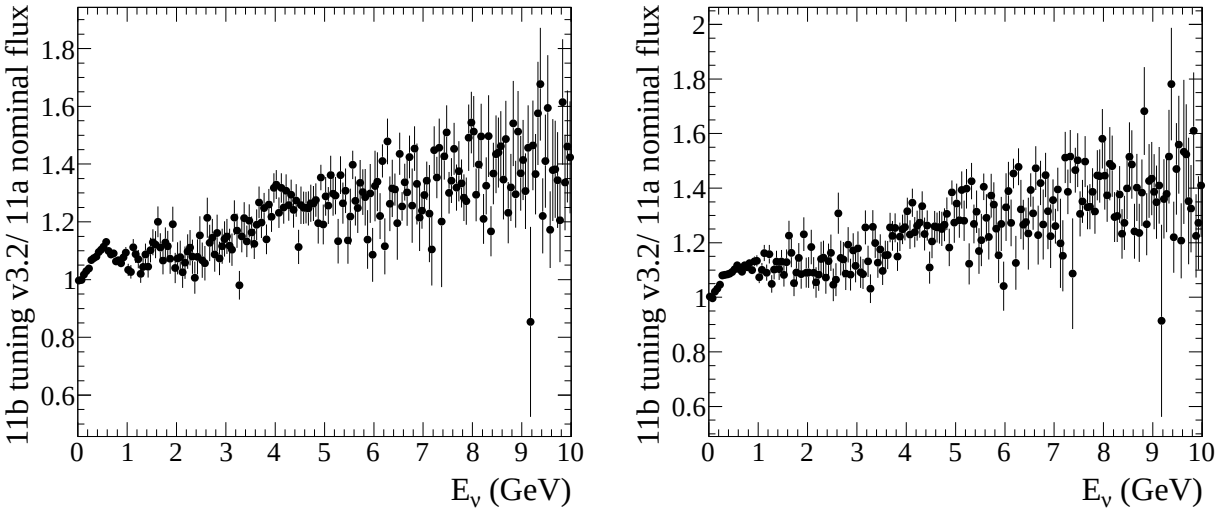
In the analysis we compared the different data run periods separately to the different MC samples generated by NEUT [9]. These generated samples consist of neutrino interactions created in the whole ND280 detector volume which includes the magnet region. NNN In addition we have used the Production 4 'Sand interaction' MC sample which contains 700 files that are equivalent to $\sim 7 \times 10^{19}$ POT and should include interactions that occur, in the surrounding pit and soil area.

MC Configuration	Number of files	Protons On Target
Run 2	1,991	99.55×10^{19}
Run 3	3,221	161.1×10^{19}
Total		260.6×10^{19}

Table 5: The available Water-in MC samples and their corresponding POT used for the Water-in analysis

3.3 Weighting by beam flux

The Production 5 MC files were generated by GENIE and NEUT, which depend on beam MC using the 11a (version 2) of Jnubeam. To account for external hadron production data from the NA61 experiment and other experiments, inputs (e.g. actual proton beam information) are incorporated by the Beam Group to reweight the neutrino beam flux as a function of neutrino energy. The Beam Group releases this information as a set of histograms that include the ratio of tuned to nominal flux histograms. We use the true neutrino energy of each MC event and the 'tuned to nominal' ratio histogram to reweight that event. The current analysis utilizes the 11b version 3.2 tuned flux ratios. These 'tuned to nominal' ratios as a function of true neutrino (E_ν) energy are shown in Fig. 1.



(a) Run 1 Ratio

(b) Run 2 Ratio

Figure 1: The 11b version 3.2 tuned to nominal flux ratios

4 Event Reconstruction

4.1 PØD and Tracker Reconstruction Algorithms

The PØD Reconstruction is a stand alone reconstruction algorithm that uses hits measured in the scintillator bars and combines them into tracks and showers. These PØD-only track objects are then passed on to the matching algorithm described later in this section. A detailed description of the PØD Reconstruction algorithm can be found in T2K-TN-072 [3].

The Tracker Reconstruction algorithm similarly reconstructs tracks passing through the Tracker region of ND280. The ND280 Tracker is consists of three TPCs and two Fine Grain Detectors (FGD) in between the TPCs. In General the Tracker the resonstruction algorithm uses TPC tracks in the YZ projection (the non-drift direction) as a seed to reconstruct tracks from hits in the FGDs. To reconstruce TPC tracks in the XZ projection (i.e. the drift direction), a T0 is extracted from either the matched FGD hits, the cathode or one of the surrounding subdetector such as the PØD, barrel ECal or PØD ECal. Greater detail on the construction of, and the algorithms responsible for reconstructing tracks passing through, the Tracker can also be found in T2K-TN-072 [3].

4.2 Tracker-to-PØD Matching Algorithm

The tracks match algorithm is base on the reconstruction output objects of both PØD and Tracker reconstruction packages. The algorithm pair a PØD track that ends in the downstream part of teh PØD with a Tracker track that start in the upstream part of TPC1 (i.e. most upstream part of the Tracker). In this section we describe the matching algorithm in more details.

4.2.1 Description of Algorithm

The matching procedure can be divided into two stages - in the first stage the outputs of PØD and Tracker reconstruction are scanned, while in the second stage a list of pair-candidates is generated for later purpose.

At the first stage the algorithm scan for candidate PØD and Tracker tracks. These tracks need to pass quality and position checks, in addition to the time window requirement of ± 100 ns from each other. This time cut is set to be able to reject tracks from different bunches on one hand, and on the other hand to take into account possible uncertainties or difference between sub-detectors calibrated time stamps.

The PØD track tests at this stage are:

- The track passes through the downstream part of the PØD.
  i.e. Last node Z position > -1016 mm.
- The track should be a 3D track.
  To determine this a check of the track position variance is done, each coordinate X_{var}, Y_{var} and Z_{var} should be smaller then the large value of 10^8 mm.

The Tracker track tests at this stage are:

- The track start position is at the upstream part of the Tracker.
  i.e. The first node Z position < -750 mm.

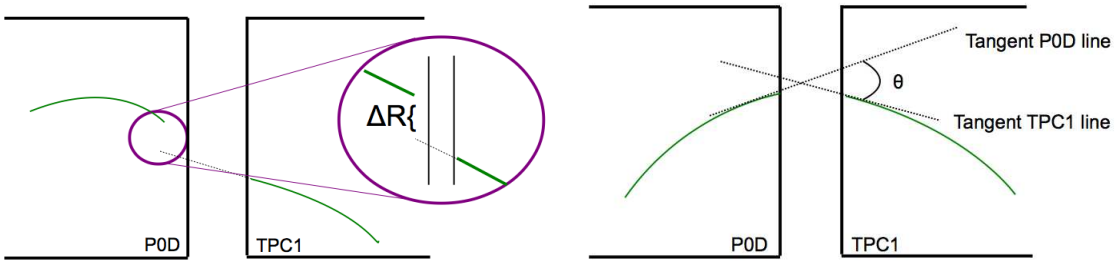


Figure 2: Diagrams of the ΔR and $\sin\Delta\theta$ matching parameters. Both parameters are extracted with the use of position and direction information of the PØD track last node and Tracker track first node. The ΔR ($\sin\Delta\theta$) parameter is illustrated on the left (right).

- The track should have more than 18 nodes.

In the second stage the algorithm identifies all possible pair-candidates. At first the algorithm creates a list of pair-candidates from all combinations of candidate PØD and Tracker tracks (outputs of the first stage). For each candidate pair the algorithm collects information for later usage (like - Bunch number, Tracker track change, PØD track start position) and calculates, in addition, the momentum of the pair-candidate at the start position of the PØD track (more details about this in Sec. 4.3) as will as the two matching parameters ΔR, $sin\Delta\theta$.
Figure 2 illustrates the above matching parameters ΔR and $sin\Delta\theta$. Where ΔR is the radial distance (in the XY plane) between the most downstream PØD track node and the Tracker track linear backwards projection to the same Z position as that last PØD track node, and $\Delta\theta$ (which we apply a sin on) is the angle between the PØD track last node direction and the Tracker track direction (see Figure 2 right side).

At the last part of this stage the algorithm sorts out pair-candidates which did not satisfy the matching cut conditions of $\Delta R < 0.76\ mm$ and $sin\Delta\theta < 0.86$ (more details on how the matching values have be obtained are laid out in Sec 4.2.2).

In addition the algorithm makes sure that any PØD or Tracker track candidate (output of the first stage) is used only once, i.e. appears only in a pair-candidate which has the smallest ΔR value.

4.2.2 Opimization of Matching Parameters

To determine the cut values used by the matching algorithm described above, we used the different MC samples. A figure of merit was calculated for a range of cut values of ΔR and $\sin\Delta\theta$ and the pair of values that yields the highest figure of merit is used for the matching algorithm. We chose the figure of merit $F = \frac{S}{\delta S}$, where S is the number of signal events and δS is the error, to minimize the error on the signal.

We begin with the same matching algorithm described above upto the final ΔR and $\sin\Delta\theta$ cuts, which are not applied.

The signal (S) is defined as any matched track where both the reconstructed PØD and Tracker track pieces originated from the same true primary trajectory (a primary trajectory is a true trajectory leaving the interaction vertex). The background (B) is the remainder of the matched tracks that do not satisfy the signal requirement. The errors δS and δB are the poissonian error on S and B respectively. For each pair of cut values (X, Y), we count up the total matched tracks that satisfy both $\Delta R <$ X and $\sin\Delta\theta <$ Y. Each matched track is then classified as either signal or background,

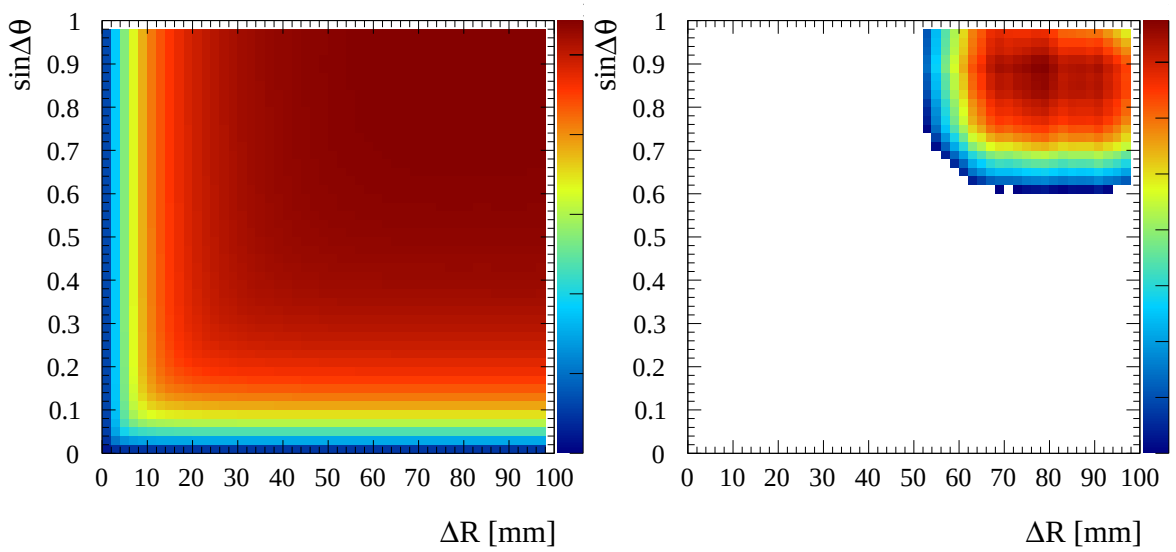


Figure 3: The Figure of Merit as a function of ΔR and $\sin\Delta\theta$ for candidate matched tracks from Water-in MC samples (Run 2). the absolute figure of merit does not vary much in the region of interest. The Z-axis zoomed version (right) shows the location of the local maxima (76 mm, 0.88).

which yields S and B. Finally we calculate the figure of merit which, using Poisson error, reduces to ($F = \frac{S^2+2B^2}{M^2}$). Figures 3 and 4 shows the figure of merit as a function of the cut values for Watre-in (Run 2) and Water-out (Run3) MC samples respectively. We find local maximas at (76 mm, 0.88) and (78 mm, 0.90) for Water0in (Run2) and Water-out (Run 3) respectively. For simplicity we use the values of $\Delta R \leq 76$ mm and $\sin\Delta\theta \leq 0.88$ for the different samples and run periods.

We also note that the local maxima is not 'sharp'. A variation in the cut values does not drastically alter the overall figure of merit. Also, as the cuts are placed in generally flat regions of the ΔR and $\sin\Delta\theta$ distributions, cut value variations do not yield large changes in total tracks reconstructed.

Finally, as Data and MC distributions of the matching parameters show excellent agreement near the cut values, variations will not affect the Data/MC ratios. These points are demonstrated in Figures 5 and 6, which shows the overall pre-cut ΔR and $\sin\Delta\theta$ distributions of matched track candidates in Data and MC. Systematic effects resulting from this matching algorithm are evaluated later in the efficiency systematic section.

4.3 Energy correction in the PØD

Our analysis selects CC inclusive events by tagging the muon candidate. To reconstruct the momentum of the muon candidate, we use the measured TPC1 momentum in conjunction with the length of the track in the PØD. Using range tables for a muon and our known material density in the PØD we can calculate the momentum of the muon at the vertex by adding in the expected energy lost. In order for the energy loss correction to be fast, we chose not to use $\frac{dE}{dx}$ tables and incrementally adjust the energy. The PØD energy loss was calculated using range data for the various superPØDules. Range is defined as $R(E') = \int_0^{E'} \left(\frac{dE}{dx}\right)^{-1} dE$, where $\frac{dE}{dx}$ is the weighted average over all materials, and is calculated for each superPØDule. This gave us a table of range vs energy/momentum for every part of the PØD and essentially allowed us to perform the integration once as opposed to for every track. To obtain the energy lost in the PØD, we simply take the TPC1 momentum, find its

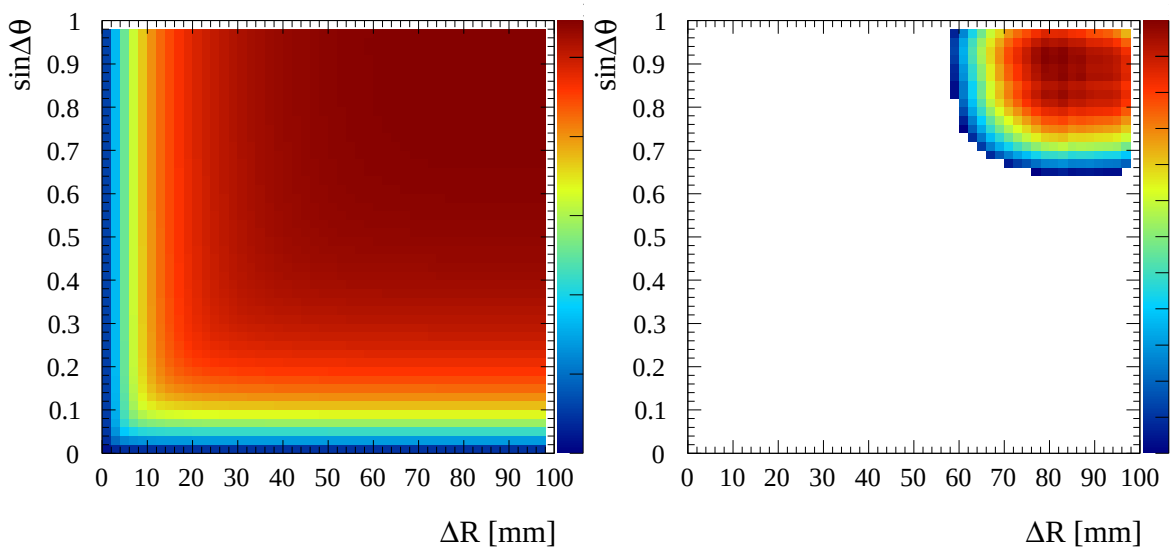


Figure 4: The Figure of Merit as a function of ΔR and $\sin\Delta\theta$ for candidate matched tracks from Water-out MC samples (Run 3). The Z-axis unzoomed version (left) demonstrates that the absolute figure of merit does not vary much in the region of interest. The Z-axis zoomed version (right) shows the location of the local maxima (78 mm, 0.90).

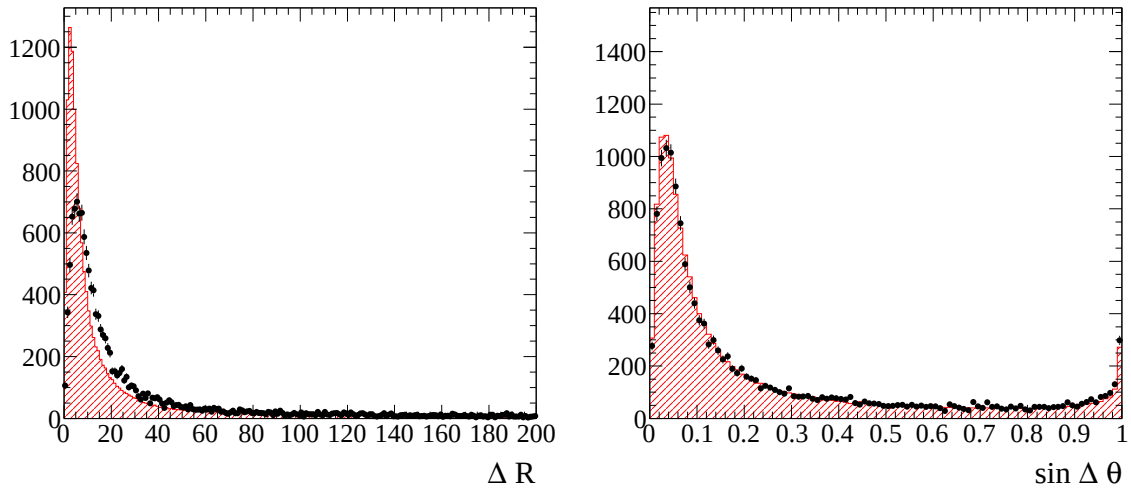


Figure 5: The ΔR (left) and $\sin\Delta\theta$ (right) distributions for Water-in samples (Run 2). The Data (black) is overlayed on MC (red). To be able to compare the MC to the Data, which includes sand interactions a cut on the start of the tracks was apply, i.e. Z > -3183 mm, $|X|$ and $|Y| < 1000$ mm.

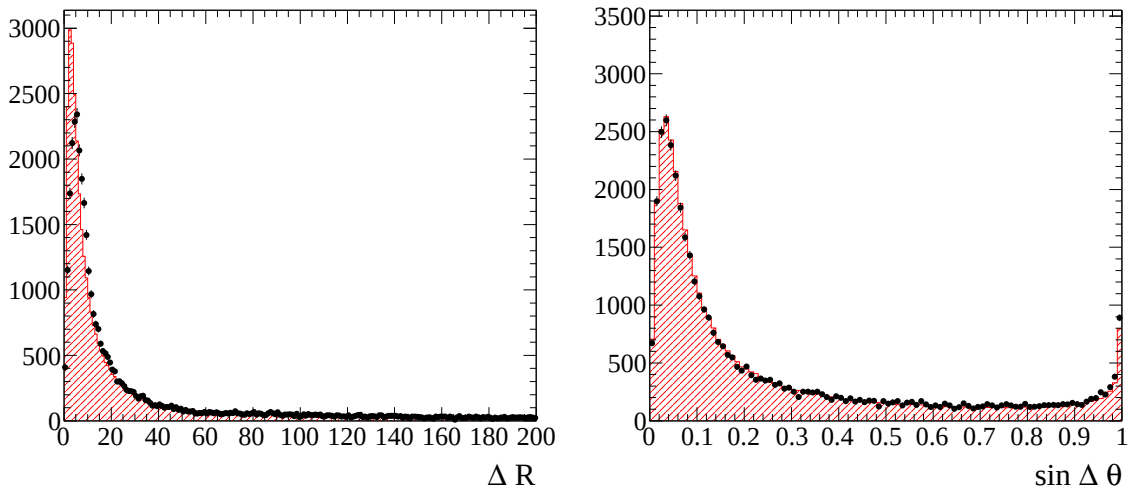


Figure 6: The ΔR (left) and $\sin\Delta\theta$ (right) distributions for Water-out samples (Run 3). The Data (black) is overlayed on MC (red). To be able to compare the MC to the Data, which includes sand interactions a cut on the start of the tracks was apply, i.e. Z > -3183 mm, $|X|$ and $|Y| < 1000$ mm.

corresponding range, add the length of the track in next superPØDule, and find the corresponding momentum (see Figure 7). We repeat this as necessary depending on whether the track traversed through more than one superPØDule.

Figure 8 shows the residuals of the calculated muon momentum. For comparison, we also provide the muon momentum residual as measured by another independent reconstruction package called 'Global Reconstruction'. The residual distribution from the outlined method above shows no bias and a marked improvement over the momentum calculation of global reconstruction. We use only the above method to calculate muon momentum in our analysis.

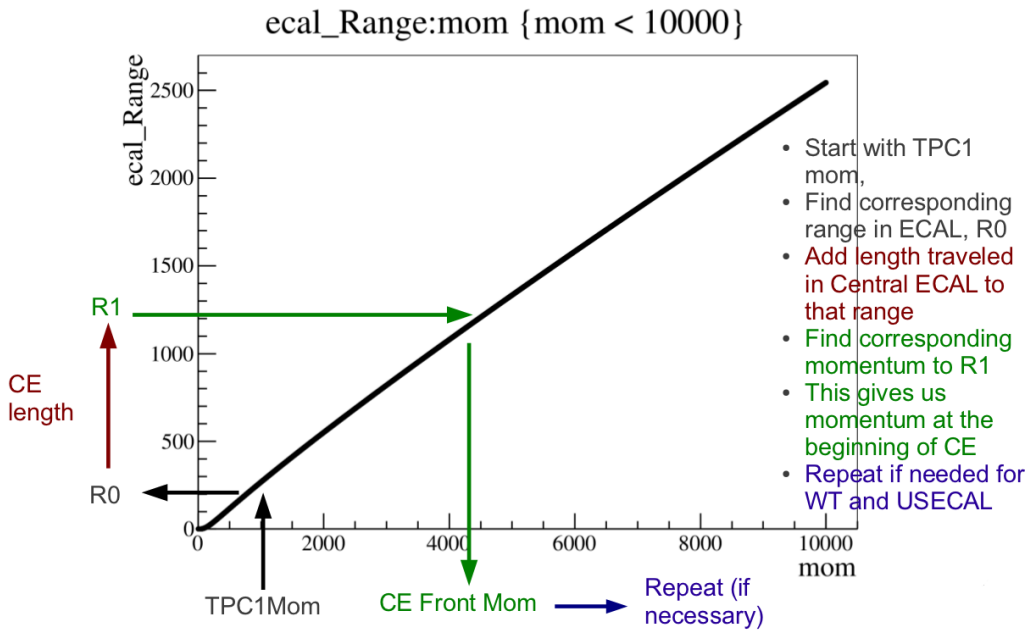


Figure 7: Calculation of PØD momentum correction.

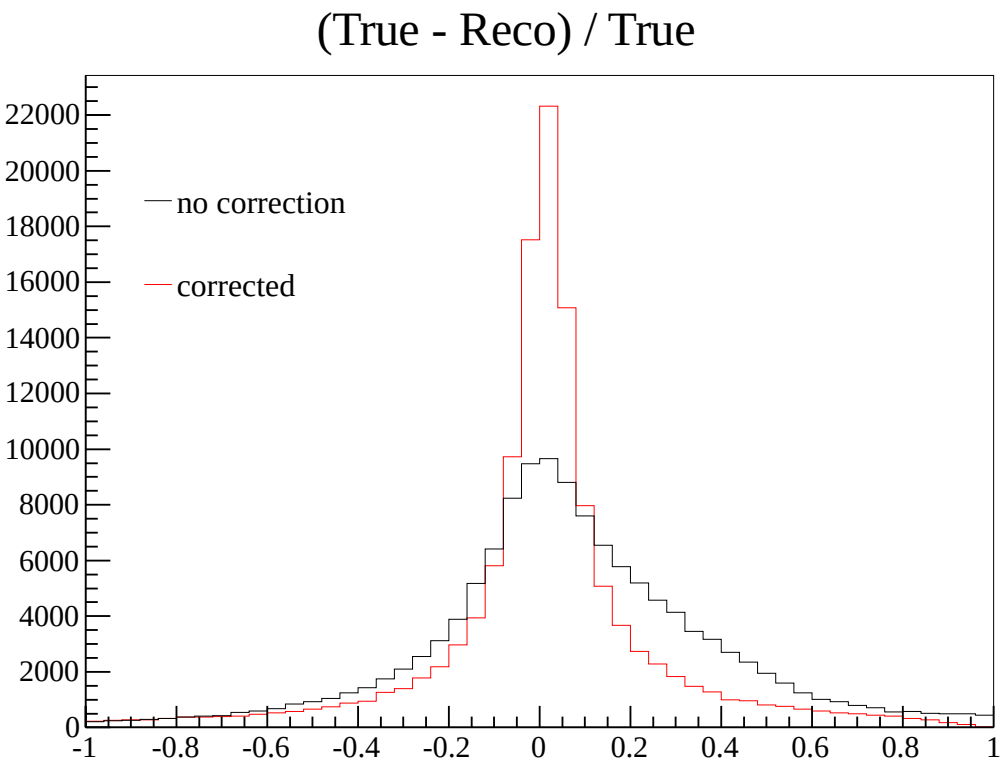


Figure 8: Comparison of the global reconstruction PØD momenta (black) vs manually corrected PØD momenta (red). Note that the manually corrected momenta are closer to the true momenta and improved compared to global reconstruction.

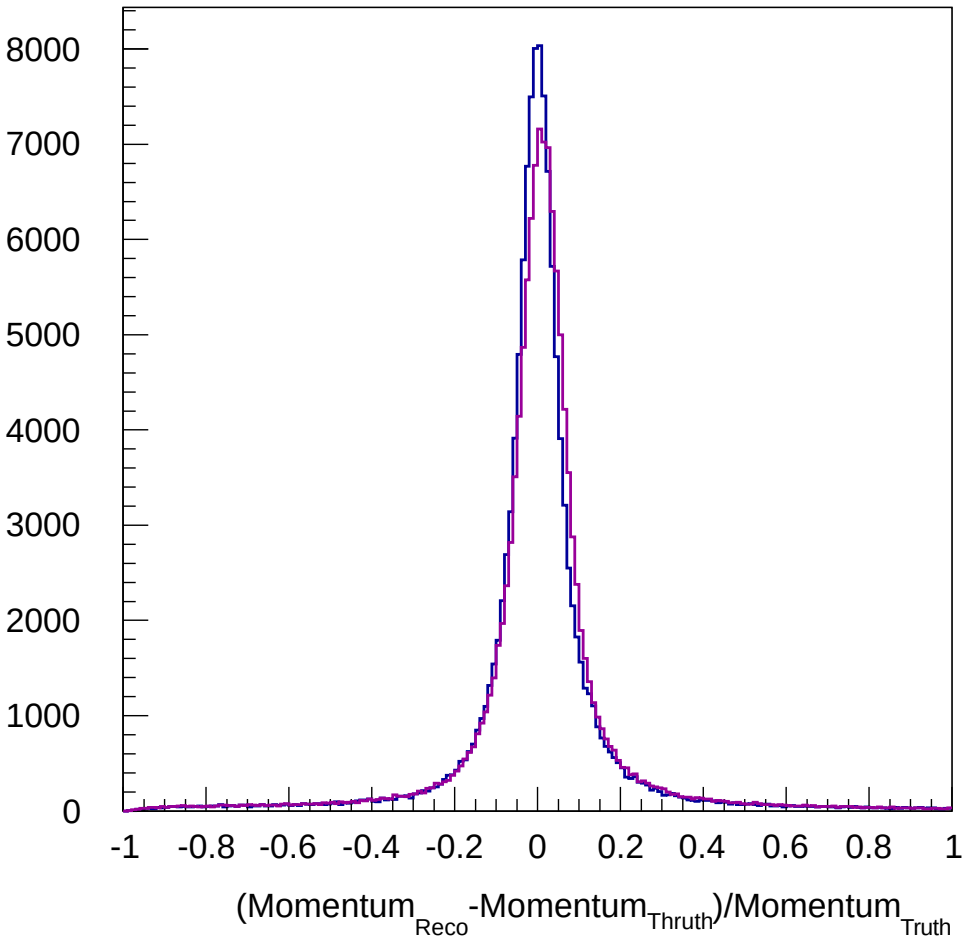


Figure 9: Comparison of the momenta - Reconstruction (Reco) vs. and Trajectory (Truth). The Run 2 (Run 3), Water-in (Water-out), period is represented by the Blue (Purpule) curve.

5 Inclusive Charged Current Event Analysis

The muon neutrino beam produced at J-PARC passes through the off-axis ND280 detector. These neutrinos can interact via Charged Current (CC) or Neutral Current with the detector material. The CC interactions will produce a μ^- track which will contain most of the time the majority of the incident neutrino momentum.
The analyses strategy is to search for the highest momentum negative track in an event and to tag it as the muom candidate track and the while event as a CC interaction one. We note that the beam spill delivered by the J-PARC facility to the experiment has a sub-structure of bunches which are separated by beam dead-time. All analyses described in this document were done at the beam bunch level instead of the ND280-DAQ event (beam spill level).

5.1 Event Selection

The CC inclusive selection procedure includes the main levels of information available (accelerator/near-complex level, track level, bunch level) which are listed below:

- Accelerator/Near-Complex level selection:
  A spill is use for analysis only if the below quality checks are satisfied
  - The beam/accelerator flag 'GoodSpillFlag' is set to the value one
  - The near detector-complex Data-Quality flag 'ND280OffFlag' is equals to zero
- Track level selection:
  The Tracker-to-PØD matching algorithm, which is outlined in Sec. 4.2, is applied.

  The algorithm inputs are the available PØD and Tracker tracks in a spill and its outputs are a set of tracks that originate in the PØD [3] and continued to pass through TPC1/Tracker sub-detector.

  To note is that the algorithm will match uniquely any PØD/Tracker track with the nearest Tracker/PØD track (by ΔR, see Sec. 4.2).
- Bunch level selection:
  All the output tracks from the previous stage are sorted into predefined bunch time windows (See Table 7).

  Then from each active bunch we identify the muon candidate track and tag that event as a CC interaction.

  A track is considered as the muon candidate if it satisfied all the requirements below.
  - The candidate track charge is negative
    (by the matched Tracker track charge property).
  - The candidate track begins in the PØD volume
    ($-988\,mm < X < 910\,mm$, $-1020\,mm < X < 1010\,mm$, $-3139\,mm < X < -900\,mm$).
    This condition is required to reject external sources tracks (cosmic, magnet or sand interactions) while keeping the largest volume of the PØD detector as a target.
  - The candidate track momentum at the start of the track is the highest in that bunch.

  Finlay, only the muon candidate tracks which originated from inside the PØD Water-Target FV (see Sec. 5.1.1) are accounted for the final results.

[3] These will include tracks that started at the edges of the PØD i.e. passing through tracks.

5.1.1 The PØD Water Target Fiducial Volume

The Water-Target (WT) Fiducial Volume (FV) cut used in this document is based on the FV description given in T2K-TN-073. The choice of the FV in the X and Y directions is made to minimize the uncertainty of the amount of water in the PØD. In the Z direction, the FV corresponds to the water target region only. This also minimizes the physics uncertainty arising from the lead radiator sheets in the ECal super-PØDules that are on the upstream and downstream ends of the water target. The FV values utilized in this note are summarized in Table 6.

Minimum				Maximum
-836	$<$	X	$<$	764
-871	$<$	Y	$<$	869
-2969	$<$	Z	$<$	-1264

Table 6: The PØD Water-Target Fiducial Volume used in the analyses, taken from Ref. [2]

5.1.2 Bunch Time Windows

The beam spill substructure is not the same for the different beam run periods, we have studied the track timing for Run 1-4 for both Data and MC samples. These studies have revealed a number of points that had to be addressed by our bunch time window determination and are listed here.

- Run 1 has six bunches in each beam spill while Run 2 - 4 have eight bunches in each beam spill.
- Run 2 includes a bunch time shift that was introduced at January 2011 (see Fig. 10).
- The MC samples have the same six/eight bunch structure as the Data samples but their bunches (means) are shifted on average $\sim$250 ns from the data ones.

In Fig. 10, which shows the track time stamps as a function of their subrun in Data, one can observe some of these effects mentioned above. A close examination of the figure shows both the six and the eight bunch structures and the time shift introduced at Jan 2011. A vertical dashed line was added to help the reader to help see this relative shift.

To account for the above conditions, we have decided to define the same set of bunch time windows for all runs periods (regardless of the run configuration). These windows were enlarged to include both the difference between Data and MC bunch means and the time shift in Run 2. Table 7 summarizes the bunch time windows used in the analyses.

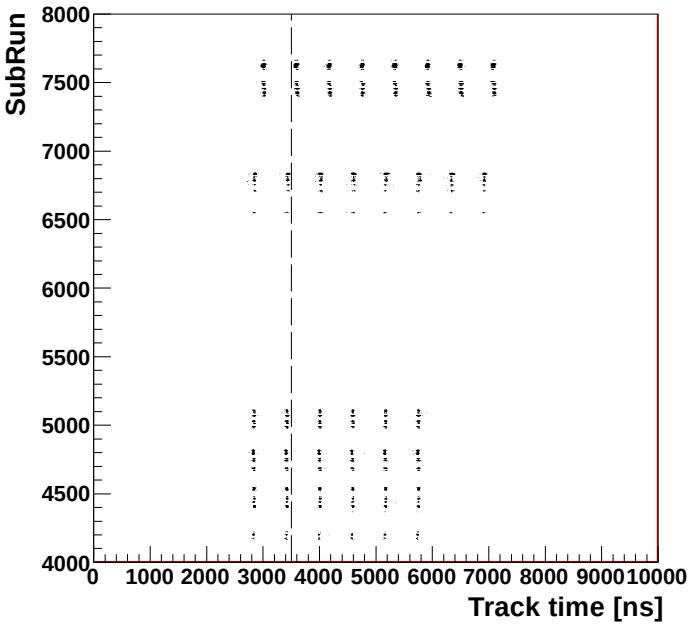


Figure 10: Track time stamp as a function of subrun in Data (for Run 1 and Run 2 run periods). The vertical dash line was added to help point out the time shift that was introduced at Jan 2011.

Bunch Number	Time window [ns]
1	2660 - 3100
2	3240 - 3680
3	3820 - 4260
4	4400 - 4840
5	4980 - 5420
6	5580 - 6000
7	6140 - 6580
8	6720 - 7160

Table 7: The bunch timing windows used in the analyses for all Run periods.

5.2 Distributions and Selection Results

In this section we present comparisons between data and Monte Carlo for timing distributions, vertex distributions, momenta, θ, and ϕ. The Monte Carlo was reweighted to the 11b version 3.2 flux. The stacked color histograms in the following plots represent the MC distributions separated into different classifications. The color scheme used is shown in Figure 11. Note that Quasi-elastic (QE), Single pion (Pi), Multi Pion (NPi), Meson, and Deep Inelastic Scattering (DIS) are all charged current interactions and are considered as signal in this analysis. Neutral Current (NC), parent Anti-Neutrino (antiNu), Other and Out of Fiducial Volume (outFV) interactions are all background.

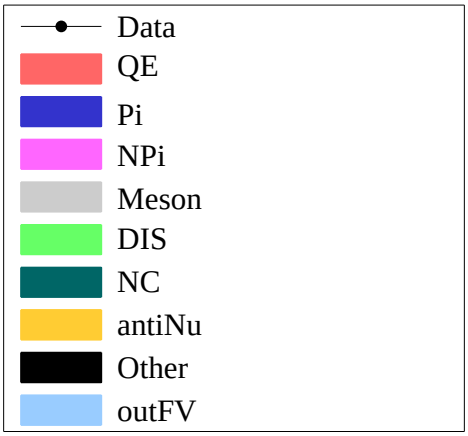


Figure 11: The color scheme used for many of the Monte Carlo plots in the document. They are separated via NEUT reaction codes, parent neutrino flavor and vertex location. Refer to the text for further detail.

5.2.1 Timing distributions

The distribution of 'Water-in' track vertex time in Run 1, Run 2 and Run 4 is shown in Figures 12. We note that there is a discrepancy between the MC track time and the Data track time. Specifically, the peaks of MC and Data are shifted within the time window. These effects are not however much of an issue in our analyses as we have very lenient timing cuts. The relative position of peaks in a given time window has a negligible effect when we integrate over the entire window. Any effects at the edge of the timing cuts will be accounted for in our timing systematic.

5.2.2 Vertex distributions

We define the start position of the muon candidate track as the vertex. This start position is determined by PØD reconstruction and corresponds to the first PØDule to have an above-noise energy deposition from the charged track. For the Z-direction, this implies that the position can only be resolved up to the thickness of an individual PØDule. The Z Vertex distributions have thus been binned by PØDule thickness, a value that varies over the different superPØDules. In the X and Y directions, though we can only get energy deposition in the discrete volumes of scintillator bars, the triangular shape of the bars allows most charged tracks to deposit energy into neighboring bars.

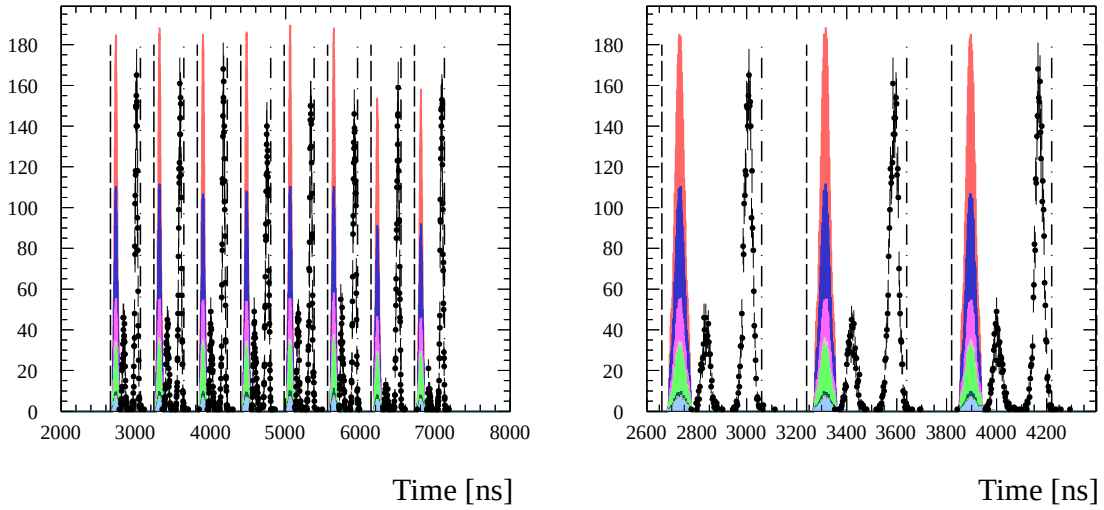


Figure 12: Time distribution of track vertices in Run 1 + Run 2 + Run 4. The MC is shown in color and the data in black crosses. The left plot shows all the beam bunches and the right plot show a zoom in of the first three beam bunches.

The locations of these two hits and the amount of energy deposited at each can be averaged to give a more accurate measurement of where the track passed through the scintillator layer. Thus, the X and Y vertex positions can be resolved to a much higher precision than the Z position.

The X, Y and Z vertex positions for the Run 1, Run 2 and Run 4 'Water-in' periods are shown in Figures 13 and 15. We note that the number of CC inclusive interactions increases as we go further downstream in the PØD. This is an acceptance effect and is expected. The muon track candidate must enter the TPC. Tracks originating further upstream lose energy as they traverse the PØD and so have a lower chance of making it all the way to the TPC. Therefore we see the shape shown in the Z vertex distributions.

5.2.3 Kinematic Distributions

We show here Data and Monte Carlo distributions of momentum, θ and ϕ. The momentum distributions of the muon candidate tracks are shown in Fig. 18 for Run 1, Run 2 and Run 4 'Water-in' periods. The markers with error bars represent data with statistical errors and the colored histogram stack shows the Monte Carlo separated into different interaction categories. There is a qualitative agreement in the shape of the overall momentum distributions between Data and Monte Carlo for Run1 + Run2. The majority of these background is in the lower momentum region. We also show the θ and ϕ distributions of Data and Monte Carlo for both running periods. Figure 19 indicates that most of our tracks are forward going, significantly more so than the FGD selections. Once again, this acceptance effect is both expected due to the TPC1 segment requirement. Figure 19 shows the ϕ distribution for the combined running periods, and we note good agreement.

5.2.4 Results

The result is reported as an integrated Data/MC ratio of the total selected CC inclusive interactions in Run 1, Run 2 and Run 4 'Water-in' combined. The total events selected in Data is 25,791 which corresponds to an exposure of 23.6×10^{19} POT. The unnormalized number of events selected in MC

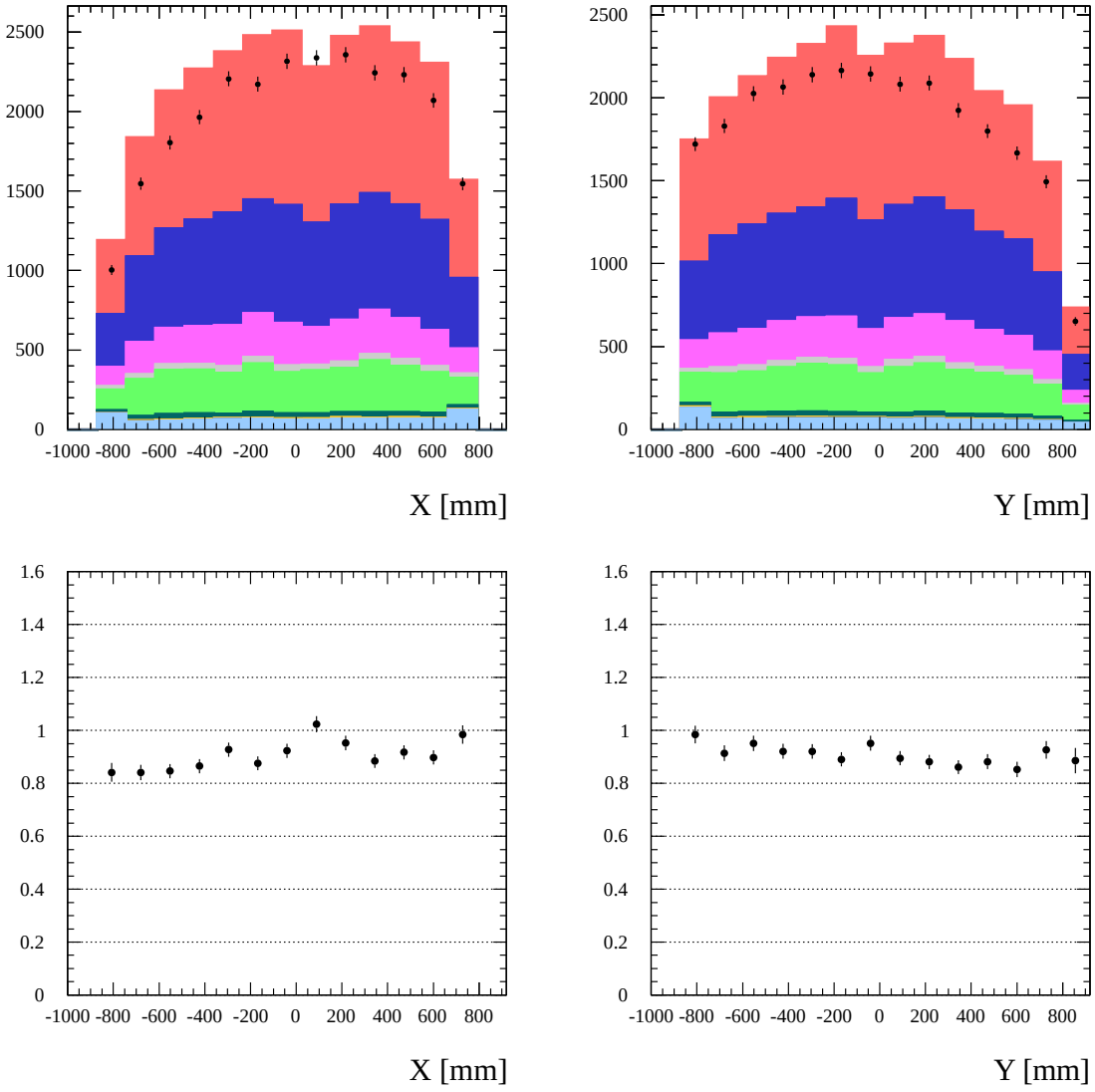


Figure 13: The start position of the muon candidate track for Run 1, Run 2 and Run 4 'water-in' period. The X position (Y position) is shown to the left (right). Below is the corresponding Data to MC ratio.

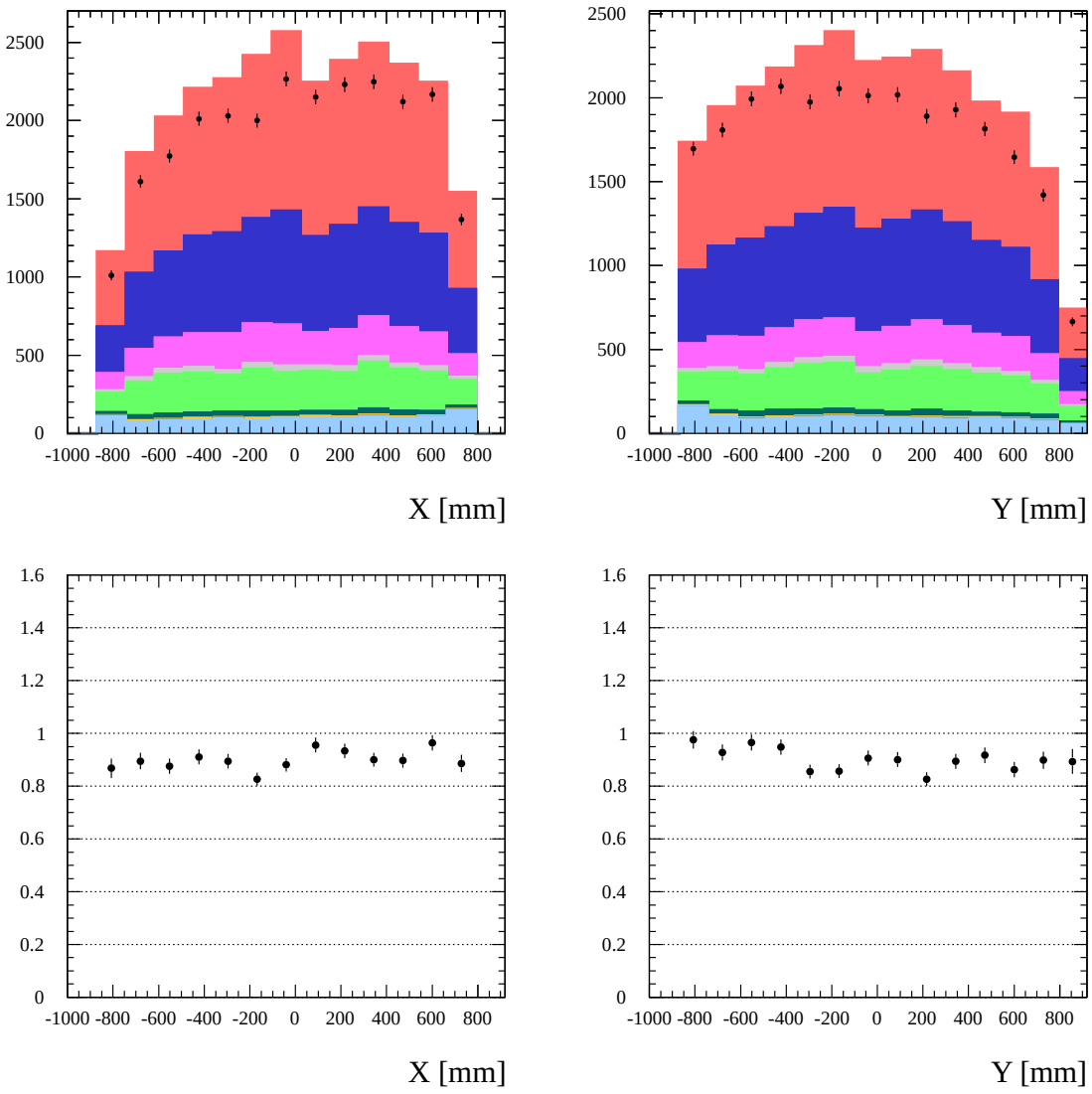


Figure 14: The start position of the muon candidate track for Run 2, Run 3 and Run 4 'Water-out' period. The X position (Y position) is shown to the left (right). Below is the corresponding Data to MC ratio.

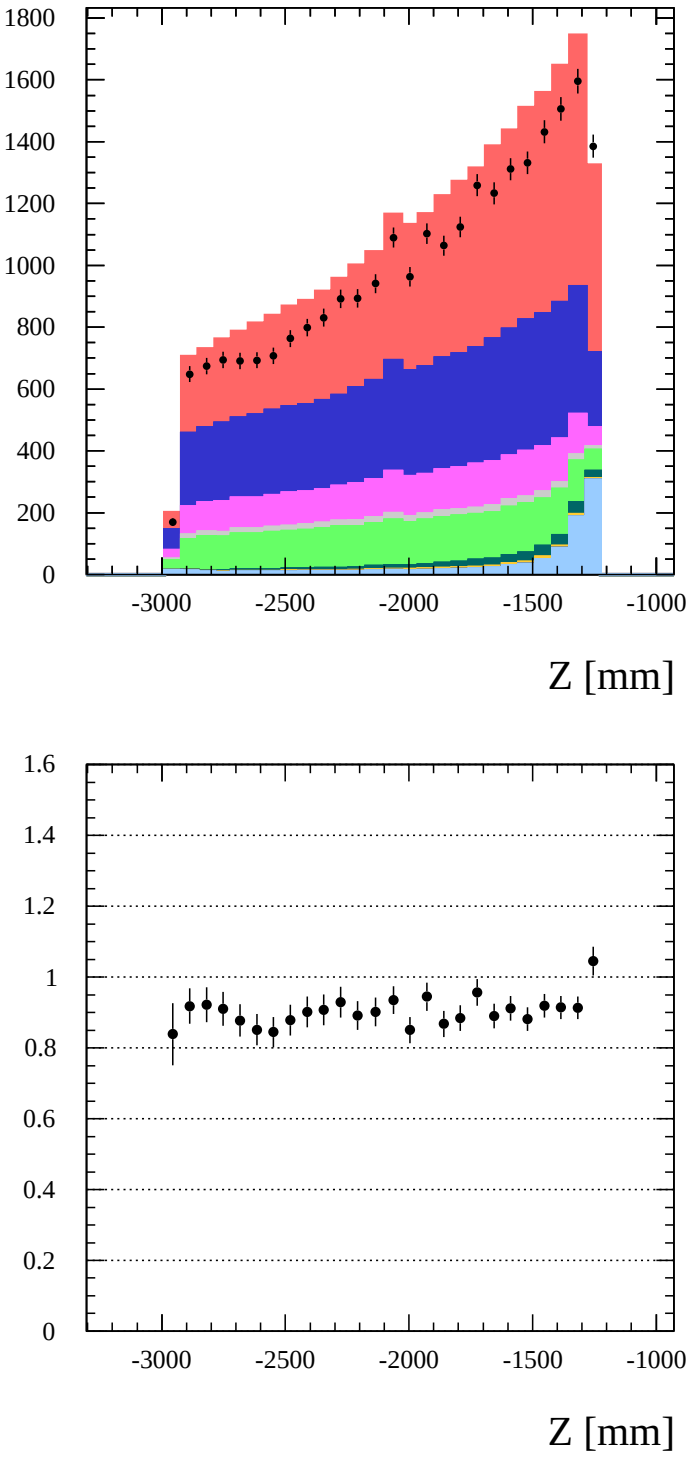


Figure 15: The Z start position of the muon candidate track for Run 1, Run 2 and Run 4 'water-in' periods. The varying bin widths equal the thickness of a p0dule in that region. Below is the corresponding Data to MC ratio.

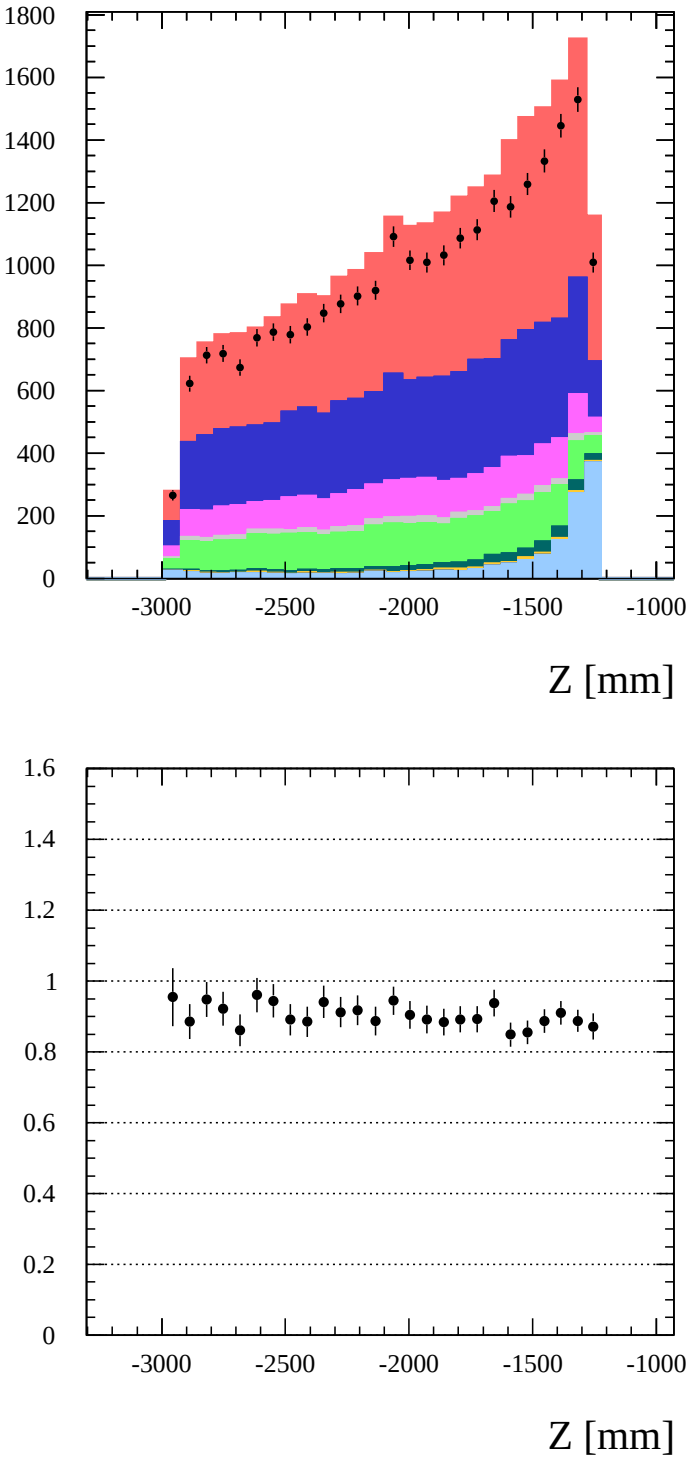


Figure 16: The Z start position of the muon candidate track for Run 2, Run 3 and Run 4 'Water-out' periods. The varying bin widths equal the thickness of a p0dule in that region. Below is the corresponding Data to MC ratio.

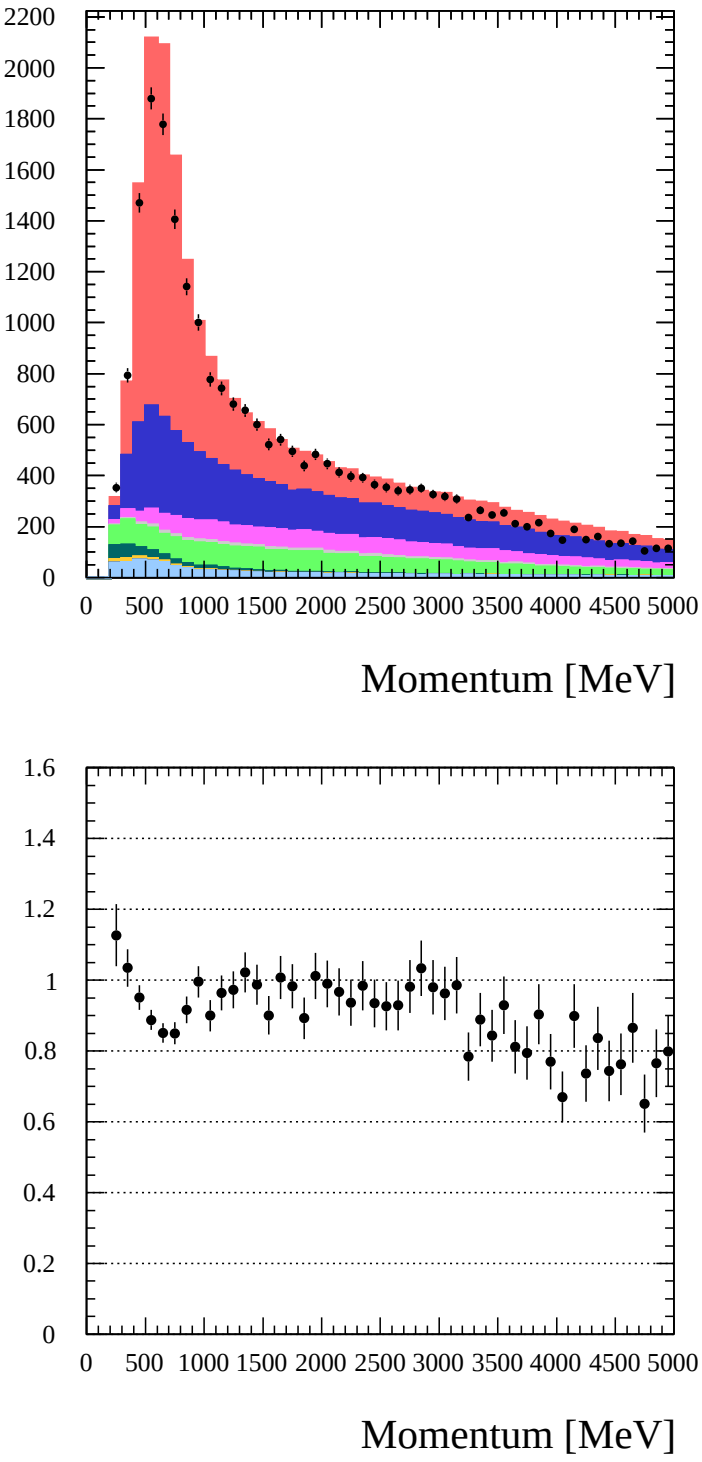


Figure 17: The momentum at the vertex of the muon candidate for Run 1, Run 2 and Run 4 'Water-in' peridos. Below is the corresponding Data to MC ratio.

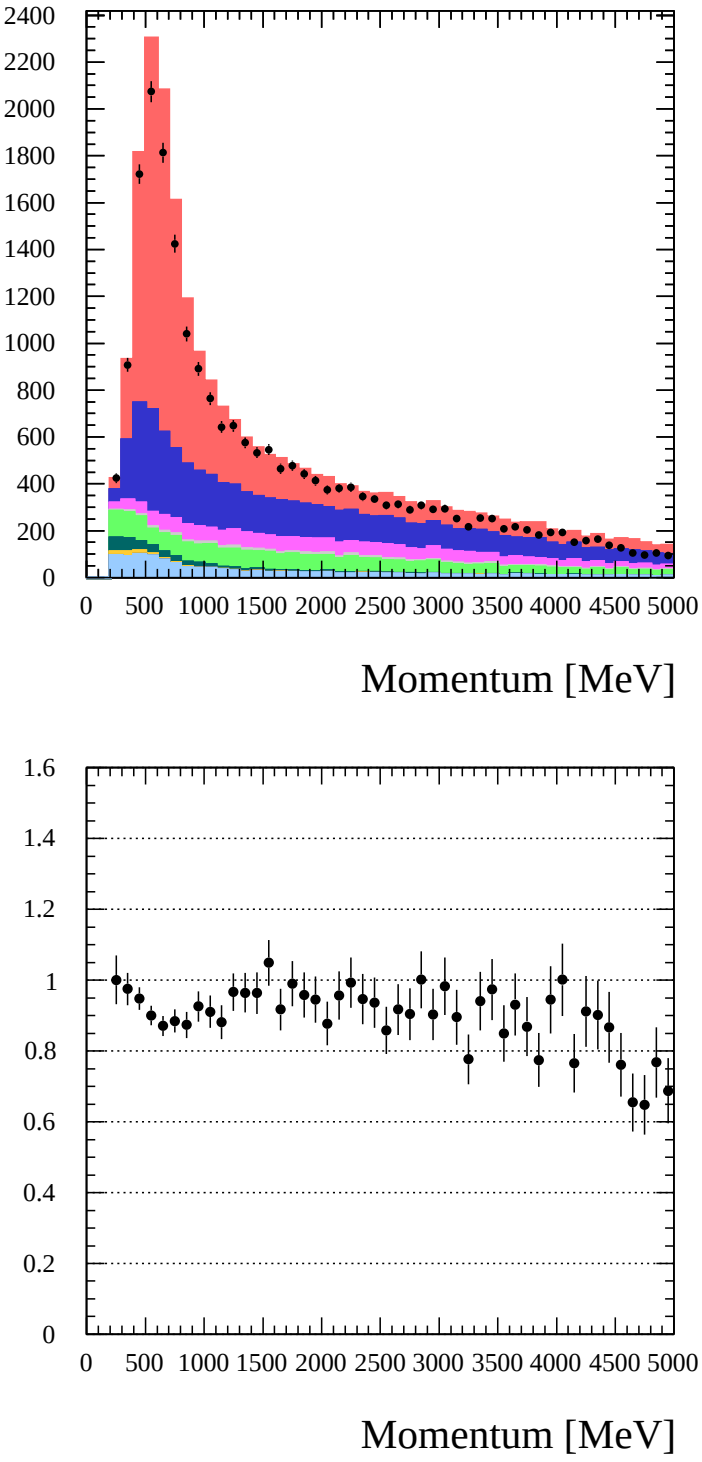


Figure 18: The momentum at the vertex of the muon candidate for Run 2, Run 3 and Run 4 'Water-out' peridos. Below is the corresponding Data to MC ratio.

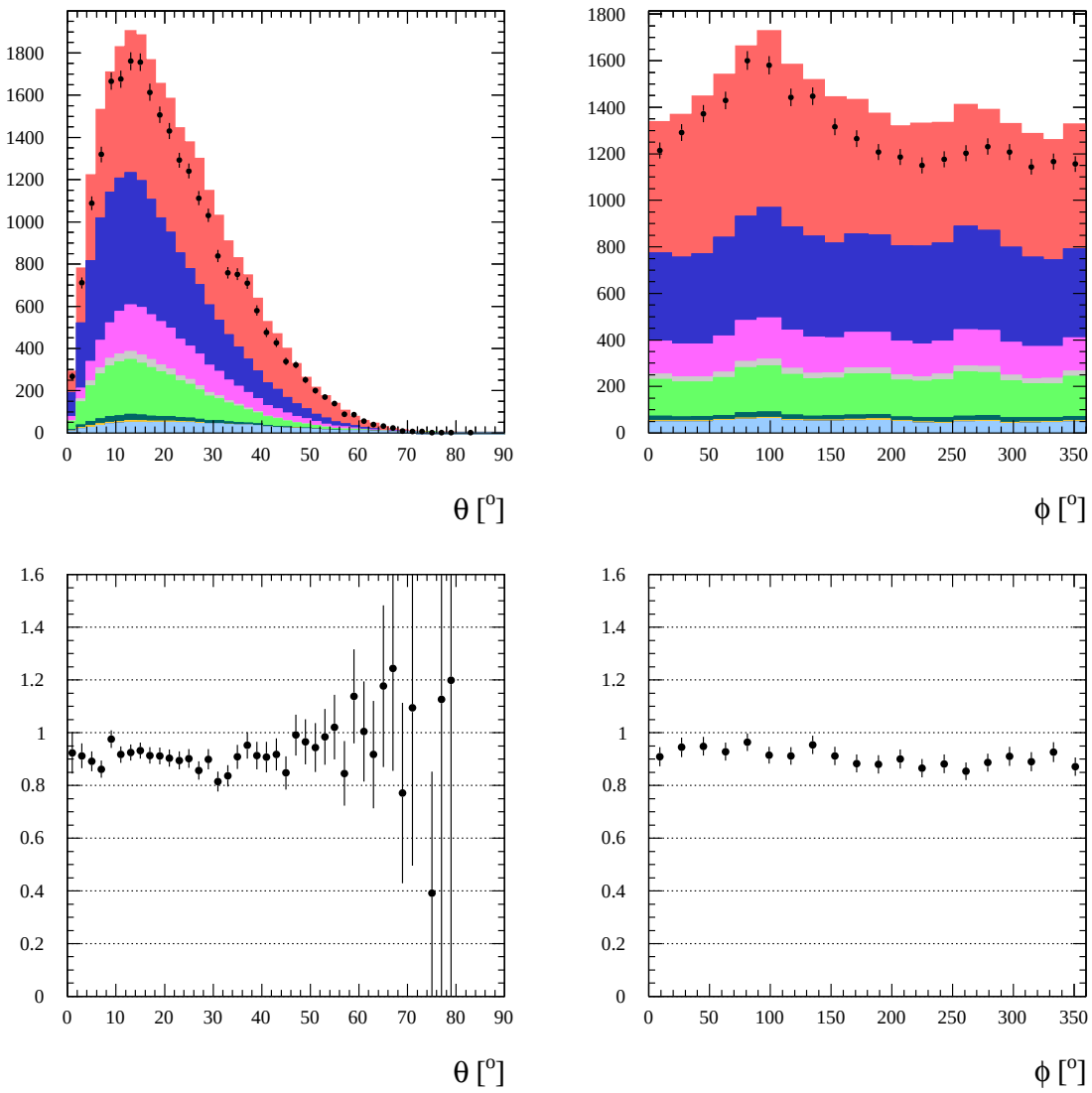


Figure 19: Theta (left) and Phi (right) of the initial direction of the muon candidate track in Run 1, Run 2 and Run 4 'Water-in' periods. Below is the corresponding Data to MC ratio with statistical errors.

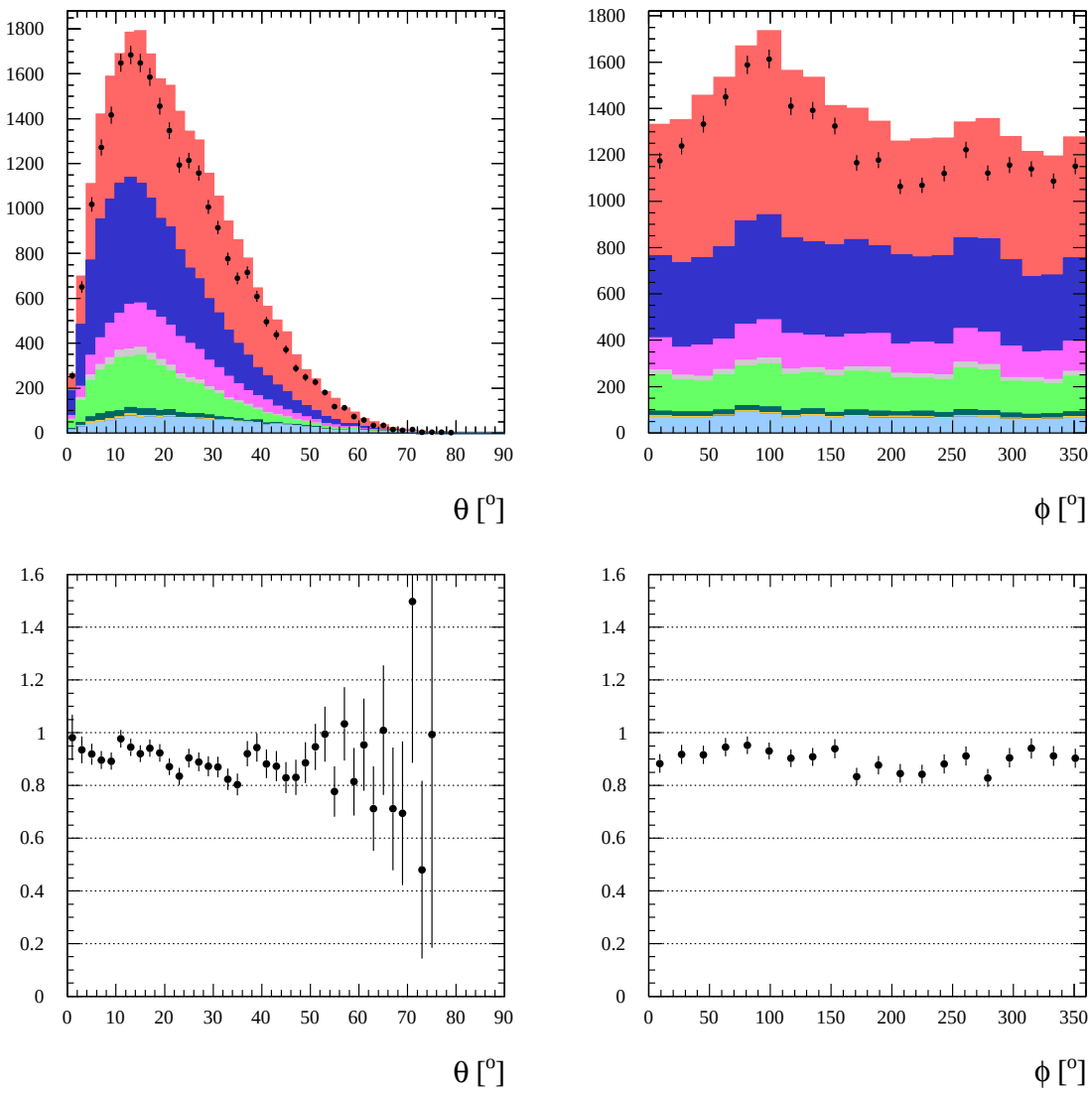


Figure 20: Theta (left) and Phi (right) of the initial direction of the muon candidate track in Run 2, Run 3 and Run 4 'Water-out' periods. Below is the corresponding Data to MC ratio with statistical errors.

is 1,031,295 for 626×10^{19} POT. After normalizing MC by the Data POT and applying the fiducial mass correction (to each of the run periods as detailed in TN-073), we have yielded the Data/MC ratio of 0.879 ± 0.006.

For completeness we summarized in Table 9 the number of events after all normalization and corrections for the different MC channels that we selected.

MC Channel	Number of events selected
QE	11,895
Pi	8,294
NPi	3,008
Meason	480
DIS	3,329
NC	381
$\bar{\nu}$	98
Other	7
Out of FV	916

Table 8: The number of events selected in the CC inclusive MC Run 1, Run 2 and Run 4 for the 'Water-in' periods combined sample, presented for the different MC channels.

MC Channel	Number of events selected
QE	11,956
Pi	7,700
NPi	2,745
Meason	440
DIS	3,078
NC	427
$\bar{\nu}$	110
Other	7
Out of FV	1,290

Table 9: The number of events selected in the CC inclusive MC Run 2, Run 3 and Run 4 for the 'Water-out' periods combined sample, presented for the different MC channels.

6 Evaluation of Systematic Errors in the CC Inclusive Selection

We investigate several sources of systematic error in our analysis. These are:

1. P0D Tracking and Matching Efficiency
2. Hit Reconstruction Efficiency
3. Fiducial Mass
4. Track Timing
5. Fiducial Volume
6. Cosmics
7. Event Pile-up per Bunch
8. Out of PØD Fiducial Volume Background
9. TPC1 track reconstruction efficiency
10. Charge Mis-ID

6.1 P0D Tracking and Matching Efficiency

The efficiency of reconstruction, and the differences between Monte Carlo and data efficiencies, are the first source of systematic uncertainty studied for the CC Inclusive selection. There are several steps involved in reconstructing the final candidate muon track as described in previous sections. In this section, we use a 'FGD Cosmic' sample to investigate the reconstruction and matching efficiencies of our analysis. Cosmics provide a sample independent of the physics we are probing, and are an excellent unbiased sample to use to evaluate the efficiency systematic. However, as Cosmics do not have the same bunched timing structure as our beam events, we also used as independent sample of sand muons from Production 4 to cross check our results. The results of the cross check are summarized in Section D. Efficiencies extracted from Production 4 cosmics and sand muons and Production 5 cosmics agree well. Sand Muons are generated by beam neutrinos interacting with material outside the off-axis near detector and so have a timing strucutre that mimics that of beam events. These two 'sideband' samples together provide a robust method for evaluating the matching and reconstruction efficiency systematic. Finally, we also performed a hand-scan of beam events in Production 4 reconstruction files to identify and quantify any reconstruction issues missed with the FGD cosmics and sand muon samples.

We use a simple method to examine the efficiency of our matching algorithm for tracks that deposit energy in the PØD, enter the Tracker and are reconstructed by the Tracker reconstruction package. We simultaneously determine how often P0D reconstruction successfully fits a track and how often we succesfully match this track with a Tracker reconstructed track. First, we pre-select a sample of quality tracks reconstructed in TPC1 that point squarely into the PØD. We then attempt to pair each of these tracks to a PØD reconstructed track. The ratio of the number of matched pairs found to the number of quality TPC1 tracks yields the 'matching efficiency'. Though this estimate is not necessarily the absolute efficiency, given that monte carlo simulates data well to the

first order, the observed difference between the monte carlo efficiency and the data efficiency gives us the systematic due to reconstruction and matching.

This definition folds in the PØD's intrinsic, lower level reconstruction efficiencies (i.e. hit finding, PØD-only tracking, etc.) as well as matching and recombination efficiencies from the matching algorithm. However, as we use reconstructed tracks in the ND280 Tracker as a baseline, we do not account for the Tracker tracking efficiency with this strategy. We determine the systematic uncertainty from the Tracker efficiency separately in a later section.

6.1.1 Cosmics Sample

Using the 'FGD Cosmics' sample, we extract the matching efficiency as defined above. The following determine the denominator and numerator. Note that for the numerator of the efficiency, the cuts closely mimic those used in Sections 4 and 5 to select muon candidates. Also, the pre-selection cuts are slightly different between the Cosmics sample as opposed to the Sand Muon sample. The necessity for this difference is discussed later.

Pre-Selection Cuts (Denominator):

1. There must be a Tracker reconstructed track in the event
2. The Tracker track must be reconstructed as beginning at the upstream face of the first TPC (Z < -750 mm)
3. The TPC must measure a momentum of at least 250 MeV
4. Project the Tracker track linearly backwards into the PØD. The projection is made to the Z = -1100 mm plane, and then just the XY fiducial cut is applied to the projected point.
5. The Tracker track has > 18 reconstruction nodes
6. The Tracker track has a 'corrected time' stamp between -4800ns and -4400ns. The time correction allows us to find the tracker time in relation to the PØD electronics. The time cut is placed 80ns from the edge of the 480ns-wide PØD integration window. This is the window where the PØD electronics are capable of properly reconstructing hits. For a more detailed discussion of the timing cut, please refer to Appendix section **??**

Matching Cuts (Numerator):

1. All pre-selection cuts are made as above
2. P0D Vertex must be reconstructed by TP0DPairwiseVertexPID algorithm and the track must be a constituent of this vertex
3. P0D Track must be exiting as defined by the last node having a Z position > -1016 mm
4. P0D Track must be 3D as defined by cutting on track position variance Evaluate ΔR, $\Delta \sin\theta$ and ΔT between the p0d track projection and tracker track as before. Apply the following cuts: $\Delta R < 86$mm, $\Delta \sin\theta < .76$, $\Delta T < 100$ ns.

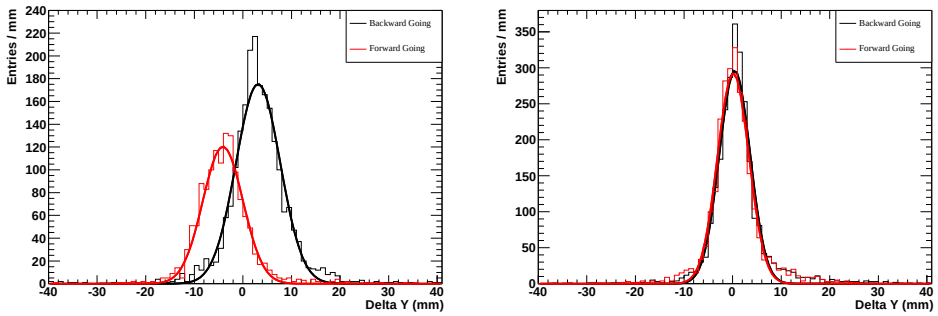


Figure 21: Matching Parameter Δ Y for Production 4 FGD cosmics in Data and MC split into forward and backward going tracks. Gaussian fits have been overlayed. The left plot shows the shift in the residual in Data cosmics when comparing forward (red) and backward (black) tracks. The right plot shows that in MC we do not see such a shift between forward (red) and backward (black) tracks.

Type and Dir.	Mean	Sigma	χ^2/NDOF
Data Forward	-4.1 ± 0.1	4.1 ± 0.1	100.0/48
Data Backward	3.1 ± 0.1	4.5 ± 0.1	155.2/58
MC Forward	0.1 ± 0.1	3.2 ± 0.1	255.1/51
MC Backward	0.5 ± 0.1	3.2 ± 0.1	263.2/59

Table 10: Gaussian fit results and fit error of ΔY for 10000 Production 4 FGD Cosmics in Data and MC. The results are divided into backwards and forward going cosmics as determined by the Tracker reconstruction.

The number of tracks passing the numerator cuts divided by the number passing the denominator cuts yields the efficiency. Figures 22 and 23 show the matching parameter distributions for ΔR, ΔX, ΔY, $\sin\Delta\theta$ and ΔT. The ΔY distribution does not agree very well and causes some disagreement in the ΔR distribution as well. This discrepancy has been isolated to a difference between forward going and backwards going cosmics tracks as determined by the Tracker reconstruction. Figure 21 shows the ΔY residuals separated by forward and backwards going tracks for FGD Cosmics in Data and MC. Note that in MC, the two different directions have a small shift in the central value of the residual. However, in Data, the shift is much greater. Using Production 4 to demonstrate the effect, we have fit gaussians to all of the residuals. The results are shown in Table 10. The root cause of this effect is a difference in the geometry used for reconstruction and the geometry of the 'as built' detector. A much more detailed investigation and the results are discussed in the Appendix in Section E. Finally, note that since the ΔR agreement between Data and MC for the Cosmics sample is actually worse than the Beam events, the evaluated uncertainty should be conservative. Similarly, though the timing distribution does not agree as well as the others, the timing cut is large enough to account for the overall shift.

Figure 24 then shows the final efficiency values for data and MC as a function of muon momentum and θ. Note that in Figure 24, the statistical errors are calculated using a probability distribution function derived with a bayesian approach. The derivation of the PDF can be found in a paper by M. Paterno [7] and is implemented in ROOT under the TEfficiency class. The central values in

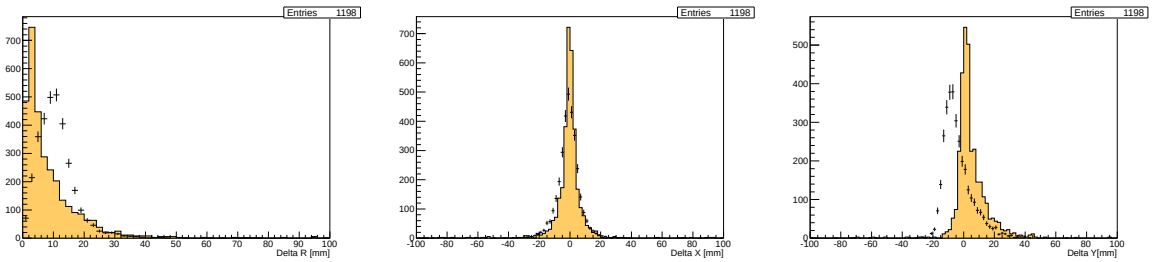


Figure 22: Matching Parameters Δ R, Δ X, and Δ Y for cosmics. The matching cut is applied only on Δ R. Black dots with error bars denote data while the orange fill is monte carlo. We note that the difference in shape in Δ R distributions is due to a shape difference in Δ Y between data and MC.

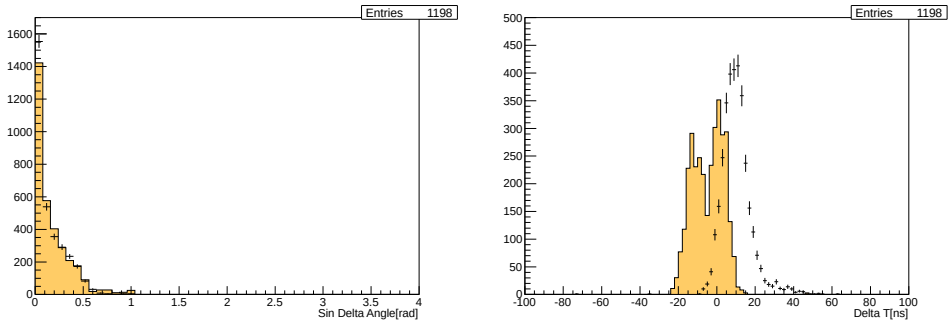


Figure 23: Cosmics matching efficiency parameters $\sin\theta$ and ΔT. The difference in ΔT is not understood, but the cut is placed wide enough to be insensitive to the overall shift.

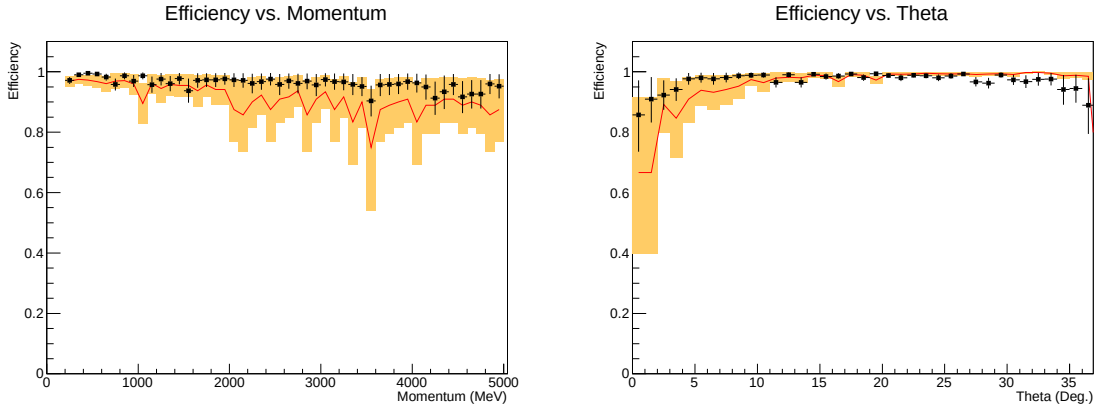


Figure 24: Cosmics efficiencies as a function of Momentum and θ. Black dots are data and red line is the MC. The orange fill shows the error on the MC. We note that the data and MC efficiencies track each other as a function of both kinematic variables.

Figure 24 are then the mean of the PDF (as opposed to the median). The statistical error for the integrated ratio is calculated more simply by approximating the efficiency as a binomial distribution. The efficiencies from the FGD cosmics sample are $99.08\% \pm 0.16\%$ for Data and $99.22\% \pm 0.24\%$ for MC.

6.1.2 Reconstruction Failures in Beam Events

We also performed a search for unclassified reconstruction failures in beam events in Production 4 to account for any possible systematic effect missed by the FGD cosmics and sand muons study. Similar to the method in Section 6.1.1, beam events with one or two quality Tracker tracks pointing directly into the P0D were pre-selected. As the Data has a large number of sand muons present, we also added a cut to veto these events. If more than two above-noise hits were found in the outer edges of the PØD coincident in time with the TPC1 track, we tag the event as a sand muon and exclude it. Of the remaining events, we filtered out those where the matching algorithm failed to find a suitable PØD track to match with the Tracker piece. Finally, the filtered events were examined by hand using an event plotting software (plot-event.exe) to identify any potential reconstruction pathologies previously missed.

When hand scanning filtered events, we founds several modes of failure, many of which were expected. For example, events which had short or partially reconstructed tracks in the downstream end of the PØD as well events with no PØD constituent were missed by the matching algorithm for obvious reasons. Furthermore, some DIS-like events had large clusters of energy deposits in the PØD and were difficult to separate properly into tracks. These were also missed by the matching algorithm as expected. However, there were three classes of failure which, by eye, looked as if they should have been successfully matched. These we examined more closely.

First, we found events with multiple clean tracks passing into TPC1 which were missed by our algorithm. Further study showed this failure mode existed both in Data and MC and was an expected effect. In Appendix Section C we explain how a reconstruction amibiguity (due to design) causes a small portion of multi track events to fail the matching algorithm. Second, we found another class of failed events where the PØD and Tracker tracks were well matched spatially, but separated widely in time (ΔT failure). Finally, the last failure mode were events where the PØD and Tracker tracks were mismatched in the XZ projection, a symptom of incorrect T0 extraction at the TPC1 reconstruction stage. An incorrect T0 calculation creates an offset in the TPC drift direction (XZ). Please see the ND280 reconstruction technical note [3] for greater detail. Figure 25 shows examples of the ΔT and $T0$ failure modes.

We hand scanned Data events corresponding to 1.056×10^{19}POT and MC events corresponding to 3.34×10^{18}POT. The two timing related failure modes, $T0$ and ΔT, were only observed in Data and never in MC. Closer examination of the failed events showed that though some were muon-like tracks originating inside the PØD, many were actually sand muon events which made it through the veto. As the timing of the TPC1 track piece and the PØD track piece were generally different in these events, our sand muon veto did not associate the hits in the outer edges of the PØD with the TPC1 track, causing the sand muons to leak through our selection cuts. In Table 11 we summarize the number of events in Data that failed via the $T0$ and ΔT modes and whether they originated from inside or outside the PØD.

An examination of the ΔT failures show that none of the tracks have FGD constitutents. As the two control samples are predominantly tracks that pass through the FGD, the ΔT failure is most likely not accounted for in the efficiencies evaluated using Sand Muons and FGD Cosmics. We use the 10 ΔT In-PØD failures as uncertainty in the Data event rate. Sand muon events which made it past our veto in this study would still be correctly rejected in the actual CC inclusive selection by

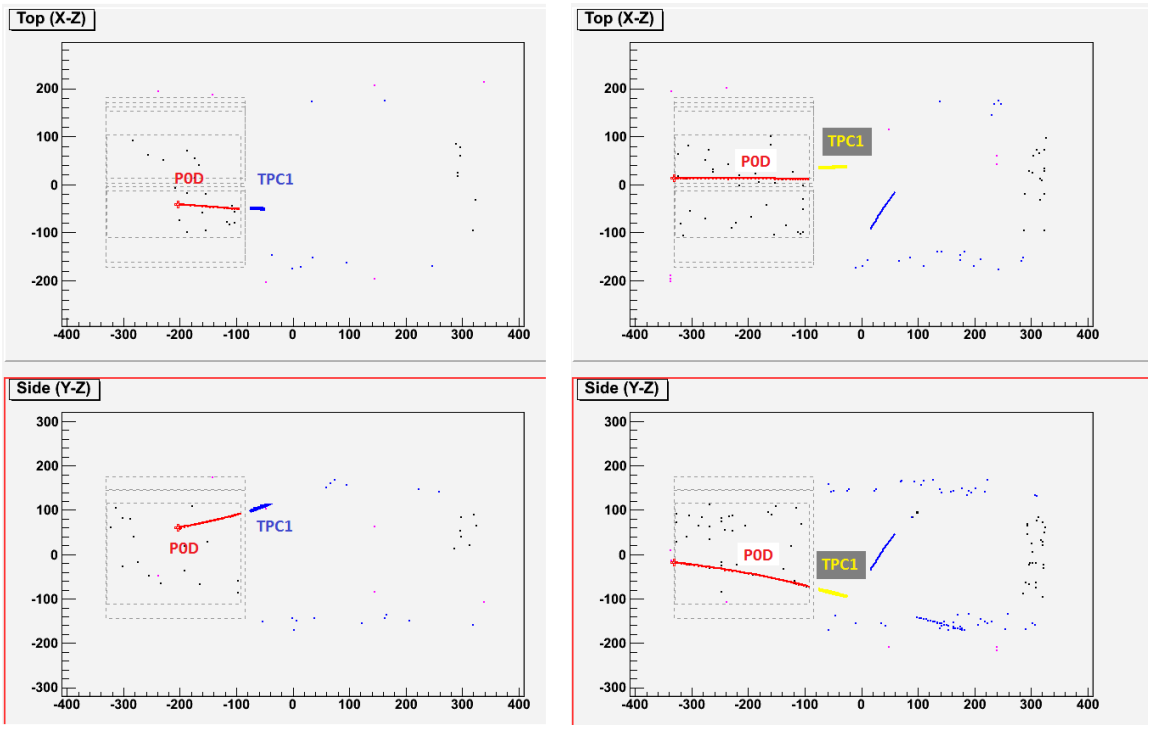


Figure 25: An example of the ΔT (left) and $T0$ (right) failure modes. In the ΔT failure, the PØD track (red) and the Tracker track (blue) match perfectly spatially, but are disjoint in time by $>$ 100ns. In the $T0$, the YZ projection has a good spatial match between the PØD track (red) and Tracker track (yellow), but the drift direction shows a symptomatic offset.

	ΔT Failures	$T0$ Failures
Sand Muon	11	20
In-PØD Muon	10	9

Table 11: Total number of events from the ΔT and $T0$ failure modes for sand muons and in-PØD muon-like events.

the fiducial volume cut. However, the T0 effect is not necessarily replicated correctly in the cosmics and sand muon samples. When relatively steep tracks pass through TPC1 without also entering an FGD, the T0 is more likely to be miscalculated. Since FGD cosmics require the tracks to pass through the FGDs and sand muons are generally lengthy tracks passing through the entire ND280, T0 problems are less likely observed. So the hand scan study also adds a matching uncertainty due to the 9 muon-like tracks corresponding to the T0 failure mode. Also, even though the study was conducted on Production 4, we expect the errors to persist in Production 5 as the effect stems from tracks lacking FGD constituents.

Including the statistical errors appropriately, we have 19 ± 4.36(stat.) events more in the Data from the T0 failure and the In-PØD ΔT failure modes combined. When normalized to the total Data POT for each run type from the inclusive analysis, we get 422.28 ± 96.88 events per 23.47×10^{19}POT for water-in running and 591.77 ± 135.76 events per 32.89×10^{19}POT for water-out running.

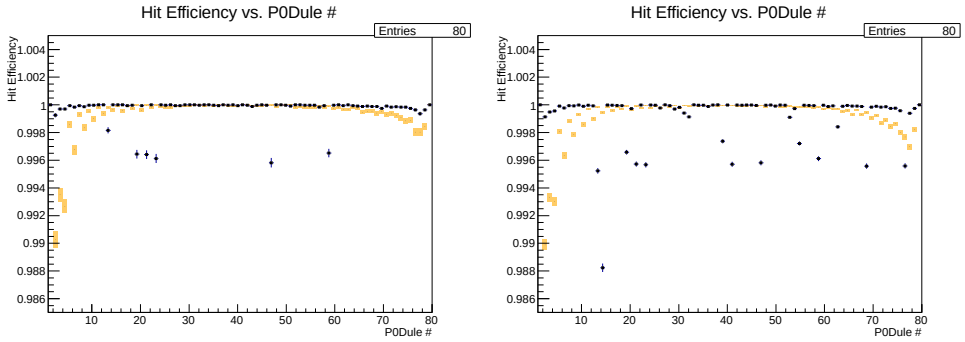


Figure 26: The hit reconstruction efficiency as a function of layer number for water-in running (left) and water-out running (right). Data (black dots with error bars) and MC (orange error bars) show very high efficiencies for all layers. Layers corresponding to the water target have almost perfect efficiency. A few layers in the water target have 0.5% inefficiency. Note that the Y-axis is zero-suppressed.

6.1.3 Results of Matching Efficiency Systematic Studies

From the FGD cosmic sample, the MC / Data efficiency ratio is $(99.22\% \pm 0.24\%)/(99.08\% \pm 0.16\%) = 100.14\% \pm 0.24\%$. This is the value we need to multiply the final Data to MC ratio by to correct for efficiency. Similarly, using the results from the hand-scanning procedure, we calculate corresponding correction factors of 1.017 ± 0.0039 for water-in and 1.023 ± 0.0053 for water-out. These correction factors are multiplicative and uncorrelated with the cosmics efficiency correction. Propagating errors in quadrature yields total correction factors of 1.018 ± 0.0049 for water-in and 1.024 ± 0.0060 for water-out. The Data to MC ratios are shifted by using this final correction factor. The corresponding errors are then $\pm 0.49\%$ for water-in runs and 0.6% for water-out runs.

6.2 Hit Reconstruction Efficiency

We use a side-band sample of beam events to evaluate the layer by layer hit reconstruction efficiency in the PØD. The sample is generated by looking at events originating in the first layer of the P0D and is not a part of the actual selection. The hit reconstruction efficiency combines both the probability of finding an above threshold hit in a PØD bar with the probability of reconstructing succesfully combining the hit into a track. As the PØD Reconstruction algorithm allows for gaps of hits in a track, we use particularly long reconstructed tracks to evaluate the rate of missed layers. Since each PØDule has two layers (an X and a Y layer), we expect any track passing through a PØDule to create two reconstructed nodes. So for each reconstructed track, we use the most upstream node and the most downstream node to calculate the number of total expected nodes. This value is compared to the number of actually reconstructed nodes. Then the efficiency per layer is given by: (# Expected Nodes - # Reconstructed Nodes)/(# Expected Nodes). To have similar levels of PØD bar coverage in both Data and MC, we also require that any tracks used in this study begin and end at similar layer ranges. The hit reconstruction efficiency as a function of layer number is shown in Figure 26.

As expected, the hit reconstruction is extremely high. The few layers in Data with small 0.5% inefficiencies are at the single bar level, which are also expected. Any systematic arising from hit reconstruction efficiency will feed in through our matching algorithm and the Fiducial Volume cut. In the matching algorithm, we require that the PØD track have a node in the last two p0dules, so if both p0dules somehow failed to reconstruct a node, then we would have a small inefficiency. This last two p0dules correspond to layer numbers 77-80. Also, when making the fiducial volume

Layer #	Data (Water).	MC (Water)	Data (Air)	MC (Air)
15	1	0.999894	0.999984	0.999857
65	0.999916	0.999718	0.99996	0.999587
77	0.999638	0.998028	0.995591	0.997677
78	0.999359	0.998028	0.999394	0.996979
79	0.999638	0.998416	0.99975	0.998247

Table 12: Data and MC hit reconstruction efficiencies by relevant layer numbers. Errors are excluded as the efficiencies are so high.

cut, we may misreconstruct out-of-FV tracks as in-FV due to missed nodes at the upstream end of the water target. Similarly, we may lose in-FV tracks in the downstream end of the water target. Even in the least efficient of our p0dules, failing to reconstruct two or more adjacent nodes from a track is negligible (less than 0.0025%), so we can estimate the effect of hit reconstruction efficiency on the fiducial cut using only the likelihood of missing a single node. The upstream end of the fiducial volume corresponds to layer 15 and the downstream end corresponds to layer 65. The hit reconstruction efficiencies for layers 15, 65, and 77-80 are given in Table 12.

From this table, we can then extract the final systematic values due to hit reconstruction efficiency. The probability of having NO nodes in the last two layers is essentially zero and therefore not included as a systematic. The probability of having gained an out-of-FV track from the upstream end of the FV cut is given by number of muon candidate tracks originating in layer 16 multiplied by the inefficiency of layer 15. Similarly, the probability of having missed an in-FV track at the downstream end of the FV cut is given by the # of muon candidate tracks orginating in layer 66 multiplied by the inefficiency of layer 65. The largest inefficiency is that from layer 65 in MC air running, and it is 0.04%. As the change in the Data/MC ratio due to gains and losses of events from hit inefficiency cannot exceed this fraction, and is realistically smaller, we can neglect any systematic effect from hit efficiency differences between Data and MC.

6.2.1 Neutral Back Scattering

To be able to apply the layer efficiency study (Sec. 6.2) on our PØD+TPC1 CC inclusive samples, we have preformed a similar layer by layer efficiency study with the CC inclusive samples. The efficiency study have yielded, as expected, very high layer efficiencies for both Data and MC samples. In addition the study found that for a certain number of times in both beam Data and MC, we observe near the start of a track a node gaps, which is larger than 1 node. We investigated these set of tracks and identified these gaps not to be related to layers inefficiencies. The investigations have found that these tracks are cases were an event had a neutral particle, which was back scattered from the interaction and converted several layers upstream of the true interaction vertex. As a result the additional hits from the converted neutral particle have been added to a forward going track. This effect have caused the track start position to migrate to a an upstream location and to add layers gap in between.

We have further studied and looked at different properties of this topology. Tables **??** and 14 shows for different samples the fraction of tracks that have more than 1 node missed normalized to the number of tracks in that sample. For Run 1 we have found that the Data and MC fractions are

0.90% and 1.38% respectively. Similar behavior between Data and MC is seen at the Run 2 period were the fractions are 1.08% and 1.40% for Data and MC respectively. We use these fractions to find the amount in which the MC should be corrected, to mimic the Data, for both run periods. The correction values we extracted are 0.13 and 0.16 events for Run 1 and Run 2 respectively.

		Fraction of tracks
Run 1	Data	1.63%
	MC	1.90%
Run 2	Data	1.40%
	MC	1.76%
Run 4	Data	1.39%
	MC	1.80%

Table 13: The fractions of tracks with more than 1 missed node for both Data and MC Water-in samples in the different run periods.

		Fraction of tracks
Run 2	Data	1.55%
	MC	1.83%
Run 3	Data	1.46%
	MC	1.80%
Run 4	Data	1.65%
	MC	1.80%

Table 14: The fractions of tracks with more than 1 missed node for both Data and MC Water-out samples in the different run periods.

Another studied aspect was the distance of missed layers with respect to the first layer, the start of the track. Fig. 27 shows the results of this study. fThe figure presents the missed layers distributions for both Data and MC samples, normalized to the number of available tracks. One can clearly identify that the second node, of the tracks that had missed layers, has the highest missed frequently. Another feature is the 'missed node' peak around layer 10 with the spread between 5 to 20 layers. It seem that the MC samples mimics the Data distribution of this feature. Followed this spread we have calculated the number of tracks with missed layers, 5 to 20 layers from the start of the track for both run periods. These rates per POT are summarized in Tables 15 and 16. From this table one can see that the rates are similar between Data and MC.

All the above show that the skipped nodes/layers due to neutral particles is a small fraction of our CC inclusive sample. Moreover the frequencies and rates per POT of these effects in the MC samples look alike the Data distributions. These leads us to a very small systematic error contribution which was not included in the final summary table.

6.3 Fiducial Mass

The calculation of the uncertainty in the PØD fiducial mass has already been discussed in Section 2. To resummarize, we use a mass uncertainty of 0.69 %. Also, since the target mass is also different in

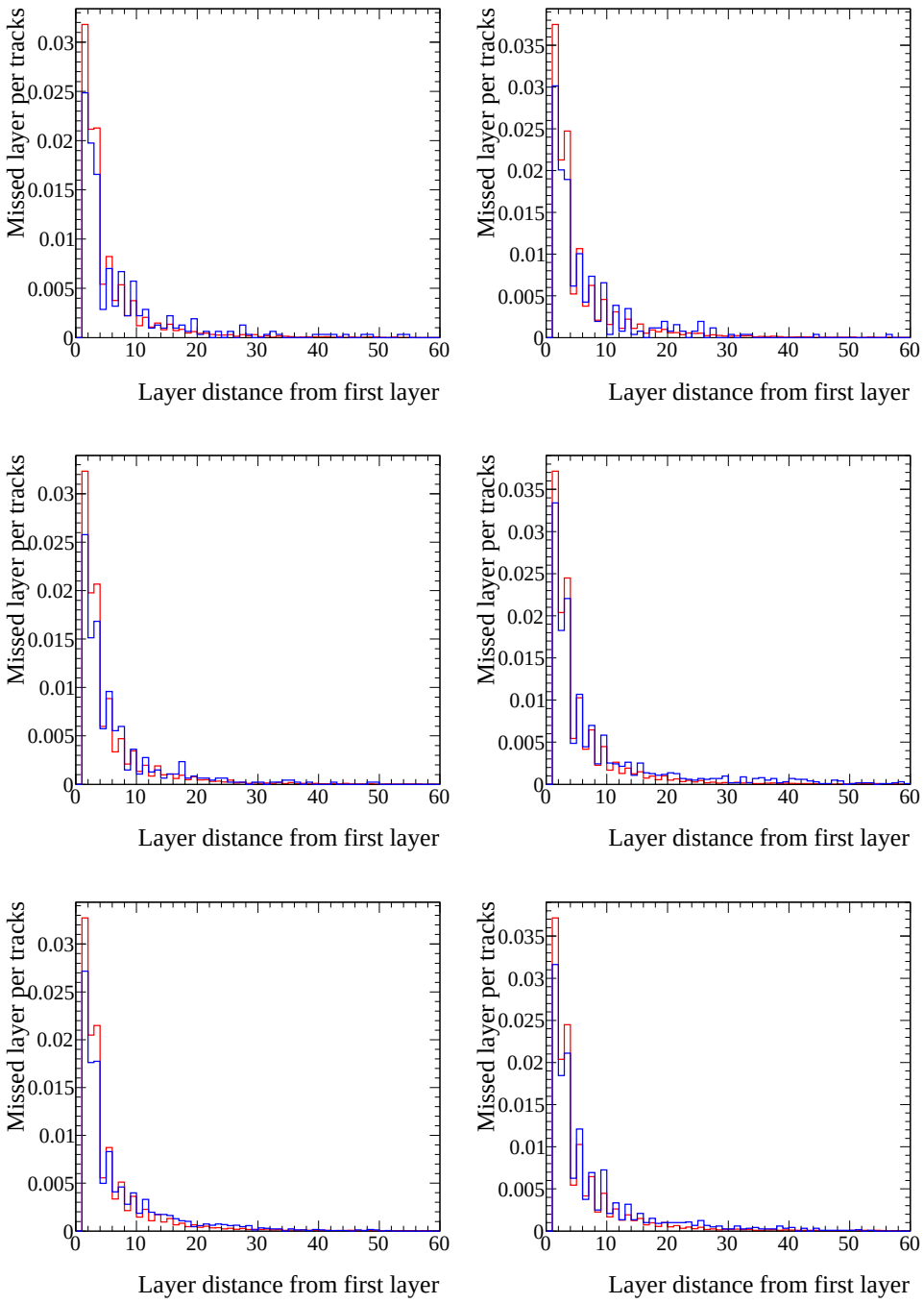


Figure 27: Layers distance from the first layer of missed nodes normalized to the number of tracks. The left columnb shows Water-in periods, from top to bottom are Run 1, Run 2 and Run 4. The right columb shows Water-out periods, from top to bottom are Run 2, Run 3 and Run 4. In all periods the blue and red histograms are generated from the Data and MC samples respectively.

		Tracks per POT [$\times 10^{-18}$]
Run 1	Data	6.53
	MC	7.92
Run 2	Data	7.13
	MC	8.08
Run 4	Data	7.52
	MC	8.12
Run 1+2+4	Data	7.32
	MC	8.09

Table 15: The rate of tracks with missed layers (layers 0 to 20 from counting from the start of the track) for the Data and MC Water-in samples as calculated for Run 1, Run 2, Run 4 and Run 1+2+4 combined periods.

		Tracks per POT [$\times 10^{-18}$]
Run 2	Data	5.38
	MC	6.34
Run 3	Data	6.14
	MC	6.34
Run 4	Data	5.80
	MC	6.34
Run 2+3+4	Data	5.90
	MC	6.34

Table 16: The rate of tracks with missed layers (layers 0 to 20 from counting from the start of the track) for the Data and MC Water-out samples as calculated for Run 2, Run 3, Run 4 and Run 2+3+4 combined periods.

the simulation when compared to the data, we correct the Data/MC Ratio by dividing the number of selected MC events by 1.0317 and 1.028 in Run 1 and Run2 respectively. This correction has already been applied in Section 5.2.4.

	Water-in
Mass correction + 1σ	-0.0079
Mass correction − 1σ	$+0.0078$

Table 17: The fiducial mass systematics for the combined Run 1, Run 2 and Run4 Water-in periods. Mass correction $+(-)$ 1σ corresponds to the case were the correction mass was increase (decrease) by its error.

6.4 Cosmics

To estimate the cosmic contamination of our data samples we have defined the time period before the first bunch in each spill as a sideband sample, for each T2K run.

These sideband samples were used to scan for tracks that would pass all of our selection cuts.

	Water-out
Mass correction + 1σ	-0.0084
Mass correction − 1σ	$+0.0085$

Table 18: The fiducial mass systematics for the combined Run 2, Run 3 and Run4 Water-out periods. Mass correction $+(-)$ 1σ corresponds to the case were the correction mass was increase (decrease) by its error.

	Run 1	Run 2
Cosmic candidate tracks	2	0
Total cosmic scanned time	2.429×10^9 ns	2.811×10^9 ns
Total analysis time	2.159×10^9 ns	3.331×10^9 ns
Expected cosmic contamination	2 tracks	0 tracks

Table 19: Cosmic contamination rates for the two run periods.

The number of tracks found [4] in each run is documented in Table 19. This number was then divided by the 'Total cosmic scanned time', i.e. number of spills × time period before first bunch in ns, to extract the rate of contaminated tracks per time unit for each run. To find the 'Expected cosmic contamination' number in each run, we multiplied the above rate by the 'Total analysis time' (which is the number of spills times number of bunches in ns). The results of these multiplications are rounded up to evaluate the number of contaminated cosmic tracks in our selected samples. Table 20 summarizes all the above findings and calculations.

We adopt a conservative approach which assumes that when a contaminated track is present in a bunch it is tagged as the muon candidate. We then recalculate the Data/MC ratios without these tracks and extract a systematic upper limit for this contamination source.

	Run 1	Run 2
Cosmic systematics	-0.00031	0

Table 20: Cosmic systematic values for the two run periods.

We note that the study in this section has been completed with the use of tracks reconstructed by the Global Reconstruction package. The CC inclusive selection outlined in Section 5 was performed with a Tracker to PØD matching algorithm. Though these two reconstruction methods are not the exact same, they are very similar. In addition, from the efficiency study performed using cosmics (Sec. 6.1.1) we find that the new matching algorithm maintains a high efficiency of reconstructing cosmics and other lengthy tracks. Since the overall contribution of the cosmics contamination is negligible, we do not expect that switching to the Tracker to PØD matching method will change this systematic uncertainty. We adopted the conservative approach and estimate our cosmic systematic contribution to be ±0.00031.

[4] In case no tracks were found, an upper limit of one track was used.

6.5 Fiducial Volume

A one-track per spill sample of beam MC has been used to extract the X and Y vertex resolution. Fig. 28 shows the X and Y coordinate resolution for Run 1 as determined by MC. The resolution is defined as the reconstructed position minus the true vertex position. The different X and Y resolution plots were fitted to a Breit-Wigner function which yielded in a FWHM that corresponds to a σ of 5.7 mm and 7.2 mm respectively.

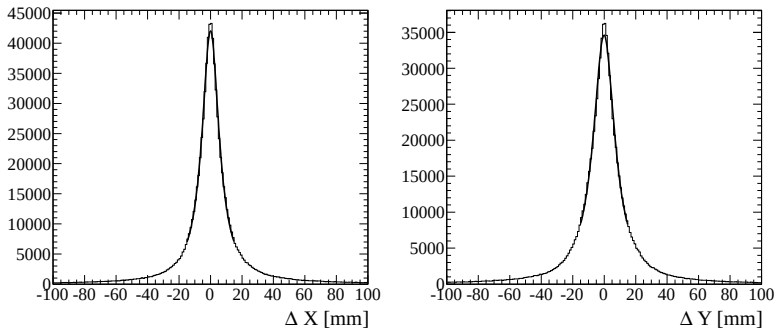


Figure 28: Resolution of start position of Reconstructed - MC truth. On the Left (Right) is the X(Y) distribution.

To extract the FV systematic values we have simultaneously varied all of our FV boundaries outside and inside, by ± 10 mm (which was a conservative $\pm\sigma$) in both X and Y coordinates and by ± 1 layer in the Z coordinate. The decision to vary the Z coordinate by one readout layer is due to the fact that each PØDule has an X layer and a Y layer, where the WT FV excludes the most upstream X layer and the most downstream Y layer.

We recalculated our Data-to-MC ratios for the outside and inside cases and extracted from these the corresponding systematic values. Tables 21 and **??** summarizes these values for the combination of both Water-in and Water-out periods respectively. It was found that the inside variation yielded a value of -0.0057 while the outside variation gave a value of $+0.0049$.

	Water-in
Outside	+0.00867
Inside	−0.00160

Table 21: The fiducial volume systematics for the combined Run 1, Run 2 and Run4 Water-in periods. Outside (Inside) corresponds to the case were the boundaries were varied outside (inside) the official ones.

There are a few other possible effects we must consider. One is the effect of hit reconstruction efficiency when making a FV cut. This has been shown to be negligible in Section 6.2. Another is the possibility that the vertex resolution in Data is different from that in MC. However, we note that small differences in the width of the vertex resolution will not translate to a systematic uncertainty. As long as the vertex resolution (i.e. the residual distribution) is symmetric, the number of true out-of-FV events that get 'smeared' into the FV will cancel out with the number of true in-FV event that get 'smeared' out of the FV. This is the case regardless of any differences between the Data

	Water-out
Outside	+0.00294
Inside	+0.00035

Table 22: The fiducial volume systematics for the combined Run 2, Run 3 and Run4 Water-out periods. Outside (Inside) corresponds to the case were the boundaries were varied outside (inside) the official ones.

and MC vertex resolution distributions. Examining the matching parameter distributions (Figures 22 and 33) from Sections 6.1.1 and **??**, we can draw some conclusions on the vertex resolution in data. The backwards projected Tracker track provides a best guess for where the true position of the PØD node should be, so the ΔY and ΔX residuals mimic the vertex resolution. We see that the Data and MC residuals have similar widths (see Table 10 for an example) and more importantly, are symmetric. This indicates that vertex resolution has negligible effect on the Fiducial Volume systematic.

6.6 Track Timing

To estimate the track timing systematic values we used a similar approach as the FV systematics, where we varied simultaneously all of our track timing window boundaries outside and inside, by ± 15 ns and ± 18 ns for Run 1 and Run 2 respectively. These variations are the average timing distribution widths for tracks in each bunch for the different run periods.

The time systematic study had shown no change in the number of selected events for Run 1. For Run 2 we found a change in the selected number of events only for the variation of the time window 'inside' by one event for the Data sample and by less than one event (after POT normalization). From the above one finds that this systematic is negligible and therefore we do not include it in the final systematic table.

6.7 Event Pileup per Bunch

The pileup[5] event contribution in the PØD CC inclusive samples should be proportional to the experimental beam power. From the fact that the estimated beam power used in the simulated MC was not the exact same as in each of the runs, we expect to have a different pileup effect between data and MC. Here we have estimated these effect difference on our final data to MC ratio of each run.

To estimate the event pileup effect we have assumed that there is an equal probability for an interaction to occur in each beam bunches. We can then evaluate this probability from the number of active bunches and the number of total bunches available for each Run both for Data and MC. These probabilities squared are our estimations to have two interaction in the same bunch, i.e. the event pileup rate. These rates were found to be very small and are summarized in Tables 23 and 24.

The systematic error on the Data to MC ratios for a specific run can be estimated with the use of the pileup event rates of both Data and MC. These were calculated for all run periods separately

[5]Coincidental neutrino interactions in the same bunch.

	Event Pileup Fractional error
Run 1	+0.0000002
Run 2	+0.000000001
Run 4	−0.0000021

Table 23: Event pileup fractional error on the Data/MC ratios as been extracted for the Water-in periods: Run 1, Run 2 and Run 4.

	Event Pileup Fractional error
Run 2	−0.0000008
Run 3	−0.0000017
Run 4	−0.0000037

Table 24: Event pileup fractional error on the Data/MC ratios as been extracted for the Water-out periods: Run 2, Run 3 and Run 4.

and yielded very small systematic errors (see Tables 23 and 24).
The fact that the observed uncertainty are negligible have caused us not to include them in the find systematics calculation.

6.8 Out of PØD Fiducial Volume Background

The current method to estimate the Out-of-Fiducial Volume (Out-of-FV) is base on the position of interactions as identified from the truth vertex information in the MC.

Our approach has 3 steps. In the first step we use the MC to define different regions and the amount of Out-of-FV background associated to each. At the second step we utilize those regions Data to MC ratios to estimate the amount of background events in Data. At the next step we subtract the estimated Out-of-FV events from both Data and MC and extract a new Data to MC ratio. In the last step we use the difference between the new ratio and the original one to extract that region Out-of-FV systematic error.

An investigation of the MC Out-of-FV samples found that the main source of backgrounds are interactions that occur in the PØD CECal, which have backwards going tracks, that were reconstructed in the Water-Target (WT) FV. As a result we defined for this systematic study three volumes outside the PØD WT [2]. This includes the PØD active volume (X or Y coordinate within ±1200 mm, −3400 mm <Z< −930 mm), which is divided into two regions, the CECal (−1266 mm < Z) and the rest of the active volume. The third volume is the region outside of the PØD active volume. The last volume includes both the surrounding UK PØD ECal and the SMRD.

The fraction of events in the different volumes for the combined MC samples of Water-in and Water-out periods are shown in Tables **??** and 26 respectively. Note that the events with no associated truth vertex were added to the volume with the largest error contribution, as a conservative approach. The total percentage of Out-of-FV is found to be $\sim 2.8\%$, where most of the contribution is from the CECal region inside of the PØD active material.

To calculate the Data-to-MC systematic error for the outer volume we used the SMRD CC inclusive Data-to-MC ratio of 1.085 [8] and the estimated MC background of that volume (see Tables

Volume region	Percent
WT Fiducial	96.76%
PØD Active: CECal (+ no Truth Vertex)	1.97%
PØD Active: not in CECal	1.01%
Outside PØD Active (PØD ECal + SMRD)	0.26%

Table 25: The Run 1 + Run 2 + Run 4, Water-in period samples truth vertex distributions for different volume regions. Please note that the values in the table are event averaged between the run periods, the actual error is calculated for each run period and average at the end of the procedure.

Volume region	Percent
WT Fiducial	96.22%
PØD Active: CECal (+ no Truth Vertex)	2.30%
PØD Active: not in CECal	1.12%
Outside PØD Active (PØD ECal + SMRD)	0.36%

Table 26: The Run 2 + Run 3 + Run 4, Water-out period samples truth vertex distributions for different volume regions. Please note that the values in the table are event averaged between the run periods, the actual error is calculated for each run period and average at the end of the procedure.

?? and 26 respectively). The errors for the 'Outside of the PØD active volume' was found to be +0.00050 and +0.00098 for the 'Water-in' and 'Water-out' periods respectively.

To be able to repeat this procedure with the CECal backgrounds we had to calculate the CC inclusive Data-to-MC ratio for that volume. This calculation followed the same analysis procedure as was done for the WT (see Section 5.1) where we changed the FV dependence in Z to be in the CECal (i.e. X and Y boundaries are set to the same values as the PØD WT[2], while the Z was taken to be between −1266 mm to −1010 mm). The yielded and utilized Data-to-MC CC inclusive ratios were respectively, 0.880 ± 0.026 and 0.915 ± 0.023 for Run 1 and Run 2 Water-in periods, and 0.859 ± 0.024 and 0.854 ± 0.012 for Run 2 and Run 3 Water-out periods respectively. With the use of these ratios and the estimated CECal backgrounds (see Tables **??** and 26) we have yielded the CECal region systematic errors of −0.00359 and −0.00074 for the 'Water-in' and 'Water-out' periods respectively. Note that the combined run periods systematic error was calculated first for each run period and then average by the number of Data events to extract the above value.

The background events in the second volume (PØD Active: not in CECal) are distributed around the other FV boundaries (except of the one interfacing the CECal). As a result these events contribution is taken into account by the FV systematics.

The total estimate Out-of-FV systematic includes both the CECal and the outer (UK PØD ECal + SMRD) volumes which yielded the error values of ±0.00062 and ±0.00123 for the 'Water-in' and 'Water-out' periods respectively.

6.8.1 Sand Muon interaction

An additional Out-of-FV backgrounds source are sand muon interactions which occur from the surrounding pit and soil area. To account for these interactions, we utilized the Production 4 'Sand

interactions' MC sample which was equivalent to an exposure of $\sim 7 \times 10^{19}$ POT. Our analysis selection procedure identified 3 events from the MC 'Sand interactions'. After normalizing by Data POT, the estimated background contamination from 'Sand interactions' becomes 3.12 events. Since this contamination is very small, its background contribution to this analysis is negligible and will not be included in the Out-of-FV systematic evaluation.

6.9 TPC1 Tracking Efficiency

When calculating the PØD tracking and matching efficiency in Section 6.1, we used a selection of tracks which had been reconstructed in TPC1. In this section, we now evaluate the systematic uncertainty arising from the efficiency of reconstrucing a track in TPC1 using sub-detector reconstruction only. The work for this was completed by Anthony Hillairet and the tracker group. The technique uses a selection of events with reconstructed tracks in the PØD and TPC2 and looks for the fraction of time where there is also a reconstructed track in TPC1. Further details can be found in the draft T2K technical note detailing tracking efficiency in the ND280 Tracker (T2K-TN-075). The efficiency is expected to be very high and the systematic uncertainty very low so we linearly add the central *in*-efficiency ratio value with the error to calculate the final systematic. The results are binned in muon track candidate momentum.

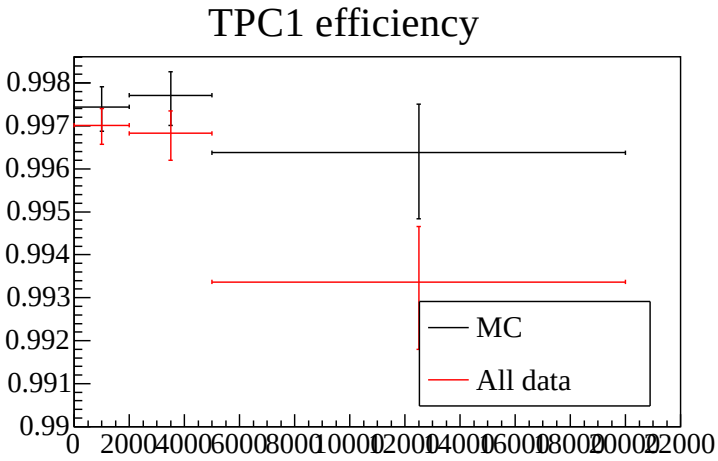


Figure 29: TPC1 Track Finding Efficiency vs. muon momentum. Black (Red) shows MC (Data). The plot is zero suppressed.

Table 27: TPC1 Track finding efficiencies and ratios of efficiency for MC and Data. Errors are also shown. All numbers are percentages.

	MC MC	Data Data	Ratio Ratio	% Data Events % Water-in	% Data Events % Water-Out
0 - 2 GeV	$99.74 \pm 0.05\%$	$99.70 \pm 0.04\%$	$99.96 \pm 0.07\%$	57.1 %	57.3 %
2 - 5 GeV	$99.77 \pm 0.06\%$	$99.68 \pm 0.06\%$	$99.91 \pm 0.09\%$	31.5 %	25.1 %
5 - 20 GeV	$99.64 \pm 0.14\%$	$99.34 \pm 0.14\%$	$99.70 \pm 0.20\%$	11.4 %	8.1 %

The TPC1 efficiency results are shown in Table 27 and Figure 29. Linearly adding the inef-

ficiency ratio and error, the TPC study obtains very small uncertainties of 0.1%, 0.2% and 0.5% for momentum bins 0-2 GeV, 2-5 GeV and 5+ GeV respectively. These uncerainties are then weight-average with the percentage of selected events in each momentum bin for each run type. The systematic yielded by this study for a single momentum bin is 0.18% for water-in and 0.15% for water-out. As expected, the overall effect from tracking inefficiencies in the TPC is extremely small.

6.10 Charge Mis-ID

We begin our study using the results of the charge mis-id study described in T2K-TN-048 v2.1 by Javier Caravaca et al. The following observations concerning our selection significantly simplify the propagation of the charge mis-ID systematic:

1. The vast majority of selected muon tracks have > 40 nodes in the TPC constituent
2. The charge of the track is extracted only from TPC1, which makes the analysis insensitive to the global mismatching effects outlined in TN-048.
3. As the tracks have a reconstructed vertex in the PØD, most are forward going.

So we use Table 2 on page 8 of TN-048, which lists the charge confusion rates for TPC tracks that satisfy the requirements listed above. To propagate the charge mis-ID systematic into our analysis, we use a simple approach. We begin with samples extracted from the entire Data and Monte Carlo sets by excluding the charge cut that is the very last CC inclusive cut. So we have quality matched tracks both positive and negative that originate in the PØD and pass through TPC1. This Data and MC sample is further classified into two mutually exclusive subsets:

1. Beam bunches where the highest momentum track is reconstructed as positive
2. Beam bunches where the highest momentum track is reconstructed as negative

If a track in subset 1 had been misreconstructed as positive when in truth it was negative, then it is a candidate muon track which we missed. To correct for this, we must add the number of charge misidentified tracks in subset 1 to the total number of selected CC inlcusive events. Similarly, if a track in subset 2 had been misreconstructed as negative when in truth it was positive, then we must remove it from the number of candidate CC inclusive events. To calculate the number of tracks in each subset that were reconstructed with incorrect charge, we simply multiply by the charge mis-ID rate. So the corrected number of events after adjusting for charge mis-ID is given by:

$$N^{corr}(i) = N(i) + P(i) \cdot \big(N^1(i) - N^2(i)\big).$$

$N(i)$ is total number of events in subset 2 in a particular momentum bin i and $N^{corr}(i)$ is the charge mis-ID corrected value in the same momentum bin. We chose to use subset 2 for the value of $N(i)$ as it is almost identical to the CC inclusive selection. $P(i)$ represents the probability of charge mis-ID in momentum bin i extracted directly from Table 2 in TN-048 except for one small change. Negative probabilities are unphysical so we set them to 0. Any error on the probability is change to match (ex:-0.2 ± 1 is changed to $0 + 0.8$). Finally, N^1 and N^2 are the total number of tracks in subsets 1 and 2 respectively. Table 28 summarizes the number of tracks in each subset for different momentum bins in both Data and MC. Table **??** then shows the charge mis-ID corrected number of CC inclusive events for Data and MC.

Table 28: Number of Data and MC events in subset 1 and 2 for both Runs. MC values are normalized by POT, not flux-reweighted and not corrected for fiducial mass discrepancy.

Data / MC	Mom. bin Water-in	Subset 1 Water-in	Subset 2 Water-out	Subset 1 Water-out	Subset 2
Data	0-1.3 GeV	667	2463	1705	5627
	1.3-2.6 GeV	220	983	486	1995
	2.6-4.0 GeV	59	568	170	1290
	4.0-5.3 GeV	21	286	62	551
	5.3+ GeV	54	359	117	719
MC	0-1.3 GeV	701	2672	1886	5959
	1.3-2.6 GeV	229	1025	547	2122
	2.6-4.0 GeV	90	689	222	1409
	4.0-5.3 GeV	34	350	83	731
	5.3+ GeV	56	404	122	891

Table 29: The charge mis-ID rate and the corrected number of CC inclusive events for Data and MC in different momentum bins for Water-in.

Mom. bin	Data Mis-ID Rate(%)	MC Mis-ID Rate(%)	N_{Data}^{corr}	N_{MC}^{corr}
0-1.3 GeV	$0 + 0.8$	$0 + 0.1$	2463.0	2672.0
1.3-2.6 GeV	1.4 ± 0.7	2.1 ± 0.1	972.3	1008.5
2.6-4.0 GeV	3.3 ± 1.3	4.5 ± 0.1	551.2	661.8
4.0-5.3 GeV	6.1 ± 2.5	4.5 ± 0.2	269.8	335.9
5.3+ GeV	12 ± 3	13.1 ± 0.2	322.4	358.5
Total	–	–	4578.8	5036.7

Table 30: The charge mis-ID rate and the corrected number of CC inclusive events for Data and MC in different momentum bins for Water-Out.

Mom. bin	Data Mis-ID Rate(%)	MC Mis-ID Rate(%)	N_{Data}^{corr}	N_{MC}^{corr}
0-1.3 GeV	$0 + 0.8$	$0 + 0.1$	5627.0	5958.7
1.3-2.6 GeV	1.4 ± 0.7	2.1 ± 0.1	1973.9	2089.1
2.6-4.0 GeV	3.3 ± 1.3	4.5 ± 0.1	1253.0	1355.7
4.0-5.3 GeV	6.1 ± 2.5	4.5 ± 0.2	521.2	701.7
5.3+ GeV	12 ± 3	13.1 ± 0.2	646.8	791.3
Total	–	–	10021.8	10896.6

To extract systematics from Tables 29 and 30, we recalculate the Data/MC ratio with the corrected values and take the difference from the nominal ratio. The nominal ratio is defined as the Data/MC ratio from the total number of events in subset 2. The errors on the charge mis-ID values are on the order of the misidentification rate itself, so we cannot ignore them. To be conservative, we linearly added and subtracted the mis-ID rate with the error in each bin. Any negative probabilities were set to 0. The change in Data/MC ratio was calculated for both the cases where all the mis-ID rates had gone up by 1 σ and gone down by 1 σ. The larger change is assigned as a symmetric systematic error. We find a systematic of $\pm 0.75\%$ and $\pm 0.72\%$ from the charge mis-ID rate for water-in and water-out running respectively.

6.11 Summary of Systematic Uncertainties

Note that systematic effects from momentum scale and resolution in the Tracker have been neglected. From a previous CC Inclusive study binned into three muon momentum bins, we calculated the error to be at a few percent level per bin. As this result is integrated over muon momentum, these sytematic sources no longer have a first order effect. If the momentum scale and resolution in Data and MC are very different, they may have had a second order effect via some momentum dependent systematics we do evaluate. The three systematics that are possibly momentum dependent are those related to matching efficiency, TPC tracking efficiency and charge mis-ID. From Figures 24 and 35 we see that the matching efficiency is relatively momentum independent. Similarly Table 27 shows that a few percent shift of event rates from one momentum bin to another would have little effect on the extremely high TPC1 efficiencies across the board. Finally, the charge mis-ID rates in Data and MC are not drastically different. Though the misID rate is momentum dependent, the systematic error is propagated in multiple bins of momentum and then integrated together. Small migrations of events between momentum bins results in negligible changes in the integrated systematic value. Therefore, any second order effects of Tracker momentum systematics have been neglected when calculating the total uncertainty.

We summarize the systematic uncertainties that have been calculated for Run 1 + Run 2 in Table 31. The uncertainties have been added in quadrature to calculate the total systematic error. We do not expect any correlations between the different sources of uncertainty listed. In calculating a systematic error from its fractional one we have used the unitary value for the data to MC ratio as this is the conservative approach. The total detector systematic errors for Run 1 + Run 2 is $^{+1.80\%}_{-1.27\%}$.

Systematic Source	
PØD Tracking and Matching Efficiency	$^{+0.00260}_{-0.02092}$
Fiducial Mass	$^{+0.00599}_{-0.00599}$
Cosmic Contamination	-0.00031
Fiducial Volume	$^{+0.00488}_{-0.00571}$
Out of PØD FV	$^{+0.00168}_{-0.00168}$
TPC1 Tracking Efficiency	$^{+0.00148}_{-0.00148}$
Charge Mis-ID	$^{+0.00729}_{-0.00390}$
Total Det. Sys.	$^{+0.02357}_{-0.00978}$

Table 31: Summary of Detector Systematic Errors for Run 1 + Run 2 combined.

7 Result of the CC Inclusive Selection

The analysis reported here contains a full systematic evaluation, which included both PØD and Tracker tracks studies and utilizes different control samples and methods.

This analysis yields the combined Run 1+ Run 2 Data/MC ratio result of $0.868 \pm 0.011(stat.)^{+0.024}_{-0.010}(syst.)$. ($0.868 \pm 1.27\%(stat.)^{+2.72\%}_{-1.13\%}(syst.)$) for PØD events that occur in the PØD WT and have the muon candidate track pass through TPC1.

References

[1] S. Assylbekov, G. Barr, B. E. Berger*et al.*, *"The T2K ND280 Off-Axis Pi-Zero Detector"*, Nucl. Instrum. Meth. A **686**, 48 (2012); [arXiv:1111.5030 [physics.ins-det]].

[2] K. Gilije, *Fiducial Mass Calculation*, T2K-TN-073.

[3] A. Hillairet, *et al.*, *ND280 Reconstruction*, T2K-TN-072.

[4] M. Paterno, *Calculating Efficiencies and Their Uncertainties*, http://home.fnal.gov/ paterno/images/effic.pdf

[5] J. Caravaca, *Charge Misidentification in local and global reconstruction*, T2K-TN-048

[6] A. Hillairet, *Track-finding inefficiency in the TPC reconstruction*, T2K-TN-043

[7] M. Paterno, *Calculating Efficiencies and Their Uncertainties*, http://home.fnal.gov/ paterno/images/effic.pdf

[8] O. Perevozchikov, *Magnet Induced background for neutrino interactions and C analysis in SMRD*, T2K-TN-088.

[9] Y. Hayato, *A Neutrino Interaction Simulation Program Library NEUT*, Acta Phys. Pol. B40, 2477 (2009).

Appendix

A Plots Normalized by Area

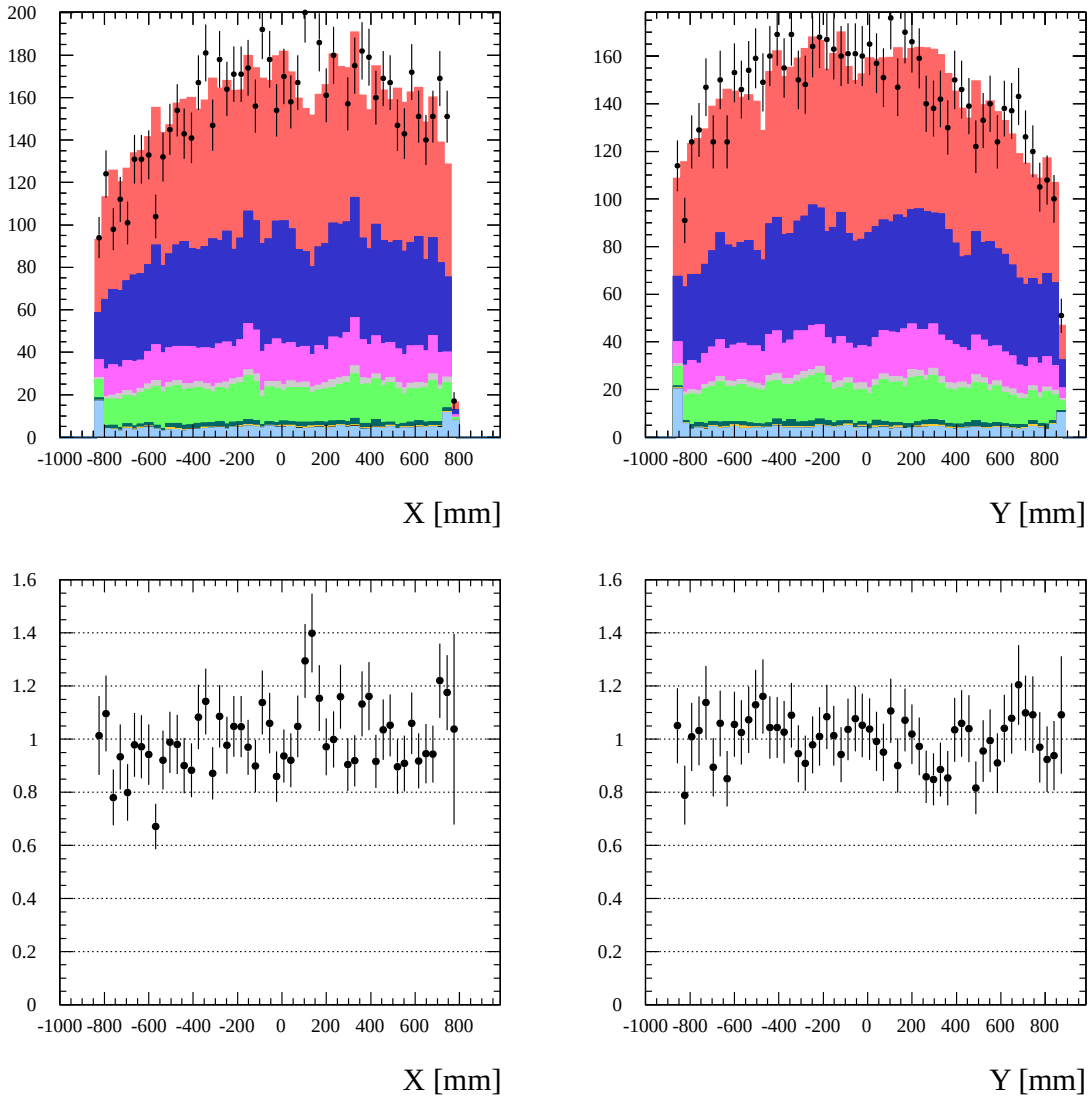


Figure 30: The candidate muon track start X and Y distributions Data to MC plots normalized by area.

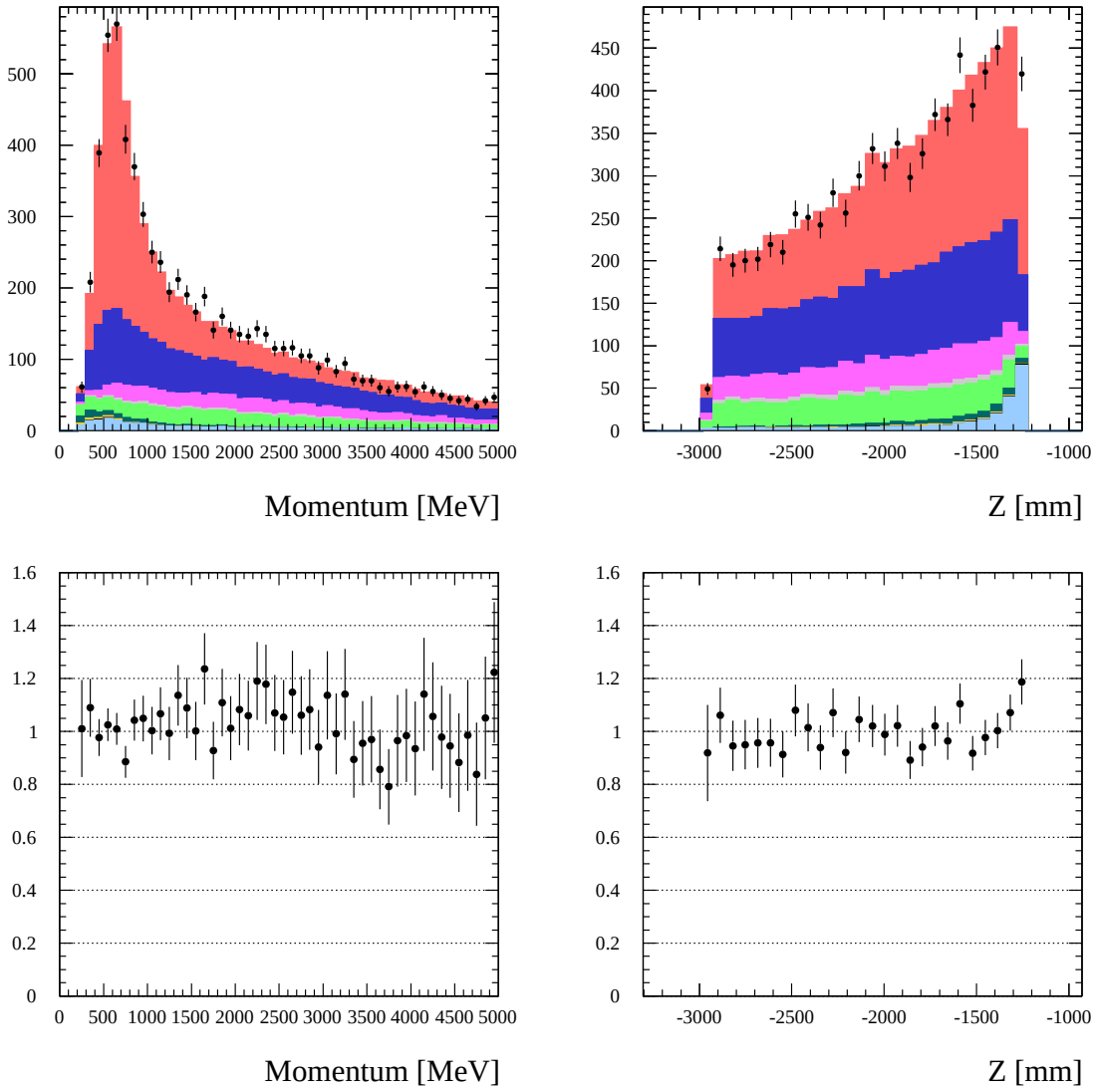


Figure 31: The candidate muon track Momentum and start Z distributions Data to MC plots normalized by area.

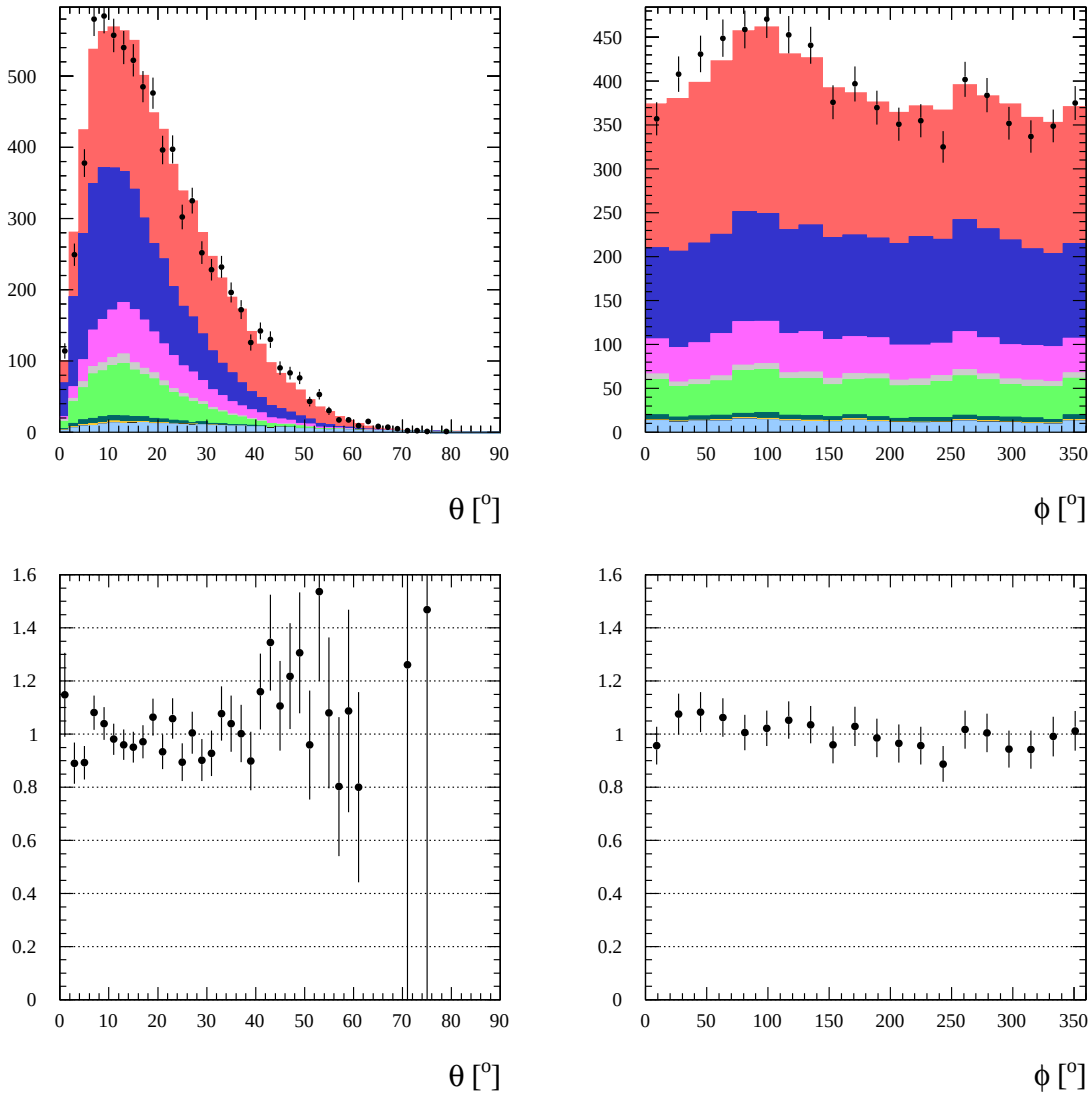


Figure 32: The candidate muon track start θ and ϕ distributions Data to MC plots normalized by area.

B PØD Electronics Structure

The PØD electronics is designed to have a live time and dead time structure that replicates the fine timing structure of the neutrino beam. The PØD cycles through 23 'charge integration windows' or periods of live time per beam spill, where 8 of the windows capture the 8 beam bunches. These integration windows are 480ns long and are interleaved with 100ns of dead time. When a beam trigger is issued, the PØD begins the cycle of 23 integration windows, and in this way is always able to keep the beam bunches beautifully centered within its live times. However, this is not the case for all types of cosmic events.

FGD cosmics are triggered via coincident activity in FGD1 and FGD2. In addition, FGD timing is asynchronous to PØD timing, so the PØD is unable to always center an FGD cosmic track within a particular integration window. This leads to some number of cosmics tracks occuring right at the edge of a PØD integration window and causing 'fringe effects'. When the PØD finds a hit near the edge of an integration window, by design it does not reconstruct the time. How close a hit can be to the edge of an integration window before it loses timing information is a function of Run number as well as whether we are using MC or Data. Values vary from 7.5ns from the edge in Run 1 Data to up to 70ns in MC. Large discrepancies such as this neccessitated cut 6 in Section 6.1.1. By requiring that the found hits are $>$ 80ns from the integration window edge, we negate any kind of fringe effects. This way we do not accidentally add a cosmics-only systematic effect to a study of beam events such as the inclusive analysis.

C PØD Matching Ambiguity

A small percentage of the time, when a multi-track event is reconstructed in the PØD with tracks entering TPC1, the matching algorithm fails to properly pair tracks together. This is an artifact of how PØD reconstruction determines hit positions in a track. The PØD extracts the X and Y positions of a hit from the scintillator bars ligned up along the X-axis and Y-axis respectively. However, When two tracks pass through the PØD simultaneously, there are two X bars in each layer with energy deposited within. This yields two X positions for hits, namely X_1 and X_2. Similarly, in the Y-layers, we have two Y positions Y_1 and Y_2. So PØD reconstruction actually has two available combinations of X and Y positions for each 3D hit, i.e. (X_1, Y_1) and (X_2, Y_2) or (X_1, Y_2) and (X_2, Y_1). Depending on whether the correct combination is chosen, the algorithm can provide a set of reconstructed tracks that is a mirror image of the proper result. If the mirror image is chosen, then when the matching algorithm attempts to pair together PØD and Tracker tracks, while it may succeed in one projection, it will always fail in the other. As this effect is essentially a combinatorial artifact, we observe the 'ambiguous' track reconstruction failure mode in both Data and MC. A possible solution is to use TPC1 tracks as seeds to select the correct combination of X and Y positions, but this algorithm does not exist in the Production 4 suite of software pacakages.

D Sand Muon Matching Efficiency Cross Check

In Production 4, Sand Muons were used as one of two independent samples to investigate and cross-check the reconstruction and matching efficiencies of the analysis. As previously mentioned, sand muons have a similar bunched timing structure to events from the beam which makes them an excellent unbiased sample to cross check the efficiency systematic. The matching efficiency has been outlined above.

The pre-selection, i.e. the denominator cuts, are exactly the same as the FGD cosmics sample except for step 6. For the sand muon sample, it is not necessary to have a 'good hit' finding al-

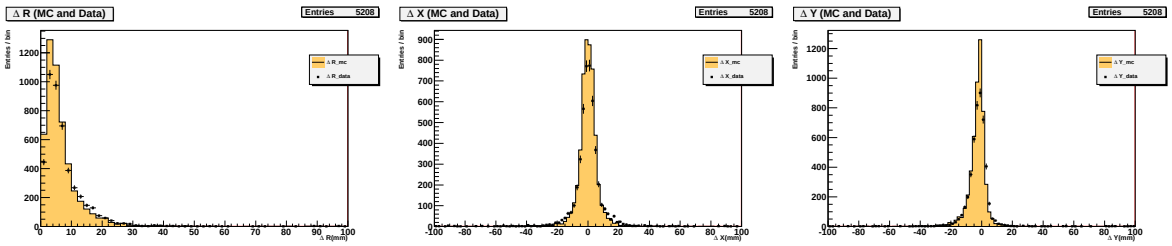


Figure 33: Matching Parameters Δ R, Δ X, and Δ Y for sand muons. The matching cut is applied only on Δ R. Black dots with error bars denote data while the orange fill is monte carlo.

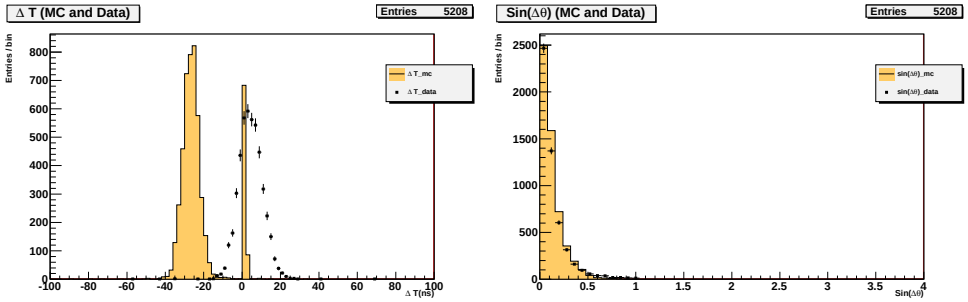


Figure 34: Sand muons matching parameters $\sin\theta$ and ΔT. The difference in ΔT is due to the the decision of the Tracker track to use the FGD or PØD as its time reference.

gorithm like the FGD cosmics sample. Again, please refer to the appendix for greater detail. The matching cuts also exactly mimic the FGD cosmics section.

Figures 33 and 34 show the same matching parameters as the cosmics sample but evaluated for the sand muon case. In Fig. 34 one does notice that the MC ΔT distribution has entries were the time difference are around zero, whereas the FGD cosmic MC distribution (Fig. 23) has all the time difference shifted from zero. The reason for this lays in the fact that the Tracker track uses FGD or PØD for its time reference. While the sand muon sample has a small fraction of events where the Tracker track utilizes the PØD as the time reference, the FGD cosmics sample uses, by design, only the FGD as its time reference.

Figure 35 then shows the data and MC efficiencies with error and have been evaluated in the same manner as Section 6.1.1. The efficiencies from the sand muon sample are $98.22\% \pm 0.19\%$ for Data and $98.31\% \pm 0.18\%$ for MC.

E PØD and TPC Alignment

Using the FGD Cosmics sample as well as the Sand Muon selection, we have found strong evidence that there is a difference between the 'as-built' detector alignment and the input geometry used for reconstruction. Specifically, the geometry has a gap between the PØD and TPC that is 10mm to 12mm larger than the 'as-built' configuration. This effect can be seen in a correlation between the matching parameter residuals (ΔY and ΔX) and the steepness of the matched track. For an explanation of the matching residuals, please see the Section describing our reconstruction procedure.

To measure the correlation between residuals and steepness, we first calculate the linear slope

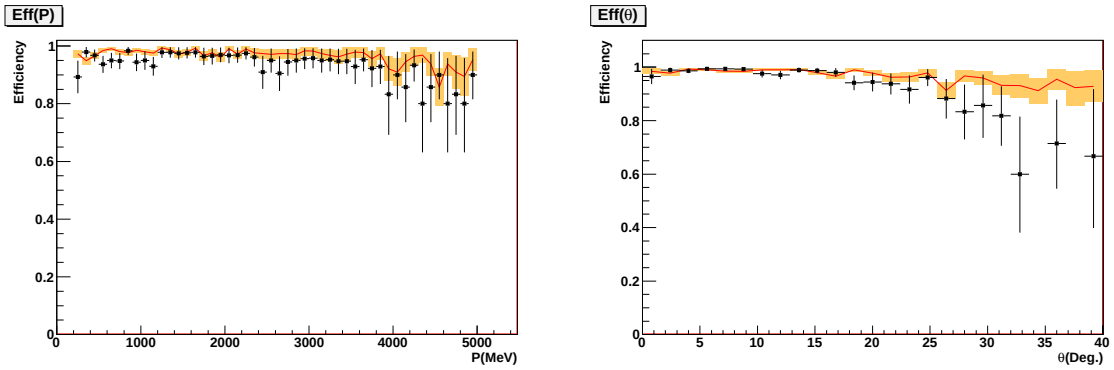


Figure 35: Sandmuon efficiencies as a function of Momentum and θ. Black dots are data and red line is the MC. The orange fill shows the error on the MC. We note that the data and MC efficiencies tracks each other as a function of both kinematic variables.

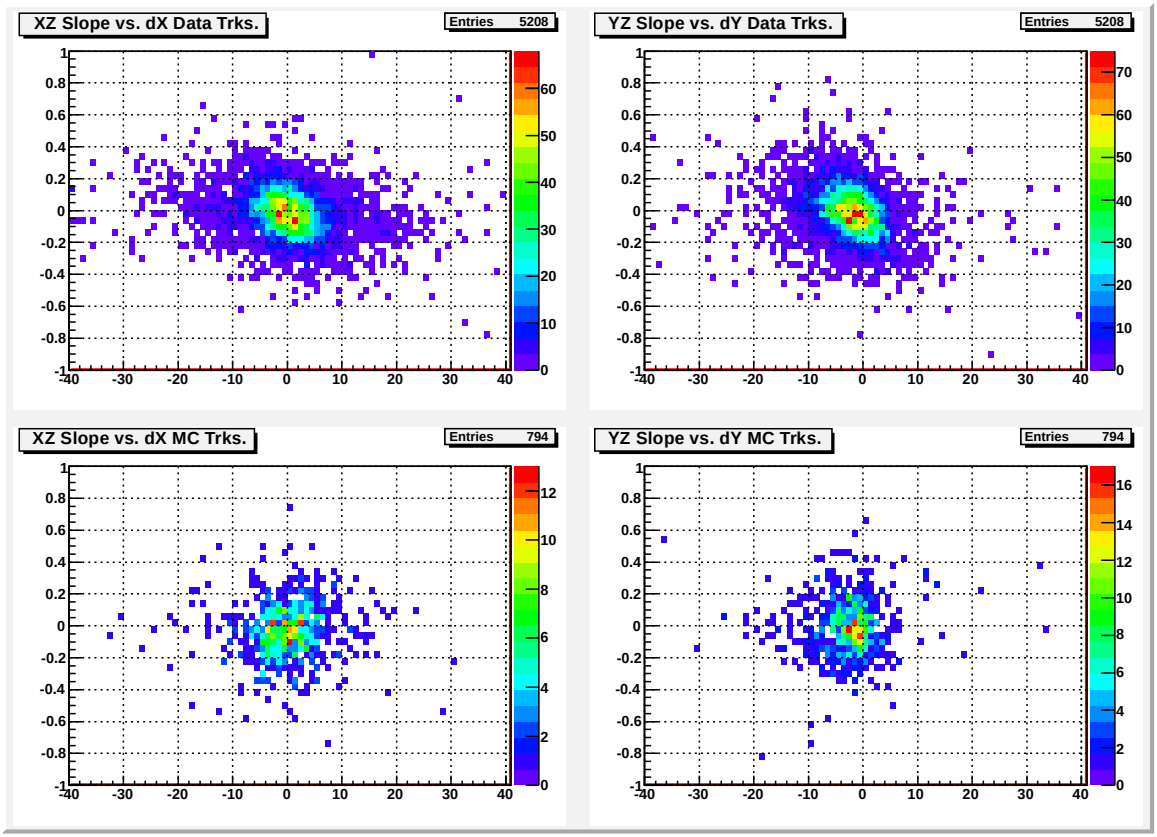


Figure 36: Top Left: XZ linear slope of track vs. ΔX residual for sand muons in Data. Top Right: YZ linear slope of track vs. ΔY residual for sand muons in Data. Bottom Left: XZ linear slope of track vs. ΔX residual for sand muons in MC. Bottom Right: YZ linear slope of track vs. ΔY residual for sand muons in MC.

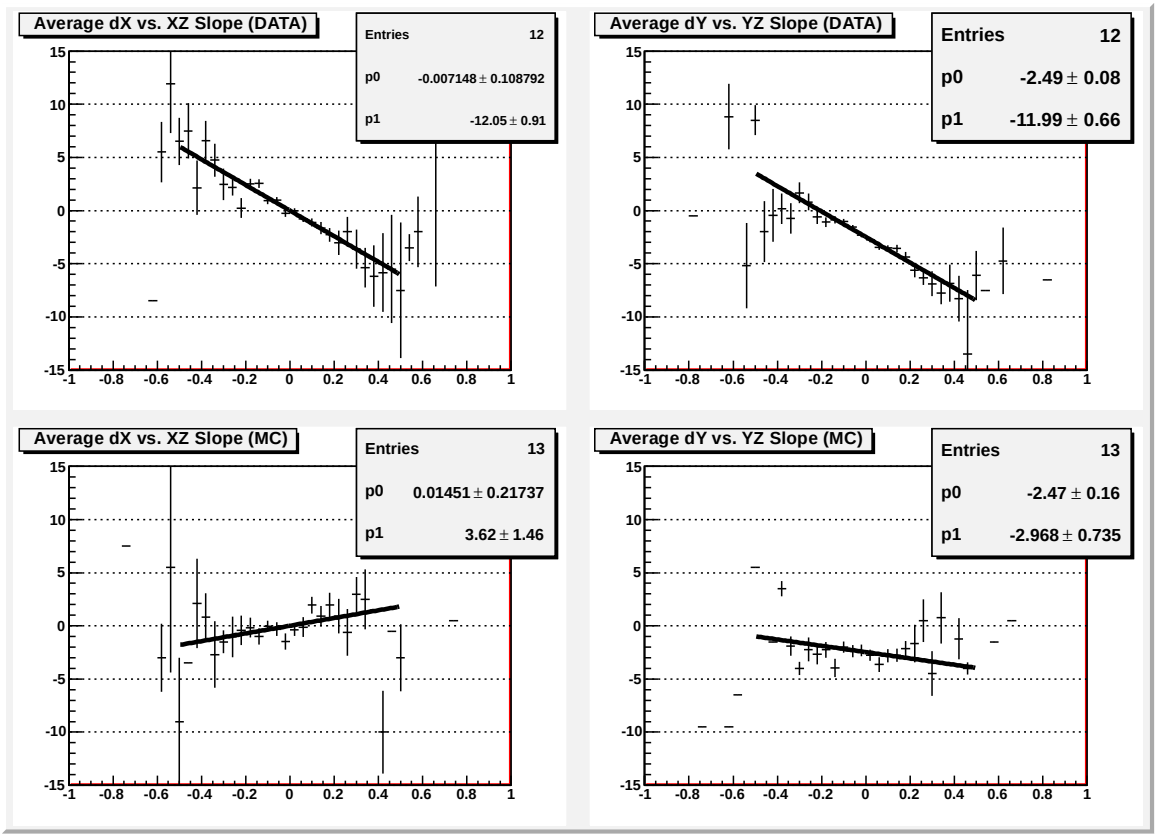


Figure 37: Top Left: Average ΔX residual vs. XZ linear slope of track for sand muons in Data. Top Right: Average ΔY residual vs. YZ linear slope of track for sand muons in Data. Bottom Left: Average ΔX residual vs. XZ linear slope of track for sand muons in MC. Bottom Right: Average ΔY residual vs. YZ linear slope of track for sand muons in MC.

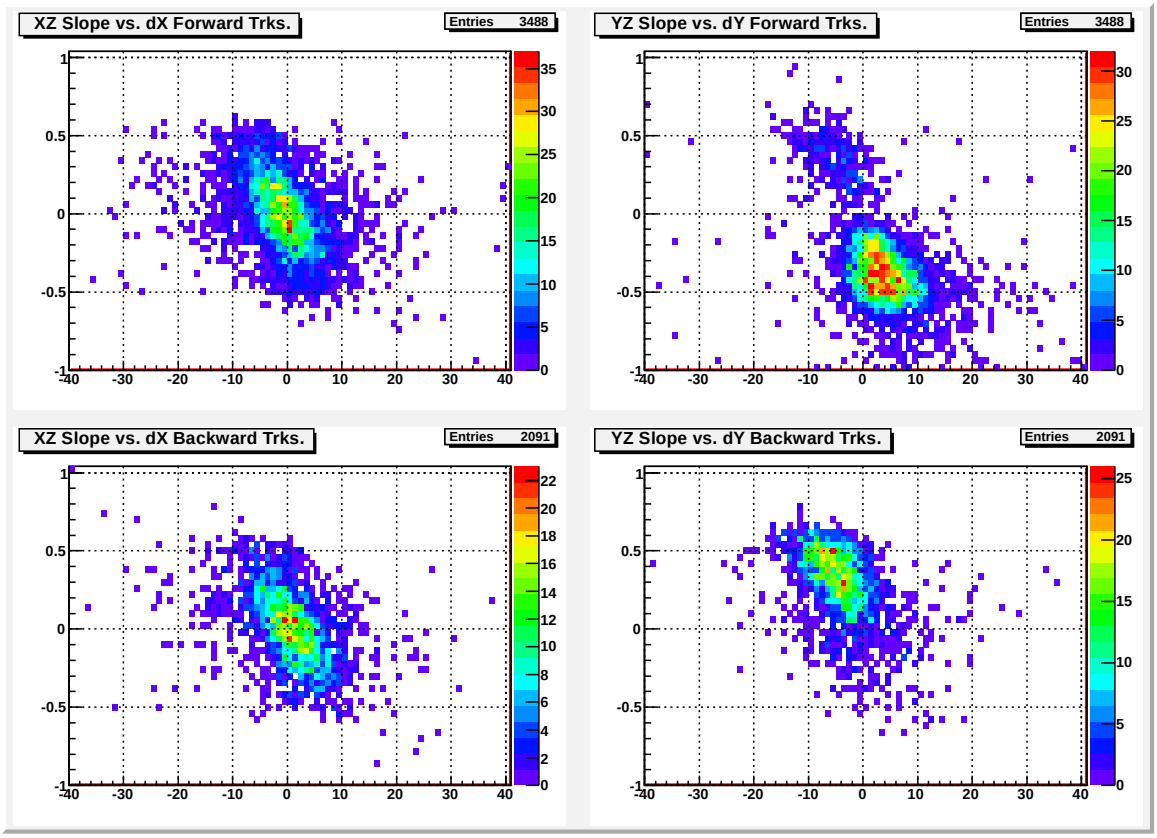


Figure 38: Plots for Cosmics Data. Top Left: XZ linear slope of track vs. ΔX residual. Top Right: YZ linear slope of track vs. ΔY residual. Bottom Left: XZ linear slope of track vs. ΔX residual. Bottom Right: YZ linear slope of track vs. ΔY residual.

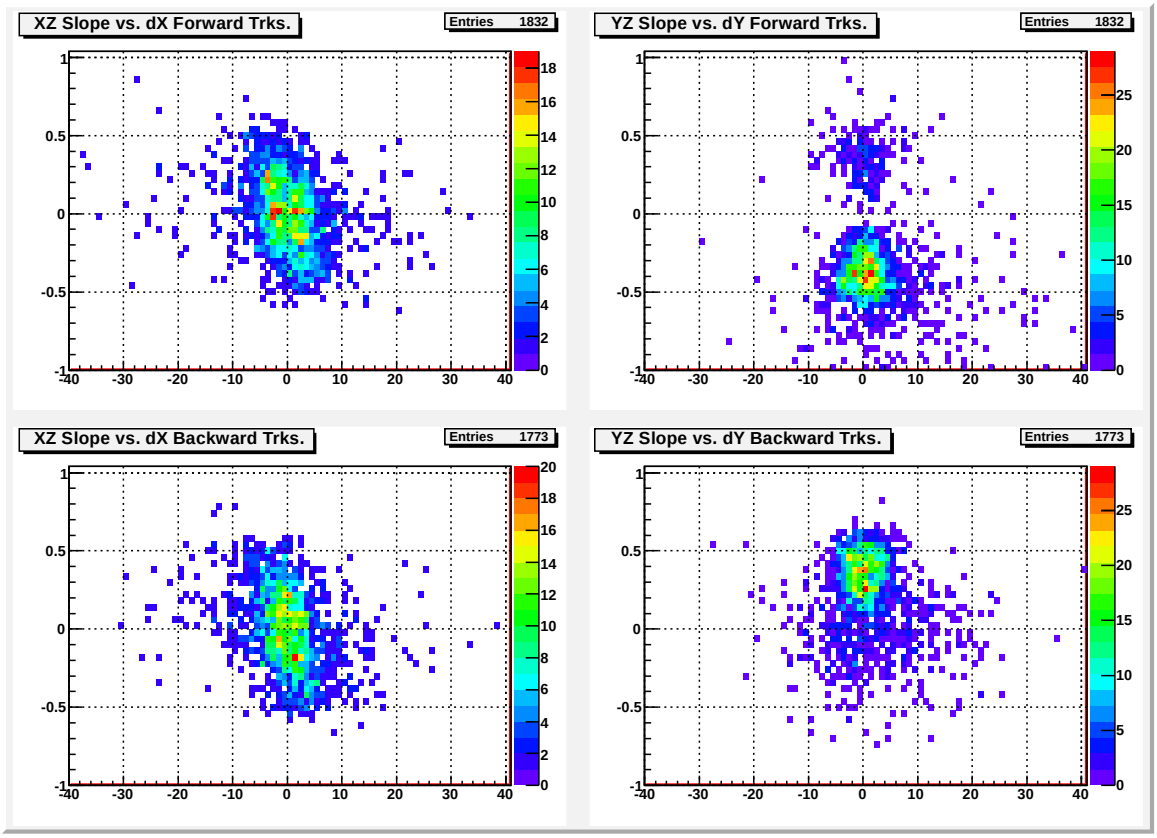


Figure 39: Plots for Cosmics MC. Top Left: XZ linear slope of track vs. ΔX residual. Top Right: YZ linear slope of track vs. ΔY residual. Bottom Left: XZ linear slope of track vs. ΔX residual. Bottom Right: YZ linear slope of track vs. ΔY residual.

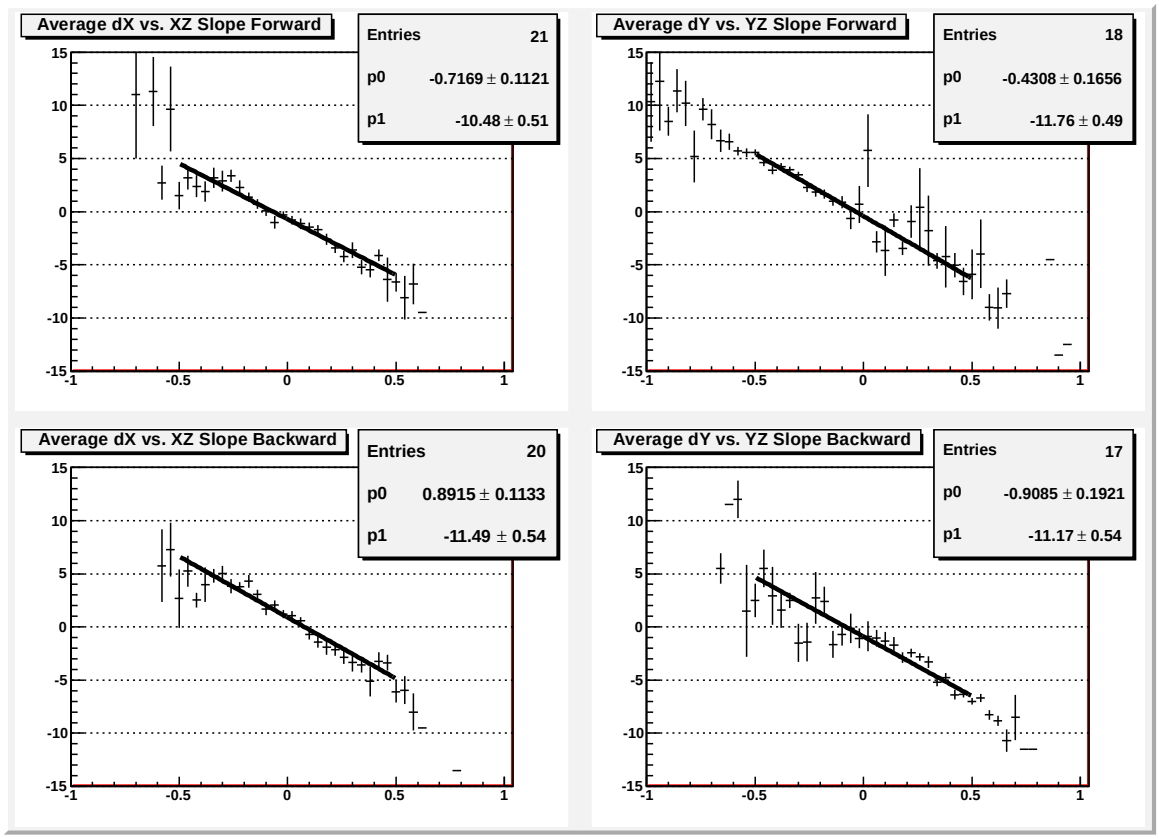


Figure 40: Plots for Cosmics Data. Top Left: Average ΔX residual vs. XZ linear slope of track. Top Right: Average ΔY residual vs. YZ linear slope of track. Bottom Left: Average ΔX residual vs. XZ linear slope of track. Bottom Right: Average ΔY residual vs. YZ linear slope of track.

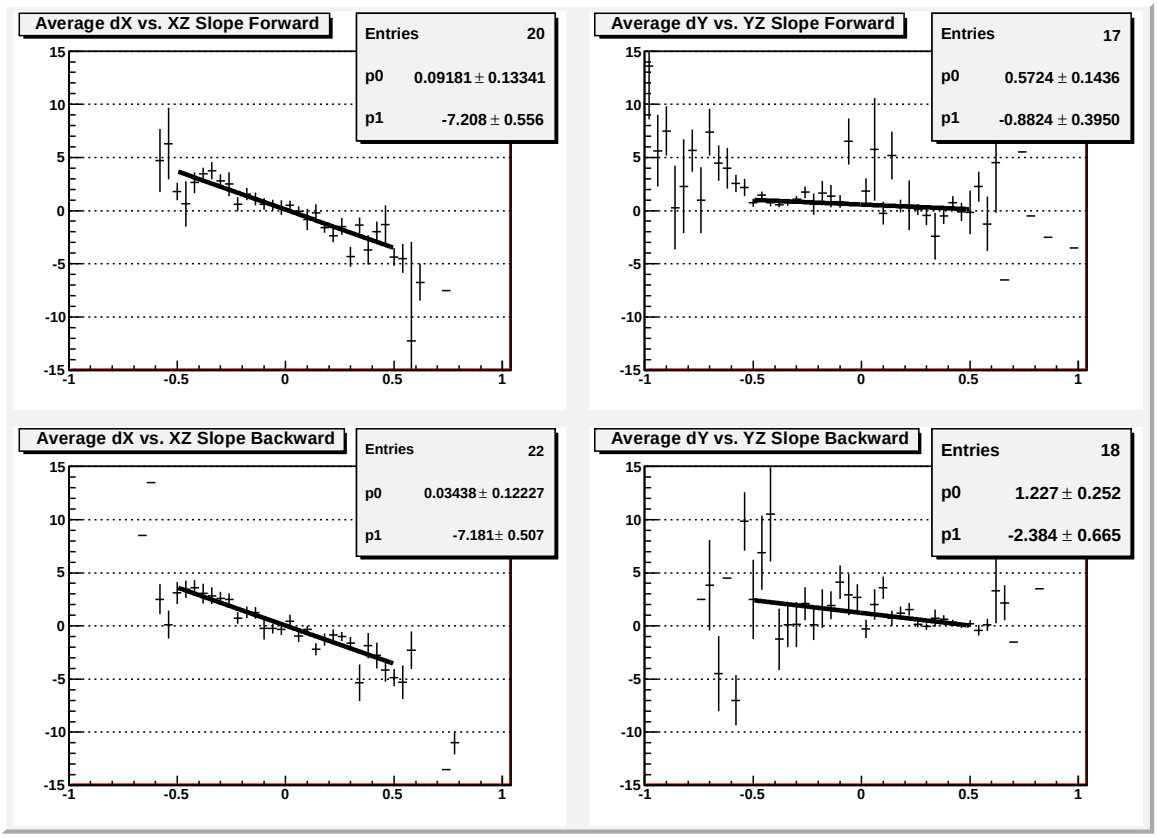


Figure 41: Plots for Cosmics MC. Top Left: Average ΔX residual vs. XZ linear slope of track. Top Right: Average ΔY residual vs. YZ linear slope of track. Bottom Left: Average ΔX residual vs. XZ linear slope of track. Bottom Right: Average ΔY residual vs. YZ linear slope of track.

of each matched track in the XZ and YZ projections. The slope is approximated by drawing a line from the start of the track to the point where the track exits the PØD. For the purposes of this exercise, the small expected curvature of the track is ignored. These linear slopes are plotted against the residuals in the XZ and YZ projections and the results are shown in Figures 36 (sand muons) and 37 (cosmics). A profile of these scatter plots then yields the Average residual vs. linear slope of the track. A quick linear fit to the profile plots allows us to extract the difference between the geometry and 'as-built' versions of the detectors. For example, a fit slope of 10mm means that for a track of slope 1 (i.e. 45 degrees) we have a shift in the residual of 10mm. Since this is a 45 degree track, this also implies that the Z axis misalignment is also 10mm. The profile plots for sand muons and cosmics for both data and MC are shown in Figures 37, 40 and 41.

To conclude, we note that both sand muons and cosmics suggest that there is a 10mm to 12mm difference between the geometry used for reconstruction and the 'as built' detector. The fit slope of the sand muon MC sample is consistent with a horizontal line, and shows no correlation between residuals and track steepness. We do note some tension in the XZ projection of the cosmics MC with the rest of our plots as there is a small correlation between the ΔX residual and track steepness. This is unexpected as we would not normally expect the reconstruction geometry to differ from the geometry used to generate events. However, upon further investigation, we found that the geometries for Production 4 have been changed several times and may be consistent with our findings. Regardless, we reiterate that the matching efficiency systematic is designed to quantify and find these exact problems. The values we have evaluated in the systematics chapter conservatively account for these alignment differences by allowing for the largest possible alignment error (see the sand muon sample and summary of the matching efficiency systematic section).