

## Test beam detector calibrations

Brandon Eberly, Rik Gran, Chris Marshall, Aaron McGowan, Aaron Mislivec, Jyotsna Otsa

March 18, 2014

Details of the calibration of the test beam detector for the 2010 run in the Fermilab Test Beam Facility for experiment T977. Particular attention is devoted to ways the test beam detector calibration is different than the full MINERvA detector installed in the NuMI beam line. Despite sounding like a dry, hyper-technical data dump, there are lots of splendid Pynchon-esque threads herein.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Material assay</b>	<b>3</b>
<b>3</b>	<b>Mapping</b>	<b>6</b>
<b>4</b>	<b>Daily calibrations</b>	<b>7</b>
<b>5</b>	<b>Underperforming hardware or run conditions</b>	<b>8</b>
<b>6</b>	<b>Pedestals</b>	<b>9</b>
<b>7</b>	<b>Gains</b>	<b>10</b>
<b>8</b>	<b>Linearity</b>	<b>12</b>
<b>9</b>	<b>Muon samples</b>	<b>14</b>
<b>10</b>	<b>Crosstalk</b>	<b>16</b>
<b>11</b>	<b>Temperature</b>	<b>18</b>
<b>12</b>	<b>Alignment</b>	<b>19</b>
<b>13</b>	<b>Strip to strip</b>	<b>21</b>
<b>14</b>	<b>Transverse and longitude position effect</b>	<b>22</b>
<b>15</b>	<b>Channel to channel smearing</b>	<b>25</b>
<b>16</b>	<b>Muon equivalent energy unit tuning</b>	<b>27</b>
<b>17</b>	<b>Energy scale and active/total fraction</b>	<b>30</b>
<b>18</b>	<b>Cross checks</b>	<b>31</b>
<b>19</b>	<b>Geant4 step size and MCHit aggregate</b>	<b>33</b>
<b>20</b>	<b>Every other side readout</b>	<b>34</b>
<b>21</b>	<b>Conclusion</b>	<b>35</b>

# 1 Introduction

This document describes the calibration of the detector we installed in the tertiary beam at the Fermilab Test Beam Facility, also known as Test Beam Experiment T977. The companion document TN017 docdb:8547 describes the calibrations of the tertiary beamline. The physics of the two are mostly but not completely decoupled.

Many aspects of the calibration are identical between the test beam mini-MINERvA detector (hereafter called testbeam detector) and the full MINERvA detector installed in the NuMI hall (MINERvA detector). To keep things short but readable, this document gives a description of how the calibrations are done with citations to the NIM article or other documents, and focuses on the results and the ways in which the technique or outcomes of the calibration are different. The most important differences between the testbeam detector data and its calibrations, compared to the MINERvA detector are:

- The planes are physically smaller, 63 strips x 107 cm.
- We did not use clear fiber to make the connection to the PMTs, WLS directly. Especially, the light yield at the PMT is about 50% higher (21 PE at muon peak, not 14 PE.)
- energy response calibrations depend high angle cosmic ray muons, and are applied to pions and protons at normal incidence. In the MINERvA detector its the opposite.
- sets of four planes (UXVX') alternate between east and west readout, instead of all from the top.
- we correct for temperature variations which are  $\pm 10\text{C}$ , much larger than in the NuMI hall.
- The air gap in the tracker and ECAL configurations is larger
- Two detector configurations, 20 ECAL + 20 HCAL (20E20H) and 20 Tracker + 20 ECAL planes (20T20E)

There are some features of the detector stability or run conditions that stand out and are noted; some justify tagging planes or data periods as bad or unsuitable for certain calibrations or analyses. It is testbeam after all.

## 2 Material assay

### testbeam Pb

The lead sheets have slightly different thickness than in MINERvA, documented in docdb:2674. The sheets are 1.99 mm thick, on average, with 2% RMS around that average. Within a typical sheet we took five measurements, which varied with an RMS of 1.2%. More important is the pseudo-density

which averaged 793 pounds / inch thickness with RMS of 11 lb/in (1.4%), which is a smaller RMS than the thickness by itself, and is more like the  $\text{g/cm}^2$  quantity of interest, for plates of uniform area. When the passage of an electron or shower crosses many planes, the bias uncertainty  $\text{RMS}/\sqrt{N}$  which is between 0.44% for ten planes and 0.31% for twenty planes. We do not have an independent estimate of the density of our Pb, so we use the official Minerva/PureLead specification in our geometry of  $11.29 \text{ g/cm}^3$ , which is less than 0.5% lower than the “book value” of  $11.34 \text{ g/cm}^3$  for pure lead.

After these measurements were taken, the Pb sheets were cut and two were joined so that the result was the correct size. In principle, there is a seam between the sheets which is thicker or thinner, and which is not simulated. These seams do not line up, so it is unlikely that a single particle travels through more than one or two seams. Also, the sheets were coated with a layer of paint. The additional material is simulated roughly by increasing the thickness of Minerva/PureLead from the measured 1.99mm to 2.03mm (the parameter `MTestEcalShieldThickness`). This gets the material  $\text{g/cm}^2$  approximately right, but at the cost of simulating it as Pb instead of the mixture that paint actually is.

Based on all of this we have the following uncertainties (bias) in the Pb material thickness: 0.4% plane variation, 0.5% density of our Pb, and 1% material combining Pb + paint, which yields a total uncertainty of 1.2%. If desired, this could be reduced somewhat by remeasuring the weight of several Pb sheets with paint.

## testbeam Fe

The testbeam steel was obtained from the Fermilab boneyard. It was tested for thickness and density, which is reported in the technical note docdb:2630. In the MC the material could be modeled with a thickness of 2.602 cm and standard steel density of  $7.847 \text{ g/cm}^3$ . However, the standard MINERvA material has in the past been listed as a density of  $7.87 \text{ g/cm}^3$ , so the equivalent thickness in this case is 2.594 cm, and currently following docdb:6016 it is listed as  $7.834 \text{ g/cm}^3$  which requires a thickness of 2.606.

The fluctuations around an average plane are 1.2%. For calorimetry in the HCAL, many events are biased toward leaving energy in the first few layers, not twenty, so we should take a value between this maximum uncertainty 1.2% and maybe four planes which implies an uncertainty of 0.6%. In our analyses we take 1.0% bias uncertainty.

## testbeam scintillator planes

The scintillator planes used in the test beam detector are smaller and square (1.07 m long, 63 strips wide) but made to nearly the same specifications and tolerances as the official MINERvA planes, so the material assay in docdb:6016 applies with very few exceptions. Unlike the full MINERvA plane made from multiple planks, testbeam planes are a single unit, and do not require any tape, eliminating 0.4% of the material. Also, the conventional wisdom is the team making the planks was at their most

experienced, which probably improved uniformity and may mean somewhat less structural epoxy, but there is no estimate for how much. The measured thickness of a single test beam plane sample was an average of  $18.55 \text{ mm} \pm 0.16 \text{ mm}$ , which is the same as the 18.50 mm spec in docdb:6016 and the MC geometry. The geometry we have actually used for the MC and reconstruction of testbeam events very closely matches docdb:6016, and potential discrepancies are accounted for in the 1.5% uncertainty.

The material specified in the MC geometry starting with v10r7p4 [?] was checked and tuned to docdb:6016 using a ray-tracing technique. There are differences related to the practical limits of building a MC geometry, so some variation in strip dimensions has been allowed in order to achieve fidelity to the total  $\text{g/cm}^2$ ; fractions of materials containing the higher Z materials Ti, Si, and Al; and the active/passive fraction.

In addition to the  $\text{g/cm}^2$ , the atomic composition of the epoxy is changed from v10r6 to the current one, and represents a chemical assay of a sample of mixed and dried epoxy. In particular, dried epoxy contains only a trace amount of the chlorine in the original resin formulation used in the v10r6 Monte Carlo material description.

**Alternate material budget** As a side-check and an exercise in understanding potential uncertainties, we have constructed an alternate approach to the material model. Two of the best measured aspects are the mass per unit length of the scintillator bars (also from docdb:6016), and the width of the fixture that holds 32 scintillator bars in place (technical drawing on last page of docdb:2822). Combining these yields  $1.754 \text{ g/cm}^2$  of material in just the scintillator bars, with only 0.6% uncertainty, but it is 3% less than the MC and the default interpretation of docdb:6016. Adding the same layers of grey epoxy and Lexan to both doesn't change the discrepancy, but adjusting the size and spacing of the scintillator implies changes in the amount ( $\text{g/cm}^2$  of scintillator).

The technical drawing calls out a 0.3 mm allowance for glue between the strips to obtain the final strip pitch; if completely filled that represents another 0.5mm of clear epoxy total, of which our MC is currently using 0.2mm. Having scintillator planes physically smaller with the same strip pitch implies a corresponding increase in the volume of clear epoxy. This is trading HC material for HC material, so we understand this discrepancy primarily appears in the active/passive fraction uncertainty, not in the total material uncertainty.

**Summary table** The following Table 1 shows the different versions of the material budget: default from docdb:6016, MC v10r6 "Titan", MC v10r7p3+ and beyond used for testbeam analysis, and the alternate estimate described above, used in consideration of the uncertainties. As will be demonstrated in a section toward the end of this document, this MC reproduces the proton range-out sample better than the intrinsic uncertainties. MINERvA is currently considering an effort to perform destructive testing on a full-MINERvA scintillator plane, and maybe a test beam plane, in order to measure directly the area and mass of several samples from the middle of an as-built plane. The 1.5% uncertainty on an assembled plane given in docdb:6016 and the NIM is taken to be the final uncertainty for us too.

	6016	v10r6	v10r7p4	alternate
PS + Fiber	1.658	1.619	1.648	
PS + TiO2	0.154	0.160	0.154	
Strip total	1.811	1.778	1.801	1.754
Lexan	0.120	0.128	0.120	0.120
Grey Epoxy	0.085	0.114	0.085	0.085
Clear Epoxy			0.021	0.021 to 0.055
Total	2.024	2.020	2.032	1.980 to 2.014

Table 1: Different material assays and MC materials. The one used for testbeam analysis is the v10r7p4 version.

In addition to the overall material, we are interested in the active/passive fraction uncertainty or better active/total fraction uncertainty. In the geometry, the WLS fiber and optical epoxy is simulated simply as more polystyrene and is not considered “active” material. Likewise the PS+TiO2 layer is also not considered active. For a complete plane, including the wrapping and structural epoxy, the active/total fraction is 79.7% in the MC and the same definition gives a fraction 80.2% in docdb:6016. The alternate model, with less scintillator material, could have up to 3% smaller active/total fraction. If the default 6% TiO2 by area was really 9%, that corresponds to an error of 2% in the active/total fraction. If the passive materials in the MC really do produce light captured by the PMT, but (for example) at 20% the rate of the active materials, that is a 3% effect. For our analyses, we will consider a symmetric 3.5% uncertainty on this fraction. This is larger than the 2% uncertainty than is described in the NIM, but has no effect on comparisons of MC to data, and cancels out and is replaced by the total material uncertainty for calorimetry measurements.

Finally, we might have mild sensitivity to the amount of moderate-Z material on the boundary between scintillator strips, coming from the titanium coating, and the silicon and aluminum in the grey epoxy. Because the primary testbeam analyses don’t isolate electromagnetic activity very well, we think we are less sensitive to this compared to the intrinsic uncertainty in electromagnetic states in hadron showers. A followup run with electrons in the MTest secondary beam could address questions like this, if our small sample of 500 MeV electrons does not.

The simulation is very close to reality. At analysis time, we are able to make first order corrections for the material model, if new information is available or if one analysis reveals evidence supporting a change.

### 3 Mapping

All testbeam detector planes were mapped on the same instrument, with the same setup used for MINERvA detector planes, except the reduced size means there were four planes on the mapper at

once. In general, the planes had few glue voids or broken fibers, in part because they were smaller and easier to produce, and in part because the crew at William and Mary doing the assembly were more experienced by the time they were made.

The mapper reveals that the most common defect are glue voids at the ends of the strips. There are a few outright broken or missing fibers, occasionally smaller glue voids in the middle of strips, and more common voids at the end of strips. The first results of the mapper were used to identify a few planes with more broken or moderately bad strips than the rest. These were preferentially placed at the front or very back of the detector.

Between the mapper data and the cosmic muon data used for the S2S calibration, there are 22 substantially underperforming strips in the detector (less than 1%) scattered here and there. Planes 1, 19, and 29 have three or four each. Also, there is a group of eight, in effect a single connector, in plane 30 that has anomalously low yield. One-off glue voids in the middle of strips are about the same 1% of the total. The response from known bad strips are masked out in both data and MC.

The mapping data is also used to give a parameterization of the attenuation correction if the uniformity along the strip is good. Instead, for the testbeam, we always use the so-called “point to point” technique from the mapper data when we have 3D information about the location of the hit and regardless of the uniformity of response along the strip. This choice of technique does not seem to make a difference for our analyses.

There is a discrepancy that we call the longitude position (LPos) effect, which is described in more detail in its own section. It is possible that this discrepancy is related to some shortcoming of the mapping procedure. Maybe not the most likely reason, and anyway we don't really know, so let's not get into a contest about whose got the shaggier dog.

## 4 Daily calibrations

Once the testbeam detector had been assembled and instrumented, the calibration procedure consisted of taking one pedestal run and three light injection (LI) runs on a daily basis. Each run (pedestal and LI) consisted of two thousand gates. This procedure was almost always carried out at the end of the data taking period, typically around 6:00 P.M. when the secondary beam to MTest was discontinued.

After the pedestal and LI runs, the trigger was switched over from the tertiary beamline trigger to a cosmic ray trigger formed from the coincidence of two walls of scintillator, one upstream and one downstream. The cosmic runs typically continued until 4:00 A.M. the next day, when it was time to re-establish beam to MTest.

Because the DAQ also used the so-called “cosmic ray trigger” mode for regular beam operations, no other operational modification was needed, only to select cosmics as the beam type from the software menu to describe this kind of data. This kind of trigger was a post-gate trigger applied to a random gate. As soon as it was able to, the DAQ started acquiring and digitizing its 16 microseconds of data

and digitizing it, a process that takes fifty to a hundred microseconds total. If an external trigger is received before the end of digitization, then the data are transferred to disk, which takes a long time, milliseconds. If no trigger is received, the buffers are flushed and reset and as soon as it is able, the DAQ starts acquiring another 16 microseconds of data. Both the cosmic trigger and the beamline trigger were connected to the DAQ in this way.

In the event of a hardware swap of a FEB, which happened a few times, a pedestal run was taken. We also changed PMTs 11 and 19, in which case both pedestal and LI runs were taken after the swap, though this happened naturally anyway because it happened when we switched from the 20E20H to the 20T20E configuration.

## 5 Underperforming hardware or run conditions

The documents in docdb:5764 catalog examples of major hardware swaps and other suspicious incidents related to the shift workers QC checks and nightly processing of the data.

The PMT's for planes 11 and 19 were replaced going between the 20E20H and the 20T20E runs. For plane 11, this eliminated two hot pixels (31 and 32), and took out an otherwise sorta-high cross-talk PMT. Primarily the swap improved the cross-talk characteristics for plane 19.

In addition to those, planes 39 and 40 had higher than normal cross talk, but on purpose we put these PMTs at the end of the detector to minimize their effect on the analysis. Planes 25 and 28 were reported to have low LI response, but after applying all calibrations they appear to behave fine.

There is a day of bad data in the 20T20E configuration, run 277 subruns 4,5,6 through to run 281 (partway through the morning of 5 July 2010 to partway through the day 6 July 2010), because these runs were loaded with a bogus HV table. Beginning with run 282 the correct HV configuration was being loaded. These settings do not correspond to the gains from the LI runs on either side, so we do not have an appropriate calibration and so we do not use these data in the analysis. There were no LI or pedestal runs on the night of 5 July, and the one cosmics run 279/1 is also excluded from analysis. This problem was caught because it produced anomalous response visible in the tails of distributions of the range-out proton samples, but a check reveals there are no other instances of this problem in the data.

Possibly caused by this HV table issue or possibly not, on 6 July the FEB for plane 34 was reported to look strange in the data quality plots, and with holding voltage. There was a period of full hall access on this day, and the FEB was replaced. There remained some issues with this FEB until the firmware version was updated for run 286/1 on 7 July. These runs (from 282 to 285) remain in the analysis, though it is possible this one plane might have an odd response, and should be treated with some suspicion for analyses that could depend on it. Some time later, checking the gains, it was further determined to turn down the voltage on this channel, which happened in time for the start of run 305/1.

## 6 Pedestals

We took pedestal data as part of the calibration runs almost every evening before switching to the cosmics trigger. With no beam and no beam trigger, readout gates are open for  $16 \mu\text{s}$  which capture noise from cosmic rays, radioactivity, electronic sources, and the PMT dark current. After outliers due to real activity are eliminated (as described in section 6.1 of the MINERvA NIM), the mean and RMS pedestal is computed for every channel of every PMT using the rest of the data for this special run and is used to set the zero for the PMT response scale.

The next plots show an example of the pedestals for a single example PMT, one that is not too good but not too bad. It was an X plane installed at electronics address 0-1-2-4, which was plane twelve, PMT serial number 1412, FEB serial number 573484, and physically was the plane 41X- as labeled and shipped from William and Mary, according to our hardware list in docdb:5006.

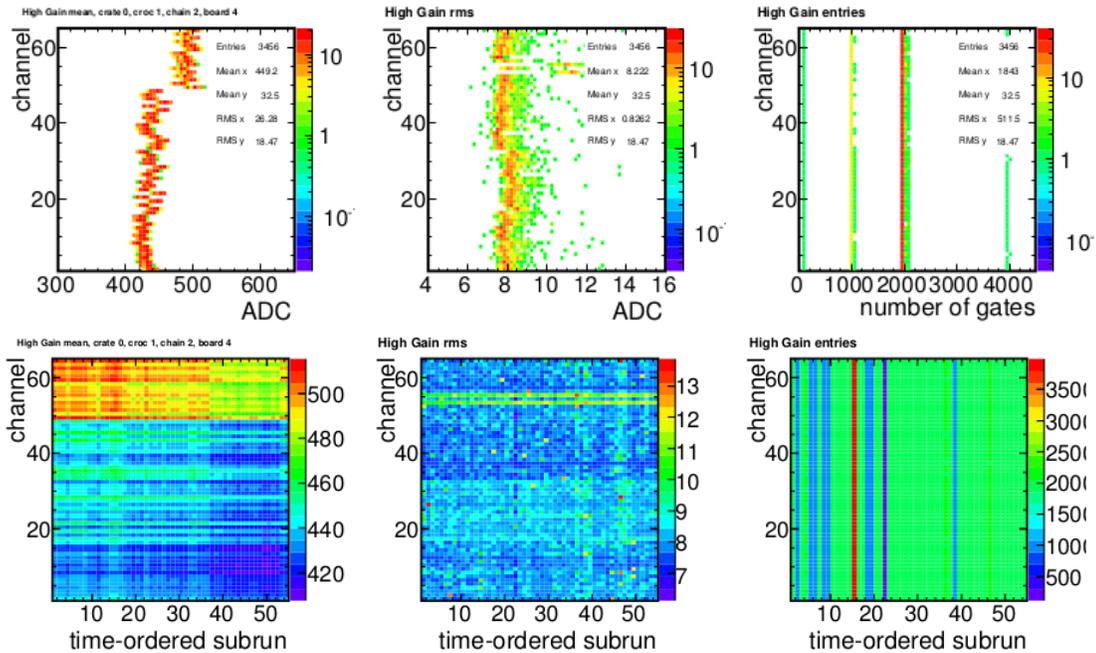


Figure 1: example of the distribution of pedestals for plane twelve of the detector. The top figure shows a histogram of the variation of the mean and RMS for pedestals across all subruns with number of entries as the color scale to show how they cluster around a stable value. The lower figure shows the same as a time series along the x axis where the color represents the mean or RMS itself. The vertical axis shows the 64 channels of the PMT. This PMT has a set of 16 channels with different pedestal, two channels with different RMS, and time stability of all channels that is typical.

When comparing testbeam and MINERvA pedestals, the mean and RMS values are not significantly different. Stability over time is comparable, but changes in the electronics configurations were

more frequent over the short data-taking period in the testbeam, more similar to the startup of, rather than the extended MINERvA data period. The visible change in color in the lower left plot of Fig. 1 at subrun 38 is the power off for the change between the 20E20H and 20T20E detector configurations.

## 7 Gains

The gains calibrations work identical to how they are done for the MINERvA detector in the NuMI beam. The light injection box fabricated for use in the test beam experiment is the same one that became the spare (?) for the MINERvA detector. Light injection data was taken most evenings, after beam running but before switching to the cosmic trigger. For comparison, the MINERvA detector data is taken with interleaved LI every ten subruns or so, such that it happens roughly three times a day.

The LI data with a one-PE spectrum was fit with a parameterization of the expected PMT and dynode chain response. The resulting measurements are used for an interval of validity (IOV) spanning from one measurement to the next, with a small number of exceptions where the LI data was applied to nearly in-time prior data.

While identical in technique to the gains calibrations for the MINERvA detector, one feature (not described in the NIM for brevity's sake, but some details are in docdb:5541 and Brandon Eberly's thesis) is much more important for the testbeam data. The LI data contains occasional noisy hits in the LI data, larger than could be created by the LI pulser. We do not know the reason, or whether this response is due to the LI system itself or due to other particles or the MTest environment, except that it is an even more rare problem for the detector in the NuMI hall. Though these are not common, the amount of charge for these hits is so large they significantly increase the width of the calibrated charge distribution, which in turn affects the gain measurement.

A technique is used to eliminate outliers in this distribution before calculating the gains. The digits in a particular channel are sorted in order of increasing charge. The algorithm then searches for the first occurrence of a consecutive pair of charges with a charge greater than the mean of the distribution as a whole (large deposited charge) and with a charge difference more than 300.0 fC (large gap between charge). For a typical PMT pixel, this is a difference of 3.75 PE, and is most likely located at the very end of the tail of the anode charge distribution. If such a pair is found, all digits with charges equal to or greater than this pair are removed, and the resulting truncated distribution is used to determine the gain.

Figure 2 illustrate the variation in the gains. The upper left plot is a distribution of gains extracted from nearly all pixels for the specific run 226, near the end of the 20E20H running, to illustrate the variation found within the detector. Upper right is the distribution of the change in gains from one run to the next, typically one day. When fitted to a Gaussian, the central peak has mean of -0.0025 and width of 1.07, meaning the change between runs is slightly more than the statistical error on the measurement. For comparison, the width of the MINERvA detector distribution is 0.97. A takeaway

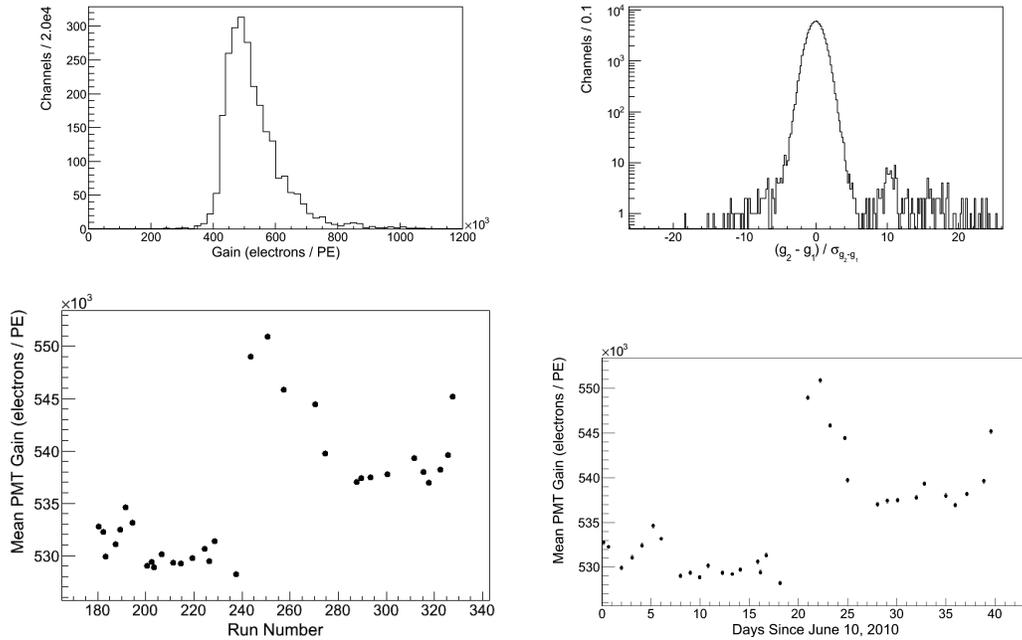


Figure 2: Illustration of the gains calibration and stability, see text for discussion. Upper left is the distribution of gains for nearly all channels of a single run 226. Upper right shows the change in gains, relative to the statistical uncertainty in the change in gains, from one calibration run to the next. The two lower plots show the average gain over all pixels for each run (the mean of the upper right plot) for every LI run with the run number and by day, respectively.

number from this analysis, the uncertainty for a typical gain measurement has a statistical error of 2% but should be 2.14% when both the statistical component and the anomalous component are combined. The entries in the tail of this distribution is primarily from the change in run configurations described next.

The mean of the distribution of the upper left plot is plotted in the lower left plot as the data point for run number 226. The two lower figures show this value for all runs, plotted vs. run number and the same data plotted vs. day since 10 June 2010. The abrupt change at run number 243 (day 20) is the change in configuration from 20E20H to 20T20E. This change involved a complete power off of the detector. When power is returned to all PMTs, the gains are higher at first. In the testbeam case, it takes three days of running for the gains to settle back into normal; for the MINERvA detector it typically takes one day. The resulting new normal is 1.5% higher than it was in the earlier configuration; planes 11 and 19 got new PMTs, while the rest of the PMT's remained with the same HV setting.

Because of the abrupt change, run 244 is excluded from the cosmics calibration samples. The first day of beam data, runs 245 to 247, is kept for analysis, but it is inspected for possible systematic effects on the analysis.

In addition to the PMT swaps, plane 34 shows up as unusual (and is omitted from the above plots) because it consistently has larger than normal gains. This behavior appeared in the QC checks during the run, and caused us to replace the FEB, though that didn't clearly fix the situation. When doing cosmics MEU tuning, we omit this plane, though it is kept for physics analysis. This plane is deep enough in the detector that it might not affect most of the physics analysis, though we will check for potential systematic effects.

## 8 Linearity

Because we have twice the light yield, the testbeam measurements will be more affected by non-linearity. Because of the nature of the data, we might be sensitive to it.

We expect the PMT to be non linear for short, high-charge pulses (high currents) due to the intrinsic saturation effect of such a PMT. This is guaranteed to happen, but when it happens, or how significant it is depends on the charge at the anode (which depends on how many PE at the cathode and the gain), and the duration of the pulse, and estimating this from first principles, or even a LI-style pulser test, is a difficult business.

There is no non-linearity effect in the simulation nor in the reconstruction as of this writing. We have an estimate from a simulation done by Howard to match work from the MINOS experiment, shown in Fig. 3. For the purpose of testbeam analysis our initial estimate is the upper curve, from Howard's simulation of a slightly longer-duration pulse than the MINOS examples also shown in the figure. In actuality, a single particle produces light at the MINERvA PMT from the combination of reflected and direct pulses, so the instantaneous current is probably roughly a factor of two smaller still.

The method to evaluate systematics due to non-linearity is to compute a non-linearity correction which can be applied to make the MC non-linear or remove the non-linearity from the data. It is constructed so that one can scale the correction away from the full correction, to whatever fraction is desired. Our standard uncertainty on this number for test beam analysis will be half of the default amount.

The implementation is done in TestBeamAna, though this is also a precursor implementation that others in MINERvA might use to evaluate a systematic. A quadratic is obtained from Howard's curve using the points at zero, 0.98 at 10pC, and 0.96 and 14pC. In TestBeamAna, every hit in a cluster has a non-linearity applied based on the charge, and then the cluster energy is resummed and compared to the default cluster energy, and the fraction (non-lin)/(default) is saved for use in analysis. This respects that, to leading order, the non-linearity is per hit (per pixel) but in practice a testbeam analysis uses attenuation corrected cluster energy. The scaling is done the usual way,  $\text{scaledcorrection} = 1 - \text{scale} * (1 - \text{correction})$ , and can then be applied to the MC. To correct the data instead, we should apply the inverse correction. There is a limit of 20%, which prevents extreme or unphysical corrections, and a flag that warns the piece of code that something unphysical happened.

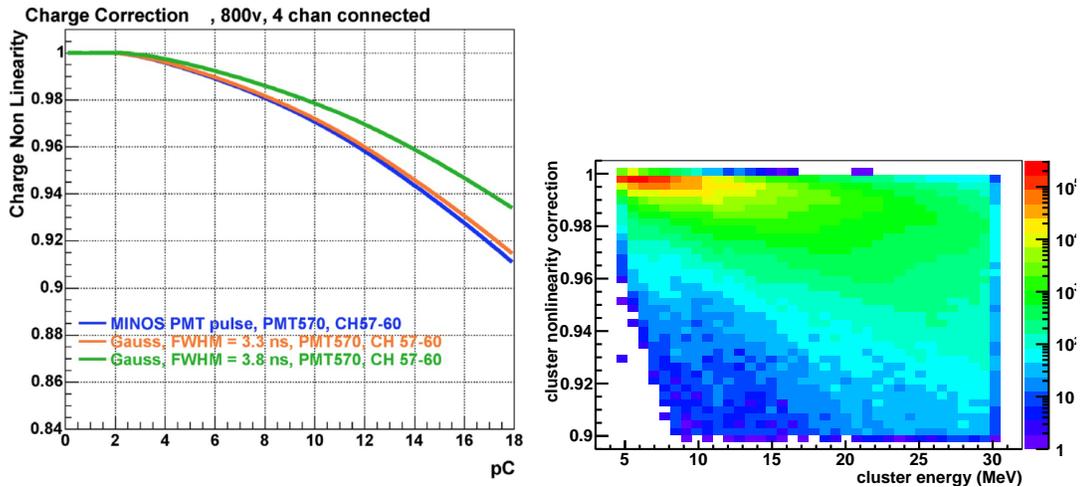


Figure 3: PMT non-linearity from Howard’s study docdb:7652 and docdb:9311. The testbeam implementation is the upper (green) calculated line. The default non-linearity is zero, and the default error for in-situ testbeam particle data is half the green line, “half” is to account for the direct and reflected pulses. As implemented, the 2D histogram on the right shows the modified cluster energy. It doesn’t follow a simple line because the (multiple) hits in each cluster can have very different energies and so very different non-linearities.

In addition to Howard’s plot, Fig. 3 has an example of the corrections applied at the cluster level from a testbeam sample that includes all species, interacting and otherwise. The scattering is because most clusters are made of multiple hits which get scaled differently, even if they add up to the same total cluster energy.

Using the range-out proton sample and the Birks’ analysis, we investigated whether we are sensitive enough to non-linearity that we can constrain it. The answer is that we may be able to see it, but between the statistical fluctuations of the data sample and the correlation with the Birks’ parameter, we can not produce measurement of this. Considering the  $\Delta\chi^2$ , we find cutting the effect in half (i.e. half of half of Howard’s line) is consistent with no non-linearity at all, and that the full effect (half of Howard’s line, same as the right plot in Fig. 3) is consistent with the best fit proton data at roughly the  $2\sigma$  level.

Less important, MINERvA has tested the linearity of the electronics chain starting with charge injection into the FEB, and is described in the NIM. From this we get parameters from a trilinear fit, which are later applied as a calibration that also removes the measured non-linearity, which is especially apparent about one-third way through the lowest gain setting on the FEB, hundreds of PE.

## 9 Muon samples

Several samples of muons were available to us for use in energy calibrations. Every evening, when we were through with our allotment of running with the beam, we changed the configuration to use a cosmic muon trigger. This trigger was made of thin scintillator panels (formerly installed for the SciBooNE muon range detector) at the front and the back of the detector. This is our largest muon sample, and the one we primarily use for calibration purposes.

Three times during the run, we put a steel beam stop to prevent a special 32 GeV tuned secondary beam from reaching the target and configured the trigger to use the downstream TOF and the cosmic muon wall. This accepts “beam halo” muons, which have a relatively flat energy spectrum from 15 to 32 GeV for muons coming from upstream pion decay. This halo is actually quite wide at the detector, though the TOF trigger localized it to select only those entering the center of the detector where most of the pions go. Finally, the same muon halo is present in the 16 GeV beam operation, and we can select those muons by looking out of time from our regular tertiary trigger.

### Cosmic muons

The cosmic muon energy spectrum is complicated, and has a strong angular dependence. To use it for calibration, we have used a MC sample to investigate its properties. The spectrum was a modification of Gaisser’s formula in the Particle Data Group document put together with source citations for MicroBooNE by Qing He and Kirk McDonald “Muon rate in the  $\mu$ BooNE TPC” (14 May 2009). These empirical modifications better describe the low angle and low energy dependence, compared to Gaisser’s parameterization.

This code fragment, from ParticleCannon.cpp, follows the prescription to modify the angle distribution following Chrikin [hep-ph/0407078] and add a modification to the spectrum for the lowest energies according to an unpublished report for Daya Bay by Guan et al. called “Muon simulation at the Daya Bay site” that was used in the He and McDonald MicroBooNE study.

```
\\For the angle modification by Chrikin
double Ca = 0.102573;
double Cb = -0.068287;
double Cc = 0.958633;
double Cd = 0.0407253;
double Ce = 0.817285;
// modified from simple costheta
double modcthz = sqrt( (cthz*cthz + Ca*Ca + Cb*pow(cthz,Cc) + Cd*pow(cthz,Ce))
                      / (1 + Ca*Ca + Cb + Cd));
if(chirkin) cthz = modcthz
\\for the energy modification by guan
double energyterm = pow( P*(1.0 + 3.64*CLHEP::GeV/(P*pow(cthz,1.29))), -2.7);
```

```

double gaisser1 = (1.0 / (1.0 + (1.1 * P * cthz)/(115.0*CLHEP::GeV) ));
double gaisser2 = (0.054 / (1.0 + (1.1 * P * cthz)/(850.0*CLHEP::GeV) ));
double constant = 0.14 / pow(CLHEP::GeV,-2.7); // convert Gaisser's formula to MeV
double dIdEdSA = constant * (gaisser1 + gaisser2);
if(guan) dIdEdSA *= energyterm;
return dIdEdSA;

```

When put together they will reproduce the Jokisch/Kiel-DESY data [H. Jokisch et al. Phys Rev D 19 (1979) p.1368] at 75 degrees incident angle down to about 2 GeV/c. The calculation overestimates the differential flux by about 30% at 1 GeV.

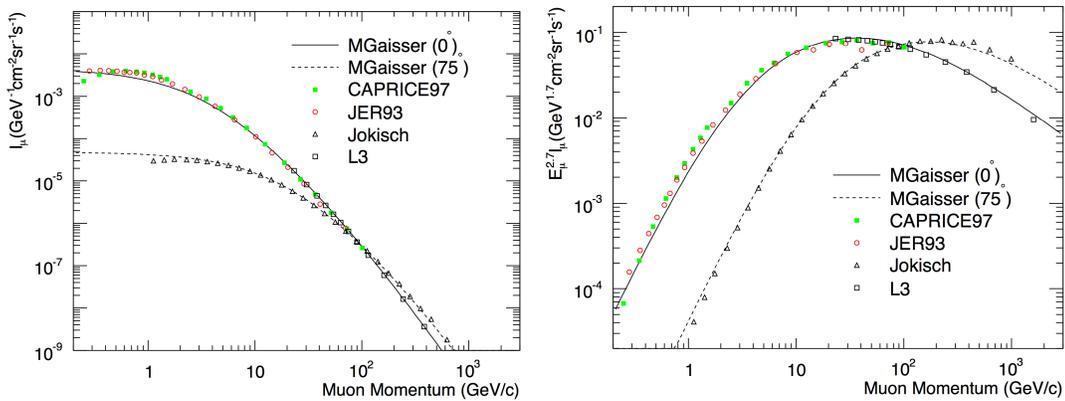


Figure 4: Spectrum of cosmic ray muons for two different angles, plots are screenshots from He and McDonald’s paper. The lines are from the full parameterization presented above, not just Gaisser’s simpler parameterization.

The agreement is not perfect, and the shape of the energy spectrum at low energy does have an effect on our calibrations. To evaluate an uncertainty, an extreme case is obtained by turning off the “Guan” low energy dependence, which has a factor of 1000 effect enhancing the 1 GeV muon rate.

To produce our simulation, we use the data driven ParticleCannon to take incident positions and directions from data muons that pass our selection criteria. These are assigned a random energy from the above parameterized spectrum. Typically we can reseed the random number generator and simulate hundreds of sets from the same muon sample. In this data-driven way, the resulting position and angle spectra respect the non-uniform acceptance of the cosmic muon trigger, both the vertical acceptance and the presence of a relatively inefficient vertical plane in the upstream muon wall.

There is a flaw in this method, the resulting simulated muons will not all pass the selection criteria again, preferentially the very lowest energy muons will be cut. But that also happened when we applied our selection in the first place; in doing so we are double counting the acceptance spectrum for some angles relative to other angles. We think this results in a bias smaller than turning off the “Guan”

dependence.

Despite these limitations and uncertainties, the LPos and MEU studies described below demonstrate that the spectrum simulated in this way does a good job describing the data, and is better than a naive mono-energetic spectrum or a spectrum devoid of the lowest energy muons, and better even than the parameterization without the “Guan” low-energy modification. The LPos and “Guan” modifications at the simulation level do not affect the MEU tuning significantly.

## Beam halo muons

We have two samples of beam halo muons. One is intrinsic to the beam we used to take data. The second is from three dedicated runs with a higher energy secondary pion beam and a beam stop to absorb the pions.

The second sample has a potential problem because the data is from just three different gain IOV’s, and so is subject to the full fluctuations of the gain calibrations. The first is more robust, but suffers from a pion and proton background in the data whose size is not uniquely known and is not in the MC sample.

This example uses the muons intrinsic to our hadron beam, but performs an analysis substantively identical to the MEU tuning. The disagreement in the PE plots is mild, and we are willing to believe it is consistent with a contamination of pions and protons passing the cuts. The same data propagated to MeV shows identical discrepancy, validating the latter portion of our energy reconstruction.

These beam halo muon samples have been analyzed roughly (in fact, they were our first choice for many calibrations), but have not been pushed to exquisite conclusions.

There is another high precision sample of protons which is used for cross checks, and is described later.

## 10 Crosstalk

Cross-talk scaling factors are determined for each PMT; in the test beam, there is one PMT per plane. To determine the scale factors, we consider the two largest hits in each plane from a cosmic muon as signal. Cross-talk hits are near to the signal hit in PMT pixel space. This space is a 2-dimensional 8x8 grid, totaling 64 pixels on each PMT. Nearest-neighbor pixels are the four which share an edge with the signal pixel. Next-to-nearest-neighbor pixels are the four which share a corner but not an edge with the signal pixel. The total cross-talk PE in an event is defined to be the sum of all observed PE on the nearest-neighbor and next-to-nearest-neighbor pixels of the two signal hits.

The cross-talk rate for each PMT is calculated in both data and simulation to be the ratio of the cross-talk PE to the signal PE, summing over the entire cosmic muon sample in data and a roughly equal-sized sample in MC. The scale factor for each PMT is then the ratio of the observed cross-talk in

MC to that in data.

The details of the analysis are in docdb:8025 and docdb:8148 and the Fig. 5 shows the difference between the default cross talk and the scaled cross talk.

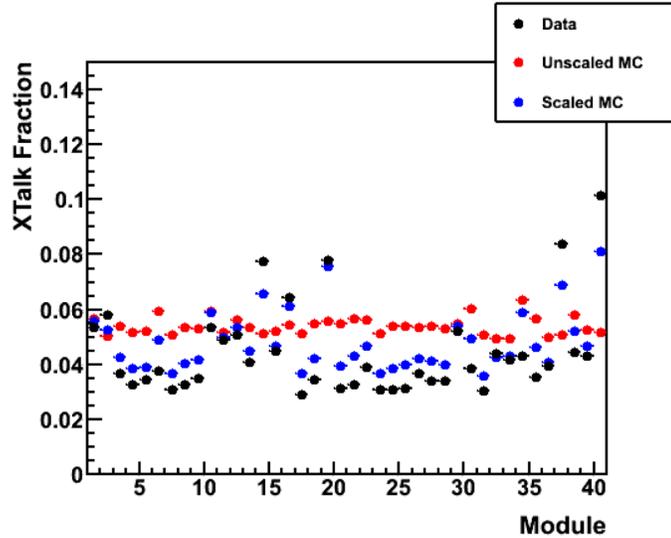


Figure 5: The cross talk in the MC was too high by default, for most pixels it is scaled down. The average for the first 25 planes is  $4.2 \pm 0.5\%$ , which will be subtracted from calorimetry analyses. The technique did not receive a final iteration, so the amount of cross talk is close to but does not exactly match the data, represented by the uncertainty.

We have a set of relatively high cross talk PMTs, because the better PMTs were prioritized for installation in the MINERvA detector in the NuMI beam. Some of the high cross talk PMTs went at the back fo the detector (planes 39 and 40), but according to this plot, some are sitting in the middle of the detector, planes 11 to 17 and plane 19. The PMT on planes 11 and 19 were swapped between 20E20H and 20T20E running.

On the other hand, we have relatively less cross-talk on average, when using this method, compared to the default  $\sim 5\%$  with no cross-talk adjustment. We think the mechanism for this has to do with the overall higher light yield and/or a difference in the distribution of PE's for cross talk.

The plots in Fig. 6 are a validation of the tuned cross talk in the proton sample. The vertical axis is the energy weighted hits in bins of distance from the proton projected path.

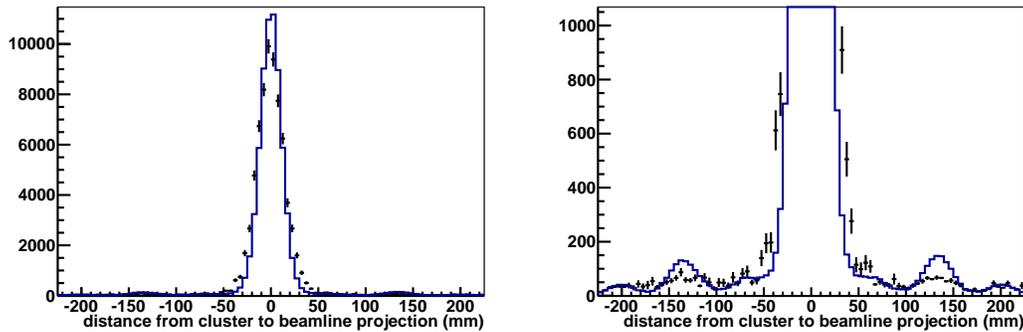


Figure 6: High level look at cross talk (and alignment). This is the sample of protons that appear to range out early in the 20T20E detector. The right plot closeup shows the MC overestimates the cross talk. The left plot shows that the position smearing in the MC is about half what it is in the data.

## 11 Temperature

Due to a combination of PMT and scintillator effects, energy response is expected to decrease with increasing hall temperature. Temperature changes in the underground MINERvA hall develop slowly. The temperature range experienced during the entire low-energy run period (January 2010 - April 2012) is about  $5^{\circ}\text{C}$  but over time periods of order a month or shorter the temperature is approximately constant. Thus, the average temperature dependence of the MINERvA detector can be absorbed into the time-varying energy scale calibration (MEU).

During the Test Beam run, the MTest facility experienced temperature fluctuations of up to  $10^{\circ}\text{C}$  in a single day. The hall was near its coolest when beam data-taking commenced at 4 A.M. and near its warmest at 6 P.M. when the beam was shut off for the night. Light injection data was taken immediately after the nightly shutoff, followed by an overnight cosmic muon run, which would typically begin before 7 P.M. and end at the start of beam data at 4 A.M. the next morning. Temperature changes occur too quickly to absorb into the MEU constant, so it is necessary to calibrate on an event-by-event basis using a hall temperature sensor downstream of the detector. The sensor used for the temperature calibration was located on the wall, about one meter north (downstream) of Module 40 and about one meter east of the east-readout PMTs.

The temperature calibration is performed plane-by-plane. Cosmic muon hits in a given plane are separated into  $0.5^{\circ}\text{C}$  bins based on the temperature recorded at the downstream sensor with a timestamp nearest to that of the gate. The temperature was recorded at approximately one minute intervals. For each bin, the truncated mean energy is computed iteratively; at each iteration, events falling between 20% and 200% of the mean from the previous iteration are considered. The resulting truncated mean is plotted as a function of temperature and fit to a straight line; docdb:7475 shows these fits for each plane.

The final temperature dependence is illustrated by these figures.

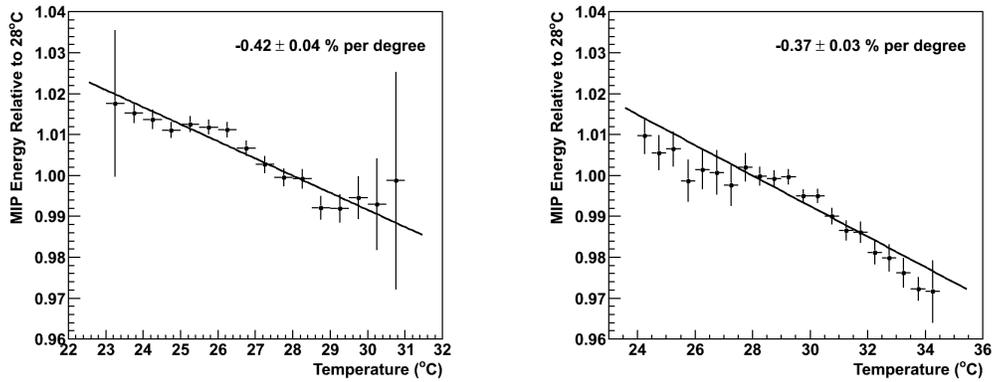


Figure 7: The temperature dependence of the energy response, based on the nightly cosmic ray muon samples. For illustration, these plots present the data integrated over all planes and after S2S was run. In practice the correction is specific for each plane and is applied before S2S is run.

## 12 Alignment

### Plane to plane alignment

For plane-to-plane alignment, refer to NIM article section 6.3. Only difference is it uses cosmic muons instead of rock muons. Within one plane, the 63 testbeam strips are treated as a rigid unit. The main offset that is corrected for is a translation in the direction of the measurement: for testbeam X planes that readout horizontally, this is a translation in the vertical direction. Rotations are also found and corrected for.

Typical shifts, relative to the nominal perfect geometry are [what, Chris will say]. The residual alignment parameter uncertainties are less than one mm and one mrad.

### Detector to beamline alignment

The plane-to-plane alignment using cosmic muons is sensitive only to relative alignment shifts and rotations between detector planes. The procedure is not sensitive to overall shifts or rotations of the entire detector. For the MINERvA detector, this is not important insofar as the angle of the neutrino beam with respect to the detector Z axis is known. For the Test Beam, projecting the beamline-reconstructed particle into the detector is valuable, and thus a millimeter-level alignment is required.

A sample with beamline-reconstructed particle mass between 1.0 MeV and 1.5 GeV is selected to include both the pion and proton peaks and reject rare events in the tails and most of the electrons.

Three-dimensional pattern recognition is not used; instead, the three detector views are treated separately and combined into an overall XY shift at the end.

The hit positions at wire chambers 3 and 4 are used to construct a vector for the particle trajectory. This vector gives the X, U and V coordinates of the track as a function of Z, assuming straight-line propagation with no multiple scattering. In each plane, the difference between the transverse coordinate (X, U or V depending on the plane view) of a cluster and the projected coordinate is plotted for all clusters within 50 nanoseconds of the trigger set by the beamline-analyzed particle. The resulting distribution is peaked at the true location of the plane in the transverse direction, with width due to multiple-scattering. This width increases with each plane, and only the first 20 planes are included. The distribution is fit with a Gaussian.

The planes are then separated by view. For each view, the mean of the Gaussian fit is plotted as a function of the plane's Z position and fit to a straight line. The linear fit is evaluated at module 1 to determine the overall shift for each coordinate, and converted from XUV to XY. The slopes represent rotations of the detector coordinate system with respect to the beamline coordinate system. The position and direction of particles at wire chamber 4 is adjusted in MTestDataAlg to account for the detector-beamline misalignment.

The following figures from docdb:7960 shows an example of the first extraction of the offset and rotation of the two coordinate systems. Subsequent changes to the beamline reconstruction required iteration and smaller adjustments than this.

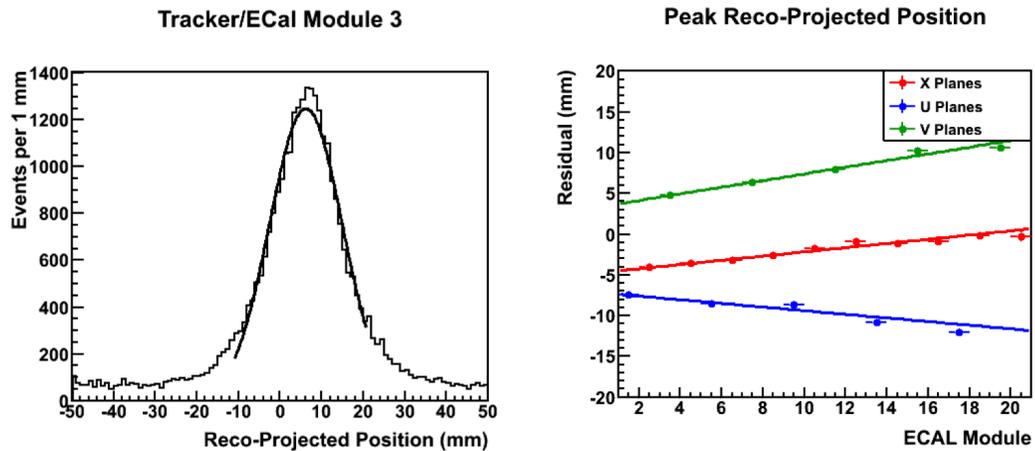


Figure 8: the alignment of the beamline and detector coordinate systems. Left plot is an example of the residual for plane 3 V-view plane showing an offset in the direction the V planes measure. The right shows the trend at different depths in the detector for all three views. From the latter, the offset and rotation of the two coordinate systems is obtained and applied to convert all beamline quantities to the detector coordinate system.

## 13 Strip to strip

Refer to NIM article section 6.4, the difference is that the testbeam uses cosmic muons instead of rock muons. Variations in light level between strips can be caused by bubbles and voids in the epoxy used to fill the fiber hole, or by couplings between the optical fibers and PMT. Strip-to-strip is done after temperature correction, and the time interval spans each detector configuration. Also, the innovations that went into the modern S2S technique was substantially developed using testbeam data, and is documented in greater detail in a technical note docdb:7601.

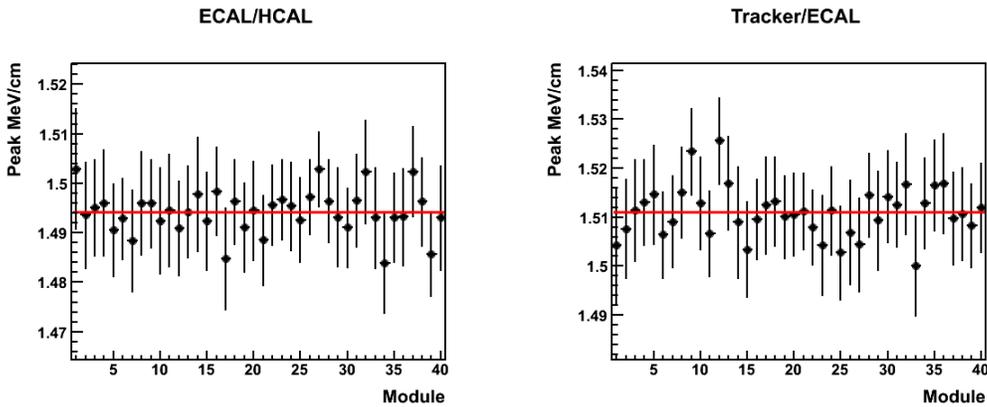


Figure 9: Validation of the final S2S adjustments constants showing the peak MeV from distributions that include all strips in a plane. As designed, the resulting response is the same (the error bars are showing the error on the fit are larger than one expects...)

Initial iterations use a truncated mean of the dEdX distribution in each strip. For better accuracy, the final iteration of the S2S calibration uses the location of the fitted peak of the dEdX distribution.

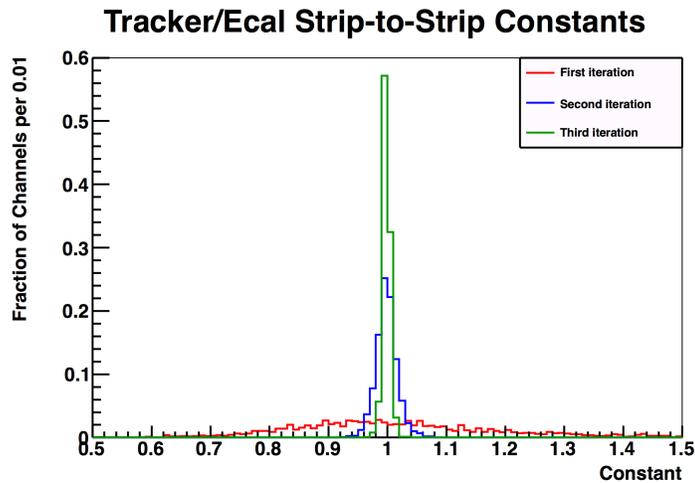


Figure 10: evolution of the S2S constants from first to third iteration. Out of the box, the variation in response from strip to strip is tens of percent.

## 14 Transverse and longitude position effect

If all is working well, strip to strip calibration accounts for the variations in the scintillator+PMT+FEB response and (when a 3D position allows for attenuation corrections) the mapper data accounts for the longitude variations in the scintillator response. In such cases, the calibration makes the response for the whole detector uniform.

In reality the calibrations themselves have a dependence on the location in the detector and we rely on the MC, much more so than in the MINERvA detector calibrations, to take out this dependence. Further there is a noticeable though not necessarily large discrepancy that remains after we are done.

The muon samples at the core of the calibrations have two kinds of geometry dependence. The simplest one to understand is that the subset of beam muon sample, triggered using the TOF scintillator, illuminates the central part of the detector, leaving little muon data to calibrate the edges of the detector.

The more important and complicated one is that the cosmic muons, triggered with upstream and downstream scintillator walls, have a feature where the planes at the edges of the detector exhibit a significant position dependence in the energy deposits coming from the ends of muon tracks for the lowest energy part of the cosmic ray spectrum. This feature manifests itself as both the transverse position (TPos) and a longitude position (LPos) dependence, and is modeled relatively well by the MC. The plots in Fig. 11 illustrate its magnitude.

To the extent that the MC does not match the data exactly, the calibration procedure will produce a bias in the S2S constants as a function of transverse position. Interestingly, we expect the LPos effect will not produce a significant relative bias among the S2S constants except where some other asym-

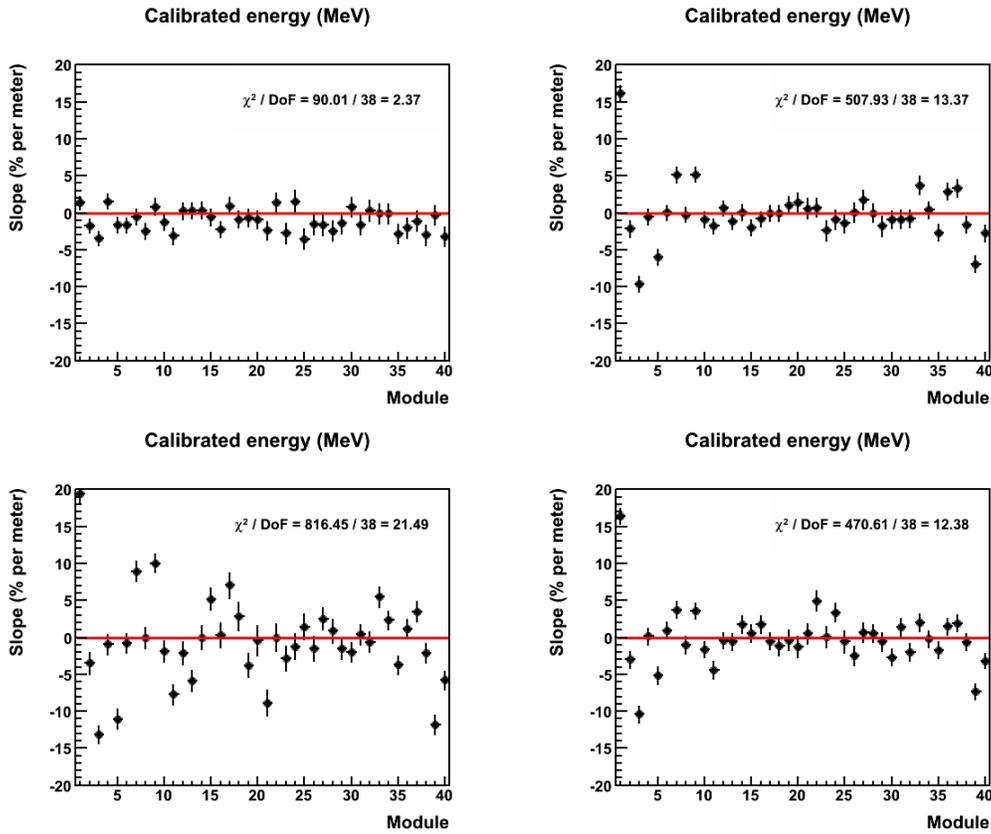


Figure 11: Four plots illustrating the intrinsic Lpos effect modeled by the MC for fully tracked muons. The vertical axis is the extracted slope of the  $dE/dx$  spectrum vs. longitude position after attenuation corrections. It is expressed as percent per meter, which is in effect percent difference between the PMT end and the mirror end. Upper left is MC with a 5 GeV horizontal muon spectrum, which (as expected) clusters near zero meaning no LPos effect. Upper right is the MC with the data-driven cosmic muon spectrum which shows a distinct LPos effect, especially in the UV planes near the front and back of the detector. Lower right shows the cosmics data, which shows a very similar LPos effect. Lower left shows the cosmics MC with the “Guan” low energy correction turned off, which does not describe the data very well, and can be taken to be one kind of extreme case for testing systematics, and reveals the source of the LPos effect is the low energy part of the spectrum. From docdb:8523.

metry comes into play. An example of such an asymmetry is the inefficient piece of scintillator in the cosmics trigger wall, though because we use a data-driven cosmic muon MC even that is accounted for in the simulation. [Some conclusion from the S2S study about how big these get] After the S2S calibration and the attenuation corrections are complete, we can reanalyze the samples. As designed, the S2S calibrations take out all transverse position effects within some residual tolerance, but we observe a discrepancy in longitude position equivalent to 3% per meter dependence, essentially consistent

across all planes.

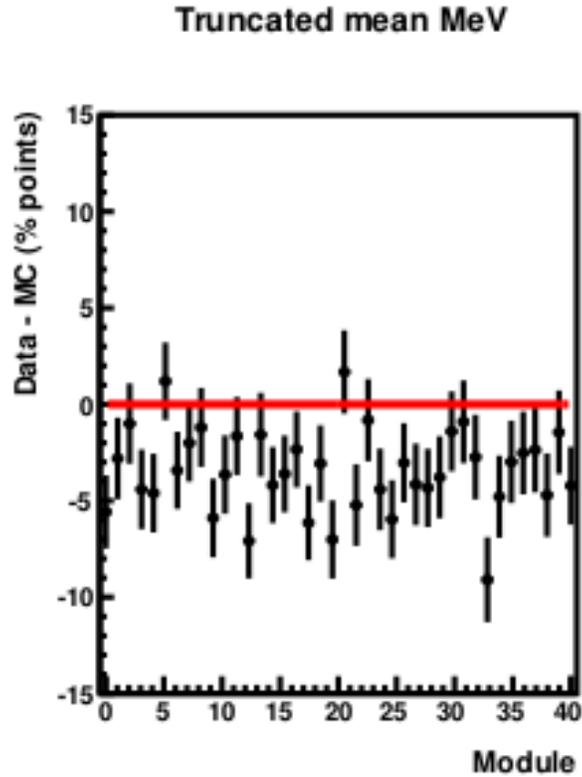


Figure 12: The residual LPos effect, the difference (data - MC) of the data shown in the previous figure. The data is 3% more negative than the MC, meaning the energy is systematically higher at the readout end compared to the mirror end. From docdb:8614.

We do not know the origin of this discrepancy, whether it is an artifact of the calibrations and attenuation corrections or something intrinsic to how the mapper information is not the same as the as-installed detector response. (Is this right – does the LPos effect only show up in samples where we are applying attenuation corrections? other guess?)

If this were a classic Icelandic saga, this would be where the Vikings are discussing whether their dogs' coats are ready for the long winter, and there would be a lot of looking around at other people's shoes, until the discussion moves on to another exciting voyage or battle.

This remaining LPos bias can have two kinds of effects on analyses. The sample of protons and pions has a correlation between horizontal position and particle energy because lower energy particles are bent more and enter the detector further to the west (and directed closer to normal to the detector). For many analyses, this effect is mitigated by the fact that we read out the planes in an east-west alternating pattern, and is half as large as it could have been. The other effect is that large samples of

tracks will have an additional smearing of individual cluster  $dE/dx$  response, on top of all other forms of fluctuations. To understand how these might affect an analysis, we have some tools that will artificially apply the residual shown in Fig. 12 of typically 3% per meter (or more) LPos effect to each plane.

## 15 Channel to channel smearing

The initial stage of the Muon equivalent energy tuning takes the  $dE/dx$  distribution from muons (a Landau-like distribution) and compares the data to the MC to determine an all-detector overall scale.

Despite incorporating variations in gains, pedestals, and other effects described above, the MC Landau distribution is narrower than the data by a noticeable amount. There is some additional variation that is not being accounted for. Because it is small, it is adequate to implement an ad hoc smearing on a channel by channel basis. The amount of smearing in the MINERvA detector is 5.57%. This smearing is applied randomly to channels in calculating the resulting PE for a hit, so a change here affects the PE and MeV (or MEU) distributions.

To measure this in the testbeam, we have three measures which are not uniformly consistent. The first is to use the MEU muon sample, as is done for MINERvA, which yields a larger amount of smearing of around 10%, shown in Fig. 13. A close inspection of the upper left plot and the bottom plot reveals that even the best smearing using the c2c method does not perfectly reproduce the smeared Landau shape, nor even as well as the method does for rock muons, though its reasonably close.

The range-out proton sample also provides a way to estimate the smearing, and the Birks fitting demands an estimate as part of the fit procedure. This has been studied two ways, once with two MC samples with different c2c smearings, and once with a simple implementation of additional smearing added to clusters.

We generated a MC sample with zero and 10% c2c smearing, and compared the widths of the  $dE/\text{plane}$  distribution near to but not at the end of the range-out proton sample. Interpolating between these two, a value of 5.75% is suggested, very close to the value used in the full MINERvA. If you look ahead to Fig. 16, the data points are fits of a Gaussian distribution to the  $dE/\text{plane}$ . Those same fits yield a Gaussian width estimate for both data and MC. An example of this is the left plot of Fig. 14.

The second way is to start with the same 0% smearing sample, but add additional smearing at the cluster level in the AnaTuple stage before filling the histogram. This isn't exactly the same as rerunning the MC, but is fast and good for late stage analysis. Scanning through a range of additional smearings, a value of  $6.0\% \pm 0.5\%$  best describes the data, and is shown as the right plot in Fig. 14.

The interpretation is that the cosmic data sample has an additional level of channel to channel smearing in it that we do not want to simulate for beam analysis. This could be because the cosmic sample averages over the full width of the strips, and so the full LPos and attenuation effects. It could be related to the way the gain IOVs are defined and calibrations and temperature corrections are done.

The other two estimates, 5.75% and 6% seem to be identical, except they describe two different

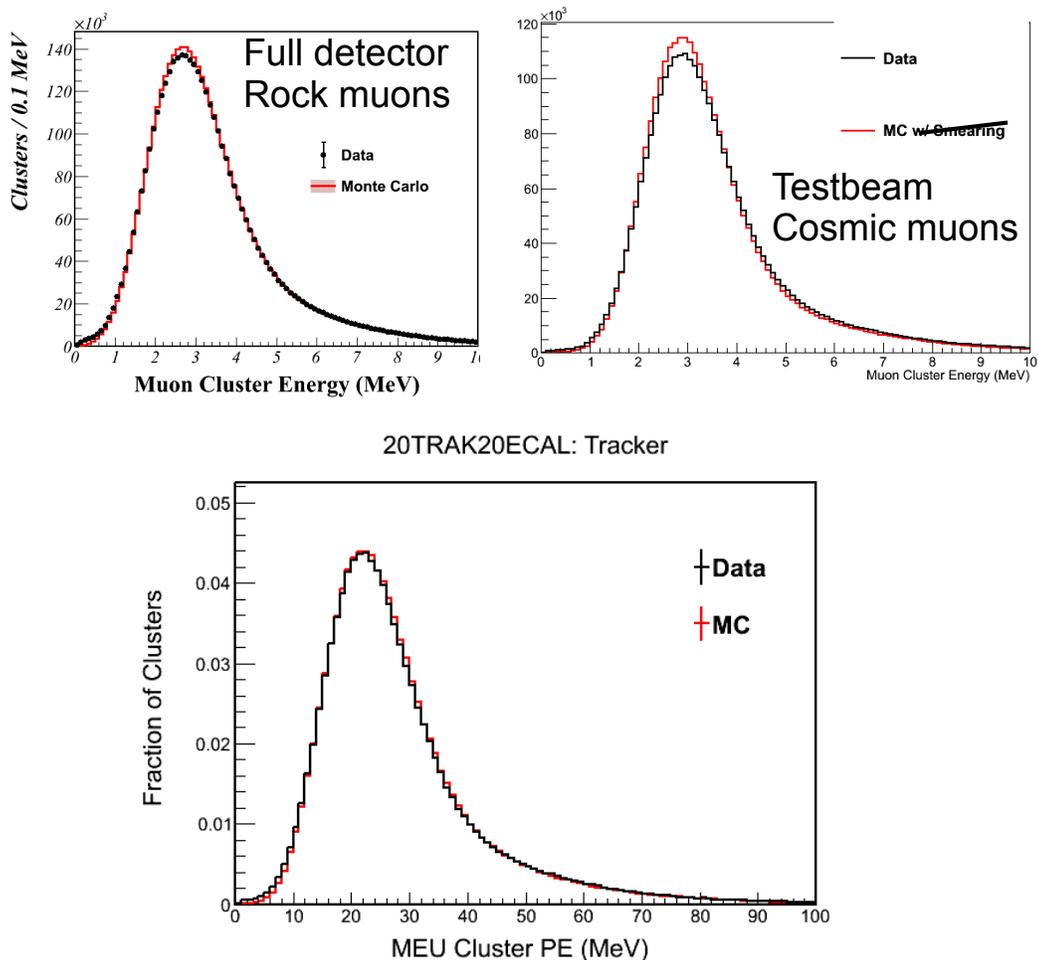


Figure 13: upper plot: evidence for adding channel to channel smearing to the MINERvA MC (left) and the test beam MC (right). lower plot: the MeV distribution from the 20T20E sample with 10% c2c smearing.

means of smearing, are qualitatively slightly different, and have different utility in that the latter can be applied at the cluster level in the analysis stage, while the former needs to be put in the front end. The qualitative difference can be described: the 10% c2c smearing and 9% cluster smearing are similar to Fig. 14, except the latter has half-error-bar larger widths in the few points on the right.

To get the smearing right for the Birks analysis, we have used the cluster level smearing. It would be possible to do a final iteration with the c2c smearing put in the front end as a check, but we are confident that there would be no effect on the analysis.

For the simulation of beam events for calorimetry and all other analysis, we will use 5.75% c2c smearing at the front end. An analysis that might be sensitive to the amount of smearing can inspect excursions from this by adding 1% or 2% using the cluster smearing method to gauge a systematic.

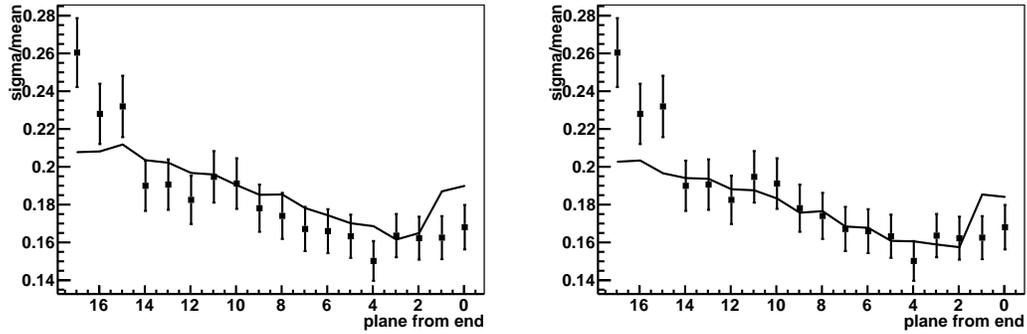


Figure 14: Left plot is the Gaussian (width/mean) ratio from a proton sample while the MC is one that has 10% c2c smearing added from the beginning. Right plot is the same, but has 6% cluster smearing added to a 0% c2c MC sample. The underlying data is an average over many physical planes, except for the three data points on the far left (which are not included in the analysis). The data points on the far right are affected by the Birks parameter and are also not included.

## 16 Muon equivalent energy unit tuning

The absolute energy scale, and just as important, the energy scale of pions and protons relative to muons is the final step in the calibration chain. The procedure is essentially identical to Sec. 6.5 of the NIM paper, but using cosmic ray muons instead of rock muons, plus other differences noted below. The procedure results in two numbers, light yield factor (LY) and muon equivalent unit (MEU). The LY converts from (post Birks constant, S2S, attenuation) photons to PE, and is obtained by matching the PE in the MC to the data. The muon equivalent energy unit (MEU) in units of MeV/PE is obtained from the MC using the the Landau peak of the dE and PE per plane distribution, and relating it to the truth energy deposited in the plastic scintillator (not fiber, not PS+TiO<sub>2</sub>). In addition, most energy scale systematics coming from the other calibrations described in the preceding sections are propagated to these two numbers.

The S2S calibrated data are compared to the cosmic MC with the channel to channel smearing applied. Good cosmic runs are used; run 244 is excluded because it was the first one after powering on in the 20T20E configuration, and the cosmic runs between runs 277 and 280 when the wrong HV settings were being used.

Because a single tuned MEU is averaged over the whole detector and then applied to the whole detector, but the whole detector is not used equally in the physics analysis, we omit some planes of data from the MEU tuning. Planes 1, 2, 39, and 40 are excluded because they are at the edge of the detector, where the cosmic sample shows more variation, and planes 39 and 40 are primarily used as a veto and not for physics analysis. Also plane 34 is excluded, because it showed consistently high gains and suspicious behavior. Planes 11 and 19 are included, though the bad channels on plane 11 are masked

out in data and MC, and the MEU tuning is done separately for the 20E20T and 20T20E samples where different PMTs were used for these planes.

## Energy scale systematics

The total uncertainty in the energy scale is one of the most important contributions to the absolute and data/mc relative uncertainty. It is evaluated in the context of the MEU tuning, and the systematics in modeling discussed throughout this document are propagated to this stage for evaluation. With a baseline MEU and LY (usually the best fit, but it could be a previous iteration), the relevant uncertainties are evaluated with systematically shifted MC or MC reweights.

New estimates for the energy response requires (usually) changing the MC, and asking what kind of shift relative to the default MC energy scale is found, not necessarily the MEU factor itself. The strategy is to make the shifted MC and evaluate how the PE peak, MeV peak, and the Ereco/Etrue change. Which combination of shifts are obtained determines the uncertainty on the energy scale. In practice, for small uncertainties, the most accurately determined of these three is the Ereco/Etrue correspondence which comes from MC only.

**Intrinsic fit uncertainty:** The peak finder for the MEU tuning uses a 5th degree polynomial fit as a function of  $(E - E_{peak})$ . The Minuit minimizer returns an uncertainty for testbeam of 0.2%. In addition, there is sensitivity to the range of energy near the peak used in the fit. Reasonable variations in this give changes of 0.4% in cluster PE and 0.2% in cluster energy. We include the 0.4% to represent the uncertainty. This applies to data/mc comparisons and the absolute energy scale both.

**Plane by plane deviation:** The MEU tuning can be done for individual planes, with a loss of statistical power. A single plane fit has an uncertainty of 2% due to fluctuations in the sample. Over all the planes, there are deviations of 3% from the sample average, but the RMS of those deviations is 1.4%, which is consistent with the individual fit uncertainties, and the error on the mean is consistent with the overall fit uncertainty of 0.2%.

**Channel to channel smearing:** MEU fits were performed with a range of smearing from 1% to 11% bracketing the central value of  $5.5 \pm 2.0\%$ . The uncertainty on the parameter that drives the amount of smearing is  $\pm 2\%$  coming from the evaluation of the test beam range-out proton sample, not the full range tested. The smearing goes into the MC prior to calculating the PE and in this method distorts the Ereco/Etrue value, more smearing pushes this higher.

**Cosmic ray energy spectrum:** To investigate the dependence on the cosmic ray energy spectrum, we turned off the low energy “Guan” modification, and found negligible change in the MEU tuning.

**LPos effect:** Because the cosmic muons used for the MEU tuning pass through nearly all the planes of the detector, they average over the different signs the residual LPos effect take, and the uncertainty comes out to be negligible. The main LPos effect is known to depend on the cosmic energy spectrum and acceptance and is modeled by the data-driven MC.

**Birks' parameter** Using MC samples with  $\pm 15\%$  shift in the best fit Birks' parameter gives an uncertainty on the MEU scale of 0.3%. This is consistent with a back-of-the-envelope estimate of the effect on a 3 MeV energy deposit. This applies to the absolute energy scale, but tuning the default MC with an incorrect Birks parameter will yield data/mc comparisons that are right for 3 MeV MIP-like energy deposits. The Birks' parameter uncertainty needs to be evaluated separately for any sample that has substantial non-MIP energy. This argument could likewise apply to non-linearity, except that it is negligible at 3 MeV effectively by definition.

**Temperature** This uncertainty is negligible for two reasons. We have a temperature calibration that takes out the leading order temperature dependence for both the cosmics and beam data samples. In transferring the MEU calibration from the cosmics to the beam data, the temperature range is roughly the same. This is much better than the opposite situation where we calibrated at a single temperature and apply the calibration to a range of temperatures.

**Whole detector** There is an uncertainty because the cosmic ray sample uses the whole detector approximately uniformly, while the pion sample does not. [Not sure if this is a sensible question to ask, or if it's implicitly what the LPos uncertainty is.]

**Consistency between detector configurations:** The MEU factor in the 20T20E detector is 0.9% higher than in the 20E20H detector, but the extracted light yield factor is not. We have not identified a reason. Some effects like scintillator aging should affect MEU and LY in a correlated way, other effects, like Geant4 modeling of muons passing through the Pb and Fe absorber would affect the Ereco/Etrue aspect of the MEU tuning. We also observed that the gains were 1% or so higher for the 20T20E for reasons beyond the change in two PMTs. This discrepancy is larger than all the above uncertainties. For the absolute uncertainty, we take the entire discrepancy as the uncertainty because we don't know its cause. For the data/mc sample, we assume that the MEU tuning procedure does accomplish a constraint, and take only half this, as if the correct MEU tune should logically be the average of the two.

**Consistency between cosmics and beam muons:** We have a second sample of muons, still independent from the pion and proton data. They are produced from pion decay far upstream, between the secondary and tertiary target. A selection of these chosen safely away from the triggered pions has been analyzed. The analysis uses only the 20E20H data, because a selection of muons that travel through the entire detector will have negligible pion background, which is not true of the thinner 20T20E data. [Results from Chris, and interpret.]

**Consistency within the pion and proton samples:** The range-out proton sample in the 20T20E shows a discrepancy where the data response per plane is 1.4% lower than the MC. However, this does not uniquely point to a discrepancy in the MEU tuning, because the energy deposits for this sample are typically between 10 and 25 MeV, so non-linearity (absent from the MC) combined with the remnant uncertainty in Birks' parameter could be the cause. Another test of pions that were selected

to have interacted very late in the detector and had their early, MIP-like hits analyzed was inconclusive, the sample selection (and muon backgrounds) allowed for a range from agreement with cosmics to data having 3% higher response. In any case, these are sanity checks, because we would not want to constrain our energy scale using the sample used to analyze calorimetry.

### Summary

systematic	absolute	data/mc	not applied
fit uncertainty	0.2%	0.2%	
fit range	0.4%	0.4%	
c2c smearing	N/A	0.2%	
cosmic ray spectrum	<0.1%	<0.1	
LPos effect	<0.1% ?	<0.1%	
Birks parameter	0.3%	N/A	
TE vs. EH	0.9%	0.4%	
Temperature			
Whole detector variations			
Single plane variations	NA	NA	0.2%
Sanity checks	TBD	TBD	
Total	1.1%	0.6%	
Active/total material	1.5%	NA	described below

Table 2: Systematic uncertainties in the energy scale, and the total.

## 17 Energy scale and active/total fraction

This is an energy scale uncertainty that has a special property for the calorimetry analysis, and we evaluate it separately from the ones described above. While the MEU tuning gives the MC the same photostatistics and single plane response as the data, great for data/MC comparisons, we are also interested that our experiment produces accurate absolute measurements. For most analyses, the final step is to convert from the cluster MeV in the (MC truth) active fraction to the MeV deposited in the whole plane.

The Erec0/Etrue conversion depends directly on the fraction of the scintillator in the MC producing photons, which has an uncertainty in the 3% to 5% range, and has not been seriously evaluated. This affects the absolute energy scale, but the tuning removes the data/mc relative uncertainty due to this effect, modulo a small second-order effect from different photostatistics.

On the other hand, the primary interest is a measurement of the calorimetric response of the detector to pions and protons. In this case, the first step is to apply a factor to that same MC to correct for the passive material. As long as this is what the analysis does, the relevant absolute uncertainty is the

smaller 1.5% uncertainty in the total material, not the larger uncertainty on the passive component. This particular uncertainty dominates all other contributions to the absolute energy scale, and is called out separately.

## 18 Cross checks

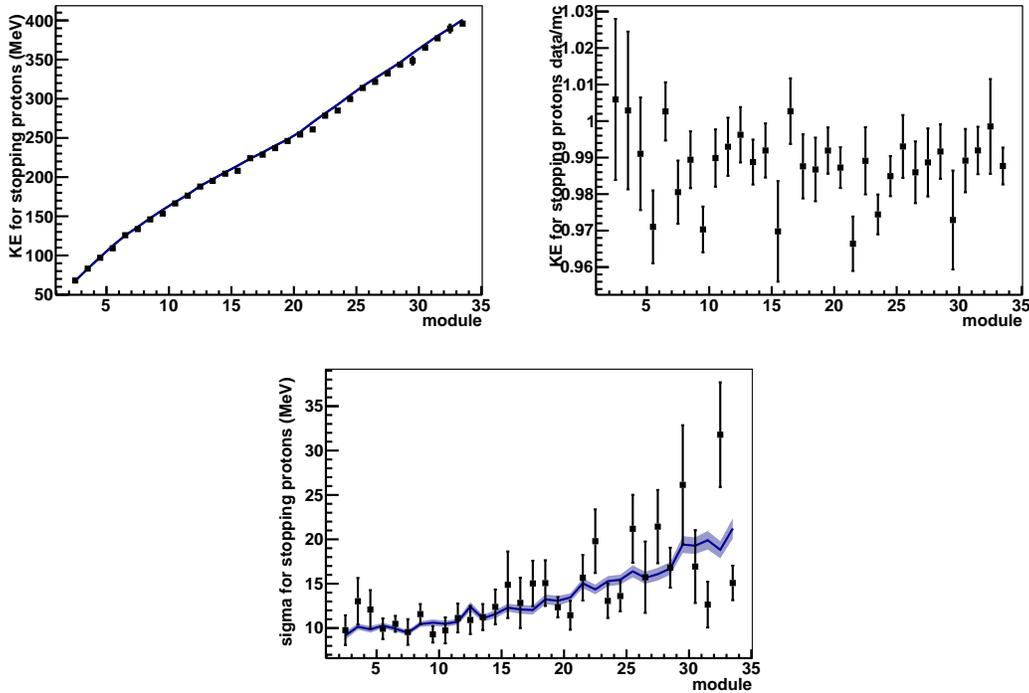


Figure 15: The mean kinetic energy for protons consistent with reaching the end of their range in a given module. First plot is the actual distribution, the second plot is the ratio of data/mc showing that the MC protons stop short by 1.3%. The third plot shows the width of a Gaussian fit to the peak of the KE distribution, and illustrates that the resolution of the beamline is well modeled.

The highest resolution beam sample is the range-out protons in the 20T20E configuration. That sample provides us with two important cross checks: proton kinetic energy vs. range and the average  $dE/dx$  between 4 and 14 planes from the end.

The range-out proton sample is formed from proton triggers by dividing the sample according to the last plane in the event and looking at the proton kinetic energy distribution for each plane. There is a peak that corresponds to the kinetic energy of the proton that would naturally range out at that location, and a high KE tail corresponding to protons that underwent an absorption process of some kind. By selecting events in that KE peak, a very high resolution sample is obtained. Full details of how the sample is selected and purified is given in the Birks' technical note TN037 docdb:9131.

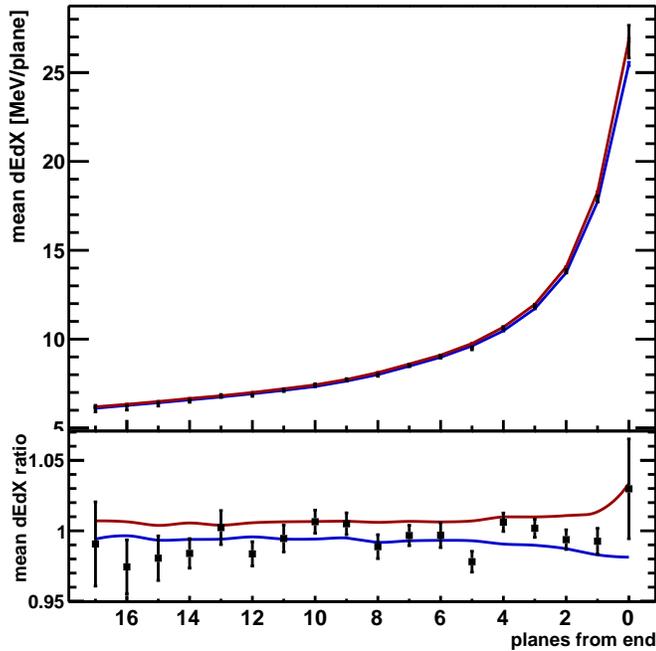


Figure 16: The  $dE/dx$  distribution for protons that are consistent with reaching the end of their range. The horizontal axis is the plane from the end of the proton event, so each data point is the mean of a Gaussian fit the peak of the  $dE/dx$  distribution that includes data from many physical planes. A constant fit to data from planes 4 from end to planes 13 from end, after correcting to the final  $c2c$  smearing and Birks parameter, are  $1.000 \pm 0.002$  (fit).

After all calibrations, it is observed that the protons in the MC stop earlier than the protons in the data; equivalently for a given end plane, the MC protons have a higher KE than the data by 1.1%. The physics bias effects that could cause this are the reconstructed energy from the beamline, detector material model, beamline material model, and the Geant4 proton energy loss model. The calibration of these quantities is such that the resulting uncertainty in quadrature is 2.2%, so an observed bias of 1.1% is less than the total uncertainty, though it is of the same size as the 1.5% uncertainty in the detector material model and larger than the 1.0% uncertainty in the beamline reconstruction. The effect is half the size in the 20E20H detector, hinting that the scintillator plane mass model is a more likely culprit than the others.

We do not take this to require a correction of any kind, but we will consider these possible sources carefully and comment on them in the rest of the analysis systematics. Final note on these plots, several U and V planes (but never an X plane) appear low in the ratio plot, we don't know if that is an artifact of an inefficient plane or geometry, or just chance.

The  $dE/dx$  distribution of these stopping protons, away from the end of range peak, is 1.4% higher in the MC than in the data in this plot [Need to update this plot? This 1.4% is the final number anyway].

Every data point is an average over many physical planes and good calibration IOV's. This calibration is a cross check of the MEU calibration, the relative energy scale between low energy protons and our spectrum of cosmic ray muons, and how Geant4 models them. It also picks up sensitivity to details of the technique because the spatial distribution of the protons and the cosmic muons is quite different, and the fit to the peak of the distribution is also different. Finally, though the non-linearity model we are working with does not predict an effect of this size at 10 MeV, this could be an indication of a non-linearity effect rather than a generic energy scale effect. The same analysis with protons ranging out early in the 20E20H shows a larger disagreement by yet another 1%: the MC is high by 2.5%.

Designing and analyzing an experiment with 2% systematic uncertainty is not so much a daunting task as an extended shaggy dog story.

## 19 Geant4 step size and MCHit aggregate

The calibrations above have been done using the standard MINERvA MC configuration. Most of the testbeam analysis, including Birks' parameter and calorimetry, is done using a configuration where the Geant4 step size is reduced for protons, neutrons, and nuclear fragments, and where the MC hit aggregator size is tuned to be quite small. This especially affects heavily ionizing particles at the end of their range, such as stopping protons and vertex activity near the neutrino interaction point. This tuning improves the fidelity of the simulation at the expense of processing time and about 20% increase in MC POOL file size. A brief summary is here, more detail is available in the Birks analysis technical note.

The options that steer the Geant4 simulation read in a configuration file called Simulation.xml. The default version of this file is a placeholder containing nothing that changes the simulation, but a second version of this file called SimulationRikSmallProtonStep.xml changes a parameter `maxAllowedStep="0.002*mm"` and specifies that this parameter takes effect only in the inner detector and only for protons, neutrons, and a variety of nuclear fragments. By default, Geant4 has an adaptive step size that looks ahead for a geometry boundary or the possibility that the particle it is stepping will encounter an interaction or scattering process, or reach the end of its range, and the step size will be set to the smallest of these possibilities. The smaller step size means Geant4 will compute energy loss and evaluate scattering processes more often.

The MCHit aggregator is a routine that Trung implemented. When multiple MChits are simulated in the same geometry volume like a single strip, they are combined into a single MC hit which is then passed to the Sim/Readout simulation of the detector response. This means we store fewer MCHits in the POOL files (about 20% savings) and spend less time simulating response to those hits (negligible time effect). This change is made in `GiGaSensDetTracker.cpp` directly (not an options file parameter at the moment) by setting the parameter `fMaximumHitLength` to 0.05mm and recompiling. The default value is an entire strip, 17mm, though there are additional rules that prevent some MCHits from being combined. For good measure, the size of `fMaximumHitSagitta` is also reduced from 1mm to 0.05mm.

In the bulk of a particles passing through the detector, this does not matter, but for the last few mm of range for a heavily ionizing particle, both of these matter at the few percent level. The resulting hits are passed to the code fragment that applies Birks suppression. If these step sizes are too big for a region where  $dE/dx$  is changing fast, then the Birks suppression is applied to some average  $dE$  over a large  $dx$ , instead of an appropriately large  $dE$  over a small  $dx$ . In this case, *not enough* Birks suppression will be applied, and the response observed in the simulated scintillator will be higher than it should be or would be in a higher fidelity simulation.

The effect is 4% in the last strip of a proton at the end of its range, and negligible everywhere else along the track. For proton calorimetry, the effect is between 0.5% and 1% depending on proton energy or the quantity and nature of any inelastic interactions along that track. There would be a similar few percent effect on vertex activity near the neutrino interaction point, from low energy protons and nuclear fragments. Finally, it has a systematic effect in the Birks fit, the smaller step size pushes the fit further from our default parameter toward an even smaller Birks' parameter, less suppression, higher MC response, in order to describe the data.

## 20 Every other side readout

One difference between the testbeam detector and the MINERvA detector is the every-other-side readout of the former. Four planes forming a UXVX set are read out on one side (beginning with east readout), with the next set read out on the opposite side. This was required so that the planes could be mounted with a spacing almost as small as the MINERvA detector, otherwise pairs of U or V plane PMTs would bump into each other. There are two vertical layers to the location of the X plane PMTs, so a given X, X' pair does not interfere with each other, but without every other side readout, the next pair would. This did not require substantially different plane designs, the U and V planes (and X and X' planes) are identical save for a flip and rotation.

There are several additional effects of this design. The software needs to know what planes get what rotation, so that the locations and strip orientation and attenuation calculations are correct. Once coded and validated, there are no significant problems, though it results in testbeam specific exceptions in the geometry and reconstruction software.

A more interesting outcome is that this detector is less sensitive to variations in the horizontal position of particles that might interplay with attenuation or geometry. The LPos effect mentioned above is an example, so is the fact that particles in the beam of different momenta travel through different parts of the detector at different angles. Significant, unsimulated LPos effects would result in anomalously high readout in some planes but low readout in the next set of four planes, so that quantities that average over many planes will have systematically less bias than they would if the detector readout was all on one side.

## 21 Conclusion

Most of the calibrations of the testbeam detector follow the procedure used for the calibrations of the main detector. Actually, some of the modern calibration procedures were first developed on the test beam detector.

There are significant differences in the use of cosmic ray muons for calibration, and the use of some aspects of the beam proton samples for validating the calibrations and understanding uncertainties. In the end, we have a calibration that gives uniformity to the reconstructed energy scale of about 3% across the geometry of the detector (LPos systematic), and about 1% for the relative data vs. MC scale, and 2% for the absolute energy scale for individual energy deposits. Analyses like calorimetry will have an additional absolute uncertainty from the beamline momentum calibration, material assay, Birks' parameter, as well as selection and background effects.

Looking at the saga through an Icelandic Spar, Odin's two ravens are visible. I see Hugin (thought) and Munin (memory) fly each day over the spacious earth. I fear for Hugin, that he not come back, yet more anxious am I for Munin.