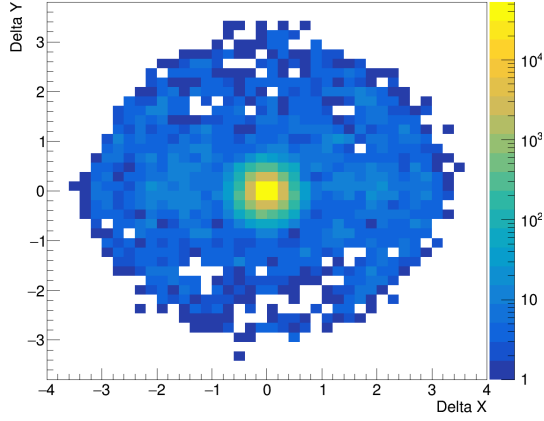


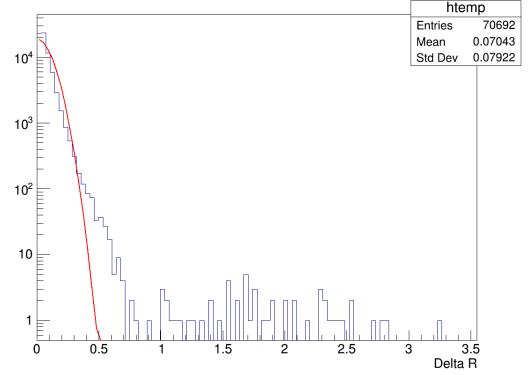
0.1 Preliminary Test and Motivation

The motivation for this work was to understand, refine, and document the *Disp* method of direction reconstruction. This method was expected to be useful to better resolve objects observed close to the horizon and facilitate studying the stability of the reconstruction to the zenith angle of observation. There was also the possibility of using machine learning algorithms to improve resolution beyond that of the standard geometric reconstruction. In particular, an improved angular resolution on the Crab may open up the possibility of resolving the spatial extent of the Crab Nebula with the VERITAS instrument.

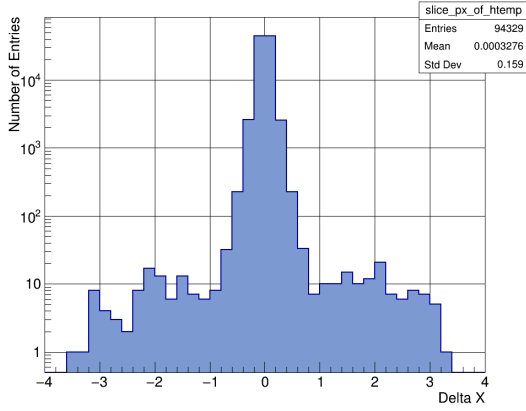
For the purposes of this work, the 68% containment radius in angular position (R_{68}) is used as a measure of angular resolution and performance of the reconstruction. This was measured in two ways - one numerically and one using a 2D Gaussian fit to the deviation of the reconstruction. From here (see Fig. ??) it was clear that the Gaussian fit was not an accurate representation of the underlying data, but could nevertheless provide some insight into it.



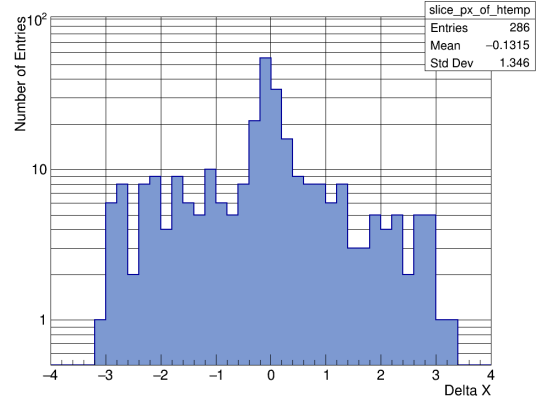
(a) A histogram of the 2D projection of the angular deviation of the reconstructed direction. The colors denote the number of entries in a given bin.



(b) The radial deviation of the reconstructed direction for a diagonal slice of the 2D histogram where $\Delta Y \in \Delta X + (-0.05, 0.05)$.



(c) The distribution (in ΔX) of the reconstructed direction for a horizontal slice of the 2D histogram where $\Delta Y \in (0.0, 0.2)$. There is a clear central Gaussian shape in the central 1° region of the plot, in addition to a shorter, wider Gaussian with amplitude $\sim 10^{-3}$ times that of the central Gaussian.



(d) The distribution (in ΔX) of the reconstructed direction for a horizontal slice of the 2D histogram where $\Delta Y \in (0.8, 1.0)$.

Figure 1: Reconstruction of the simulation direction using the *Disp* Method for simulations at 45° zenith angle with $10^3 \text{ GeV} < E < 10^{3.5} \text{ GeV}$. From (b) it is evident that a R_{68} based on the width of the best-fit Gaussian would not be an accurate measure of the R_{68} , and would consistently overestimate the resolving power of the method. However, it is also clear that the vast majority of statistics fall in the central region, and that the width of this region can provide a first-order approximation to the R_{68} .

Based on the distribution of the deviation of reconstructed direction of the simulations (Fig. ??) from the simulated direction, it was determined that a fit to the central Gaussian could provide some useful information about the R_{68} , and to this end, the 2D distribution (Fig. ??) was fit with the superposition of two Gaussians, with the 68% of this superposition as a first-order approximation to the R_{68} . Additionally, it was determined a numerical integral would provide a more accurate measure of this quantity and so the distribution of radial deviations was integrated from the tail inwards, until 32% of the events were counted and this value of the radial deviation from the simulated direction was used as a more accurate measure of the R_{68} ; a representative sample of the results is shown in Fig. ??.

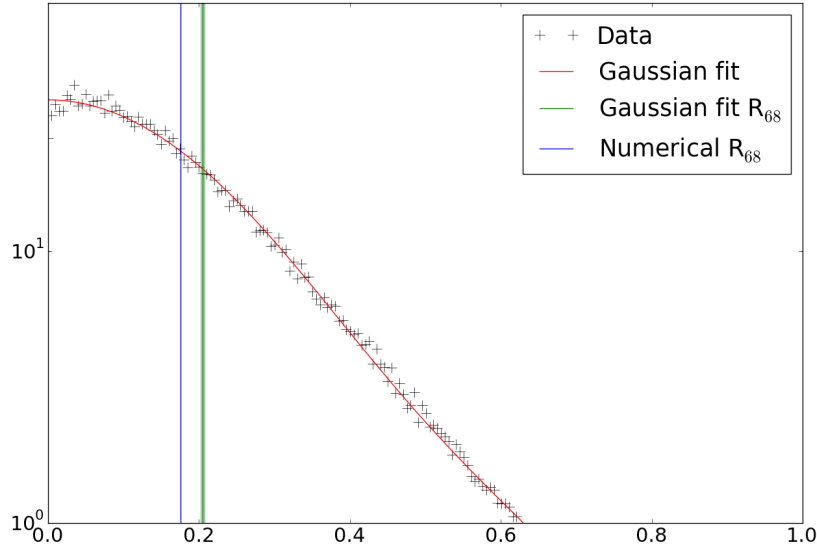


Figure 2: The R_{68} calculated for a simulation using a fit to a superposition of two Gaussians (green vertical line) and a numerical integration from 1° inwards (blue vertical line), overlaid on a radial slice for Monte-Carlo simulations.

A preliminary test for the R_{68} of the data was enabled by a set of runs where the Crab was tracked from horizon to culmination and back to the horizon (from the nights of Jan 12 2018, Jan 13 2018, and Jan 04, 2019). This provided a reference data set with high significance to track the energy and zenith dependences of the direction reconstruction in stable (and therefore directly comparable) weather conditions. The differential resolution

for the Crab data was measured in several energy and zenith bins and is presented in Fig. ??.

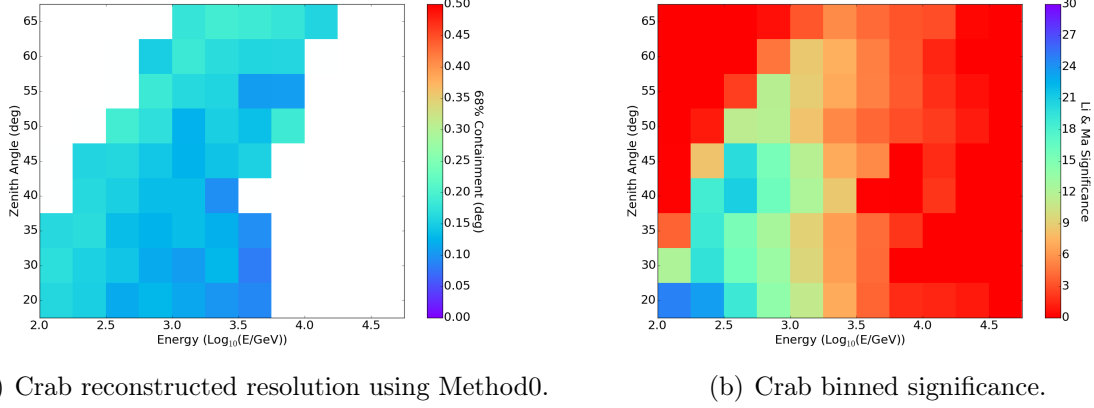


Figure 3: Reconstruction of the Crab direction using Method0 (standard geometric reconstruction from *VEGAS*). On the left, the regions of white denote regions where Li & Ma significance $< 3\sigma$ (for a description of this, see section ??).

0.2 *Disp* Table and R_{68} Dependencies

The BDT weight tables were generated independently of the old *Disp* method in order to have well-understood documentation of the underlying effects and dependencies.

0.2.1 Zenith Angle Dependence

A set of *Disp* tables was generated with a small sample of simulated events ($n \approx 1.9 \times 10^6$) across the range of zenith angles of interest ($20^\circ - 65^\circ$). This was compared to the regular *Disp* method. Since there was no record of the training sample size for the standard tables, this test sample was useful in determining the resolution of the *Disp* method with a relatively small computational footprint. Additionally, it allowed for some simple tests of dependence of the *Disp* tables on parameters not explicitly in the *Disp* tables.

These small *Disp* tables and the standard *Disp* tables were used to reconstruct simulation events and compared with Method0 (the standard method of direction reconstruction). In

both cases, the *Disp* method performs better than Method0 at the largest zenith angles ($\geq 55^\circ$, see Fig. ??-?? for R_{68} of the *Disp* methods divided by that for Method0 at the same zenith angle).

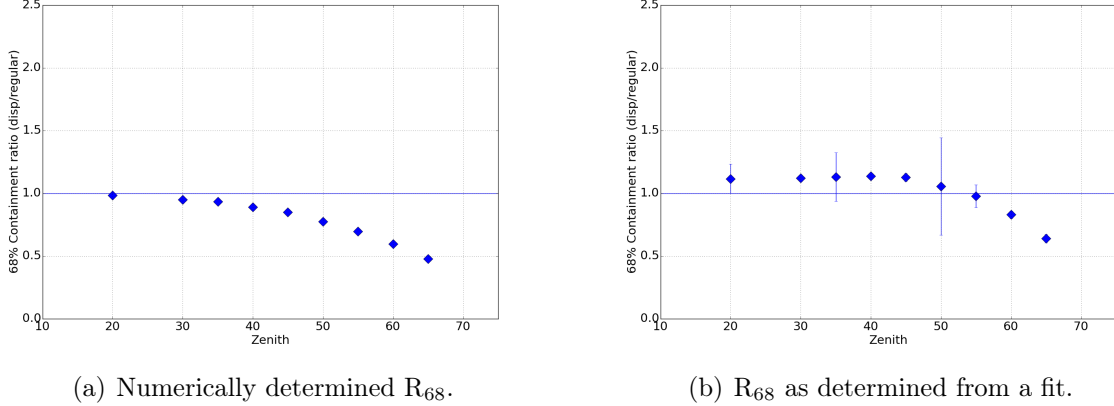


Figure 4: Ratio of R_{68} of the “standard” *Disp* table to that from Method0, with the numerically determined R_{68} (left) and that found from the fit to two Gaussians (right). Since this is the ratio of methods, the horizontal blue line denotes the performance using Method0; and the *Disp* method definitively performs better than Method0 for zenith $\gtrsim 50^\circ$.

From Fig. ??, we can see that the *Disp* method outperforms the standard method across zenith angles when using the numerical integration, i.e. the tails of the distribution from the geometric method are better behaved than those found using the *Disp* method. However, when using the 2-Gaussian fit, which provides an understanding of the central region where most events lie, the *Disp* method outperforms the geometric reconstruction only in the LZA range ($\phi \gtrsim 50$). This is considered the regime of interest for the *Disp* method.

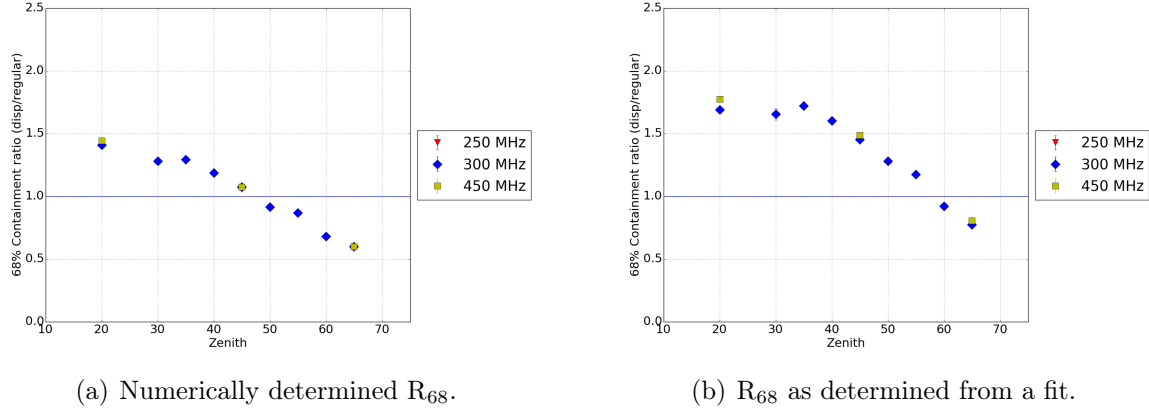
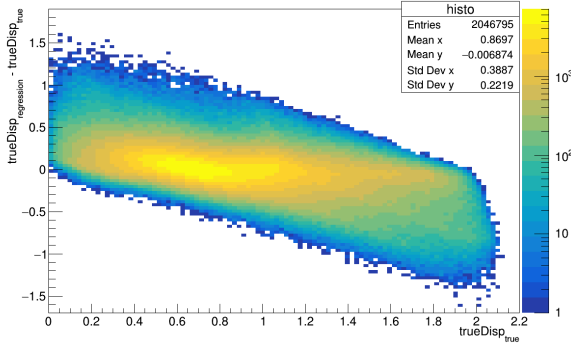


Figure 5: Ratio of the R_{68} from reconstruction using the small *Disp* table ($\sim 1.9 \times 10^6$ events all at noise = 250 MHz) and that from Method0 with the numerically determined R_{68} (left) and that found from the fit to two Gaussians (right). Note the horizontal blue line denotes the performance using Method0, so this method performs better than Method0 for zenith $\gtrsim 50^\circ$.

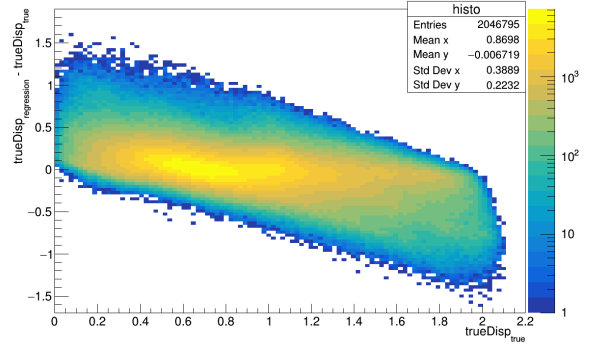
0.2.2 Over-training

BDT-based regression is quite robust under non-linear correlations between discriminating parameters. The primary vulnerability of this method is that to over-training - where the decision tree starts to be informed by noise and nuisance parameters in the training sample rather than relevant effects. This results in substantially different reconstruction efficiencies between training and testing samples. The ROOT TMVA package includes a test for over-training where it randomly selects a given fraction of the supplied events (for the purposes of this work, this fraction was taken to be 50%) to use for testing. These events are then not used to train the regression trees and are instead used only to generate a measure of the over-training. This check of the over-training for one of the test tables (noise = 450 MHz), shown in Fig. ??, demonstrates that there was no meaningful over-training of the table, at least based on effects present only in the training sample.

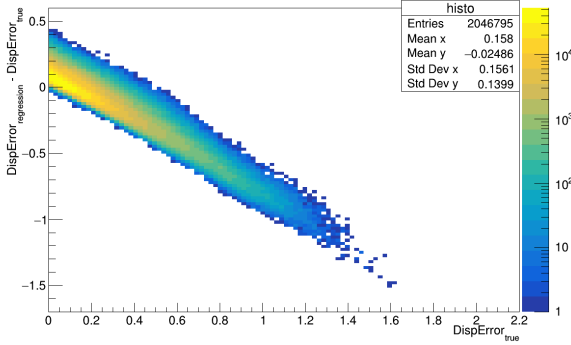
There remains however, the possibility of effects related to noise level that might appear in the training *and* testing samples (which are generated separately at each noise level), but not in observational data sets, which would be overlooked by this measure of over-training.



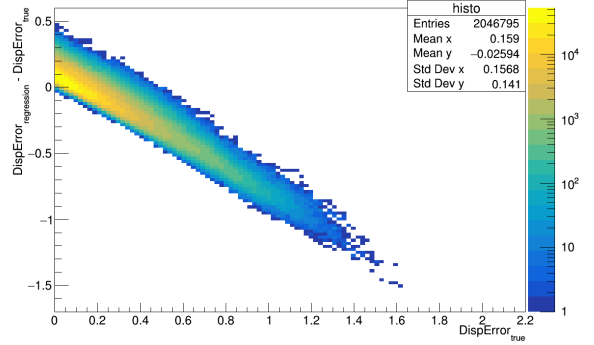
(a) Deviation of reconstructed *Disp* parameter from Monte-Carlo *Disp* parameter in training sample.



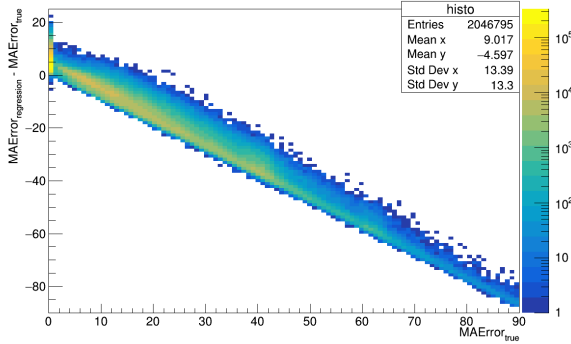
(b) Deviation of reconstructed *Disp* parameter from Monte-Carlo *Disp* parameter in testing sample.



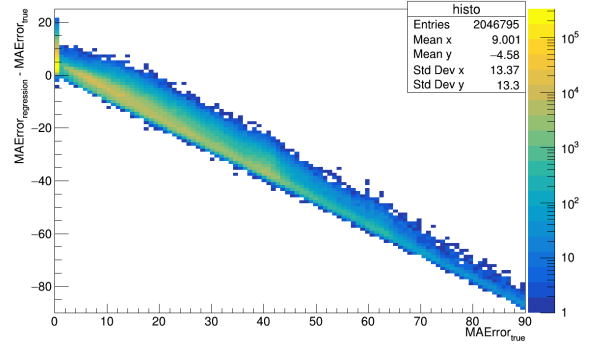
(c) Deviation of reconstructed *Disp* Error parameter from Monte-Carlo *Disp* Error parameter in training sample.



(d) Deviation of reconstructed *Disp* Error parameter from Monte-Carlo *Disp* Error parameter in testing sample.



(e) Deviation of reconstructed MAError parameter from Monte-Carlo MAError parameter in training sample.



(f) Deviation of reconstructed MAError parameter from Monte-Carlo MAError parameter in testing sample.

Figure 6: Over-training check on reconstruction using a *Disp* table generated at a single noise level. The left column shows the deviation of the reconstructed parameter from the true value in the training sample and the column on the right shows the same in the testing sample. The difference between the two columns is small which suggests there is little or no overtraining.

0.2.3 Noise Related Effects

The first set of *Disp* tables was also generated at a single noise level (250 MHz), allowing us to test the dependence of the resolution of this method (as measured by R_{68}) on noise level in the testing sample – some kind of noise-dependent effect would suggest over-training that would not be evident from the testing sample in the ROOT TMVA method since all the data provided to the package would have been at the same noise level. A comparison of angular resolution across noise levels revealed no significant dependence of the R_{68} on noise (see Fig. ??, ?? and ??).

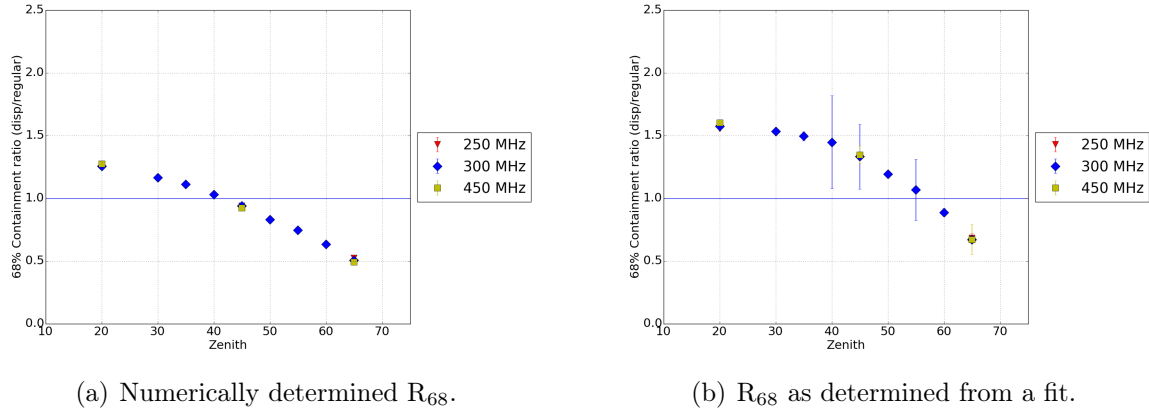


Figure 7: Ratio of R_{68} of the noise=450MHz *Disp* table ($\sim 2.1 \times 10^6$ events) to that from Method0 with the numerically determined R_{68} (left) and that found from the fit to two Gaussians (right). Note that the result is largely independent of the noise level in the simulation sample being reconstructed.

A second set of test *Disp* tables was generated using a single noise level (noise = 450 MHz) to test for over-training related to noise in the training sample (see Fig. ??). These tables performed slightly better than the first test tables and comparably to the standard *Disp* tables, but not significantly so. Since the noise-related effects did not seem to play a significant role in reconstruction, noise was dropped as a discriminating parameter for further analysis. Together, these two tests confirm that we do not expect the *Disp* method reconstruction to be dependent on the noise levels in either the training sample or the data being reconstructed and therefore this reconstruction should not be sensitive to different

NSB models.

To generate a set of *Disp* tables with better angular resolution (as measured by R_{68}), another set of *Disp* tables was trained on a larger number of simulations across zenith angles (as before) as well as across the noise spectrum. Since noise was determined not to impact the reconstruction, the events at different noise levels were used only to create samples with greater statistics.

0.2.4 Higher Statistics Tables

Once it was determined that there was no significant over-training in the small sample *Disp* tables, and the noise level had little bearing on the R_{68} measure of the reconstruction, it was determined that different noise level simulation events could be used as independent training events to have a higher statistics *Disp* table, and make small improvements on the statistical uncertainty on the reconstruction. The simulation data from across the noise spectrum and zenith range was used to generate a *Disp* table that sampled the entire parameter space more exhaustively.

A new set of *Disp* tables was generated (Fig. ??) with a training sample four times that of the initial test tables. The improvements in resolution due to change in sample size were modest, and confined to the range of zenith angles ($\phi < 45^\circ$) where the standard method outperforms the *Disp* method quite considerably.

As expected from the small statistical uncertainty on the R_{68} values, this increase in sample size did not lead to any meaningful improvements and a training sample of $\sim 2 \times 10^6$ was determined to be sufficient to achieve the desired resolution with small uncertainties.

0.2.5 Acceptance Correction for Offset from Camera Center

Showers arriving further from the camera center have a larger fraction of the shower arriving outside of the camera and therefore being lost. These showers are therefore reconstructed with a lower efficiency and resolution than showers arriving closer to the camera center. To compensate for this in the BDT training, so that the training sample does not mis-

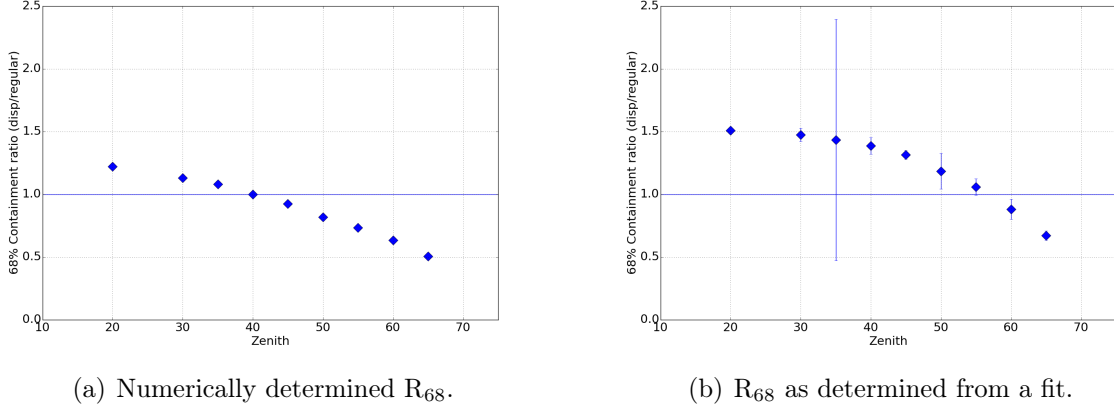


Figure 8: Ratio of R_{68} of the noise=450MHz *Disp* table ($\sim 8.4 \times 10^6$ events) to that from Method0 for a higher statistics *Disp* table with the numerically determined R_{68} (left) and that found from the fit to two Gaussians (right).

characterize the overabundance of events closer to the camera center as an anisotropy in incoming gamma rays, we fold in an acceptance correction by assigning a larger weight to events that are further away from the camera center.

The acceptance correction used here affects the training sample and therefore might be assumed to affect the resolution in a zenith dependent way, perhaps explaining the difference in performance between the new *Disp* tables and the old ones. This was tested by using a number of different correction functions in the training sample, as shown in Fig. ??

These tests reveal small changes in the R_{68} value despite large changes in the correction function (Fig. ??). This suggests that the *Disp* tables are not sensitive to changes in acceptance and therefore the different acceptance functions are unlikely to be the reason why the new *Disp* tables perform worse at smaller zenith angles.

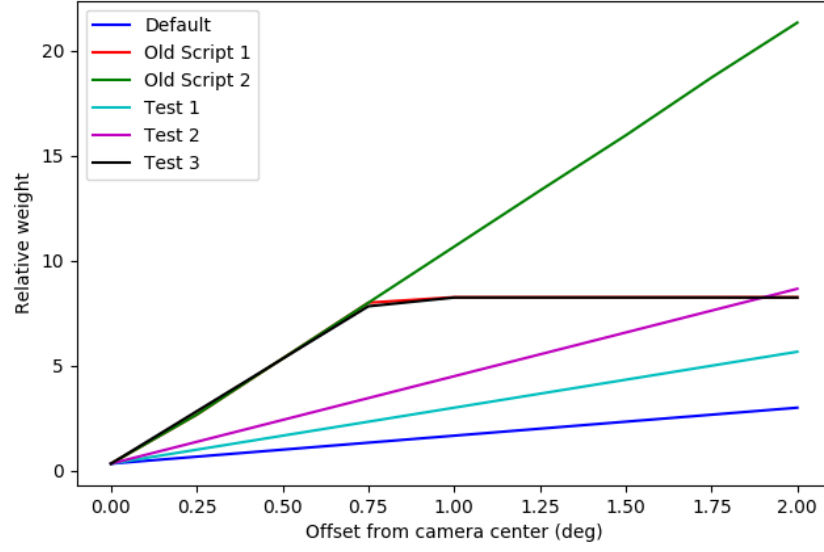
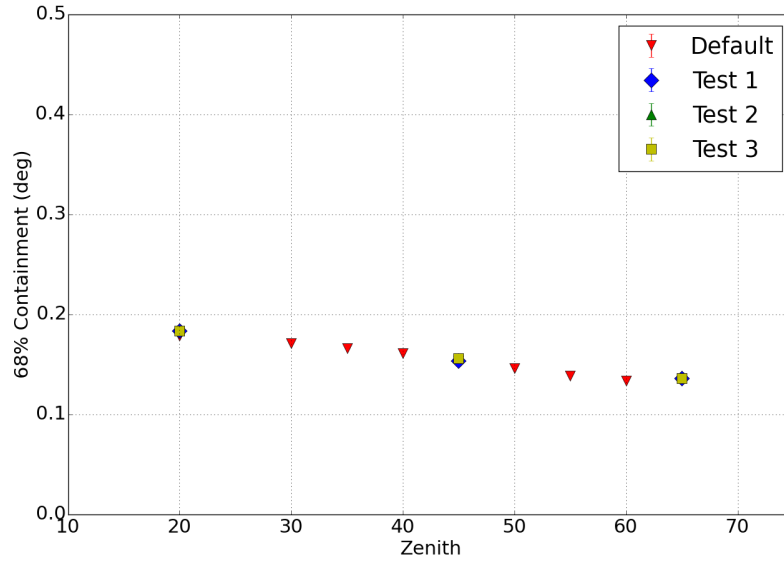
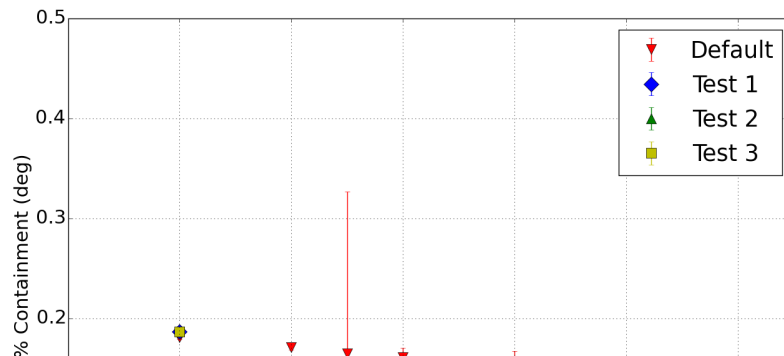


Figure 9: The weight functions used in the training samples in the new tables (Default, Test 1, Test 2 and Test 3) and those found in the scripts used to generate the older tables (Old Script 1, Old Script 2).



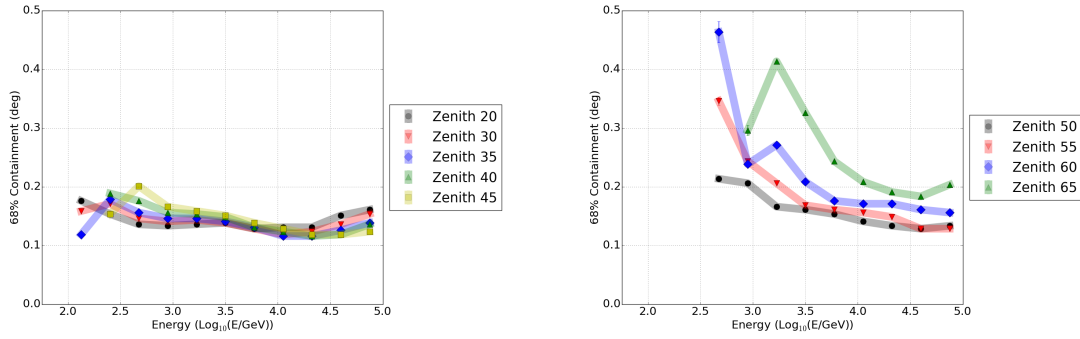
(a) Numerically determined R_{68} .



0.2.6 Energy Dependence

Another important dependence of the reconstruction resolution (and therefore the R_{68}) is that on energy. Higher energy photons are expected to comprise a larger fraction of LZA photons because lower energy showers suffer more absorption in the atmosphere. Conversely, high energy photons make up a small fraction of SZA photons (and more generally, all cosmic photons) due to the $\sim E^{-2}$ shape of the spectrum. A better resolution at higher energy would also be expected to contribute to the improved resolution at LZA.

Due to the slicing in energy in addition to zenith and therefore smaller number statistics, this is expected to lead to relatively sparse data in each bin. For this reason, in this section we address only the R_{68} determined from the numerical integral, and not that from the fit.



(a) Energy Dependence of Method0 at small and (b) Energy Dependence of Method0 at large medium zenith angles.

Figure 11: Energy Dependence of Method0 with R_{68} for small zenith angles (left) and that for large zenith angles (right). Colored bands are intended to guide the eye and do not represent data points.

The Method0 energy dependence (Fig. ??) follows the same trend as in Fig. ?? – seeing the best resolution for all zenith angles in the 3-30TeV range as well as a low energy improvement likely driven by higher statistics. The energy dependence for the older *Disp* tables (Fig. ??) appears to have a minimum in R_{68} close to 1 TeV. These old *Disp* tables also see a degradation in resolution at the highest energies. The newer *Disp* tables (Fig. ??) on the other hand appear to do **better at higher energies for all zenith angles**. At

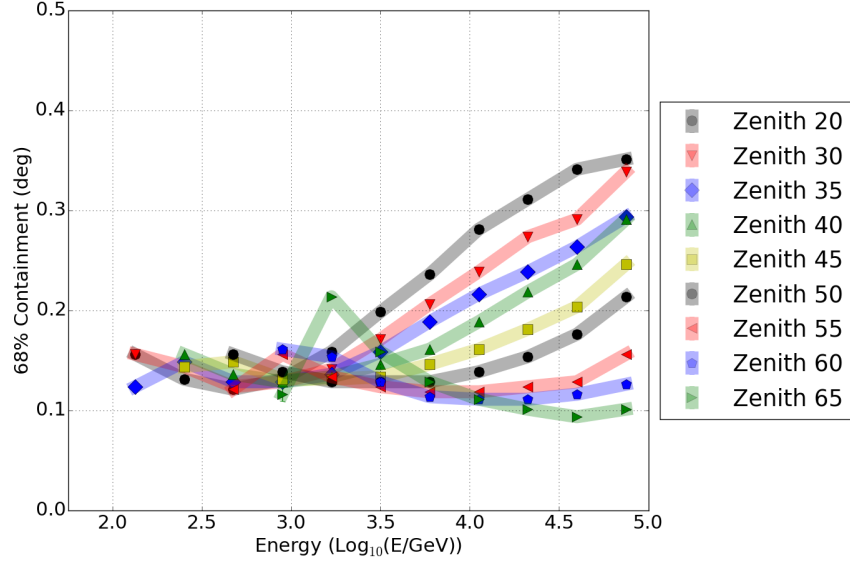


Figure 12: Energy Dependence of the old Method5t with the numerically determined R_{68} . Colored bands are intended to guide the eye and do not represent data points.

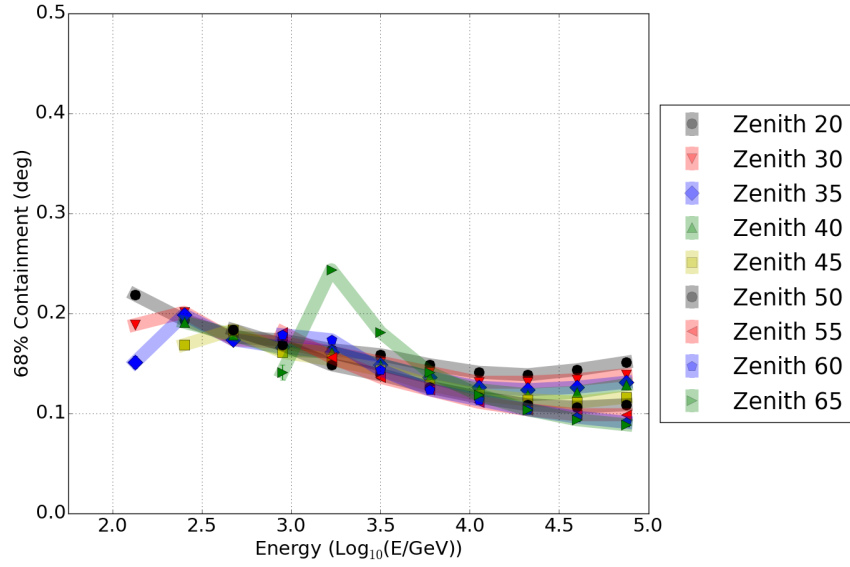


Figure 13: Energy Dependence of the new Method5t with the numerically determined R_{68} . Colored bands intended to guide the eye and do not represent data points.

energies above ~ 1 TeV and zenith angle greater than 30° , the R_{68} is at or better than 0.3° (see Fig. ??). Compared to the geometric method this provides major improvements in the LZA region (top rows of Fig. ?? and ??) and compared to the older *Disp* tables, we see major improvements in the high energy-medium zenith regime (lower right corner of Fig. ?? and ??)

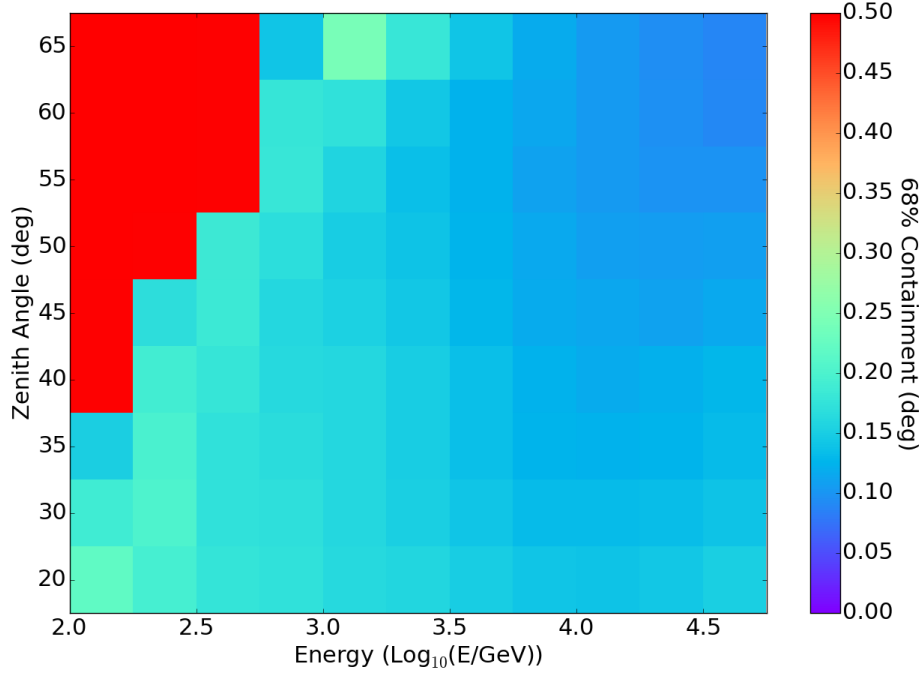


Figure 14: Energy and zenith dependence of the new Method5t, with colors denoting the R_{68} and red denoting $R_{68} \geq 0.50^\circ$. The upper-left corner shows regions of loss in resolution in the large-zenith low-energy region due to low statistics.

0.3 R_{68} for Observational Data and Known Objects

In order to test the validity of the results from the simulations, a known point-source with sufficient data collection at LZA and a hard spectrum (and therefore high statistics in the TeV range) was needed. Since our initial test was performed on the Crab Nebula and, in the VERITAS archival data it is the known object with the longest total exposure time at

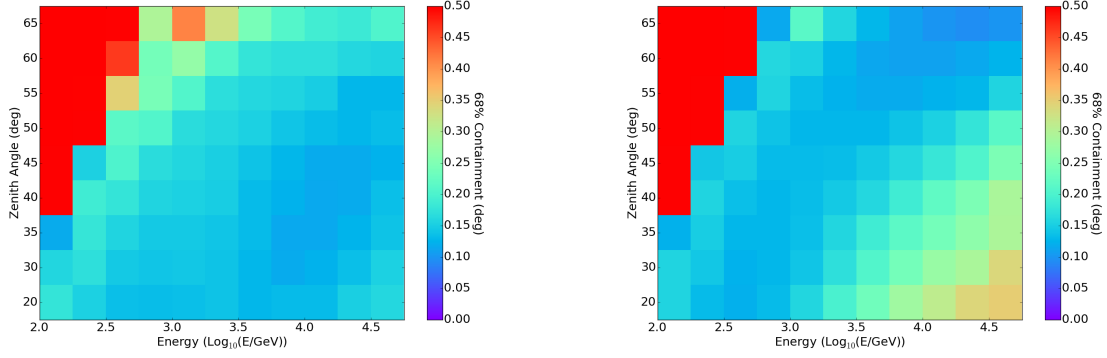


Figure 15: Energy and zenith dependence of the geometric reconstruction (left) and the old Method5t (right), with colors denoting the R_{68} and red denoting $R_{68} \geq 0.50^\circ$.

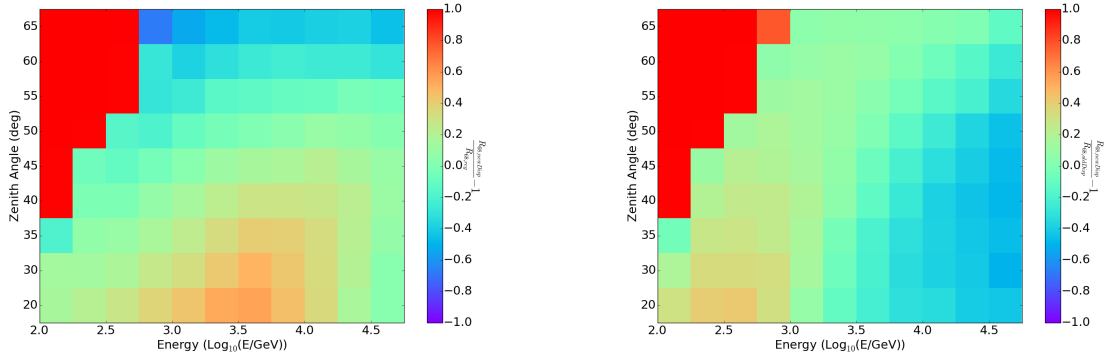


Figure 16: Performance of the new *Disp* method compared to Method0 (left) and the old *Disp* method (right), with purple at -1.0 denoting 100% improvement over the older method and bright red (in the upper left corners) at 1.0 denoting 100% worse performance in that bin or regions of insufficient statistics. The range of usefulness for Method5t (relative to Method0) now extends to $E \geq 1\text{TeV}$ and $\phi \gtrsim 55^\circ$ (upper-right corner).

LZA, the Crab was used to generate some benchmarks. Additionally, Mrk421 was used as a known point source to test the power of discrimination between point and compact sources. Each of these objects and the corresponding analysis will be discussed in greater detail in the following sections.

0.3.1 Data Constraints and Li & Ma Significance

For the data analysis, unlike with the simulations, a “true” direction was not knowable. Instead, the true direction must be assumed to be that of the known object, and we must restrict our analysis to those showers determined to be from the source object, and where the source object is in fact detected.

As a measure of detection, we use both the number of excess events – defined as the number of events determined to be from the source minus the number of events from the off-source regions scaled by the ratio of the exposures of the on and off-source regions. For a number of events from the source region N_{on} , a number of events from the off-source region N_{off} , and a ratio of exposure $\frac{A_{on}}{A_{off}} = \alpha$, the number of excess events is given by $N_{on} - \alpha N_{off}$. This corresponds to the number of events observed above the number expected given the background observation, and statements regarding the small- or large-number statistics refer to this number being small or large.

As a measure of the robustness of the R_{68} determined for the data, we also use the gamma-ray astronomy standard Li & Ma? significance, with a threshold significance of 3σ . A high significance measurement of the object in a given energy and zenith bin should provide a measure of R_{68} in that bin that is relatively stable against statistical fluctuations.

Lastly, as a measure of how the data analysis performs relative to expectations from the simulation analysis we also use a direct bin-wise comparison of the resulting numerical R_{68} defined as $\xi = \frac{R_{68, \text{data}}}{R_{68, \text{sim}}} - 1$. This provides a numerical quantity that is expected to be zero and provides a simple percentage deviation from expectation. A ξ of 0.3 means the R_{68} from data in a given bin was 30% larger than that from simulations, with positive values signifying “under-performance” and negative values signifying “over-performance”.

0.3.2 R_{68} for the Crab Nebula

The Crab Nebula is measured to have a GeV-TeV extension of $\sim 0.03^\circ$???. To observe this, a R_{68} of at least 0.03° is required which, based on the simulations, we do not achieve (see Fig. ??). This however provides another metric by which to quantify our resolution, as well as to measure possible future gains in resolution.

Fig. ?? shows skymaps for zenith cuts and energy cuts for the Crab data set with $\phi \geq 40^\circ$. This zenith cut was applied in order to only look at Crab data where it would be relevant to this analysis, since the Crab has several hundred hours of total data in the VERITAS archival dataset. This serves also to demonstrate the change between the dataset with $E \geq 1$ TeV and $\phi \geq 50$ and that with $E < 1$ TeV and $\phi < 50$ without the otherwise overwhelming effects of larger statistics in the low-energy small zenith angle region. With the colors providing a log-scale representation of the number of photons at a given 2D deviation from the source location, we can see that Fig. ?? has more events closer to the center of the skymap than Fig. ?. It is however also clear that a more robust demonstration of this effect requires a higher significance dataset in the regions of interest.

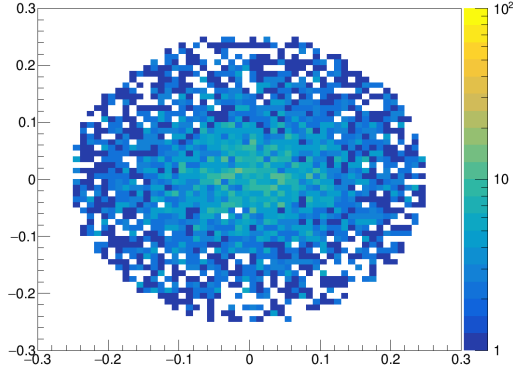
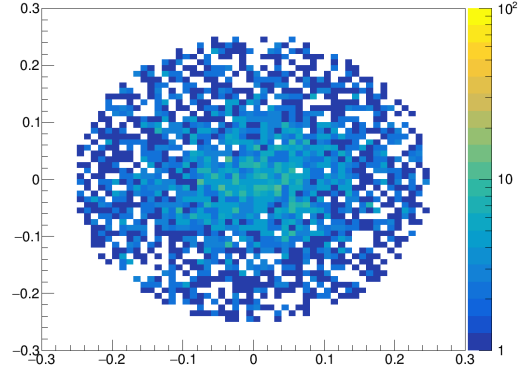
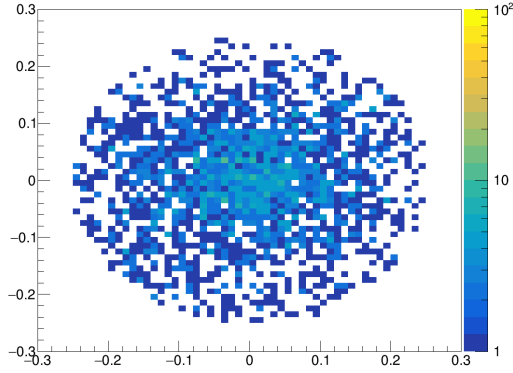
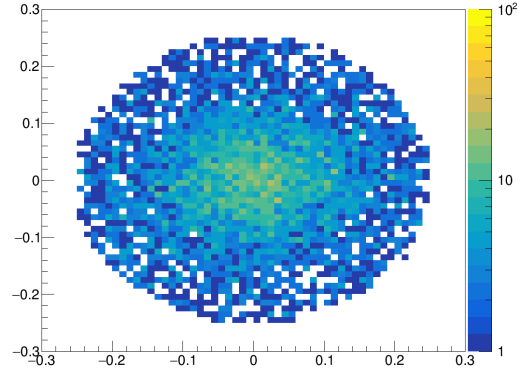
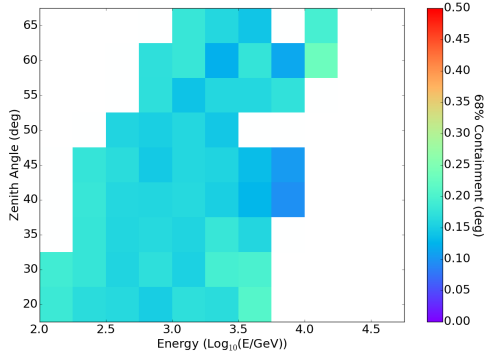
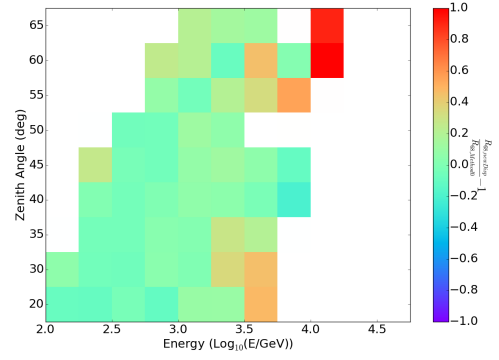
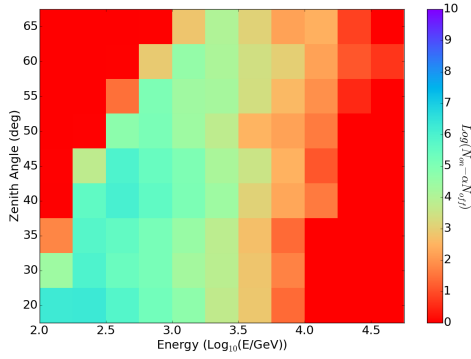
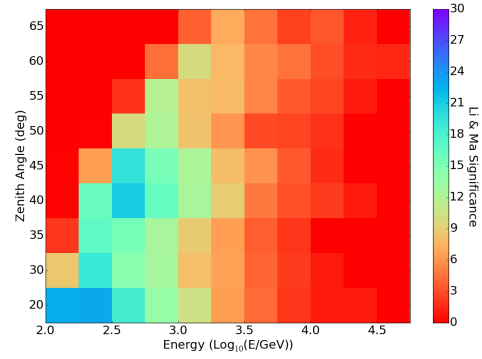
(a) Crab skymap for $\phi < 50^\circ$ and $E < 1$ TeV.(b) Crab skymap for $\phi \geq 50^\circ$ and $E < 1$ TeV.(c) Crab skymap for $\phi < 50^\circ$ and $E \geq 1$ TeV.(d) Crab skymap for $\phi \geq 50^\circ$ and $E \geq 1$ TeV.

Figure 17: Reconstruction of the Crab direction using Method5t and with data runs where the mean zenith angle $\phi \geq 40$. With the colors providing a log-scale representation of the number of photons at a given 2D deviation from the source location, we can see that (d) has more events closer to the center of the skymap than (a), and while a similar effect may be visible in (b) and (c), those effects are washed out by small statistics and low significance.



(a) Angular resolution using Method5t.

(b) Deviation from simulations (ξ).(c) Number of excess events ($\log(N_{on} - \alpha N_{off})$).

(d) Li & Ma significance.

Figure 18: Reconstruction of the direction of the Crab Nebula horizon-to-horizon runs using the new *Disp* tables. In the top two plots, white denotes regions of significance $< 3\sigma$. Relative to Fig. ??, this provides better resolution at the highest zenith angles, and is able to reconstruct higher-energy events in the $45 - 55^\circ$ region.

Relative to Fig. ??, this provides better resolution at the highest zenith angles and is able to reconstruct higher-energy events in the $45 - 55^\circ$ region. This provides a proof-of-principle that the *Disp* method is an improvement on the standard method at the highest energies and at large zenith angles. In the region $\phi \geq 50$ and $E \geq 1$ TeV the Crab nebula has a R_{68} of $\sim 0.15^\circ$.

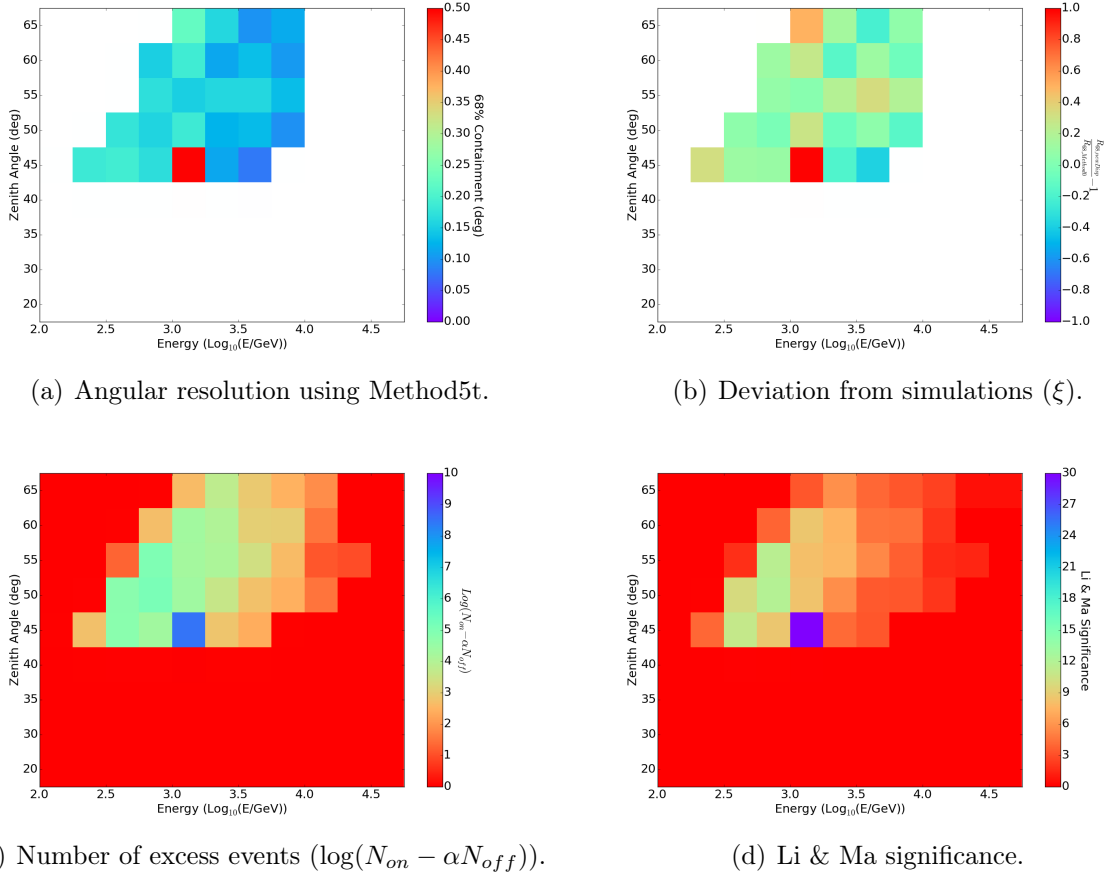


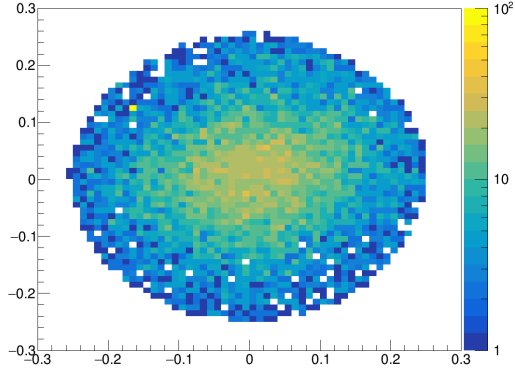
Figure 19: Reconstruction of the direction of the Crab Nebula LZA runs using the new *Disp* tables. In the top two plots, white denotes regions of significance $< 3\sigma$.

The numbers for the Crab Nebula in Fig. ?? suggest that we can not yet detect the extension of the Crab, which is consistent with the simulation analysis. Also expected from the simulation results, the R_{68} decreases with increasing energy. In the regime of interest, ($E > 1$ TeV and $\phi \gtrsim 50^\circ$), the data reconstruction under-performs relative to simulations with $\xi \sim 0.3$, and the dependence of the R_{68} (see Fig. ??) on zenith angle is smaller than expected. This is perhaps due to lower statistics in the region (see Fig. ??), since for most of our parameter space, $|\xi| \leq 0.15$ (Fig. ??).

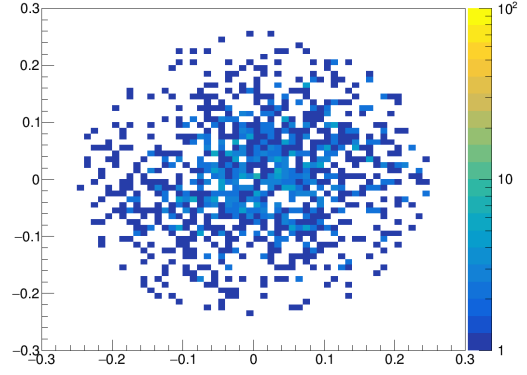
0.3.3 R_{68} for Mrk421

Mrk421 is an extra-galactic source ($z = 0.031$) with a hard spectrum (spectral index $\Gamma = 2.2$), but less than 5 hours of observation in the zenith range of interest ($\phi > 45^\circ$). Even with a short total duration of observation, because Mrk421 is a high-flux object, it is still observable at high significance. In the region $\phi \geq 50$ and $E \geq 1$ TeV Mrk 421 has a R_{68} of $\sim 0.11^\circ$.

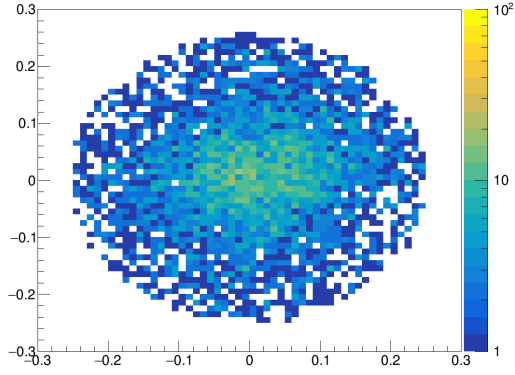
The result of this analysis was that the smallest values of R_{68} for Mrk421 ($\sim 0.09^\circ$, Fig. ??) are found in a region of parameter space where the Li & Ma significance is small (Fig. ??) and more generally that the high values of ξ are found where Li & Ma significance is low, suggesting that a more robust analysis would require more statistics. Consequently this point-source analysis does not definitively validate the results from the simulations but it does demonstrate the simulation result that the point-source reconstruction in the regime of interest is $< 0.2^\circ$ (Fig. ??). Fig. ?? shows skymaps for zenith cuts and energy cuts for the Mrk421 data set with $\phi \geq 15^\circ$. This zenith cut was not as tight as that used for the Crab data because the Mrk421 data was used to validate the simulation results across a larger range of zenith angles. This plot is meant to demonstrate the change in R_{68} between the dataset with $E \geq 1$ TeV and $\phi \geq 50$ and that with $E < 1$ TeV and $\phi < 50$, but the low significance in the high-zenith and high-energy regions means that statements about trends in R_{68} are not very robust. With the colors providing a log-scale representation of the number of photons at a given 2D deviation from the source location, we can see that Fig. ?? has more events closer to the center of the skymap than Fig. ?. It is however also clear that a more robust demonstration of this effect requires a higher significance dataset in the regions of interest.



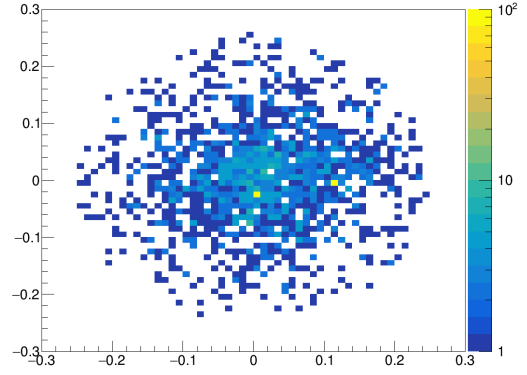
(a) Crab skymap for $\phi < 50^\circ$ and $E < 1$ TeV.



(b) Crab skymap for $\phi > 50^\circ$ and $E < 1$ TeV.

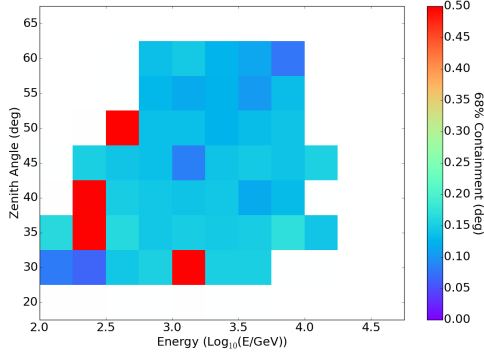


(c) Crab skymap for $\phi < 50^\circ$ and $E > 1$ TeV.

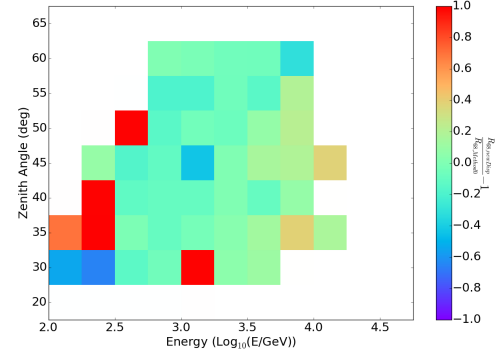
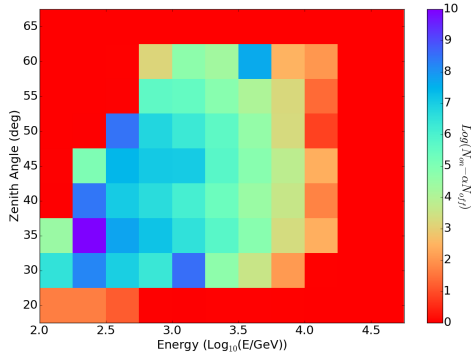
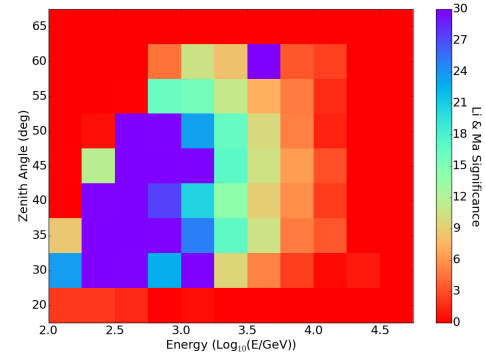


(d) Crab skymap for $\phi > 50^\circ$ and $E > 1$ TeV.

Figure 20: Reconstruction of the Mrk421 direction using Method5t.



(a) Angular resolution using Method5t.

(b) Deviation from simulations (ξ).(c) Number of excess events ($\log(N_{on} - \alpha N_{off})$).

(d) Li & Ma significance.

Figure 21: Reconstruction of the direction of Mrk421 using the new *Disp* tables. In the top two plots, white denotes regions of significance $< 3\sigma$.