1 Hough Transform

1.1 Abstract

We applied a Hough Transform to look for arcs of events in the sky. Such an arc structure is expected when the charged cosmic rays pass through the galactic magnetic field. We then looked for potential correlation of the arcs found in the data with the locations of the BL Lacertae objects. No statistically significant arcs nor significant correlations were found in this study.

1.2 Introduction

The Hough transform is a class of duality transformations commonly used to identify structures in graphical images [1]. The goal of this study is to use a Hough transform to look for arc structures in HiRes stereo data set. Here we use the term "arc" to describe segments of great circles in the dome of the sky. When cosmic rays pass through the galactic magnetic field they are deflected. The amount of deflection depends upon their energy. For small deflections, the pointing directions of a series of cosmic rays may form an arc in the sky as illustrated in Figure 1 [2]. These are the types of structures we would like to identify.

The Hough Transform is a duality transformation. In our case, it is applied on a sphere, where the pointing direction of each event is mapped to a great circle that is perpendicular to it. For events that lie on an arc, their associated great circles would intersect at two points, as illustrated in Figure 2. We then divide the sky into equal sized and shaped bins. A coarsely binned example (for illustration purposes) is shown in Figure 3. The data analysis consists of looping over all events. For each event, we increment those bins through which the corresponding great circles passes. The bin values therefore represent the total number of great circles that lie within its boundaries. The arcs are then identified by significant excess in a particular bin.

Consider, for example, an arc containing five points. Each point is transformed into its perpendicular great circle as shown in Figure 2. The sky is then binned into 6°, equal area bins (a total of 4584 bins). This bin size was not optimized. It was arbitrarily chosen. The main idea is that the bin size is not so small that we would loose signal or so big that we would over estimate one. Each bin is then arbitrarily numbered in latitude from north pole to south and around the axis. Finally, the bins values are plotted and an excess is found in the two bins were the great circles meet.

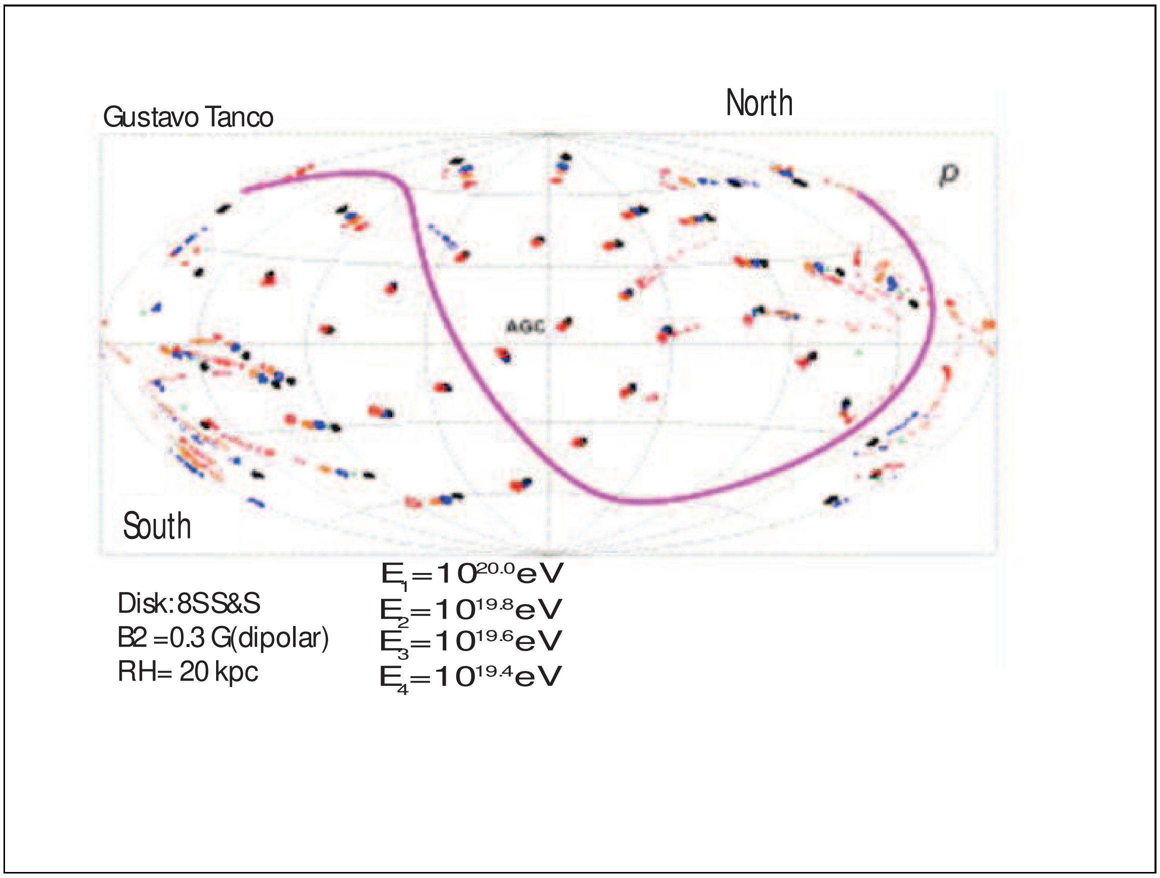

Figure 1: Simulated polarization plot by Gustavo Medina-Tanco [2]. It shows cosmic rays with energies between $10^{19.4}$ and 10^{20} eV. The events are deflected while passing through the galactic magnetic field forming arcs shapes. While the events originate at point sources, the magnetic field deflects the events to form the arc shapes which one sees in the plot.

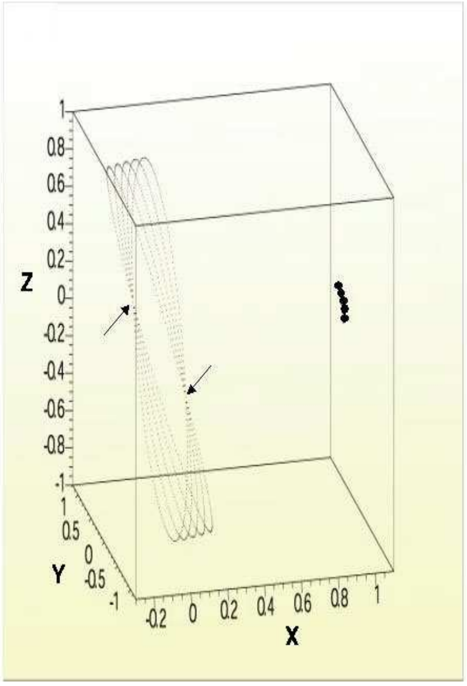

Figure 2: When there are points in an arc, the great circles perpendicular to them will all meet in the same place. The positions at where they meet is pointed indicated by the arrows in this example.

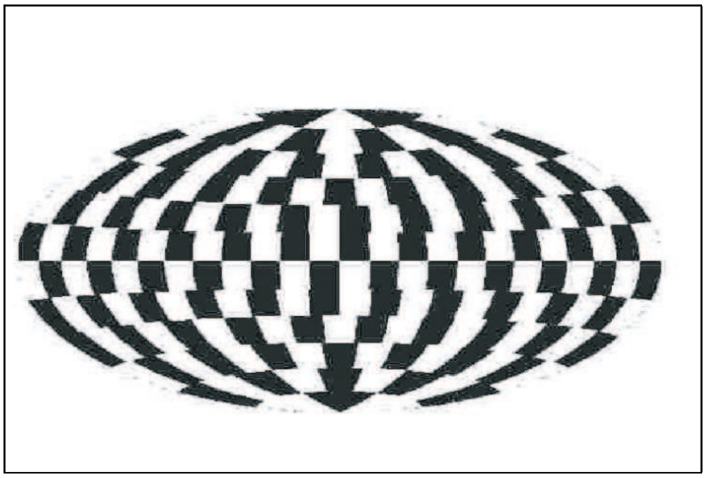

Figure 3: An illustration of the sky divided into equal-sized and equal-shaped bins. This illustration is coarsely binned to demonstrate the technique. In practice, during this analysis, the sky gets binned much more finely.

1.3 Applying The Hough Transform.

We apply the Hough transform method to the stereo HiRes data to search for arcs of events in the data. The HiRes stereo data has excellent angular resolution of $\sim 6^{\circ}$. The data selected are those events with energies $> 10^{19}$eV (particles with energies $< 10^{19}$ eV are deflected by larger angles). This data set contains 271 events collected between December 1999 and 31 January 2004. More details about this data set are given in [3].

Figure 4 shows the particle's arrival direction (in equatorial coordinates) for the 271 events used in this analysis. The same data is also shown in Hammer-Aitoff projection in Figure 5. After applying the Hough Transform to the data set, the bin occupancy for the great circles is calculated. The resulting histogram is shown in Figure 6.

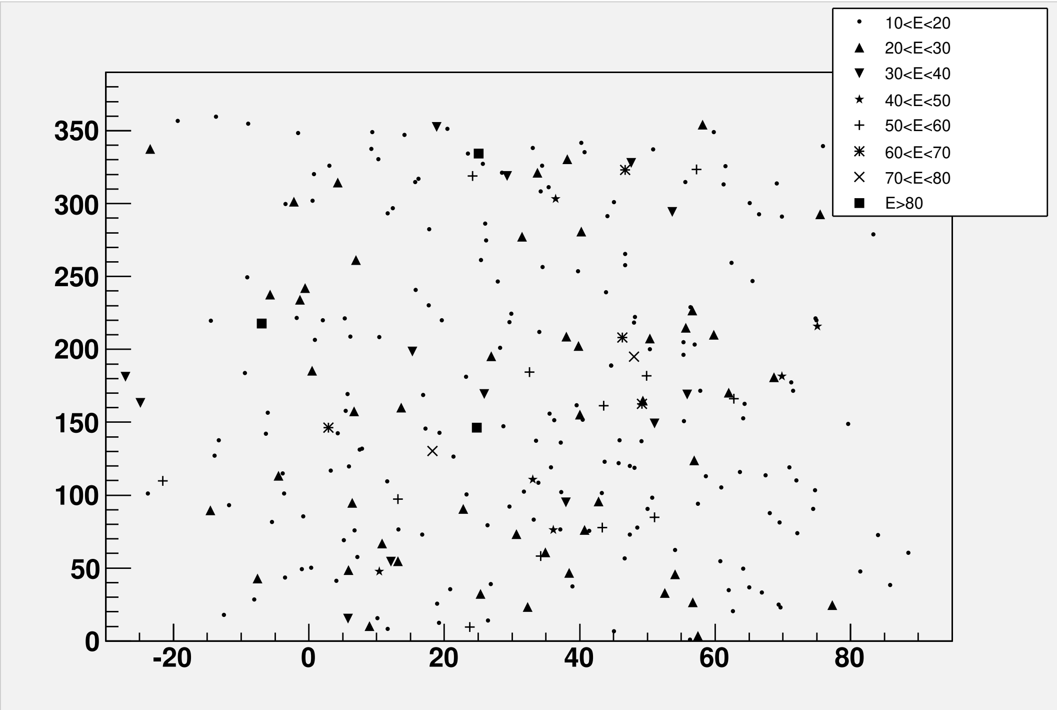

Figure 4: The pointing direction, right ascension vs. declination in degrees, of the 271 events used in this analysis. The box on the top right of the figure shows the color key for the energies (in EeV) of the events.

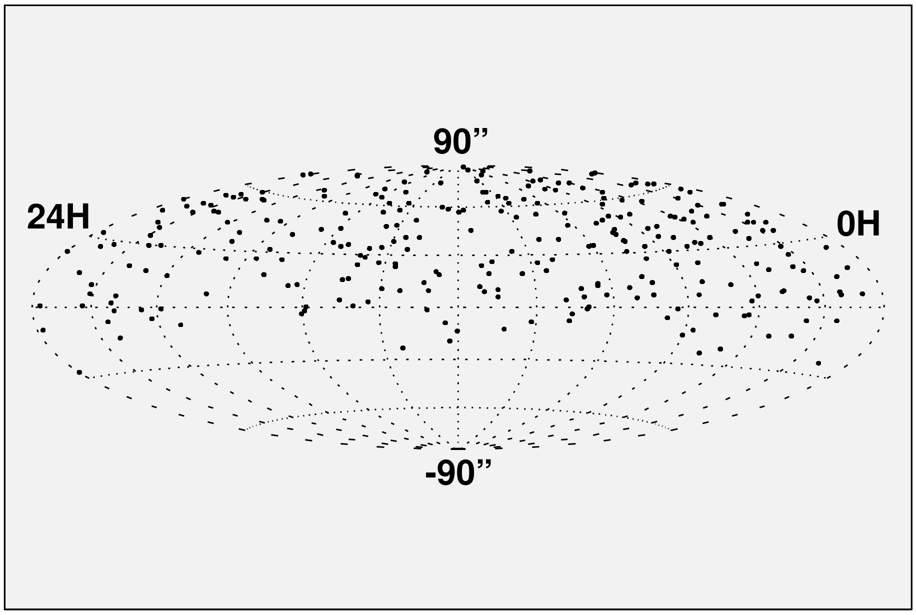

Figure 5: The 271 event stereo data set in Hammer-Aitoff projection.

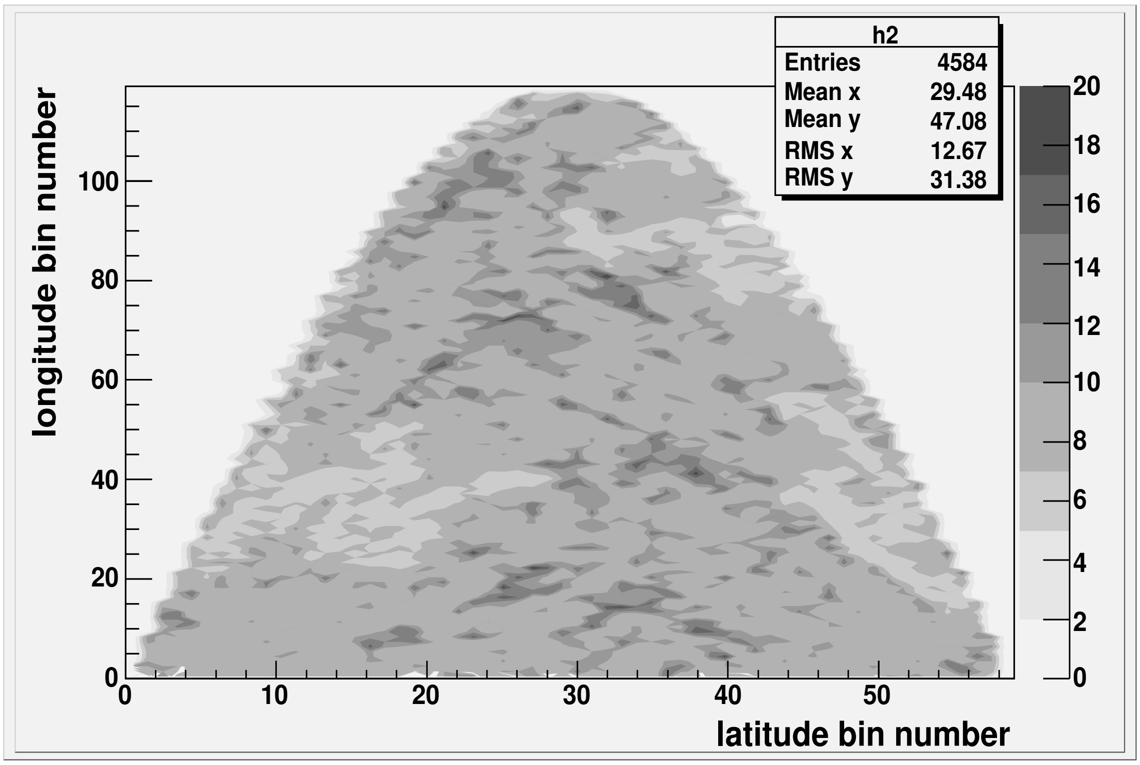

Figure 6: The bin occupancy values of normal great circles to the stereo data set . The sky was divided into 4584 number of bins of equal size (6°) with the north pole at the zenith. The bins are arbitrarily numbered in latitude from north pole to south and around the axis.

1.4 Search for significant excess.

In order to identify arcs, we look for apparent excess over what is expected from an isotropic background. To estimate the significance of such excess in each bin, we compared the data distribution to that of 2×10^4 isotropic simulated data sets. Each simulated data set contained 271 events distributed according to the actual HiRes stereo exposure (taking into account missing mirrors, etc) and assuming an isotropic initial distribution.

From the Monte Carlo data sets, we calculate the probability for each bin to have a simulated occupancy greater than or equal to that found in the data. We define the significance, S, for a bin to be given by $S = log(\frac{1}{P})$ and $P = \frac{M}{N}$, where N is the total number of simulated data sets, and M is the number of MC sets which recorded an occupancy greater or equal to that seen in the data bin.

The resulting significance values are shown in Figure 7, where each bin is indexed by both a longitudinal (RA) and a latitudinal (DEC) bin number. Bins with S > 2.5 designated as "high significance". This cut was chosen *a priori*. In addition to the significance and energy ($E \geq 10^{19}$eV) selection, we also required the opening angle between any pair of events in an arc to be less than $5°$. This requirement was added to reduce the background from long-range random overlaps.

The HiRes stereo data showed seven "high significance" bins. The arcs corresponding to these bins are shown in Figure 8. The arc with the greatest significance had a value of S=3.5.

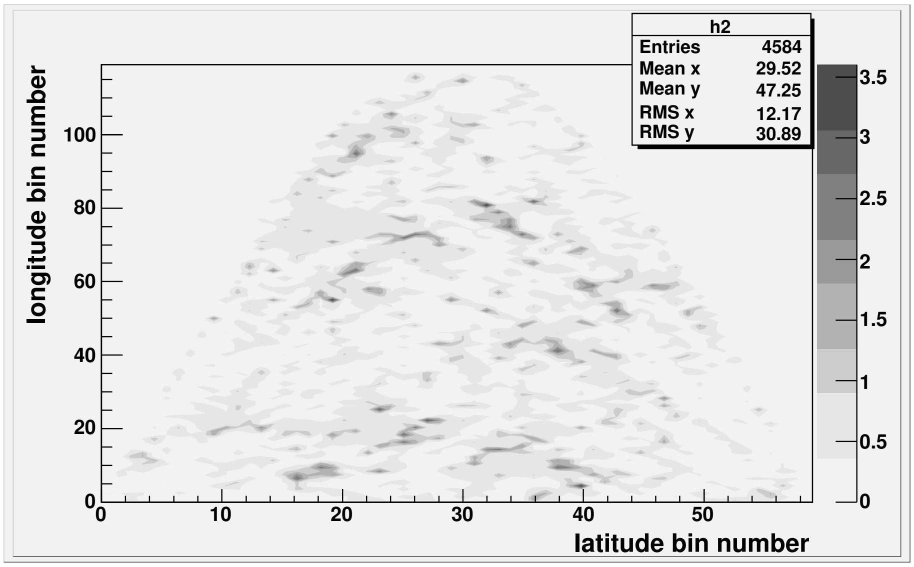

Figure 7: Significance values for each bin.

For the seven high-significance bins we show the occupancy histogram for the 2×10^4 isotropic MC set, Figure 9 shows the histograms of the logarithm of the occupancy level vs. the number of events in the arc. A vertical line is drawn on each plot at the bin occupancy level found in the data. The probability of randomly finding at least this occupancy level in an isotropic data set is then the ratio of the area under curve to the right of the line to the whole area (2×10^4). The significance value is than calculated and recorded in the boxed number for each bin.

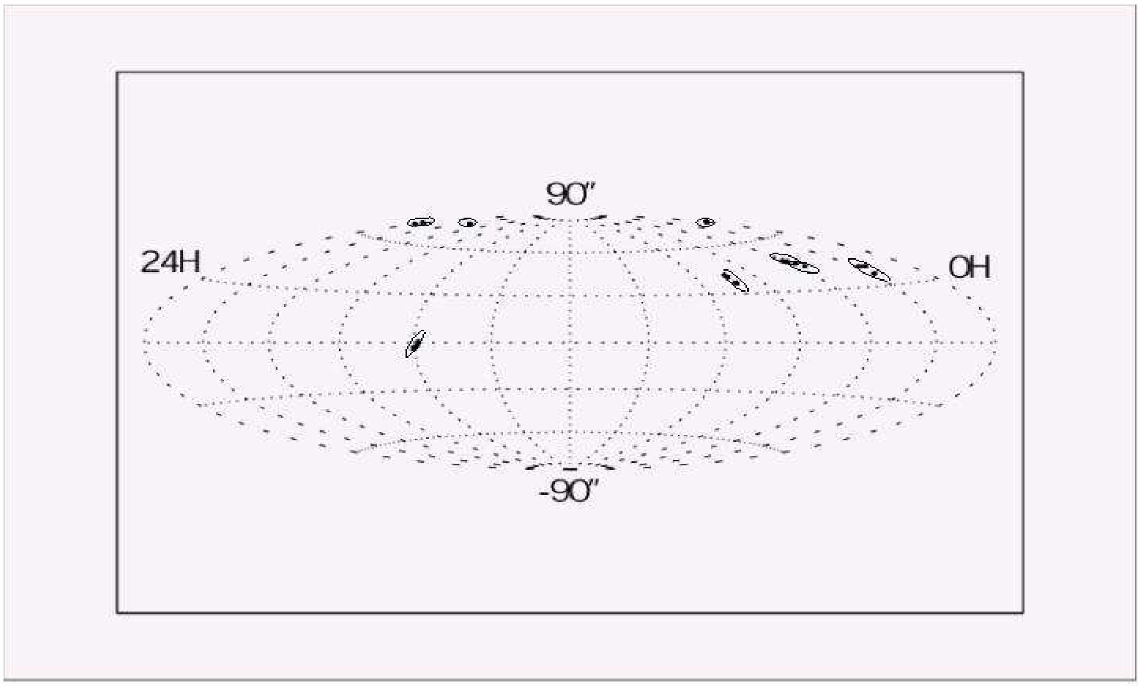

Figure 8: The arcs found in the stereo data set (the actual arrival direction) in Hammer-Aitoff projection (for clarity each arc has been enclosed with an ellipse).

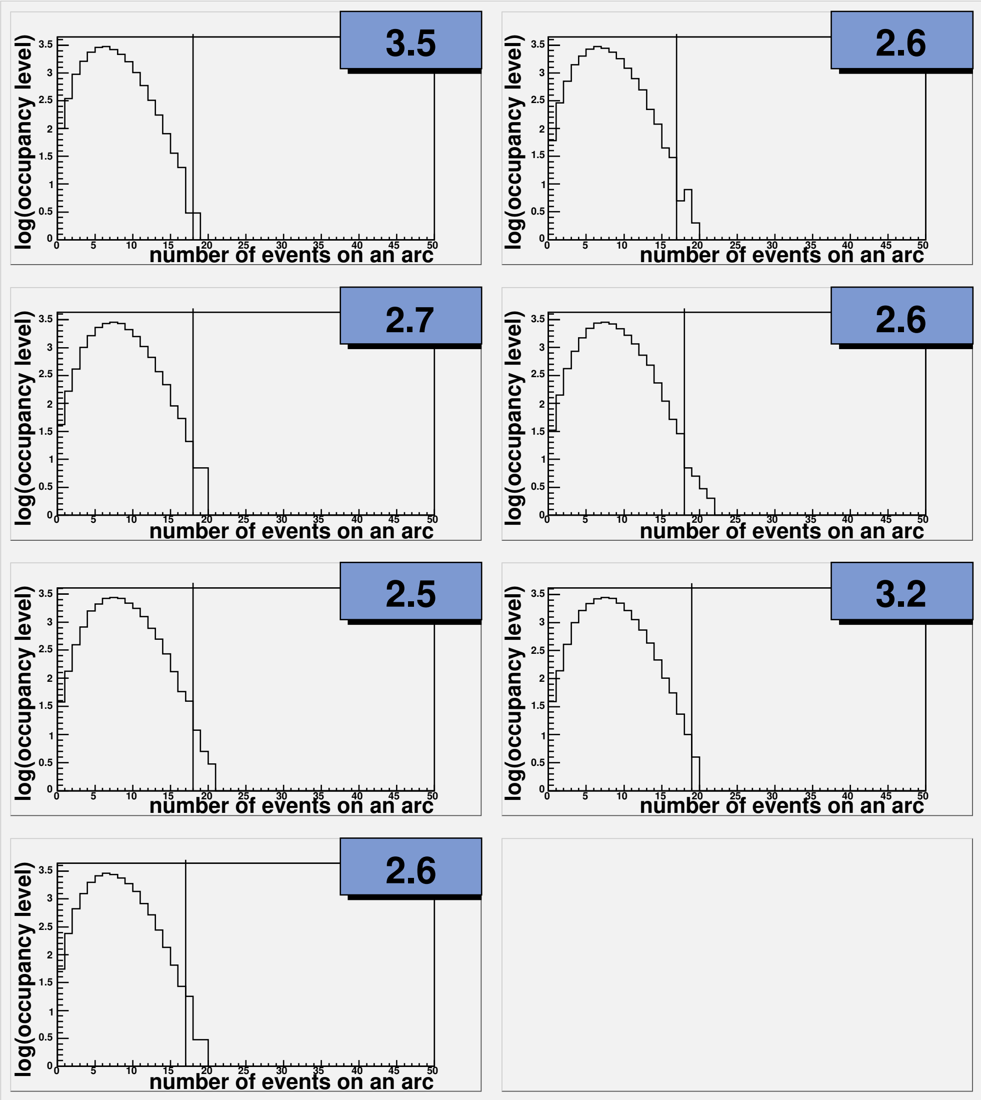

Figure 9: For the 2×10^4 isotropic MC sets, we histogramed the occupancy level for the bins which showed "high significance" in the data. The histograms above show the logarithm of the occupancy level vs. the number of events in the arc. A vertical line is drawn on each plot at the number of events in the arcs found in the data. The probability of randomly finding at least this many events in an isotropic data set is then the ratio of the area of the curve to the right of the line to the whole area (2×10^4). The significance value is than calculated and recorded in the boxed number for each bin.

1.5 Global Significance of data excess

The next study we made was to determine whether the arcs found in the data in consistent with random fluctuations over the full sky, not just in the location in which they occurred. To do this, we compare our 2×10^4 isotropic simulated sets to a new set of 1000 isotropically simulated data sets, each with 271 events. We applied the same analysis and cuts as we did to the real stereo data set. Including the "significance cut" of S$>$ 2.5. We found that 99.6% of the time, the maximum significance of the isotropically generated data sets is greater than 2.5. Also, in 41.2% of the samples, the simulated data sets recorded a significance ≥ 3.5, which is the highest significance found in the real data set.

Finally, the greatest number of events found on an arc in the real data is 19, as shown in Figure 9. This was also found to be quite common in the simulation, since arcs with S$>$ 2.5 and at least 19 events were found in 90.4% of the simulated isotropic sets.

1.6 Conclusion

The Hough transform has been applied to the HiRes stereo data and seven arcs were found. However, when a global significance analysis was performed, similar results were commonly reproduced within isotropic data sets of equivalent size. Hence, the observation of seven arcs in the data is not significant.

2 Cross-correlation analysis with BL Lacertae.

BL Lacertae or BL Lacs are a type of Active Galactic Nucleus (AGN). An AGN has a super-massive black hole in it's center and an accretion desk around it. The material in the accretion disk contains gas and dust; it tumbles around the black hole and in some cases gets shot away in twin luminous jets along the spin axis of the AGN. BL Lacs are blazars, which are characterized by their featureless spectrum (in radio, infrared and optical wavelengths), were blazars are a subclass of AGNs with their luminous jets aligned to our line of sight.

BL Lacs are an established source of TeV Gamma rays [4]. They have also been a suspected source of Ultra High Energy Cosmic Rays (UHECRs). Many studies have been made to search for correlations between these galaxies and cosmic rays events measured by Akeno Giant Air Shower Array (AGASA), Yakutsk and HiRes. The results of those have been mixed, with a number of papers reporting hints of weak correlation ([5], [7]), although more of these claims have been verified by independent data sets. These results have also been criticized for the *a posteriori* nature of their analysis cuts([8], [9]). In addition, the apparent correlation which have been reported are consistent with the angular resolution of the respective experiments. This would require the primary cosmic rays to be (exotic) neutral particles. In a more conventional picture, one would expect cosmic rays to be bent by magnetic fields. If the fields are sufficiently coherent then the events should fall into arcs near their sources. Because of this controversial results, we are trying to answer the question if BL Lacs are the sources of the UHECR's.

In an attempt to answer this question a correlation analysis between the arcs (obtained when applying Hough Transform to the stereo data set) and the BL Lacs positions has been applied to two sets of BL Lac data. The comparison was first performed for BL Lacs with magnitude > 18 and with classification type BL. This set was defined by (Ref. [6]). We call this Data Set One, DS1. A second comparison was made after including the (so called) Highly Polarized (HP) BL Lacs, again requiring a magnitude > 18. We call this Data Set Two, DS2.

Figure 10 shows a histogram for the cosine of the angle between the BL Lacs and the closest event in the arc to the BL Lac in DS1. One can notice what appears to be an excess in the last two bins. To study the significance of this excess, a comparison was made with 10^4 Monte Carlo data sets. Each simulated set contained 271 events (same size as real data) isotropically simulated according to HiRes stereo detector exposure. For each MC set, we find the arcs in the set and we calculate and histogram the cosine of the angle between these arcs and each BL Lac.

The values of the cosine histogram of each of the 10^4 simulated sets is then normalized to the area of the real data distribution shown in Figure 10. Figure 11 shows the distribution of the occupancy level of the simulated sets for the right most bin of the cosine histogram. A vertical line is drawn at the number of pairs found in the real data at that bin. When the distribution was fitted to a Gaussian it showed an agreement. Therefore, The probability of finding at least this many pairs in an isotropic data set is then calculated by dividing the area of the curve to the right of the vertical line to the whole area of the distribution. In other words,

$$P_{last\ bin} = \frac{\#simulated\ data\ sets \geq real\ data\ value}{\#total\ simulated\ data\ sets} \tag{1}$$

For DS1, we obtain a value of P=0.174.

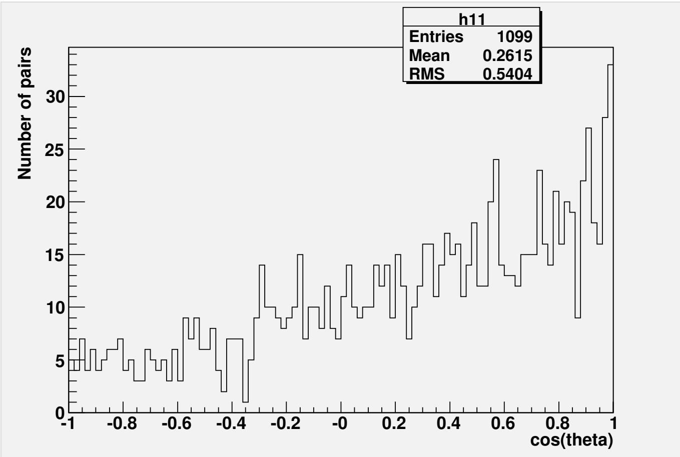

Figure 10: Histogram of the cosine of the angle between the BL Lac and the closest event in the arc to the BL Lac for data set one, DS1.

Figure 12 shows the histogram for the cosine of the angle between the BL Lacs and the closest event in the arc to the BL Lac for data set two, DS2. The same process applied to DS1 is used here. The probability of the value of the last bin being ≥ that of in Figure 12 is then calculated using Equation 1 to be 0.318.

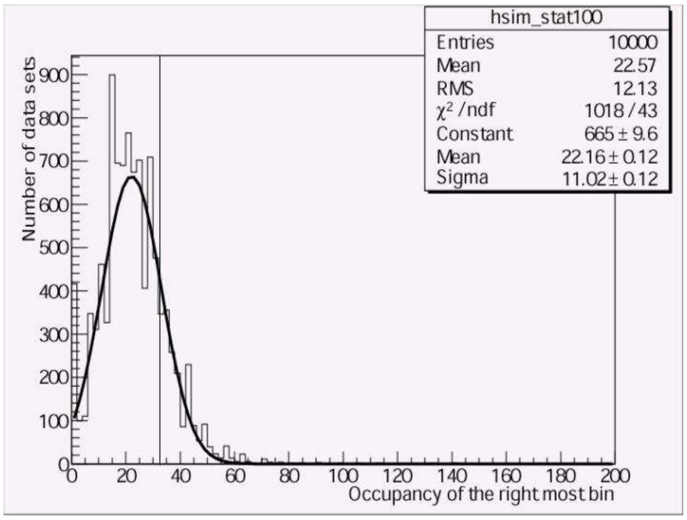

Figure 11: This figure shows the distribution of the occupancy level of the 10^4 simulated sets for the right most bin from the cosine theta histogram. A vertical line is drawn at the number of pairs found in the real data set at that bin. The probability of finding at least this many pairs in an isotropic data set is then calculated by dividing the area of the curve to the right of the vertical line to the whole area of the distribution.

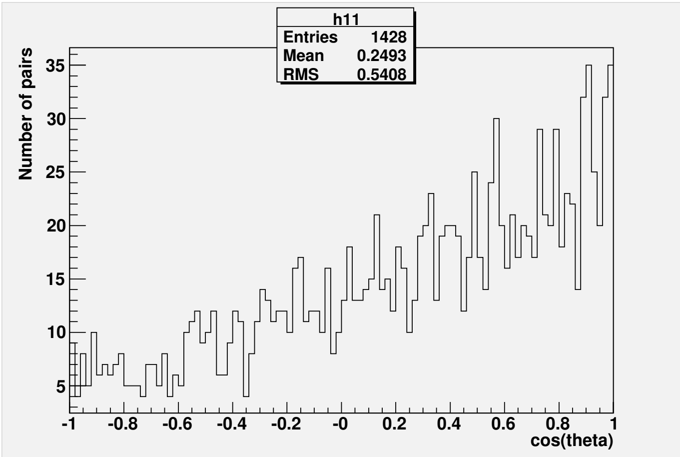

Figure 12: Histogram of the cosine of the angle between the closest point in the arc to the BL Lac and the BL Lac for data set two, DS2 (BL & HP BL Lacs).

2.1 The arc average point correlation

In Figure 11 we noticed a peak in the two bins near $\cos\theta = 1$. This suggests that the use of the closest event to the BL Lac to calculate the event-BL Lacs pairing produce a bias at small angles.

To get rid of this effect, we repeated the same study except that instead of studying the correlation between the closest point in the arc and the BL Lac source we examined the correlation between the centroid (the average vector of the events on the arc) in the arc and the BL Lac source.

Figure 13 shows the histogram of the cosine of the angle between the central point on the arc and the BL Lacs for data set one, DS1. The same study applied previously to finding the probability of the value of the last bin in Figure 10 is applied here. Except the fact that we rebinned the cosine histogram into 50 bins, because the excess in Figure 10 was noticed in the last two bins, and since the cosine histogram is not poisonian distribution, we would not be able to add the probabilities of the last two bins. Therefore, we decided to bin the histogram to 50 bins instead of 100 bins (as binned in the previous study). The probability is then found to be 0.095.

The same method of study has also been applied to data set two, DS2 - BL Lacs (BL & HP types). The cosine histogram between the stereo data set and the DS2 BL Lacs is shown in Figure 14 and the probability of the last bin histogram having a value $\geq$ that found in the real date (Figure 14) is found to be 0.168.

We performed one additional study to examine the excess in the last bin

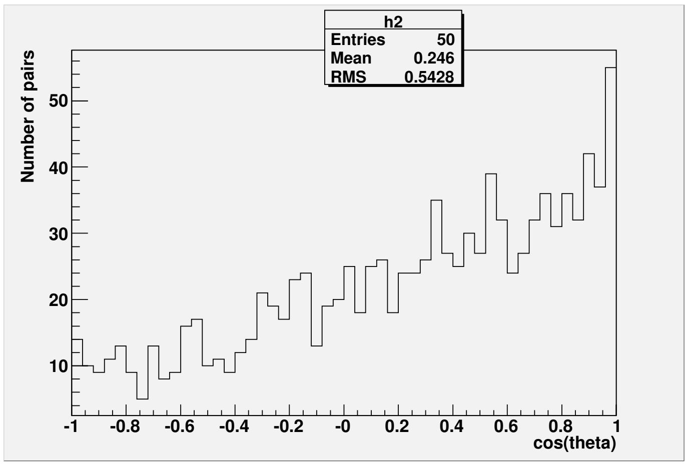

Figure 13: Histogram of the cosine of the angle between the average point in the arc and the DS1 BL Lacs.

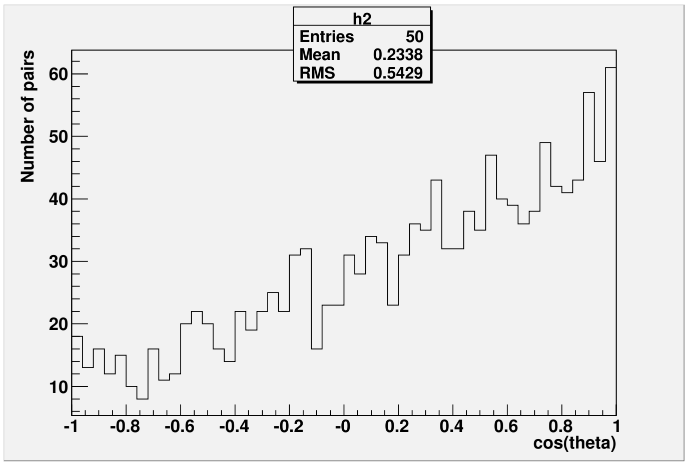

Figure 14: Histogram of the cosine of the angle between the average point in the arc and the DS2 BL Lacs (both BL & HP).

of the cosine histogram in the real data. We compared the cosine histogram in the real data to that of the simulated data using only those sets which had seven arcs (the same number of arcs that was found in the real data set). We simulated additional data sets and accepted 2500 data sets. Using data sets with seven arcs, the probability of the excess in the last bin in Figure 13 is found to be 0.084.

2.2 Conclusion

After applying Hough Transformation to the HiRes stereo data set, we looked for a correlation between the arcs found by Hough Transform and BL Lacs locations as defined in Veron 10th Catalog was applied. We found small excess in the histogram of the cosine of the opening angle between the arcs and the BL Lac sources. When this excess was investigated by comparison to isotropically simulated data sets, we then found that it is not statistically significant. When we repeated this study for simulated data sets that contained only seven arcs, we also found the excess in the last bin to be not statistically significant.

References

[1] V.F.Leavers, Shape Detection in Computer Vision Using the Hough Transform, London, New York:Springer-Verlag, 1992.

[2] G.M. Tanco, M. Teshima & M.Takeda, Constrains on the Galactic magnetic field from the two dimensional correlation function of AGASA events, 28th International Cosmic Ray Conference, Tsukuva, pp. 747-750, Universal Academy Press, Inc. 2003.

[3] R.U. Abbasi *et. al.*, Astrophys. J.610, L**73**, 2004.

[4] Hartman *et. al.*, Astrophys. J.Suppl.,123,**79**, 1999.

[5] Tinyakov P.G. & Tkachev I.I. JTEP Lett.,74,445, 2001.

[6] Tinyakov P.G. & Tkachev I.I. Astropart. Phys.,18, 165, 2002.

[7] Gorbunov D.S., Tinyakov P.G., Tkachev I.I.,& Troitsky S.V. Astrophys. J.,577,L**93**, 2002.

[8] Evans,N. W.,Ferrer,F.,& Sarkar, S. Phys. Rev.,103005, D**93**, 2003.

[9] Torres,D. F. Reucroft,S.,Reimer,O.,& Anchordoqui, L. A. Astrophys. J.,595,L**13**, 2003.