Two-point Angular Autocorrelation Function and the Origin of the Highest-energy Cosmic Rays

R. W. Clay[1], B. R. Dawson[1], L. Kewley[2] and M. Johnston-Hollitt[1]

[1]Department of Physics and Mathematical Physics,
University of Adelaide, SA 5005, Australia
rclay@physics.adelaide.edu.au
[2]Research School of Astronomy and Astrophysics, Australian National University,
Weston Creek PO, Weston, ACT 2611, Australia
lkewley@mso.anu.edu.au

Abstract: Construction of the Pierre Auger Observatory for the study of the highest-energy cosmic rays is about to begin. Prior to the availability of data from that experiment, decisions should be made on techniques for the analysis of the directional properties of those data. We examine here one possible analysis tool, the two-point angular autocorrelation function. As a concrete example, data from the SUGAR array are examined in this way. Possible clustering of the data is observed, and the identification of such clustering with candidate astronomical objects in a purpose-developed catalogue is investigated.

Keywords: methods: data analysis—cosmic rays—galaxies: intergalactic medium

1 Introduction

The Sydney University Giant Air Shower Recorder (SUGAR) was built to detect the cosmic-ray air showers produced by the highest-energy cosmic rays, including a significant number with energies above 10^{19} eV. Until the Pierre Auger Observatory comes into operation, SUGAR will remain the only source of information on the highest-energy cosmic rays from the southern skies. The SUGAR dataset is thus uniquely valuable in pointing the way to future southern-hemisphere, high-energy cosmic ray studies.

In preparation for the arrival of data from the Pierre Auger Observatory, we wished to examine a way of looking at SUGAR data that is likely to be used in the analysis of those future data. We expected this approach to give us some insight into the practical use of such an analysis, and to be of interest in giving us a new perspective on the astrophysics that already exists within the SUGAR dataset.

The approach we took was through determination of the two-point angular autocorrelation function for the array directional data. We found that there was already suggestive evidence in the SUGAR data for correlations over tens of degrees in the sky, and we investigated how those correlations may then be related to the sources of the highest-energy cosmic rays.

2 The SUGAR Array

The SUGAR array has been discussed in detail by Winn et al. (1986) and references therein. The array was located in northern New South Wales, Australia, and was operated during the period January 1968 to February 1979. At its peak development, it consisted of 47 sites covering an effective area of 100 km^2. Each station consisted of two liquid scintillator tanks buried 50 m apart. The detectors were sensitive to the muon component of cosmic ray showers. An array event consisted of coincident local events from three or more stations.

Because of the large collecting area of the array, a significant number of events were recorded with energies above 10^{19} eV. These are of particular interest to us here, as such events are the primary interest of the Pierre Auger Observatory, which will begin operations early in the new millennium and will have a collecting area of 3000 km^2 at each of a southern and a northern location. We note that the Pierre Auger Observatory will incorporate an operational philosophy pioneered in the SUGAR array, such that detector sites operate independently and have their data combined at a somewhat later stage to provide information for shower analysis. SUGAR thus provides a vehicle for assisting us to prepare for the arrival of the new data.

3 Cosmic Ray Directions and Sources

Cosmic rays are charged particles with very high energies. As such, there is an expectation that, at some energy, they will have a tendency to travel in approximately straight lines through intergalactic magnetic fields. The magnitude of those fields is uncertain. Within clusters of galaxies, the fields may have strengths as high as tens of microgauss but they could be a thousand times weaker in intercluster space. Clay et al. (1998) have discussed such propagation on the assumption that those fields are turbulent, and found that energies approaching 10^{20} eV are likely to be sufficient for approximately linear motion to occur in most fields of the type just mentioned.

An important component in that conclusion is the interaction of the highest-energy cosmic rays with the 3 K microwave background (e.g. Lampard et al. 1997a;

Muecke et al. 1999). Above about 6×10^{19} eV, there is a photopion interaction between cosmic ray protons and blueshifted 3 K photons which severely restricts the possible distances to sources of particles at those energies (Lampard et al. 1997b). The interaction is not catastrophic but source distances are limited to a volume with a radius of the order of 100 Mpc, which depends somewhat on the hardness of the source spectrum.

Catalogues of arrival directions for cosmic rays with energies above 10^{20} eV contain directions from all over the sky. There are no clear point sources from which the majority of particles appear to originate. There are some suggestive directions in the north, where there is apparent clustering on angular scales that are comparable with detector angular uncertainties, but, even if those are real, there are still many remaining directions which are spread over the sky. One is forced to conclude that there are either many sources spread over the sky and within a distance of a few hundred megaparsecs, or a limited number of sources over the same volume but with directions that are spread by scattering in intergalactic magnetic fields.

Unless the production of these particles is episodic, it appears that we can eliminate galaxies such as our own as sources at the highest energies, since there is no evidence for a cosmic-ray version of the Milky Way, although there is evidence that, up to 10^{18} eV, our Galactic centre may be a cosmic ray source (e.g. Hayashida et al. 1999; Clay et al. 2000). In the same way, Centaurus A, our nearest active galaxy, shows no evidence for cosmic ray acceleration to these energies, although there is some evidence at lower energies (e.g. Allen et al. 1993; Clay, Dawson & Meyhandan 1994). If this applies as a general rule for FR I sources (centrally concentrated, double-lobed, extended radio sources, as opposed to FR II sources, which have widely separated regions of maximum brightness; see Jackson 1999) such as Cen A, we have to look to very esoteric objects for the origins of the highest-energy cosmic rays. All powerful sources, such as Cen A, within a few hundred megaparsecs are known and have been catalogued. They are rather few and, in general, are not identified with the arrival directions that have been measured. It appears that many of the highest-energy cosmic rays must come from distances large enough that they suffer appreciable intergalactic scattering, and that we should look for rather broad-scale directional spreads from a limited number of very unusual sources.

The search for sources of the highest-energy cosmic rays can then be pursued by looking in cosmic ray directional data for preferred directions, or preferred areas of the sky, and then looking in suitable astronomical catalogues for candidate sources in those areas. The requirements for candidate sources have been discussed in many studies but Hillas (1984) has encapsulated the arguments in a readily accessible exposition. It is noteworthy that the energy of the detected particle is not the one for consideration when examining possible source mechanisms, since there is a progressive energy loss due to the photopion mechanism. The source criteria are thus more stringent than might appear at first sight.

We usually assume that acceleration occurs through some form of progressive shock acceleration, although rotating objects with strong magnetic fields may alternatively provide a single voltage of the required magnitude. In the case of shock acceleration, large objects are required, since the particles must be contained for long periods and scattered many times while energy is gained. Large volumes containing clusters of galaxies or the hot spots of radio lobes associated with FR II radio sources are possible source candidates.

4 Two-point Angular Autocorrelation Function

The two-point angular autocorrelation function has been described by Peebles (1980). It involves the analysis of an ensemble of data points with known angular locations. The two-point angular autocorrelation function is then the joint probability of finding an object in both of two elements of solid angle placed at an angular separation of θ as a function of θ. In practice, the probability can be found by comparing the number of ensemble pairs having a particular separation (in a given separation interval) with the number derived for a random ensemble that contains the same number of directions. We can express the function (w) and its uncertainty (dw, derived from a consideration of Poisson statistical uncertainties) through the ratio of the number found in the experimental data to the number in the random data (at a particular angular separation and within an angular range) minus one. A positive value at small angular separations indicates a clustering of the data points.

A significant observational difficulty in applying this procedure is in determining a suitable random dataset for the comparison data. The experimental dataset may contain clustering information but it will also contain the effects of non-uniformity of sky coverage, which might have changed over the period of the data acquisition.

In order to determine the function for our complete catalogue of SUGAR data, we created a randomised SUGAR dataset through a procedure that shuffled components of the real SUGAR data. The SUGAR dataset available to us did not contain the altitudes and azimuths of the events. We determined these from the event times and the celestial coordinates of the event arrival directions. A new dataset was then generated, using randomly chosen Julian dates for the events from the ensemble of dates within the real catalogue. The catalogue was then used to assign horizontal coordinates for each of these events, and corresponding celestial coordinates were derived as though events from those directions had been observed at the randomly selected Julian dates. In this way, the new ensemble contained both the real on-time properties of the array and the altitude and azimuthal properties of the array. In each

case for our analysis of a subset of SUGAR events, we generated 1000 shuffled datasets so that it was possible to determine the number of real events with a particular angular separation together with the probability that such a number would have been exceeded (the number of times the random sets gave a larger value divided by 1000, the number of random trials).

5 Application to the SUGAR data

We have derived the two-point angular autocorrelation function for the SUGAR data above three threshold energies; these data are shown in Tables 1–3. It is evident from the tables that, for the data available for SUGAR, any evidence for clustering is statistically weak. However, both of the higher-energy datasets (Tables 2 and 3) show positive correlations on the whole out to angular distances of the order of 40°. If that angle had been chosen *a priori* for testing as a possible scattering scale, we would have regarded even the SUGAR data as significant. Since that decision had not been taken, the result is strongly suggestive but not usefully significant. As a pilot study for the future Pierre Auger dataset, this emphasises that we must make a number of *a priori* decisions about the analysis procedures.

Rather than picking a likely spread for any clustering, our choice might have been to determine optimum angular intervals to be used in deriving the two-point function with the hope that our function would be statistically significant at each angular spacing. In the case of our SUGAR data, this involved an appropriate *a priori* choice of the angular interval, which we selected as 5° for the tables. If clustering exists out to a certain angle, we can then determine the angular deviation, to this precision, for propagation from various sources with different spreads, provided that the data are dominated by a small number of sources.

For instance, in the case of the SUGAR data, we might have chosen 15° intervals rather than 5°. The data for both of the higher-energy datasets would then have been arguably significant at the smaller angular spacings, but the 4×10^{19} eV dataset would have spread to 45° compared to 30° for the 6×10^{19} eV data. This might seem to be physically consistent, due to the expectation of reduced scattering at the higher energies and, if a source could be identified, would yield valuable information on intergalactic magnetic fields.

Table 2. The two-point autocorrelation function w for SUGAR data as a function of angular spacing for showers with energies above 4×10^{19} eV

The fourth column indicates the number of times the given value of w is exceeded in 1000 randomised data sets

Angular interval	$w \times 10^{-2}$	$dw \times 10^{-2}$	Number
0–5	−47.9	14.,2	766
5–10	6.0	22.6	357
10–15	25.8	18.2	109
15–20	9.9	14.7	253
20–25	−5.2	12.2	569
25–30	5.4	11.9	333
30–35	17.9	11.8	105
35–40	−20.3	9.3	908
40–45	4.9	10.2	339
45–50	1.1	9.9	432
50–55	−14.1	9.0	837
55–60	−22.3	8.4	952
60–65	−14.2	8.9	851
65–70	−27.0	8.1	978
70–75	7.1	9.9	263
75–80	20.7	10.6	43
80–85	11.0	10.0	183
85–90	14.5	10.5	118
90–95	−5.7	9.6	635
95–100	10.7	10.7	200
100–105	−3.2	10.3	553

Table 3. The two-point autocorrelation function w for SUGAR data as a function of angular spacing for showers with energies above 6×10^{19} eV

The fourth column indicates the number of times the given value of w is exceeded in 1000 randomised data sets

Angular interval	$w \times 10^{-2}$	$dw \times 10^{-2}$	Number
0–5	−100	undef.	585
5–10	86.4	59.0	47
10–15	22.3	35.3	212
15–20	24.6	31.1	182
20–25	39.7	29.8	100
25–30	24.6	25.4	165
30–35	10.8	22.6	180
35–40	−40.1	16.0	910
40–45	2.9	20.2	371
45–50	−37.3	15.7	896
50–55	−3.2	15.7	884
55–60	−29.3	15.8	851
60–65	−27.3	16.2	816
65–70	−22.0	16.6	757
70–75	−14.0	17.5	642
75–80	22.4	21.0	160
80–85	18.1	20.9	209
85–90	23.8	22.6	151
90–95	2.6	20.9	387
95–100	78.4	28.2	1
100–105	−1.8	22.0	433

Table 1. The two-point autocorrelation function w for SUGAR data as a function of angular spacing for showers with energies above 2×10^{19} eV

The fourth column indicates the number of times the given value of w is exceeded in 1000 randomised data sets

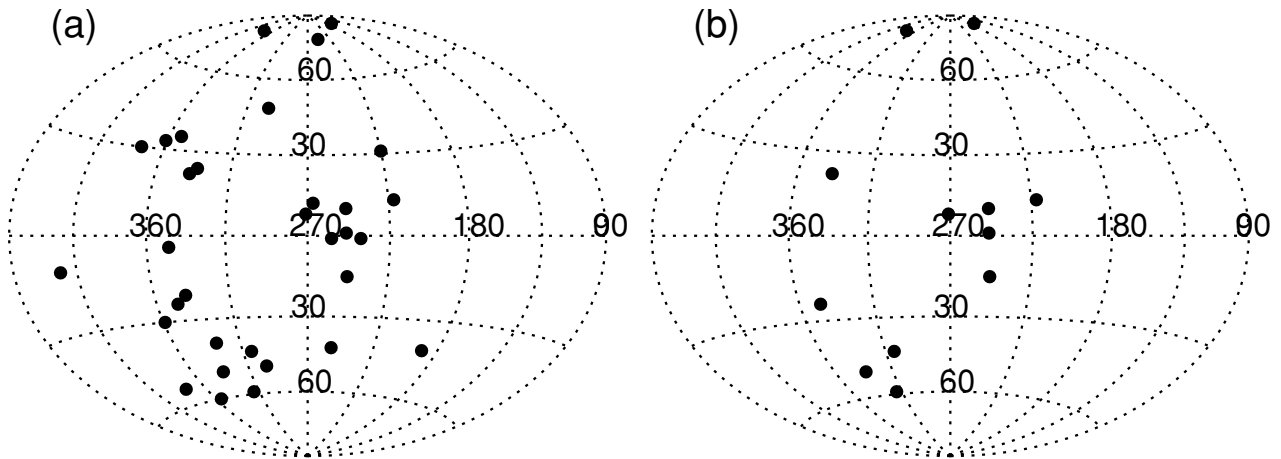| Angular interval	$w \times 10^{-2}$	$dw \times 10^{-2}$	Number
0–5	8.4	14.2	278
5–10	−5.3	8.0	670
10–15	8.0	6.5	161
15–20	−7.3	5.2	830
20–25	6.8	5.0	148
25–30	3.5	4.5	255
30–35	−0.7	4.1	541
35–40	−1.5	3.9	600
40–45	5.9	3.9	122
45–50	−3.6	3.6	747
50–55	−1.9	3.6	654
55–60	3.0	3.6	259
60–65	−8.5	3.3	957
65–70	−3.0	3.5	732
70–75	2.5	3.6	286
75–80	−1.4	3.5	590
80–85	5.3	3.7	131
85–90	−7.2	3.6	937
90–95	−4.6	3.7	812
95–100	1.6	3.9	369
100–105	2.9	4.0	297

Figure 1—Distribution in galactic coordinates of SUGAR events with energies above (a) 6×10^{19} eV and (b) 8×10^{19} eV.

6 Interpretation of Two-point Angular Data

Data such as those discussed in the previous section are studied to investigate the possibility of there being information in air shower directional data pertaining to possible cosmic ray source directions. Figure 1a shows the SUGAR directional distribution for the highest energy included in Table 3. As we discussed above, based on the two-point angular autocorrelation function, there may be underlying evidence for clustering on a scale of a few tens of degrees. In examining Figure 1a, one can subjectively see the origin of this effect in an apparent non-uniform distribution within the accessible range of the sky on scales of tens of degrees. Without our shuffling method or some equivalent, this is difficult to quantify. It appears that the data show some groupings within the overall non uniformity, together with an admixture of non-grouped directions. These groupings are arguably at galactic longitudes of about 240°, near the Galactic equator (RA 116°, dec –24°) plus a region at about 330° of Galactic longitude and –40° of Galactic latitude (RA 320°, Dec –64°). Figure 1b shows the same data but with a rather higher energy threshold; here the apparent clustering is clearer.

We hope that by studying the arrival directions of the highest-energy cosmic rays, we might derive directional information such as this which will enable us to identify some sources. Such sources are believed to be extragalactic because, despite there being no significant clustering of arrival directions along the Milky Way, known Galactic magnetic fields are believed to be incapable of randomising those directions from such nearby sources.

Our directional data will be limited by the angular uncertainty of the detection array and by any intergalactic magnetic deviations. Even with an angular uncertainty limited only by the array to the order of one degree, many astronomical objects will be within the 'beam'. It will be necessary to have a selected catalogue of potential extragalactic sources to limit such confusion.

In order to investigate this process, we have chosen some catalogue selection criteria and used these to select potential sources from a complete, but less selective, catalogue. We took the catalogue of strong radio sources produced by Robertson (1973). This is complete above 10 Jy. We are thus assuming that strong radio sources are potential sources of the highest-energy cosmic rays. This is consistent with a number of proposals, such as those discussed by Hillas (1984) as potential source regions. For instance, the catalogue includes FR II sources and clusters of galaxies. The selection criterion of strong emission should ensure that the catalogued sources have a combination of closeness and strong activity.

Of course, if the sources were to be active galactic nuclei, which are variable, possibly short-lived and beamed (such that we only observe a subset of all possible sources), this process could reasonably be regarded as hopeless, since magnetic deviations result in the cosmic rays lagging the presently observable light or radio waves by millions of years. Such sources are intrinsically variable but they are also likely to beamed, and a natural result of long-period precession of the central black hole would be to move the most intense part of the photon beam away from the observer for large periods of time (e.g. Hartle, Thorne & Price 1986; Jackson 1999). It could be that the appropriate sources of the highest-energy cosmic rays are not even visible as bright photon sources at the present time.

In developing our catalogue from that of Robertson, we assumed that any potential source must have a redshift below 0.07. This is a generous allowance for the distance limitation due to photopion energy losses. This reduces the number of catalogued objects from 160 to 42. Based on the NASA Extragalactic Database, there are two sources in the catalogue within 5° of the centre of the first clustering noted above. These are IRAS 06343–2032 and Hydra A. There are also two sources for the second cluster centre but they are both within the same structure, the cluster of galaxies A3667. The IRAS

source is an FR II object and Hydra A is a cluster of galaxies with significant X-ray luminosity (Ikebe et al. 1997). At the centre of Hydra A is 3C218, a complex, strong radio source with the highest Faraday rotation ever measured for a radio galaxy. A3667 is a source that has been mentioned as a likely type of source for the highest-energy cosmic rays. It is a huge galactic cluster with strong X-ray emission and two very large shock structures. All of these sources have redshifts that are a little over 0.05. If one were to use the propagation model of Clay et al. (1998), a fit to the angular spreads would result from a characteristic intergalactic magnetic field strength of 0.1 µG together with a largest turbulence scale of 100 kpc.

It is clear that the process we have described, of defining a general potential source direction and then examining a purpose-developed catalogue of potential sources, can select a manageable number of plausible cosmic ray sources for further study. It is then possible for further interesting astrophysical information to become apparent.

7 Conclusions

We have examined the use of the two-point angular autocorrelation function in the analysis of directional data concerning the highest-energy cosmic rays. We have done this with data from the SUGAR air shower array in anticipation of data becoming available, and requiring interpretation, from the Pierre Auger Observatory within a few years. The process was to identify any possible clustering in the directional data. We recognise that in the future, this identification must be based on predetermined criteria. The next step is to examine the direction of any clustering in comparison with a preselected catalogue of potential sources.

We find that there is some evidence for clustering in the arrival directions contained in the SUGAR data, and that clustering is related in direction to three potential source objects in a catalogue which we had chosen. These were an FR II radio source, a cluster of galaxies containing a strong radio source, and a very large cluster of galaxies that has already been suggested as a possible source of the highest-energy cosmic rays.

References

Allen, W. H., et al. 1993, Astroparticle Physics, 1, 269

Clay, R. W., Dawson, B. R., & Meyhandan, R. 1994, Astroparticle Physics, 2, 347

Clay, R. W., et al. 1998, Astroparticle Physics, 9, 221

Clay, R. W., et al. 2000, Astroparticle Physics, 12, 249

Hartle, J. B., Thorne, K. S., & Price, R. H., 1986, in Black Holes: The Membrane Paradigm, ed. K. S. Thorne, R. H Price & D. A. McDonald (New Haven: Yale University Press), p. 160

Hayashida, N., et al. 1999, Astroparticle Physics, 10, 303

Hillas, A. M. 1984, ARA&A, 22, 425

Ikebe, Y., et al. 1997, ApJ, 481, 6601

Jackson, C. A. 1999, PASA, 16, 124

Lampard, R., Clay, R. W., & Dawson, B. R. 1997a, PASA, 14, 258

Lampard, R., Clay, R. W., & Dawson, B. R. 1997b, Astroparticle Physics, 7, 213

Muecke, A., et al. 1999, PASA, 16, 160

Peebles, P. J. E. 1980, in The Large Scale Structure of the Universe (Princeton: Princeton University Press), p. 174

Robertson, J. G. 1973, Aust. J. Phys., 26, 403

Winn, M. M., et al. 1986, J. Phys. G: Nucl. Phys., 12, 653