Correlation function of ultra-high energy cosmic rays favors point sources.

P.G. Tinyakov[a,c] and I.I. Tkachev[b,c]
[a] *Institute of Theoretical Physics, University of Lausanne, CH-1015 Lausanne, Switzerland*
[b] *CERN Theory Division, CH-1211 Geneva 23, Switzerland*
[c] *Institute for Nuclear Research of the Academy of Sciences of Russia, Moscow 117312, Russia*

We calculate the angular two-point correlation function of ultra-high energy cosmic rays (UHECR) observed by AGASA and Yakutsk experiments. In both data sets, there is a strong signal at highest energies, which is concentrated in the first bin of the size of the angular resolution of the experiment. For the uniform distribution of sources, the probability of a chance clustering is 4×10^{-6}. Correlations are absent or not significant at larger angles. This favors the models with compact sources of UHECR.

1. The measurements of the flux of UHECR at energies of order 10^{20} eV [1–5] provide compelling evidence of the absence of the Greisen-Zatsepin-Kuzmin (GZK) cutoff [6]. The resolution of this puzzle seems to be impossible without invoking new physics or extreme astrophysics. All models suggested so far can be classified in three groups, according to the way the GZK cutoff is avoided: i) "one nearby source" ii) non-standard interaction with CMB iii) bump in the injection spectrum.

The possibility i) assumes that a substantial fraction of the observed UHECR comes from a relatively nearby source(s) and thus is not subject to the GZK cutoff. This idea may be realized in different ways, examples being the models of decaying superheavy dark matter [7], or the model where UHECR are heavy nuclei accelerated by young NS in our galaxy [8]. It may also happen, as in the model of ref. [9], that an extragalactic astrophysical source is — by pure chance — so close to us that it produces a flux comparable to the rest of the Universe. Two last examples require diffuse propagation of UHECR in the magnetic fields. Although models of this type generically predict large-scale anisotropy [10,9], they might still work.

In the option ii) the GZK cutoff is eliminated (or shifted to higher energies) by assuming non-standard interaction of primary particles with the cosmic microwave background. This may happen, for instance, if primary particles are hypothetical light SUSY hadrons [11], or if the Lorentz invariance is slightly violated at high energies [12]. Finally, the possibility iii) can be realized in models which involve topological defects [13] or in the model where primary particles are neutrinos [14].

Existing experimental data hint also at another important feature of UHECR, namely, the clustering at small angles [15]. The AGASA collaboration has reported three doublets and one triplet out of 47 events with energies $E > 4 \times 10^{19}$ eV, with chance probability of less than 1% in the case of the isotropic distribution [16]. The world data set has also been analysed; 6 doublets and 2 triplets out of 92 events with energies $E > 4 \times 10^{19}$ eV were found [17], with the chance probability less than 1%.

If not a statistical fluctuation, what does the clustering imply for models of UHECR? There are two possible situations: either clustering is due to the existence of point-like sources, or it is a result of large variations in the number density of sources over the sky (if there are regions in the sky with higher density of sources, these regions are more likely to produce clusters of events). In the first case, the models which involve the diffusive propagation of UHECR, like those of refs. [9,8], would be excluded. Moreover, in that case one could estimate the number density of sources [18] (see also ref. [19]). The observation of point-like sources would also imply that there is no defocusing of UHECR in the magnetic fields during their propagation. Such defocusing occurs even in the regular field (e.g., in the galactic magnetic field), since different events in a cluster have different energies. Thus, one can put bounds on the charge of the primary particles. For the extra-galactic rays, knowing primary particles are charged would imply direct bounds on the extra-galactic magnetic fields.

In the second case, the regions of higher luminosity may reflect the actual higher density of sources, as in the models of superheavy dark matter (they would correspond to dark matter clumps in the halo). Alternatively, they may result from the defocusing of the point sources in the random magnetic fields, or may be due to magnetic lensing [20].

In order to determine which of these two cases fits the present experimental data better, it is not enough to know the probability to have a certain number of clusters. In this respect previous analyses [15–17] are not sufficient. One has to find the angular correlation function. This is the approach we accept in this paper.

2. The two-point correlation function for a given set of events is defined as follows. For each event, we divide the sphere into concentric rings (bins) with fixed angular size (say, the angular resolution of the experiment). We count the number of events falling into each bin, sum over all events and divide by 2 to avoid double counting, thus obtaining the numbers N_i. We repeat the same procedure for a large number (typically 100 000) of randomly generated sets and calculate the mean Monte-Carlo value

$N_i^{\rm MC}$ and the variance $\sigma_i^{\rm MC}$ for each bin in a standard way. We then define the correlation function as

$$f_i = \frac{N_i}{N_i^{\rm MC}} - 1.$$

A deviation of f_i from zero indicates the presence of the correlations on the angular scale corresponding to i-th bin.

The correlation function f_i fluctuates. In order to see whether its deviation from zero is statistically significant, one has to compare it to $\sigma_i^{\rm MC}/N_i^{\rm MC}$. Thus, we define the ratio

$$r_i = \frac{f_i N_i^{\rm MC}}{\sigma_i^{\rm MC}} = \frac{N_i - N_i^{\rm MC}}{\sigma_i^{\rm MC}},$$

which shows the excess in the correlation function as compared to the random distribution in the units of the variance. With enough statistics, this quantity becomes a good measure of the probability of the corresponding fluctuation.

The Monte-Carlo events are generated in the horizon reference frame with the distribution

$$dn \propto \cos\theta_z \sin\theta_z d\theta_z,$$

where θ_z is the zenith angle. Coordinates of the events are then transformed into the galactic frame assuming random arrival time. We restrict our analysis to the events with zenith angles $\theta_z < 45°$ for which the experimental resolution of arrival directions is the best [17].

We also make a cut in energy. Since there is no *a priori* reason to choose some particular value, it makes sense to look for the energy cut which gives a maximum signal. For instance, one may imagine a situation when the spectrum consists of two components, the clustered component taking over the uniform one at a certain energy. The cut at an energy at which the clustered component starts to dominate should give the most significant signal.

Motivated by these arguments, we calculated the probability of chance clustering as a function of energy cut. For these simulations we took the bin size equal to 2.5° and 4° for AGASA and Yakutsk, respectively, which is the quoted (see e.g. [16,17]) angular resolution of each experiment multiplied by $\sqrt{2}$. The results are summarised in Fig. 1, which shows the probability to reproduce or exceed the observed count in the first bin, as a function of the energy cut. AGASA curve starts at $E = 4 \times 10^{19}$ eV because the data at smaller energies are not yet available. Yakutsk has much lower statistics. Both curves rapidly rise to 1 in a similar way when the statistics becomes poor. They suggest that the optimum energy cut is higher than can be imposed at present statistics.

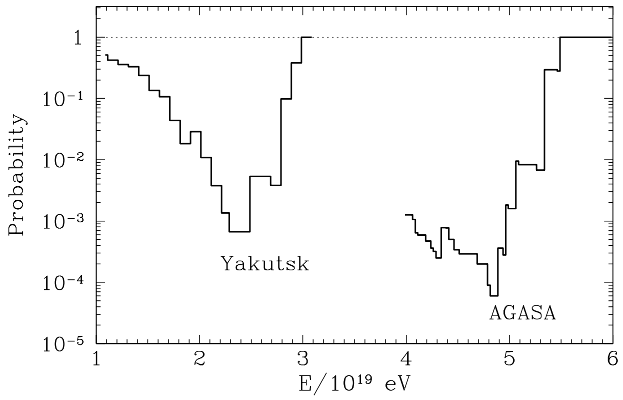

FIG. 1. Probability to have observed count in the first bin as a function of energy for the random distribution of arrival directions.

The difference between our results and those of Ref. [16] (cf. Fig. 1 of this paper and Fig. 12 of Ref. [16]) is due to two reasons. First, 10 more events with $E > 4 \times 10^{19}$ have been observed [21] which bring a new doublet. Second, and more important, we calculate a different probability. The difference arises when there is a triplet or higher multiplets in the data. In our approach a triplet is equivalent to 3 or 2 doublets, depending on the relative position of the events (compact or aligned), while higher multiplicity clusters effectively have larger "weight". In Ref. [16] the probabilities of doublets and triplets are calculated separately; the probability of doublets is defined in such a way that a triplet contributes as 3/2 of a doublet. The drawback of this method is that the probabilities of doublets and triplets are not independent, and it is not clear how to combine them. Triplets and higher multiplicity clusters are better accounted for in our method, and the probability of chance clustering which we get is lower than in Ref. [16].

Correlation functions calculated with the energy cut corresponding to the lowest chance probability is shown in Fig. 2. Both AGASA and Yakutsk correlation functions have substantial excess in the first bin. The peak in AGASA curve corresponds to 6 doublets (of which 3 actually form a triplet). The peak in Yakutsk curve corresponds to 8 doublets (of which 3 also come from a triplet).

Since the number of events in the first bin for AGASA and Yakutsk is not large, the probability to have this or larger number of events in the random set is not well approximated by the Gaussian distribution, and the excess in the units of the variance is not a good measure of the probability. We calculated the probability directly by counting, in the Monte-Carlo simulation, the number of occurances with the same or larger number of doublets in the first bin. Probabilities in the minima are small (see Fig. 1), so we have recalculated them with 10^6 random Monte-Carlo sets.

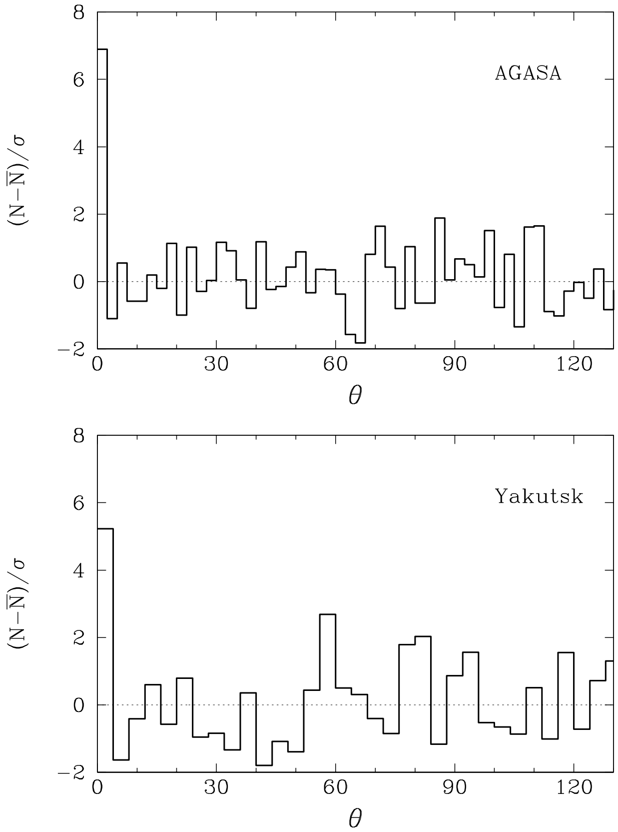

FIG. 2. Angular correlation functions of UHECR with binning angles and cuts in energy quoted in the text.

As the lowest probabilities were obtained by scanning over the energy, one may argue that they have to be multiplied by the number of steps in the scan. This is not, however, correct because the results at different energy cuts are not independent: higher energy set is a subset of the lower energy one. As can be seen from Fig. 1, the chance probability for AGASA is lower than 10^{-3} in the whole energy range $(4-5)\times 10^{19}$ eV, regardless of the number of steps in the scan. There still may be a correction factor. To estimate it we performed the following numerical experiment. For 10^3 randomly generated sets of events we have performed exactly the same procedure as for the real data, i.e. scanned over energies and obtained 10^3 different minimum probabilities. We found that the probability less than 10^{-2} occurred 27 times, while the probability less than 10^{-3} occurred 3 times. Thus we conclude that the correction factor is of order 3. This factor is included in the final results which are presented in Table.1. It does not depend strongly on the number of steps in the scan.

experiment	bin size	$E_{\min}$	probability of chance clustering
AGASA	$2.5°$	4.8×10^{19} eV	3×10^{-4}
Yakutsk	$4°$	2.4×10^{19} eV	2×10^{-3}

It is interesting to estimate the combined probability of clustering observed in AGASA and Yakutsk data sets. One may approximate probabilities listed in the Table 1 as given by two independent Poisson distributions. Considering combined probability as given by a Poisson distribution which is a sum of them we find that the combined probability is 4×10^{-6}.

We now turn to the determination of the angular size of the sources. To this end we calculate the dependence of the probability to have the observed (or larger) number of events in the first bin on the bin size. This dependence is plotted in Fig. 3. Jumps in the curves occur when a new doublet enters the first bin. Despite fluctuations, one can see that the minimum probability corresponds roughly to $2.5°$ and $4°$ for AGASA and Yakutsk, respectively. Remarkably, these numbers coincide with the angular resolutions of the experiments, as is expected for sources with the angular size smaller than the experimental resolution. Thus, we conclude that the data favour compact sources with angular size less than $2.5°$.

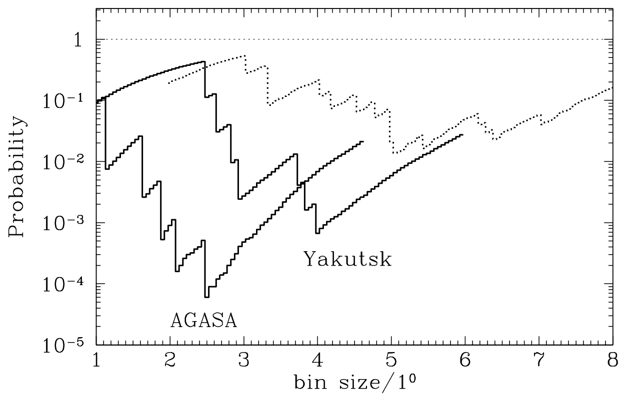

FIG. 3. Probability to have observed count in the first bin as a function of the bin size. Cuts in energy are $E > 4.8\times 10^{19}$ eV and $E > 2.4\times 10^{19}$ eV for AGASA and Yakutsk, respectively. Dotted line shows the same quantity for Yakutsk events with energies $1.5\times 10^{19}\,\mathrm{GeV} < E < 2.2\times 10^{19}$ GeV

There is another interesting possibility which is not excluded at present. If there exists a defocusing of UHECR in magnetic fields, it is inversely proportional to the energy. Since the bulk of the Yakutsk events have energies roughly 2 times smaller than the AGASA events, their typical defocusing angle should be 2 times larger. This is in agreement with Fig. 3 if the angular size of the sources is attributed to defocusing in magnetic fields. Interestingly, this picture is consistent with our result for the remaining low-energy Yakutsk events with energy $1.5\times 10^{19}\mathrm{GeV} < E < 2.2\times 10^{19}$ GeV where we have found a signal with chance probability 2% at the bin size $\sim 5°$ (dotted line in Fig. 3). AGASA data at low energies may help to distinguish between these two possibilities.

We have also studied the correlation function of UHECR with arrival directions corrected for the Galactic magnetic field. With unit electric charge of primary particles ($Z = 1$), the peak in the first bin does not change significantly. One can not discriminate between this case and the case of neutral particles. The peak becomes small at $Z = 2$, and disappears at larger Z.

3. It is extremely unlikely that the clustering observed by AGASA and Yakutsk experiments is a result of a random fluctuation in an isotropic distribution. Rather, the working hypothesis should be the existence of some number of compact sources which produce the observed multiplets (and, consequently, the excess in the first bin). Is this hypothesis consistent with the fact that Haverah Park experiment does not observe significant clustering (with the data of Ref. [3] we find the probability 4% at the bin size 4° and $E_{\rm min} = 5 \times 10^{19}$ eV), even though it looks at the same sources?

If clusters are due to compact sources and their number is small, the relevant quantity which should be common (in average) for all experiments is the fraction of events in clusters [18]. The number of really observed doublets has the Poisson distribution at small clustering. Let us take the AGASA data as an example. At $E > 4.8 \times 10^{19}$ eV there are $N = 6$ doublets in the first bin, while only 0.6 is expected for the random distribution. So let us assume all 6 doublets are due to real sources. Thus, in the case of AGASA, one could have observed with similar probabilities, any number of real doublets between $N_- \sim N - \sqrt{N} \sim 4$ and $N_+ \sim N + \sqrt{N} \sim 8$ doublets. With 4 doublets, the probability to reproduce them with isotropic distribution would be $\sim 1\%$, and one would say that there is no significant excess. With 8 doublets, however, the probability of chance clustering would be as small as 10^{-6}! Therefore, with the same set of sources it is (almost) equally probable to detect clustering at the confidence level ranging from 10^{-2} to 10^{-6} depending on the number of doublets one was lucky to observe in a particular experiment. At the present level of statistics different experiments may see a signal with totally different significance. We find that the fraction of 1 doublet per 10 events on average is roughly consistent with all observations.

The simulations show that the mean expected numbers of chance doublets are 0.6 for AGASA (at bin size 2.5°, $E_{\rm min} = 4.8 \times 10^{19}$ eV) and 1.5 for Yakutsk (at bin size 4°, $E_{\rm min} = 2.4 \times 10^{19}$ eV). Therefore, most of the clusters we have found are likely to be due to real sources. In Fig. 4 we plot these clusters in the Galactic coordinates (small and large circles correspond to AGASA and Yakutsk events, respectively). Positions of triplets are indicated by arrows.

To summarise, the clustering of UHECR is statistically significant and favours compact sources. This places further constraints on models which can resolve the puzzle of the GZK cutoff. Those models which involve large extragalactic magnetic fields, Ref. [9], as well as models with heavy nuclei as primaries, e.g. [8] are disfavoured because they assume total isotropisation of original arrival directions of UHECR. If violation of the Lorentz invariance is the solution of the GZK puzzle, and primaries are protons, our results place extremely strong limit on the extragalactic magnetic fields. Regarding the models of decaying superheavy dark matter, it is important to calculate [22] the angular correlation function predicted by these models and compare it to Fig.2 in order to see if the clumping on subgalactic scales can be responsible for the clustering of UHECR.

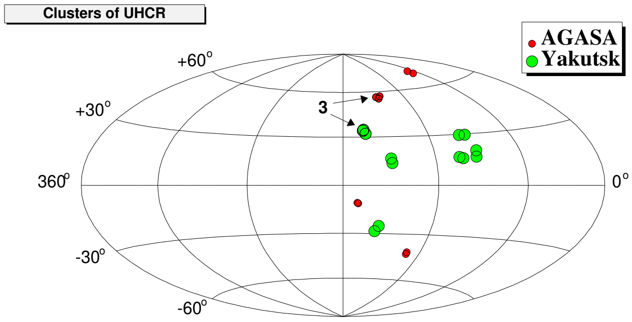

FIG. 4. Observed clusters in Galactic coordinates.

ACKNOWLEDGMENTS

We are grateful to S.I. Bityukov, S.L. Dubovsky, A.A. Mikhailov, V.A. Rubakov, M.E. Shaposhnikov, D.V.Semikoz and M. Teshima for valuable comments and discussions. This work is supported in part by the Swiss Science Foundation, grant 21-58947.99.

---

[1] M. Takeda *et al.*, Phys. Rev. Lett. **81** (1998) 1163, astro-ph/9807193.

[2] D. Bird et al., Ap. J. **441** (1995) 144.

[3] Catalogue of Highest Ehergy Cosmic Rays, No. 1, May 1980, ed. M. Wada, World Data Center C2 for Cosmic Rays, Institute of Physical and Chemical Research, Itabashi, Tokyo, Japan.

[4] Catalogue of Highest Ehergy Cosmic Rays, No. 2, February 1986, ed. A. Inoue, World Data Center C2 for Cosmic Rays, Institute of Physical and Chemical Research, Itabashi, Tokyo, Japan.

[5] Catalogue of Highest Ehergy Cosmic Rays, No. 3, June 1988, ed. A. Inoue, E. Sakamoto, World Data Center C2 for Cosmic Rays, Institute of Physical and Chemical Research, Wako, Saitama, Japan.

[6] K. Greisen, Phys. Rev. Lett. **16** (1966) 748;
G.T. Zatsepin and V.A. Kuzmin, Pisma Zh. Eksp. Teor. Fiz. **4** (1966) 144;

[7] V. Berezinsky, M. Kachelriess and A. Vilenkin, Phys. Rev. Lett. **79** (1997) 4302; V.A. Kuzmin and V.A. Rubakov, Phys. Atom. Nucl.**61** (1998) 1028, Yad.

Fiz. **61** (1998) 1122; V.A. Kuzmin and I.I. Tkachev, Phys. Rept. **320** (1999) 199.
[8] P. Blasi, R.I. Epstein and A.V. Olinto, Astrophys.J.**533** (2000) L123.
[9] G.R. Farrar and T. Piran, astro-ph/0010370.
[10] S.L. Dubovsky and P.G. Tinyakov. JETP Lett.**68** (1998) 107.
[11] D.J.H. Chung, G. R. Farrar and E. W. Kolb, Phys. Rev. **D57** (1998) 4606.
[12] S. Coleman and S. L. Glashow, Phys. Rev. **D59** (1999) 116008.
[13] V. Berezinsky, P. Blasi, and A. Vilenkin, Phys. Rev. **D58** (1998) 103515.
[14] T. J. Weiler, Phys. Rev. Lett. **49** (1982) 234; G. Gelmini and A. Kusenko, Phys. Rev. Lett. **84** (2000) 1378.
[15] X. Chi et al., J. Phys. **G18** (1992) 539; N. N. Efimov and A. A. Mikhailov, Astropart. Phys. **2** (1994) 329.
[16] M. Takeda et al., Ap. J. **522** (1999) 225.
[17] Y. Uchihori et al., Astropart. Phys. **13** (2000) 151.
[18] S.L. Dubovsky, P.G. Tinyakov and I.I. Tkachev, Phys. Rev. Lett. **85** (2000) 1154;
[19] Z. Fodor and S. D. Katz, Phys. Rev. **D63** (2001) 023002.
[20] M. Lemoine, G. Sigl and P. Biermann, astro-ph/9903124.
[21] N. Hayashida et al, astro-ph/0008102.
[22] S. Colombi, P. Tinyakov and I. Tkachev, in preparation.