Clustering of ultrahigh energy cosmic rays and their sources

N. W. Evans
Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 OHA, United Kingdom

F. Ferrer and S. Sarkar
Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, United Kingdom
(Received 23 December 2002; published 22 May 2003)

The sky distribution of cosmic rays with energies above the ''GZK cutoff'' holds important clues to their origin. The AGASA data, although consistent with isotropy overall, shows evidence for small-angle clustering, and it has been argued that such clusters are aligned with BL Lacertae objects, implicating these as the sources. It has also been suggested that such clusters can arise if the cosmic rays come from the decays of very massive relic particles in the galactic halo, due to the expected clumping of cold dark matter. We examine these claims and show that both are, in fact, unjustified.

DOI: 10.1103/PhysRevD.67.103005 PACS number(s): 98.70.Sa, 95.35.+d

I. INTRODUCTION

The mystery of the ultrahigh energy cosmic rays (UHECRs) with energies exceeding $E_{\rm GZK} \simeq 4\times 10^{19}$ eV—the Greisen-Zatsepin-Kuzmin (GZK) ''cutoff'' [1,2]—continues to deepen. This energy sets the threshold for photomeson production on the cosmic microwave background so the observation of such UHECRs (assumed to be protons or heavier nuclei) would indicate that the sources are relatively nearby, within the local supercluster of galaxies [3]. Recent observations by the HiRes air fluorescence detector [4] are however inconsistent with previously published data from the Akeno Giant Air Shower Array (AGASA) which ruled out such a cutoff with a significance $\gtrsim 5\sigma$ [5,6]. HiRes has reported only 1 event above 10^{20} eV, whereas about 20 would have been expected on the basis of the AGASA spectrum. The two spectra can be made to agree *below* this energy if the energies of the AGASA events are systematically lowered by 20% (within the quoted uncertainty); however, 5 of them still remain above this energy [7]. Subsequently the AGASA Collaboration have carefully assessed their energy measurement uncertainties and reaffirmed that their observed spectrum does extend well beyond the GZK energy [8]. To resolve this situation requires making simultaneous measurements using both the air shower and air fluorescence methods; such measurements are underway at the Pierre Auger Observatory being constructed in Argentina [9,10].

Another development has been the AGASA observation that the UHECR arrival directions, although consistent with isotropy overall, exhibit clustering on small angular scales [11]. Among the 59 AGASA events above 4×10^{19} eV, there are 5 ''doublets'' and 1 ''triplet'' with separation angle less than the estimated angular resolution of 2.5° [12].[1] The probability for this to arise by chance from an isotropic distribution is less than 0.1%. However, this probability is very sensitive to the assumed angular resolution [13], e.g. increasing to $\sim 1\%$ if the angular resolution is 3° [14]. Moreover, adding data from three other air shower experiments (Volcano Ranch, Haverah Park, and Yakutsk) *dilutes* the significance. In an earlier such analysis [15], 8 doublets and 2 triplets were found in a dataset of 47 AGASA plus 45 other events with $E > 4\times 10^{19}$ eV, taking the effective angular resolution of the dataset to be 4°. The chance probability for this to arise from an uniform distribution is $\sim 10\%$, thus statistically not significant.

Nevertheless, the existence of such clusters has been linked to the possibility of (repeating) point sources of UHECR [16,17], specifically cosmologically distant BL Lacertae [18]—a sub-class of active galactic nuclei (AGN) which have been long discussed as possible accelerators of UHECRs [19]. However, the expected deflections of UHECRs (assumed to be charged particles) by galactic and intergalactic magnetic fields ought to smear out such tight source correlations [20–22].[2] Contrary to these results, it has been claimed recently that the correlations with BL Lacs are preserved, even improved, if the UHECRs are protons, after allowing for deflections by the galactic magnetic field [26]. Little is known about the intergalactic magnetic field [27]; requiring rectilinear propagation of protons over the attenuation length of $l \sim 1000$ Mpc at $E > 4\times 10^{19}$ eV (decreasing to ~ 100 Mpc at $E > 10^{20}$ eV [28]) would imply that its homogeneous component on such scales is extremely weak: $B < 2\times 10^{-12}(l/1000\ {\rm Mpc})^{-1}$ G [29]. It has also been claimed [30,31] that such clustering is predicted in a model where the UHECR arise from the decay of superheavy relic particles accumulated in the galactic halo [32,33], due to the expected clumping of halo dark matter.

In this paper we examine both these claims in detail, using as our basic statistical tool the two-point correlation function. Our intention is to determine whether the claimed correlations are meaningful, given the present limited event statistics.

---

[1] For simulated events with $E > 4\times 10^{19}$ eV, 68% have a reconstructed arrival direction within 1.8° of the true direction and 90% within 3°; the corresponding angles for all events above 10^{19} eV are 2.8° and 4.6°, keeping in mind that the energy resolution is $\pm 30\%$ [6,8].

[2] In fact focusing effects by such fields can give rise to apparent clustering even when the arrival directions are random [23–25].

II. UHECR CLUSTERING AND CORRELATIONS WITH POSSIBLE SOURCES

It is natural to look for correlations between the observed UHECR arrival directions and plausible astrophysical sources; however, it is essential to take care not to generate spurious correlations by introducing biases. For example, it has been claimed that the 5 highest energy events with $E > 10^{20}$ eV are all aligned with compact radio-loud quasars (CRQSOs) having redshifts between 0.3 and 2.2, and the chance probability for this coincidence was estimated to be 0.5% [34,35]. However, this rises to 3% when the event used to formulate the hypothesis itself (the previously noted [36] alignment of the quasar 3C147 with the 3.2×10^{20} eV Fly's Eye event [37]) is excluded from the sample [38]. A careful recent analysis [39] based on an updated event list (5 AGASA [6], 4 Haverah Park [40,41] and 1 Fly's Eye [37]) demonstrates that there are *no* significant correlations between UHECRs and CRQSOs. These authors show also that another recent claim [42] of significant correlations with CRQSOs is based on inadequate data, and, in addition, that there are *no* significant correlations with an interesting subgroup of these sources, viz. γ-ray blazars [39]. A correlation between events with $E > 4\times10^{19}$ eV and nearby galaxies likely to host quasar remnants (QRs) has also been found at the 3σ level, although this disappears if attention is restricted to events above 10^{20} eV [43].

What has revived interest in the possibility of such correlations is the claimed clustering in the arrival directions of UHECRs [11]. This may arise for example if the sources are "compact" (i.e. smaller than the experimental angular resolution) with the clusters corresponding to more than one UHECR being received from the *same* source. Since the number of events in such clusters is much smaller than the total number of events, the majority of such sources have clearly not been seen at all. However, it is possible to estimate their number density using Poisson statistics. Allowing for the attenuation of UHECRs from distant sources due to GZK energy losses, the observed occurrences of "triplets" and "doublets" relative to the number of single events was used to estimate the spatial density of such sources to be 6×10^{-3} Mpc^{-3} [16]. This would obviously place stringent constraints on candidate astrophysical sources, e.g. γ-ray bursts (GRBs) have a spatial density of only $\sim10^{-5}$ Mpc^{-3}. However, a more careful analysis [44] shows that the uncertainties in this estimate are very large. The true number is $2.77^{+96.1(916)}_{-2.53(2.70)}\times10^{-3}$ Mpc^{-3} at the 68% (95%) C.L.; moreover relaxing the assumptions made, viz. that the sources all have the same luminosity and a spectrum $\propto E^{-2}$, increases the allowed range even further, e.g. to $180^{+2730(8817)}_{-165(174)}\times10^{-3}$ Mpc^{-3} for a Schechter luminosity function and a spectrum $\propto E^{-3}$ [44]. Clearly the present limited event numbers do *not* permit any candidate class of sources to be definitively excluded. Note that the observed clustering may also arise because of a higher density of local sources in certain directions, e.g. due to the clumpiness of halo dark matter in the decaying dark matter model.

The next step taken was the construction of the angular autocorrelation function of UHECRs [17]. For the AGASA data this displays a clear peak at separation angles less than 2.5°, consistent with the point spread function [12]. Moreover the chance probability estimated by Monte Carlo to match or exceed the observed count in the first angular bin, when plotted versus energy, is seen to have a minimum at E_{GZK}; the peak in the autocorrelation function for $E > 4\times10^{19}$ eV is stated to have a significance of 4.6σ [12]. An equally significant autocorrelation was claimed using a different method of analysing the data in which a "triplet" was taken to correspond to three or two "doublets" depending on whether the events are bunched together or linearly aligned [17]. These authors found the probability for chance coincidences to be minimum for events above 4.8×10^{19} eV in the AGASA data [6], and above 2.4×10^{19} eV in the Yakutsk data [45]. Restricting attention to events above these energies, the chance probability for the observed clustering in the first angular bin was quoted as 3×10^{-4} for AGASA and 2×10^{-3} for Yakutsk, taking 2.5° and 4°, respectively, for the size of the first bin, corresponding to the respective experimental angular resolutions [17].

A. Correlation with BL Lacs

Motivated by the results quoted above which implicate compact sources for UHECRs, Tinyakov and Tkachev [18] have proposed that the sources are in fact BL Lacs. The physical motivation they provided for this is that only AGNs in which the central jet points towards us—"blazars"—are likely to be UHECR sources (since particles accelerated in a relativistic jet are strongly beamed), and among all blazars, BL Lacs in particular have few emission lines in their spectra, indicating low density of ambient matter and radiation, thus presumably more favorable conditions for particle acceleration.

Tinyakov and Tkachev [18] used a catalogue of AGNs and quasars containing 306 confirmed (out of 462 listed) BL Lacs [46]. They asserted that since the ability of BL Lacs to accelerate UHECRs may be correlated with optical and radio emissions, it would be appropriate to select the most *powerful* BL Lacs by imposing cuts on redshift, apparent magnitude and 6 cm radio flux. In fact the redshift is unknown for over half of all confirmed BL Lacs but they assumed that all such BL Lacs are at $z > 0.2$ and included them in the sample anyway. By imposing the cuts $z > 0.1$ (or unknown), $m < 18$, and $F_6 > 0.17$ Jy, they selected a sample of 22 BL Lacs.

They considered 39 AGASA events with $E > 4.8\times10^{19}$ eV and 26 Yakutsk events with $E > 2.4\times10^{19}$ eV, the energy cuts being motivated by their earlier autocorrelation analysis [17] which had indicated that the small-angle clustering of UHECRs is most pronounced above these energies in the respective data sets. Assuming that the event energies are not important for small angle correlations, they combined these into one set of 65 events. Then they computed the correlation between the arrival directions of these UHECRs and the selected 22 BL Lacs, finding a significant number of coincidences. Eight UHECRs were found to be within 2.5° of 5 BL Lacs, the chance probability of which is only 2×10^{-5}.[3] The authors acknowledged that the imposition of the arbitrary cuts made on the BL Lac catalogue can affect the significance of this result and estimated the "penalty factor" to be about 15; however, the significance of the coincidences taking this into account was then quoted as 6×10^{-5} (implying a penalty factor of only 3) [18]. This was the basis for their claim that BL Lacs are the probable sources of UHECRs.

Since these are cosmologically distant sources, it is pertinent to ask how the UHECRs get to the Earth. Initially Tinyakov and Tkachev [18] inferred that the primaries have to be neutral, i.e. photons or neutrinos (unless the GZK effect is inoperative because of violation of Lorentz invariance). However, in subsequent work [26] they found that the correlations are *improved* if the primaries are assumed to be protons, whose trajectories are modified by the galactic magnetic field (GMF). In this work they used the full set of 57 AGASA events with $E>4\times10^{19}$ eV (but no Yakutsk events) and allowed for deflections by the regular component of the GMF (but ignored the fluctuating component which is in fact of comparable strength [47]), while assuming that deflections by intergalactic magnetic fields (IGMF) are *negligible*. The same BL Lac catalogue [46] was used but this time no cuts were made on redshift or the 6 cm radio flux, only on the apparent magnitude ($m<18$) since this maximized the correlations. It was found that 18 BL Lacs then lie within 2.5° of the reconstructed arrival directions of 22 UHECRs, if these mainly have charge +1 (however, 8 might alternatively be neutral and 4 must be neutral). Of these 18 BL Lacs, only 6 have measured redshifts and these authors now proposed [26], contrary to their previous supposition, that the rest must in fact have redshifts *less than* 0.1 in order that the protons they emit can overcome the GZK losses and reach the Earth.[4] They asserted further [26] that their success at finding significant correlations between BL Lacs and UHECRs in this manner confirmed that BL Lacs are the sources, *as well* as validating their adopted model of the GMF, *and* their assumption that there are no significant deflections due to the IGMF.

Moreover, Tinyakov and Tkachev [26] noted that many of these 22 BL Lacs are x-ray sources. In subsequent work [48], an updated catalogue of QSOs containing 350 confirmed BL Lacs [49] was examined for correlations with the third EGRET catalogue of γ-ray sources [50] and 14 were identified as strong γ-ray emitters (of which 8 were known to be so already). Correlations between these 14 BL Lacs and the set of 39 AGASA plus 26 Yakutsk events selected earlier [18] were then studied, again allowing for deflections by the GMF modelled as in earlier work [26]. It was found that there are 13 possible coincidences within 2.7° (for charge 0 or +1) with a chance probability of 3×10^{-7} [48]. Leaving out the 2 BL Lacs that are invisible to the Northern Hemisphere cosmic ray experiments, 8 of the remaining 12 are found to be along the (reconstructed) arrival directions of UHECRs. It was concluded [48] that γ-ray emission is the physical criterion for a BL Lac to be a UHECR source. However, UHECRs are known *not* to correlate with γ-ray blazars [39]. It was stated that there is no contradiction because the BL Lacs considered display a low degree of polarization whereas γ-ray blazars are highly polarized [48].

Given this set of interesting claims we wish to ascertain to what extent the strong correlations found depend on the selection criteria used. To do so we calculate the correlation function in the same manner as Tinyakov and Tkachev [17,18,26] and use Monte Carlo simulations to determine the probability of chance coincidences. We consider four cases using the AGASA data [6]; we do not consider the Yakutsk data [45] because the events which contribute dominantly to the correlations found earlier [17,18,48] have energies *below* the GZK cutoff, where the uncertainty in the arrival directions exceeds 4°, so the correlations found at smaller angles cannot be meaningful.

Our 4 models correspond to considering:

(1) the 39 UHECR with $E>4.8\times10^{19}$ eV [6] and 22 BL Lacs selected by the Tinyakov and Tkachev criteria [18],

(2) the full set of 57 UHECR with $E>4\times10^{19}$ eV [6], but retaining the cuts on the BL Lacs [18],

(3) the full set of 57 UHECR with $E>4\times10^{19}$ eV [6] and the full set of 306 BL Lacs with no cuts [46],

(4) the full set of 57 UHECR with $E>4\times10^{19}$ eV [6] and 915 GRBs [51].

The last case is a control to determine whether there are equally significant correlations with other suggested sources of UHECRs at cosmological distances, which are *not* expected to contribute events beyond the GZK cutoff [52].

In Fig. 1 we plot the positions on the sky (Hammer-Aitoff projections in equatorial coordinates) of both the UHECRs and the selected objects in order to give a visual impression of how the coincidences arise, particularly for the "doublets" and "triplet" in the AGASA data. The two-point correlation function for the four cases are shown in Fig. 2, calculated according to the Tinyakov and Tkachev prescription [17,18,26], adopting the same angular bin size of 2.5°. To determine the significance of these correlations we run 10^5

[3]The most significant correlation listed is the alignment of a BL Lac (1ES 0806+524) with a "triplet" in the Yakutsk data having energies of $(3.4,2.8,2.5)\times10^{19}$ eV. Note that these events are all *below* the GZK cutoff. Moreover, the uncertainty in the arrival direction of Yakutsk events is 4° at 4×10^{19} eV, and even higher at lower energies, so these close alignments are unlikely to be physically significant.

[4]However, of the 6 BL Lacs at known distances, only B2 0804+35 ($z=0.082$) is near enough to be a possible source of the 4.09×10^{19} eV event (assumed to be charge +1) it is associated with; the other object (TXS 0806+315, $z=0.22$) which is also closely aligned with this event is probably too far on the basis of UHECR propagation calculations [28]. The remaining 4 pairings are also implausible—these are RX J1058.6+5628 ($z=0.144$) aligned with a "doublet" (charge 0) having energies (5.35,7.76) $\times10^{19}$ eV, TEX 1428+370 ($z=0.564$) aligned with an event (charge +1) of energy 4.97×10^{19} eV, 1ES 1853+671 ($z=0.212$) aligned with an event (charge +1) of energy 4.39×10^{19} eV, and EXO 1118.0+4228 ($z=0.124$) aligned with an event (charge 0 or +1) of energy 7.21×10^{19} eV. We have estimated distances from the redshifts using the Mattig formula, $d\sim4500$ Mpc$[z-0.2z^2]$, indicated by measurements of cosmological parameters.

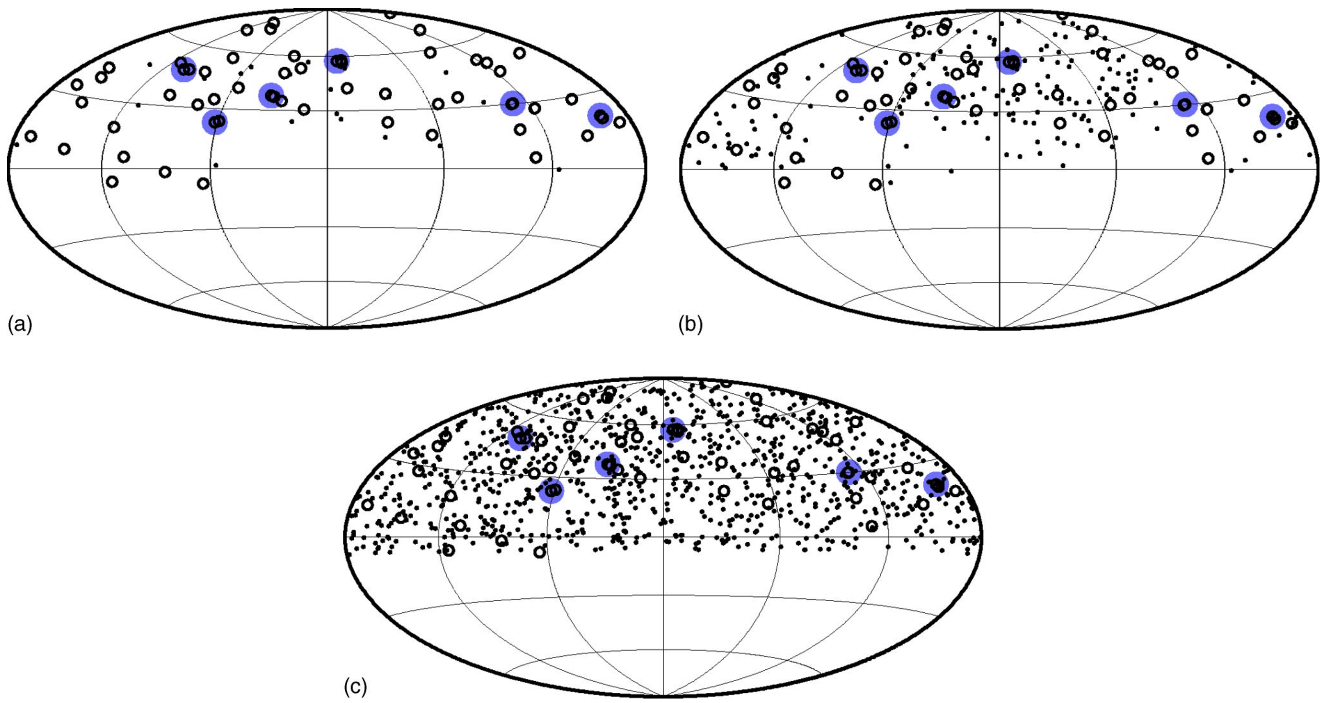

FIG. 1. (Color online) The sky distribution of 57 UHECRs (circles) with $E>4\times10^{19}$ eV observed by AGASA [6] is shown, with the 5 ''doublets'' and 1 ''triplet'' marked with blue circles. Panel (a) shows also the 22 BL Lacs (dots) satisfying the cuts on redshift, magnitude and 6 cm radio flux imposed by Tinyakov and Tkachev [18], while panel (b) shows all 306 BL Lacs in the catalogue [49]. Panel (c) shows instead the 915 GRBs observed by BATSE [51].

Monte Carlo simulations, as they did, to calculate the probability of chance coincidences. It is evident that while there is indeed a $\sim3\sigma$ correlation between UHECRs and BL Lacs if suitable cuts are employed [18], the significance weakens to $\sim2.7\sigma$ if the energy cut is relaxed, and *disappears* if the cuts on BL Lacs are also relaxed. Thus there is *no* basis for the claim that BL Lacs are the sources of UHECRs; indeed cosmologically distant GRBs correlate just as well with post-GZK UHECRs as do BL Lacs.

We now give details of the four cases considered, with reference to Table I.

Model 1: Of the 22 BL Lacs with $z>0.1$, $m<18$ and $F_6>0.17$ Jy considered by Tinyakov and Tkachev [18], 20 are visible to AGASA [6] which reported 39 events with $E>4.8\times10^{19}$ eV. Adopting an angular width of 2.5° for each bin (corresponding approximately to the experimental angular resolution) only 0.8 coincidences are expected on average while 5 are observed, the chance probability for which is 0.15%. It is the coincidences of 2 BL Lacs (RX J10586+5628 and 2EG J0432+2910) with UHECR doublets which contributes most of this signal, which has a significance of $\sim3\sigma$ if the statistics are Gaussian. No coincidences with triplets are seen either in the data or in the Monte Carlo; note that the triplet coincidence emphasized by Tinyakov and Tkachev [18] was composed of Yakutsk events well below the GZK energy, having arrival directions uncertain by $>4°$ [45].

Model 2: Retaining the cuts in the BL Lacs, we now consider all 58 events with $E>4\times10^{19}$ eV in the AGASA catalogue [6]. The probability of clustering has now decreased by a factor of ~5 and has a significance of only $\sim2.7\sigma$, being still mostly due to the coincidences of the 2 BL Lacs with UHECR doublets.

Model 3: Next we consider all 172 BL Lacs [49] which are visible to AGASA. The correlation has now disappeared completely. The 2 coincidences of BL Lacs with UHECR doublets now has an accidental probability of 6.3%.

Model 4: Finally we consider the correlation with 915 GRBs in the BATSE catalogue [51] which are visible to AGASA. The (lack of) correlation is just as significant as for the full set of BL Lacs. There are 4 coincidences of GRBs (4B 920617C, 4B 931211, 4B 950131 and 4B 960128) with 3 UHECR doublets [having energies $(21.3,5.07)\times10^{19}$ eV, $(5.47,4.89)\times10^{19}$ eV, and $(5.50,7.76)\times10^{19}$ eV]—the first and third GRBs coincide with the same doublet, while the last GRB coincides with two of the three events forming the AGASA triplet.

In Fig. 3 we show the dependence of the correlation probability on the angular bin width for all 4 models. There is indeed a minimum at 2.5° corresponding to the AGASA angular resolution if the cuts employed by Tinyakov and Tkachev [18] are imposed (model 1). However, we see that the significance decreases as we relax the cut on AGASA events (model 2) and disappears altogether if we remove the cuts on the BL Lacs (model 3). There is similarly no minimum for correlations with GRBs (model 4). We conclude that Tinyakov and Tkachev [18] vastly underestimated the ''penalty factor'' corresponding to the arbitrary cuts they imposed *post facto* on the data.

III. CLUSTERING AND HALO DARK MATTER

In cold dark matter cosmogonies, galaxies are built up from the merging and accretion of smaller structures. This

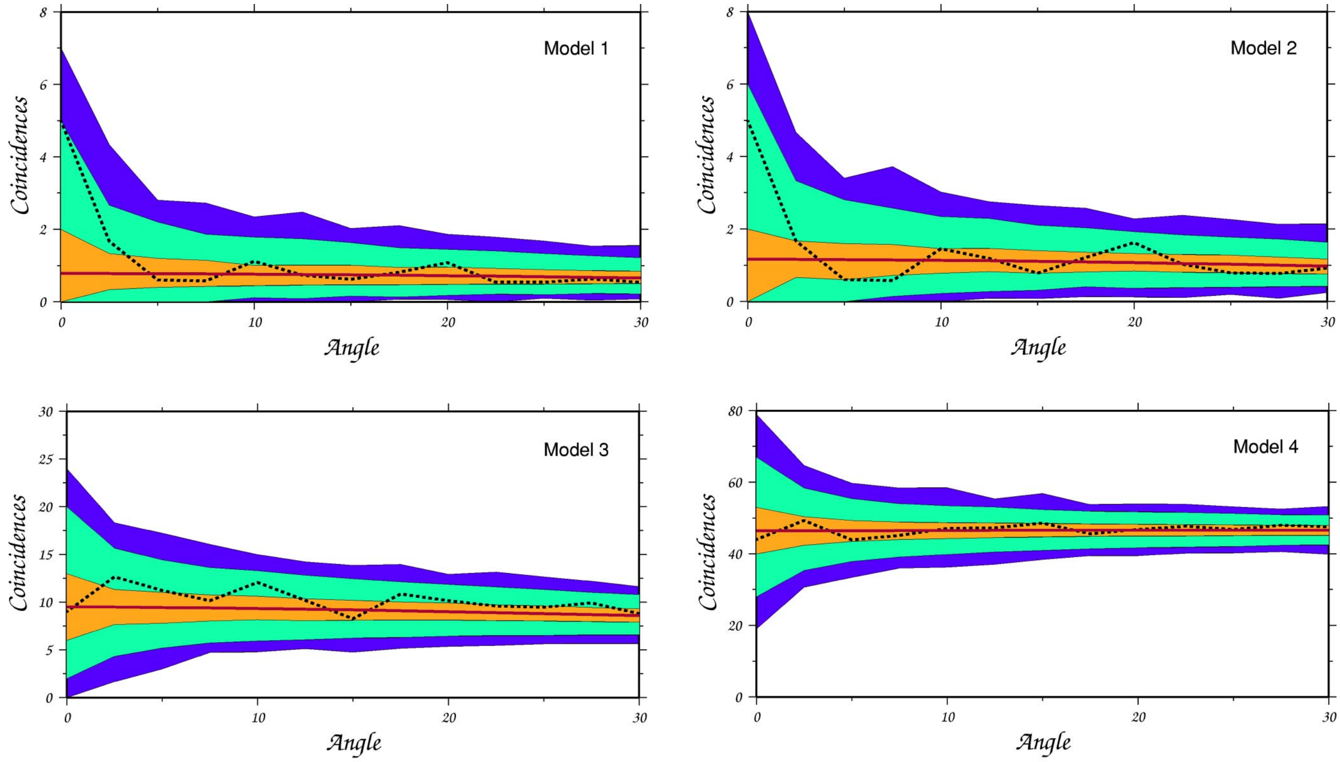

FIG. 2. (Color online) The two-point correlation function between UHECRs observed by AGASA and BL Lacs for the 4 models discussed, viz. using (1) the energy cuts and selection criteria employed by Tinyakov and Tkachev [18], (2) relaxing the cut in energy but retaining the selection criteria for BL Lacs, (3) with no cuts at all, and (4) using a catalogue of GRBs instead of BL Lacs. The regions shaded light, medium and dark correspond to chance probabilities less than 31.73%, 0.27% and 0.0057% (chosen so as to correspond to 1σ, 3σ and 5σ significance for Gaussian statistics), as estimated by running 10^5 Monte Carlo simulations.

leaves phase space substructure in the form of clumps and streams imprinted in galaxy haloes today. Numerical simulations of galaxy formation suggest that $\lesssim 10\%$ of the total halo mass may be in the form of such substructure [53,54]. However, care is needed in interpreting the results of such simulations, as very high resolution is required to resolve the substructure left over from merger events. More directly, there is unambiguous evidence of substructure in the stellar populations of both the galactic [55,56] and the Andromeda (M31) haloes [57]. These are probably the remnants of smaller galaxies engulfed by larger neighbors. The anomalous flux ratios of quadruplet lenses have also been claimed as evidence of substructure [58,59], but this evidence is not clear-cut, as microlensing, differential extinction or scatter-broadening may also be affecting the flux ratios [60].

In the inner parts of the galactic halo (say $r \lesssim 25$ kpc), dynamical friction and tides are efficient at erasing substructure [54]. In the outer parts, clumps and streams can preserve their identity for longer. Blasi and Sheth [30,31] have suggested that the clumping of halo dark matter will give rise to clustering on the UHECR sky if the UHECRs themselves arise from the decay of superheavy dark matter particles [32,33]. Blasi and Sheth developed models in which the clumps occur with masses m at a distance r from the center according to the joint probability distribution

$$n_{\rm cl} \propto \left(\frac{1}{m^{1.9}} \right) \left(\frac{1}{1+(r/r_{\rm c})^2} \right)^{3/2}, \qquad (1)$$

where $r_{\rm c}$ is a constant, which they take as ~ 10 kpc. They argue that this is a good fit to the simulation data. Nonetheless, it runs counter to physical intuition, as the number density of clumps peaks in the center rather than the outlying portions of the Galaxy halo. Blasi and Sheth assume that the clumps themselves have the singular isothermal sphere (SIS) density law $\rho_{\rm cl} \propto r_{\rm cl}^{-2}$, where $r_{\rm cl}$ is the radial coordinate measured from the clump center. The singular isothermal spheres are truncated at their tidal radius in the Galaxy halo, which is assumed to take the Navarro-Frenk-White (NFW) form [61]

$$\rho_{\rm NFW} \propto \frac{1}{r(1+r/r_{\rm s})^2}. \qquad (2)$$

The choice of the isothermal law for the clumps has no physical basis whatsoever. In fact, the results of numerical simulations are usually claimed to be *self-similar*, in the sense that superclusters, clusters, galaxy haloes and subhaloes all have NFW profiles [61].

To test the robustness of the Blasi-Sheth results, we develop a different model of the substructure. First, the clumps are distributed homogeneously in the Galaxy halo, so the number of clumps within a radius r increases like r^3. Second, the masses of the clumps are chosen so that there is equal mass in equal decades [i.e., $n(m) \propto m^{-2}$]. Third, the clumps themselves are chosen to have NFW profiles in the parent NFW halo of the Galaxy. This is motivated by the

TABLE I. The observed number of coincidences (within 2.5°) between BL Lacs (models 1, 2 and 3) or GRBs (model 4) and UHECRs detected by AGASA is tabulated in total, as well as separately as coincidences with single events, doublets and triplets of UHECRs. These are compared with the results of 10^5 Monte Carlo simulations assuming random UHECR arrival directions. The table gives the mean number of coincidences in the simulations, as well as the 95% C.L. upper bound, and the probability $P(>\text{Expt.})$ of attaining results at least as correlated as the actual data.

	Experimental value	Monte Carlo mean	95% C.L. upper bound	$\mathcal{P}(>\text{Expt.})$
		Model 1		
Clusters	5	0.78335	2	0.00152
Singlets	1	0.75253	2	0.53514
Doublets	2	0.01522	0	0.00019
Triplets	0	0.00014	0	1
		Model 2		
Clusters	5	1.16664	3	0.00713
Singlets	1	1.09889	3	0.67605
Doublets	2	0.03315	0	0.00061
Triplets	0	0.00047	0	1
		Model 3		
Clusters	9	9.50677	15	0.63391
Singlets	5	8.96987	14	0.94949
Doublets	2	0.26078	1	0.03164
Triplets	0	0.00499	0	1
		Model 4		
Clusters	44	46.3406	58	0.65727
Singlets	36	43.8211	55	0.90202
Doublets	4	1.22408	4	0.05606
Triplets	0	0.02351	0	1

scale-freeness of the results of the simulations. In fact, in the inner parts of the Milky Way, there is ample evidence that the halo does *not* have the NFW form [62]. However, our main motivation here is to understand the results obtained by Blasi and Sheth, so we have only changed the substructure properties and not the underlying smooth halo model. (For reference, the NFW concentration parameter c is ~ 10 for the Galaxy halo and ~ 5 for the clumps.) From the simulations, it is found that most clumps fall in at a redshift of $z \sim 4$ and their concentration remains largely frozen after this. The overdensity at formation is calibrated with respect to the critical density at that epoch. In selecting parameters for the clump distribution, our canonical choice is to distribute a generous 10% of the total halo mass in clumps between 10^7 and 10^{10} $M_\odot$. Given the mass of the clump, the virial radius follows. The virial radius and the concentration determine the NFW lengthscale r_s.

Let us start by presenting different properties of the clump populations generated according to the recipes of Blasi and Sheth and ourselves. For a Galaxy halo, we have typically the same numbers of clumps (~ 1600) in both models. With our prescription, the clumps tend to be smaller in extent which makes their detection more difficult. However, the main difference is in the distribution of clumps with galactocentric radius, as shown in Fig. 4. In the Blasi-Sheth model, the peak in the radial distribution of the clumps occurs at ~ 20 kpc, which is close to the solar circle. This is of course optimum for causing visible consequences. In our model, the peak occurs at 220 kpc, very much in the outer parts of the Galaxy halo. Let us remark that the numerical simulations clearly show that most of the surviving substructure *is* in the outer parts. Both NFW and SIS profiles are

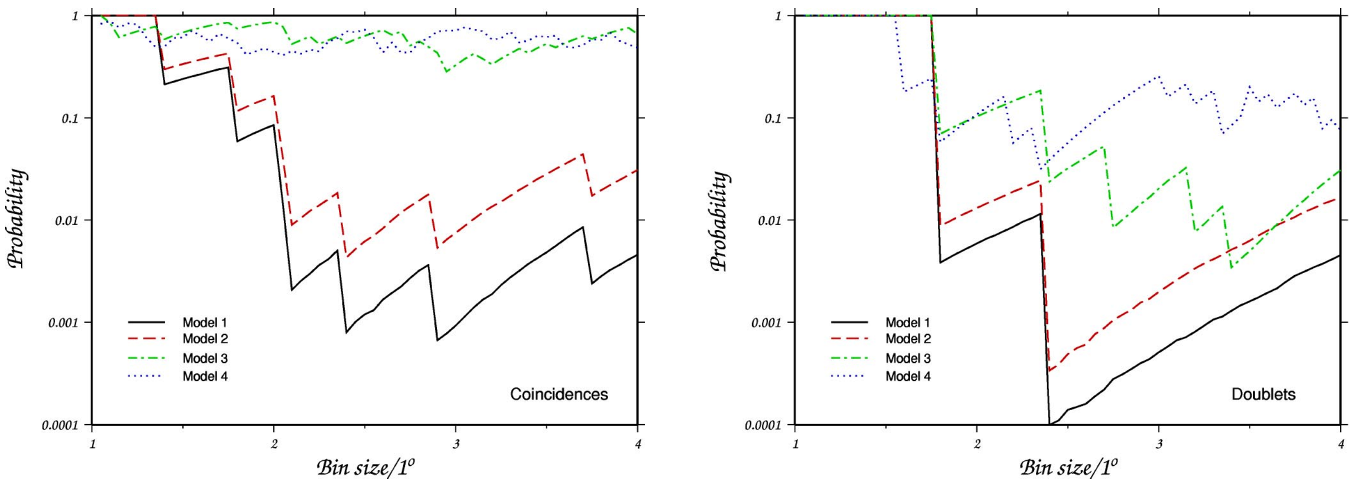

FIG. 3. (Color online) Dependence on the angular bin width of the probability for coincidences with all UHECRs (left), and doublets of UHECRs (right), for model 1 (solid line), model 2 (long-dashed line), model 3 (short-dashed line), and model 4 (dotted line).

singular at the origin, so a small regularizing core radius must be included in the computations. For the NFW profiles, we take our cue from our earlier paper [63] where the regularizing core r_ϵ was chosen as ~ 0.5 kpc (the resolution limit) for a halo of ~ 250 kpc extent. For the NFW clumps, r_ϵ is scaled to be the same fraction of the extent. For the SIS profiles, we adopt $r_\epsilon = 3\times 10^{-7}$ kpc, as suggested by the limit imposed by particle dark matter self-annihilation [64,65].

Figure 5 shows the incoming UHECR flux in a Hamer-Aitoff projection in equatorial coordinates folded with the response of the AGASA detector. The flux from the underlying smooth model is calculated by integrating the emissivity density along the line of sight (see e.g., Ref. [63]). However, the angular size of clumps can be smaller than the angular resolution of the instrument, in which case the flux is computed by the volume integral of the emissivity over the clump, divided by the square of the distance of the clump from Earth. There are four panels showing random arrival directions, a smooth NFW galactic halo, a smooth NFW halo plus NFW clumps and a smooth halo plus SIS clumps. The gray region corresponds to no detection, as AGASA is a Northern Hemisphere experiment. The lower right panel shows irregularities, caused by the SIS clumps which are brighter than any contribution from the underlying smooth halo. The lower left panel corresponding to NFW clumps shows much less evidence for irregularities. The flux is much more uniform with less clustering. We cannot verify the claim made by Blasi and Sheth that a smooth NFW halo alone is able to provide almost half the observed clustering

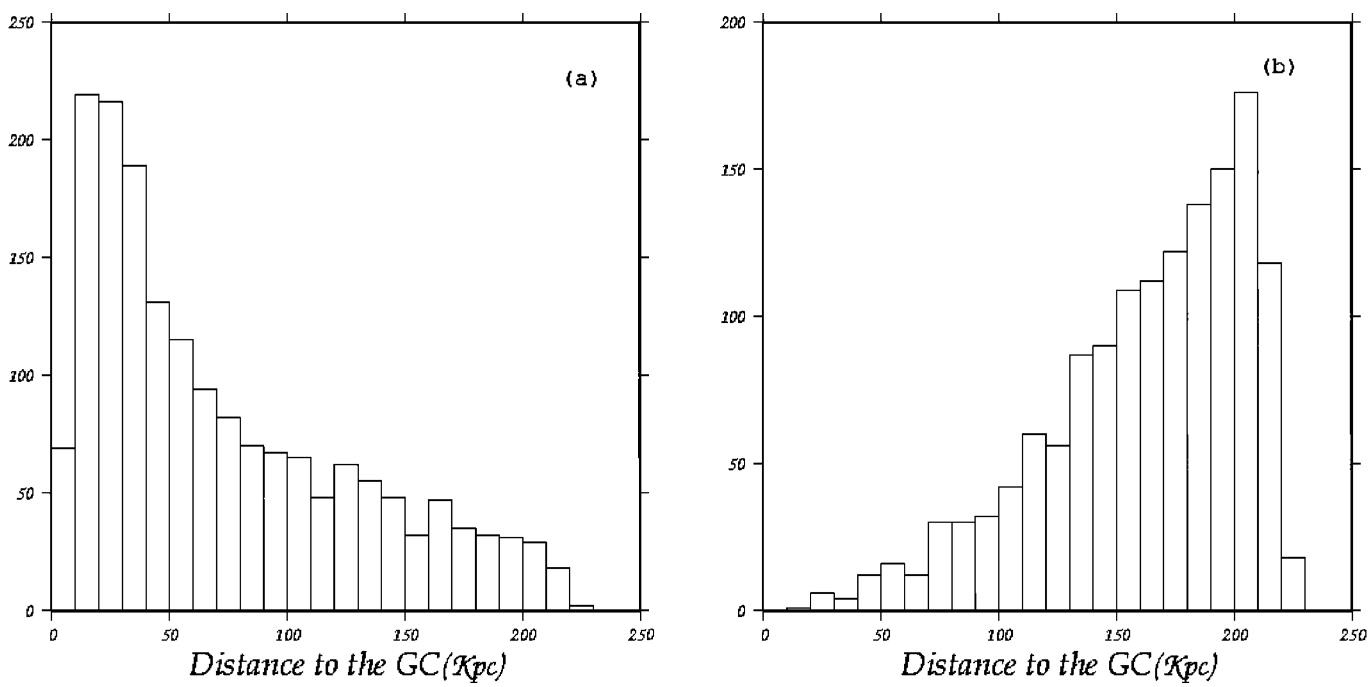

FIG. 4. Distribution of the clumps with radius adopted (a) by Blasi and Sheth [30,31] and (b) in this work.

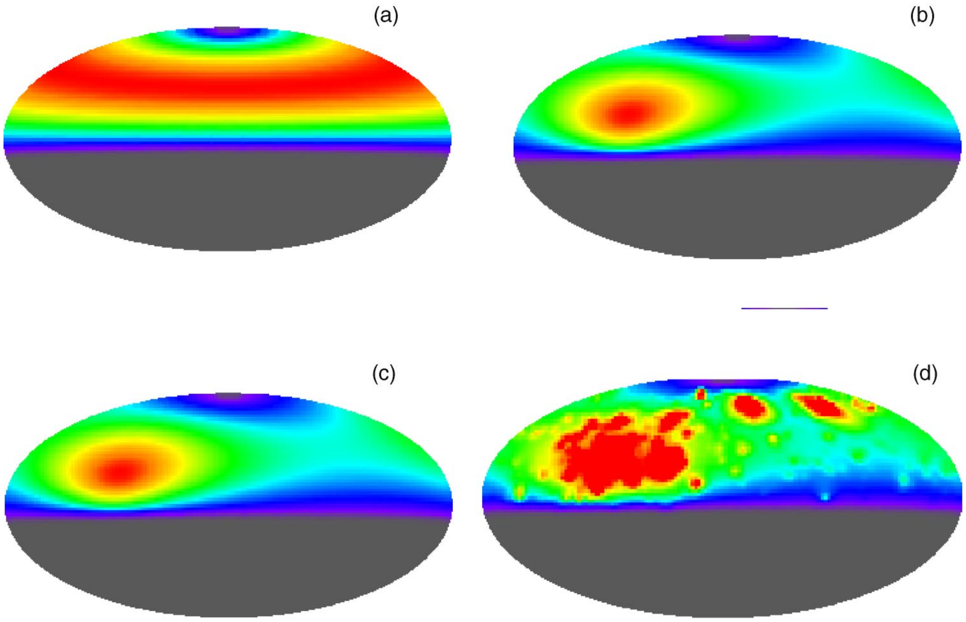

FIG. 5. (Color online) The UHECR sky as it would be seen by AGASA—top left: Isotropic model; top right: Smooth NFW galactic halo; bottom left: Smooth NFW galactic halo with 10% of the mass in NFW clumps; bottom right: Smooth NFW galactic halo with 10% of the mass in SIS clumps. The highest flux in each panel corresponds to the darkest (red) regions, the lowest flux to the lightest (lilac).

[30,31]. In fact, it is hard to see how a smooth halo can be responsible for small scale flux variations.

To quantify this, we use the 58 events in the AGASA experiment above 4×10^{19} eV as our dataset. The two point autocorrelation function for the four cases is shown in Fig. 6. Samples of 58 UHECRs, with the AGASA response function folded in, are generated, and the average autocorrelation is compared to the one found for the experimental data. The clustering in the experimental dataset is not well reproduced. Even when SIS clumps are present, the disagreement is at the

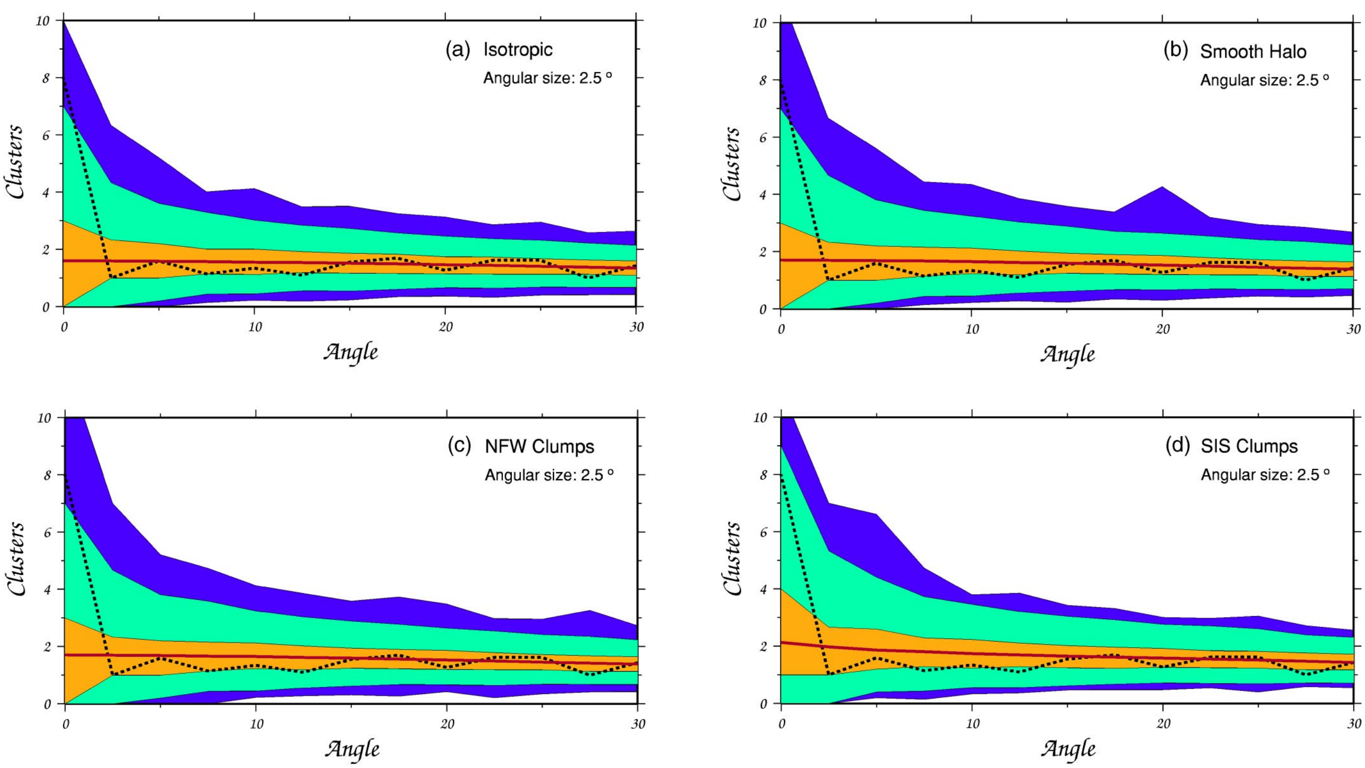

FIG. 6. (Color online) The two point correlation function for UHECRs generated with (a) isotropic arrival directions, (b) in a smooth NFW halo model, (c) in a smooth NFW halo with 10% of the mass in NFW clumps or (d) in SIS clumps. The angular size on which correlations are sought is $\Delta\theta = 2.5°$. The thick dashed line corresponds to the autocorrelation function of the experimental data. The thick unbroken line is the average autocorrelation function of 10^5 Monte Carlo simulations. The regions shaded light, medium and dark correspond to chance probabilities less than 31.73%, 0.27% and 0.0057% (chosen so as to correspond to 1σ, 3σ and 5σ significance for Gaussian statistics), as estimated by running 10^5 Monte Carlo simulations.

TABLE II. The experimentally observed numbers of clusters, doublets and triplets are compared with the results of 10^5 Monte Carlo simulations. The table gives the mean numbers of clusters, doublets and triplets for the simulations, the 95% upper bounds, together with the probability $P(>\text{Expt.})$ of obtaining results at least as clustered as the data.

(a) Isotropic arrival directions	Experimental value	Monte Carlo mean	95% C.L. upper bound	$\mathcal{P}(>\text{Expt.})$
Clusters	8	1.60217	4	0.00038
Doublets	8	1.59989	4	0.00029
Triplets	1	0.01661	0	0.01642

(b) Galaxy halo modeled by a smooth NFW profile	Experimental value	Monte Carlo mean	95% C.L. upper bound	$\mathcal{P}(>\text{Expt.})$
Clusters	8	1.70778	4	0.00094
Doublets	8	1.71073	4	0.00076
Triplets	1	0.02003	0	0.01985

(c) Galaxy halo modeled by a smooth NFW profile plus NFW profile clumps	Experimental value	Monte Carlo mean	95% C.L. upper bound	$\mathcal{P}(>\text{Expt.})$
Clusters	8	1.70712	4	0.00083
Doublets	8	1.70282	4	0.00057
Triplets	1	0.02023	4	0.02001

(d) Galaxy halo modeled by a smooth NFW profile plus SIS clumps	Experimental value	Monte Carlo mean	95% C.L. upper bound	$\mathcal{P}(>\text{Expt.})$
Clusters	8	2.1306	5	0.056
Doublets	8	2.2589	5	0.029
Triplets	1	0.057	1	0.055

2.8σ level and it is beyond 3σ for the other models. We also note from panels (a), (b) and (c) that the smooth NFW profile (with or without clumps) is almost indistinguishable from the isotropic model. Table II gives the explicit probabilities deduced from 10^5 Monte Carlo calculations. We see that in the first three models, we fall short of obtaining the 8 clusters required as less than 2 are expected on average. The probability of obtaining 8 or more clumps is of order $10^{-3} - 10^{-4}$. When SIS clumps are introduced in the halo, however, the probability increases to 0.056. Note that, in every case, the discrepancy between model predictions and experiment is always less than 5σ.

The angular width of each bin was taken to be 2.5° in the calculations reported above. If this is enlarged to 5°, then the discrepancies between model predictions and data are smaller. This is a reasonable angular size on which to look for correlations as it corresponds to the typical deflection of a 4×10^{19} eV proton in the galactic magnetic field [20,21]. Figure 7 shows the autocorrelation functions for the cases of isotropic arrival directions and for a smooth NFW halo with NFW clumps. We expect typically 6 and 7 clusters for these two cases. The probability of obtaining as many clusters as the experimental data (11) rises to 0.06 and 0.07 respectively. Note that, even for the isotropic model, the discrepancy is now at only the $\approx 2\sigma$ level. The dependence of the probability on the bin width is shown in Fig. 8 for the model with NFW clumps in an NFW halo. It reaches a minimum around 2.5°. This is where the discrepancy with the data is at a maximum. Either decreasing or increasing $\Delta\theta$ gives model predictions in better agreement with the data. So, this figure shows how sensitive the clustering results are to the bin width. For example, changing 2.5° to 3° causes almost an order of magnitude change in the probability.

Using the 92 events from the combined datasets of AGASA, Volcano Ranch, Haverah Park and Yakutsk [15], Blasi and Sheth estimated that the probability of attaining more doublets than the data from SIS clumps is 12% (for 3° bin widths) and 47% (for 4°). These high numbers appear to be primarily a consequence of placing the SIS clumps nearby. Using our model for the mass and spatial distribution of the clumps, we obtain corresponding probabilities of 7% (for 3°) and 29% (for 4°). If the clumps have an NFW profile as is more likely, these numbers are reduced further to 3.5% and 15%, respectively.

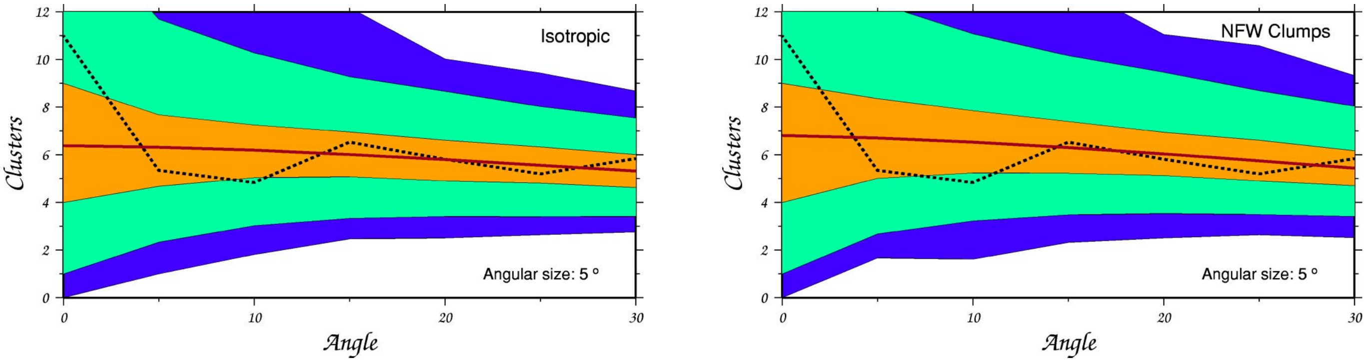

FIG. 7. (Color online) As Fig. 6, but changing the angular size $\Delta\theta$ to $5°$. Only two models are shown, namely isotropic arrival directions (left) and a smooth NFW model with NFW clumps (right).

We conclude that clustering is *not* a generic prediction of the decaying dark matter model. Even though clumps are indeed expected in the dark matter distributions in Galaxy haloes, any clustering of UHECRs depends sensitively on the density profile of the clumps. In particular, NFW clumps do not give rise to much clustering, but SIS clumps may do so. However, SIS profiles are not very natural, and almost all the signal comes from the $r_{\rm cl}^{-2}$ singularity. In fact, the self-similarity of the structure formation process suggests that NFW clumps are more natural. We note that the 10% of mass that we placed in clumps is already generous, and so it is not clear whether any real clustering of UHECRs can be expected from dark matter substructure. However, it is also not clear that there is any real clustering of the UHECRs at all, as the signal is less than 2σ at the most natural angular scale of $\Delta\theta=5°$.

IV. DISCUSSION

It has been a general expectation that it should be possible to identify the long-sought sources of cosmic rays at energies exceeding $\sim 10^{19}$ eV, when their arrival directions can no longer be randomized by galactic magnetic fields. However, the sky distribution has remained consistent with isotropy up to the highest energies observed. It is clear that at such energies the sources cannot be in the disk of the Galaxy and all experiments indicate that the energy spectrum of such sources is significantly flatter than the component at lower energies. However, it is not clear whether the isotropy of arrival directions implicates a relatively *local* population of sources in the galactic halo (e.g. decaying supermassive dark matter) or a cosmologically *distant* population of astrophysical accelerators (e.g. active galaxies or γ-ray bursts). Support for the latter possibility has come from the new HiRes data which shows a cutoff in the spectrum beyond $E_{\rm GZK} \simeq 4\times 10^{19}$ eV, as has long been expected for extragalactic sources. However, the AGASA Collaboration has reaffirmed that there is no GZK cutoff in their data, which strongly favors the former possibility.

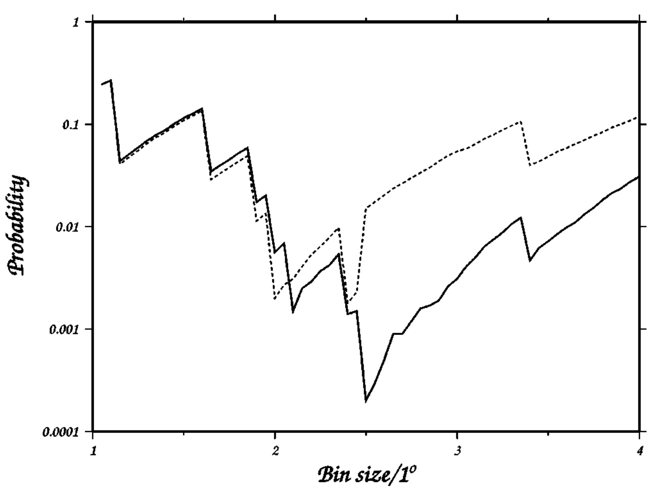

FIG. 8. The probability of obtaining as many clusters (solid line) and doublets (dashed line) as the data given as a function of the angular size $\Delta\theta$ of the correlations.

Given this confusing situation, the indication of small-angle clustering in the AGASA data has naturally been seized upon as a possible further clue as to the nature of the sources. It has been argued both that the clusters coincide with a specific class of extragalactic objects, viz. BL Lacs [18,26], and, alternatively, that such clustering might arise due to the expected clumping of halo dark matter [30,31]. The BL Lac hypothesis is in fact *inconsistent* with the absence of the GZK cutoff in the same AGASA data since many of the identified objects are at very large distances. Hence it appeared more plausible that the sources are clumps of dark matter in the galactic halo (composed in part of supermassive decaying particles) which would explain the absence of the GZK cutoff.

We have shown that the correlations claimed between BL Lacs and the observed clusters of UHECRs are spurious, being entirely due to selection effects. This is not the first time that such correlations with a particular class of astrophysical accelerators has been claimed; the moral is clearly that care must be taken to not become intrigued by weak accidental correlations and then make arbitrary cuts on the dataset to emphasize them further. We have also found that the extent to which dark matter may be clumped in the halo is not sufficient to generate the observed small-angle clustering, if the UHECRs indeed arise from decaying dark matter. Here the proponents have been misled due to the use of an unphysical density profile for the clumps, as well as a radial distribution in the Galaxy which is inconsistent with the general expectations for hierarchical structure formation.

The net result of our investigations is thus rather negative. The claimed small-angle clustering in the arrival directions of post-GZK UHECRs does not definitively implicate either extragalactic compact sources such as BL Lacs, or decaying clumps of dark matter in the galactic halo. On the positive side, the forthcoming increase in statistics from the Pierre Auger Observatory will enable us to identify the expected signal from dark matter decays if this is indeed the source of UHECRs. Moreover, Auger will also definitively resolve the current contradiction between the air shower and atmospheric fluorescence methods for energy measurement, thus clarifying whether the spectrum does have a GZK cutoff. If so, searches for coincidences with cosmologically distant candidate sources such as active galaxies or γ-ray bursts would be of interest. Again the increased statistics provided by Auger would enable the significance of such coincidences to be meaningfully assessed. We will soon know whether the mystery of UHECRs implicates astrophysical sources or new physics beyond the standard model.

ACKNOWLEDGMENTS

We are grateful to Motohiko Nagano and Alan Watson for discussions and for providing UHECR data sets, and to Ben Moore for clarifications of the results obtained in numerical simulations of halo substructure. We also thank Pascuale Blasi, Ravi Sheth and Igor Tkachev for helpful correspondence.

[1] K. Greisen, Phys. Rev. Lett. **16**, 748 (1966).
[2] G.T. Zatsepin and V.A. Kuzmin, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 114 (1966) [JETP Lett. **4**, 78 (1966)].
[3] F.W. Stecker, Phys. Rev. Lett. **21**, 1016 (1968).
[4] High Resolution Fly's Eye Collaboration, T. Abu-Zayyad *et al.* astro-ph/0208301.
[5] M. Takeda *et al.*, Phys. Rev. Lett. **81**, 1163 (1998).
[6] N. Hayashida *et al.*, Astrophys. J. **522**, 225 (1999).
[7] D.R. Bergman, hep-ex/0208024.
[8] M. Takeda *et al.*, astro-ph/0209422.
[9] A.A. Watson, astro-ph/0112474.
[10] http://www.auger.org
[11] M. Takeda *et al.*, Astrophys. J. **522**, 225 (1999).
[12] M. Takeda *et al.*, in Proceedings of the 27th ICRC, Hamburg, Germany, 2001, Vol. 1, p. 341.
[13] H. Goldberg and T.J. Weiler, Phys. Rev. D **64**, 056008 (2001).
[14] L.A. Anchordoqui, H. Goldberg, S. Reucroft, G.E. Romero, J. Swain, and D.F. Torres, Mod. Phys. Lett. A **16**, 2033 (2001).
[15] Y. Uchihori, M. Nagano, M. Takeda, M. Teshima, J. Lloyd-Evans, and A.A. Watson, Astropart. Phys. **13**, 151 (2000).
[16] S.L. Dubovsky, P.G. Tinyakov, and I.I. Tkachev, Phys. Rev. Lett. **85**, 1154 (2000).
[17] P.G. Tinyakov and I.I. Tkachev, Pis'ma Zh. Eksp. Teor. Fiz. **74**, 3 (2001) [JETP Lett. **74**, 1 (2001)].
[18] P.G. Tinyakov and I.I. Tkachev, Pis'ma Zh. Eksp. Teor. Fiz. **74**, 499 (2001) [JETP Lett. **74**, 445 (2001)].
[19] P.L. Biermann and P.A. Strittmatter, Astrophys. J. **322**, 643 (1987).
[20] T. Stanev, Astrophys. J. **479**, 290 (1997).
[21] G.A. Medina Tanco, E.M. de Gouveia Dal Pino, and J.E. Horvath, Astrophys. J. **492**, 200 (1998).
[22] J. Alvarez-Muniz, R. Engel, and T. Stanev, Astrophys. J. **572**, 185 (2001).
[23] G.A. Medina Tanco, Astrophys. J. Lett. **495**, L71 (1998).
[24] D. Harari, S. Mollerach, and E. Roulet, J. High Energy Phys. **02**, 035 (2000).
[25] D. Harari, S. Mollerach, E. Roulet, and F. Sanchez, J. High Energy Phys. **03**, 045 (2002).
[26] P.G. Tinyakov and I.I. Tkachev, Astropart. Phys. **18**, 165 (2002).
[27] L.M. Widrow, Rev. Mod. Phys. **74**, 775 (2002).
[28] T. Stanev, R. Engel, A. Mucke, R.J. Protheroe, and J.P. Rachen, Phys. Rev. D **62**, 093005 (2000).
[29] V. Berezinsky, A.Z. Gazizov, and S.I. Grigorieva, astro-ph/0210095.
[30] P. Blasi and R.K. Sheth, Phys. Lett. B **486**, 233 (2000).
[31] P. Blasi and R.K. Sheth, astro-ph/0108288.
[32] V. Berezinsky, M. Kachelriess, and A. Vilenkin, Phys. Rev. Lett. **79**, 4302 (1997).
[33] M. Birkel and S. Sarkar, Astropart. Phys. **9**, 297 (1998).
[34] G.R. Farrar and P.L. Biermann, Phys. Rev. Lett. **81**, 3579 (1998).
[35] G.R. Farrar and P.L. Biermann, Phys. Rev. Lett. **83**, 2472 (1999).
[36] J.W. Elbert and P. Sommers, Astrophys. J. **441**, 151 (1995).
[37] D.J. Bird *et al.*, Astrophys. J. **441**, 144 (1995).
[38] C.M. Hoffman, Phys. Rev. Lett. **83**, 2471 (1999).
[39] G. Sigl, D.F. Torres, L.A. Anchordoqui, and G.E. Romero, Phys. Rev. D **63**, 081302(R) (2001).
[40] M.A. Lawrence, R.J. Reid, and A.A. Watson, J. Phys. G **17**, 733 (1991).
[41] M. Ave, J.A. Hinton, R.A. Vazquez, A.A. Watson, and E. Zas, Phys. Rev. Lett. **85**, 2244 (2000).
[42] A. Virmani, S. Bhattacharya, P. Jain, S. Razzaque, J.P. Ralston, and D.W. McKay, Astropart. Phys. **17**, 489 (2002).
[43] D.F. Torres, E. Boldt, T. Hamilton, and M. Loewenstein, Phys. Rev. D **66**, 023001 (2002).
[44] Z. Fodor and S.D. Katz, Phys. Rev. D **63**, 023002 (2001).
[45] *Catalogue of Highest Energy Cosmic Rays, No.3*, edited by A. Inoue and E. Sakamoto (World Data Center for Cosmic Rays, Institute of Physical and Chemical Research, Saitama, 1988).
[46] M. P. Veron-Cetty and P. Veron, A catalogue of Quasars and Active Galactic Nuclei, 9th ed., 2000 [http://www.obs-hp.fr/www/catalogues/veron2_9/veron2_9.html].
[47] R. Beck, Space Sci. Rev. **99**, 243 (2001).
[48] D.S. Gorbunov, P.G. Tinyakov, I.I. Tkachev, and S.V. Troitsky, Astrophys. J. Lett. **577**, L93 (2002).
[49] M.P. Veron-Cetty and P. Veron, Astron. Astrophys. **374**, 92 (2001).
[50] R.C. Hartman *et al.*, Astrophys. J., Suppl. Ser. **123**, 79 (1999).

[51] W.S. Paciesas *et al.*, Astrophys. J., Suppl. Ser. **122**, 465 (1999).
[52] J.N. Bahcall and E. Waxman, Phys. Lett. B **556**, 1 (2003).
[53] B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel, and P. Tozzi, Astrophys. J. Lett. **524**, L19 (1999).
[54] S. Ghigna, B. Moore, F. Governato, G. Lake, T. Quinn, and J. Stadel, Astrophys. J. **544**, 616 (2000).
[55] R. Arnold and G. Gilmore, Mon. Not. R. Astron. Soc. **257**, 225 (1990).
[56] R. Ibata, G. Gilmore, and M. Irwin, Nature (London) **370**, 194 (1994).
[57] R. Ibata, M. Irwin, G. Lewis, A. Ferguson, and N. Tanvir, Nature (London) **412**, 49 (2001).
[58] C.S. Kochanek and N. Dalal, Astrophys. J. **572**, 25 (2002).
[59] R.B. Metcalf and H. Zhao, Astrophys. J. Lett. **567**, L5 (2002).
[60] N.W. Evans and H.J. Witt, astro-ph/0212013.
[61] J.F. Navarro, C.S. Frenk, and S.D.M. White, Astrophys. J. **462**, 563 (1996).
[62] N.W. Evans, in *IDM 2000: Third International Workshop on the Identification of Dark Matter*, edited by N. Spooner and V. Kudryavtsev (World Scientific, Singapore, 2000), p. 85, astro-ph/0102082.
[63] N.W. Evans, F. Ferrer, and S. Sarkar, Astropart. Phys. **17**, 319 (2002).
[64] C. Tyler, Phys. Rev. D **66**, 023509 (2002).
[65] P. Blasi, A.V. Olinto, and C. Tyler, Astropart. Phys. **18**, 649 (2003).