Large-Scale-Structure in the Pierre Auger Observatory Data Directions

"The Long Dark Teatime of the Soul[1]"

J. Sorokin

B.Sc. (Honours), Physics

A thesis presented to the University of Adelaide

for admission to the degree of

Doctor of Philosophy.

School of Physical Sciences

Department of Physics

April 2016

[1] with thanks for the kind permission of Jane Belson

I wish to thank

my supervisor, Professor Roger Clay for hanging in there

Dr Jonathan Woithe for his kind and obliging help with my various idl problems

Dr Benjamin Whelan for his thoughtful edits

my friend Ramona Adjoran for fixing my hardware problems

Dr Millie Vukovic and Dr Elizabeth Heath for their support

and

my legion of fans

My wonderful mother Yolande

my unexpected husband Michael and my singular brother Adam

fabulous Millie Butler and interesting Cyrus Masters

Thankyou, thankyou,

I love you all

Abstract

This thesis presents a method of analysis of Pierre Auger Observatory Cosmic Ray (CR) directions. I look for evidence of large-scale-structure within these CR directions. I have associated directional events by virtue of the angular proximity of their arrival directions, and within three energy ranges around the Greisen-Zatsepin-Kusmin (GZK) energy limit. I design graph theoretical algorithms to grow minimum spanning trees for these directional events and then 'cut' the trees along certain galactic longitudes and latitudes into 'branches', where I expect the galactic magnetic fields or cosmic ray point-sources to exhibit behaviours or patterns, which can be interpreted by branch features. 1,200 simulated CR directions in each energy range provide some statistical context for the Pierre Auger Observatory branch features which may be considered significant. The thesis is a preliminary study of a method of analysis of 'regions of interest' which later may be optimized with a full statistical 'tuning' analysis.

Statement of Originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree. I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. I also give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Contents

List of Figures

List of Tables

Chapter 1

An Introduction

High energy cosmic rays (CRs), extra terrestial particles that are detected within the energy ranges $\sim(10^9 - 10^{21})$ eV entering the Earth's atmosphere, have long been a topic of conjecture[1].

Their discovery in 1912 by Victor Hess, elicited some disbelief and intense speculation which continues to present times. What source or interaction could accelerate protons, neutrons and particles as substantial as the actinides (15 radioactive heavy metals including uranium with atomic number 92 [73]) to such powerful and unlooked-for energies?

The Pierre Auger Observatory, situated in Argentina, exists to collect a statistical database that will clarify our understanding of these particles. Searches - for structure, from the point-like to the very large; in particle directions and the emergence of associations between particle energies and the reach of their source types - serve to better detail and delimit not only our home galaxy, but neighbouring large-scale objects.

This thesis concerns directional data collected from the Pierre Auger Observatory, with an analysis that concentrates on large-scale directional structure through the application of Minimum Spanning Trees (MSTs). MSTs are optimised structures that construct graphs by ordering a collection of measurements from the smallest measurement first to the largest measurement last. The analysis method involves aspects of graph theory which are briefly reviewed in Chapter 5. The analysis method itself is described in greater detail in Chapter 6 and Chapter 7. There are many

[1]CR particles with energies below $\sim$ 10 GeV are considered to be mainly sourced and modulated from our Sun [92],one reason being that their flux has been observed to correlate with the 11 yr cycle of our Sun's magnetic heliospheric fields.

references about the use of MSTs and filamentary structures in other aspects of astronomy. Studies include [9, 10].

The layout of this thesis is as follows:

This chapter briefly discusses the cosmic ray background.

Chapter 2 contains sections on :

- Extensive air Showers (EAS), CR composition and the energy spectrum.
- The Pierre Auger Observatory - referencing some of the EAS observables measured by Pierre Auger Observatory instrumentation. These observables are covered in more detail in section 2.1 on page 8 of Chapter 2.
- The Pierre Auger Observatory prescription concept.

Chapter 3 provides an introduction to Galactic Magnetic Field (GMF) models, galactic cosmic ray anisotropies, and extra-galactic anisotropies.

Chapter 4 provides a review of some recent anisotropy methods and applied results for energies, E > 10^{17} eV.

Chapter 5 reviews the graph theory underpinning the MST.

Chapter 6 discusses the generation of simulated data sets.

Chapters 7 presents the results of this method for studying data in 3 ultra-high-energy-cosmic-ray (UHECR) energy cuts.

Chapter 8 is a discussion of these results.

Appendix A on page 233 provides plots of MSTs and plots of simulated cosmic ray parameter distributions in the energy range 40 EeV < $E \leq$ 50 EeV. Appendix B on page 251 and Appendix C on page 269 provide plots of MSTs and simulated cosmic ray parameter distributions in the energy ranges 50 EeV < $E \leq$ 60 EeV and E > 60 EeV. All simulations are compared with the Pierre Auger Observatory data results.

Data which are simulated are referred to as simulated data, and real data are called data.

1.1 Background

The turn of the 20th century was spanned by the Second Industrial Revolution (1870 – 1914) and saw the rise of the industrial power of the US and Germany. Mathematicians quantified graph theory, abstract algebras with broken symmetries and strange equivalence relations were found – Gödel destroyed the 'Principia Mathematica'[2]. Special Relativity was expressed in the language of algebraic invariance[3], radiation could be a consequence of the atomic nature of matter and quantum mechanics set boundary conditions.

The measurement, by Victor Hess in a balloon in 1913 with 3 Wulf electrometers [179], of an ionization **increase** with atmospheric height was the start of the investigation into CRs.

The progression of the 20th century into the 21st has been a gathering of understandings of the physical world and developments of new hardware. The body of knowledge on CRs is finally extensive enough to provide solutions to questions long outstanding. The index of CR observables includes their direction, energy and composition. At sufficiently high energies, anisotropies in CR directions can now be compared against current astronomical catalogues and correlations quantified. These energies need to be high, so that large-scale magnetic fields are transparent and particle deflection is minimized. The nature of these energetic particle primaries can now be back engineered, to deduce their properties before they shatter into what are termed extensive air showers (EAS) in our upper atmosphere.

We now possess a detailed compositional table of CRs for at least the lower energies($E < 10^{15}$ eV). CRs are particles - virtually all the naturally occurring ionized nuclei and their isotopes, γ-rays, neutrons, electrons, anti-particles, and neutrinos. Their proportions share much in common with the interstellar medium (ISM) composition of gas and dust, aside from several important abundances - the so-called secondary nuclei.

An important component is the presence of unstable isotopes. It is these isotopes and their half-lives that signpost how far some CRs have travelled, and their containment time in the galaxy. The nuclear processes responsible for their formation help isolate the taxonomy of sources. γ-rays

[2]The central paradox of recurrence requires initial conditions for **any** meta theory.

[3]Developed by Arthur Caley who also was responsible for many of the first minimum spanning trees.

have proved useful in point-source detection. Synchrotron radiation from Supernova Remnant (SNR) electrons has a photon fraction in the radio to X-ray frequencies, and SNR hadrons are supposed to produce TeV γ-rays in their interaction with the ISM. H.E.S.S. may have found some examples [79], as did HEGRA, which sought out TeV γ-ray bursts [89].

The electromagnetic property of lower energy charged CR particles has them in harness with the geo-magnetic field and their arrival flux has a resulting bias to the West – the East-West effect. Known properties of hadronic interactions now yield a clearer picture of how CRs reach us through our atmosphere. The overall CR flux has a simple energy spectrum, a falling sweep of flux punctuated by what is believed to be 2 major transitional points. The 'Knee' at $\sim 3 \times 10^{15}$ eV and the 'Ankle' at $\sim 5 \times 10^{18}$ eV. Below the Knee, where the CR flux is strongest and the energies are at their lowest, the sources are taken chiefly to be our local Sun and other stars from within our local galaxy. Just above the Knee, the 'standard' sources are taken to be SNR activity into the ISM. Approaching the Ankle[4] , the sources, although still *perhaps* galactic, may be more exotic (galactic superbubbles [204], the galactic bulge [203], hyper-novae into stellar winds [7],have also lately been posited as galactic CR sources). These galactic CRs are reckoned to be no older than $\sim 10^7$ years - the time calculated to have lapsed piloting the magnetic fields between source and observer [73].inzberg and Syrovatskii's galactic CR diffusion model [204, 99],notes a coupling between charged CRs, this confinement time, and the galactic magnetic field (GMF) as seen in the similar energy densities of the local ISM and a GMF of $B \approx 3\,\mu$G. Fermi's [205] paper on first and second order shock acceleration was the first record of this relationship.

Beyond the Ankle, there is believed to be an incursion of extra-galactic particles of such extreme energies that, to date, there exists no credible model of source or acceleration process. These turning points in the energy spectrum are plainly read in the relationship, such that for the number of particles, N_A, above the threshold energy, E,

$$N_A(> E) \propto E^{-\gamma}$$

[4]It is important to acknowledge there are no physical constraints preventing CR origins to be from anywhere in the universe at energies below the ankle [29].

Shown in Figure 1.1 is the CR energy spectrum. Note how very slight are the changes in γ.

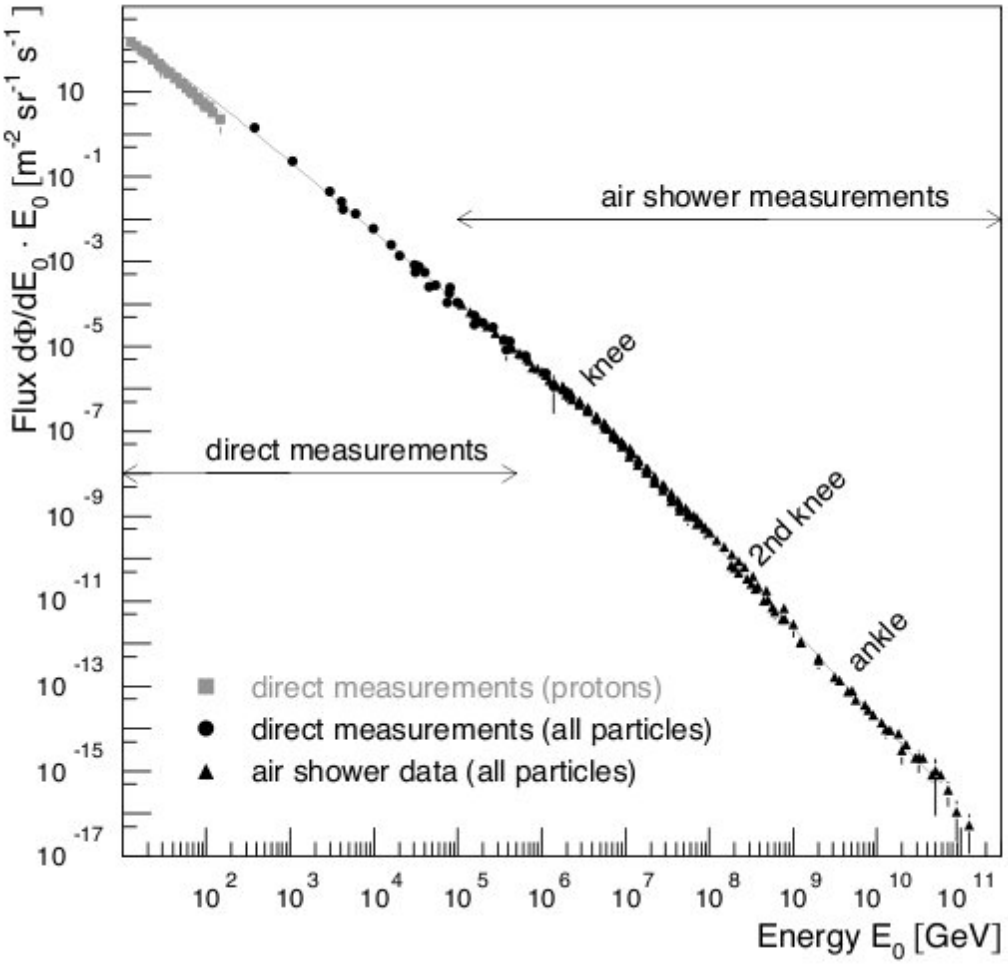

Figure 1.1: **Cosmic ray all-particle energy spectrum. The primary proton flux and some of the all particle flux for low energies were recorded by direct spacecraft measurements. Higher energy flux measurements relied on EAS reconstruction techniqes. From [93].**

The value of γ changes at the critical points of Knee and Ankle, and what is termed a secondary knee. The break in the power law where there is an appreciable suppression of flux, the suggested Greisen-Zatsepin-Kusmin (GZK) proton limit, ~(4-6)×10^{19} eV, has been treated as either a further change in γ, or γ is fitted with a smoothing function [110].

For the ultra high energies (UHE's) which the Pierre Auger Observatory has been designed to capture (energies > 10^{18} eV), we must look to the properties of extensive air showers (EAS).

EAS are complicated and have significant associated uncertainties because of their complicated makeup, and because of the limitations of the experimentally backed theories regarding hadronic reactions. Standard general models of EAS properties give us information on their direction, energy and composition. Directions are from fast timing of the EAS front, energy is found in the integrated particle signal density, and mass composition is deduced from EAS development distributions [126].

Atmospheric depth, X, in units of g cm^{-2}, is often used in descriptions of a shower's development. The depths of stages in EAS are measured, i.e. the depth at points, h, along the shower axis, l, is integrated over density to the atmospheric ceiling such that

$$X = \int_{h}^{\infty} \rho(l) dl \tag{1.1}$$

This atmospheric depth perspective is used to clarify the EAS measurements.

Chapter 2

EAS and Two Key Descriptors of Cosmic Rays

Although this research topic features Pierre Auger Observatory CRs at ultra-high energies, it is interesting to briefly review CRs over their full energy range. It is possible that the pattern of composition detected in lower energies and associated with local sources may well be a help to understand compositions repeated at higher energies with larger, more distant sources. This may assist us when we hazard interpretation of incomplete theoretical models and large experimental uncertainties.

CR composition and its energy dependence between critical points in the energy spectrum are strong indicators of sources and acceleration mechanisms. These critical points form a natural range for each section of this chapter on the composition and energy spectrum, although some latitude is allowed the measurements at each of these points. This latitude is due to differences between experimental procedures, hardware, and calculation parameters such that it is often difficult to compare experiments across their normalizations [86]. The third key descriptor of CRs, anisotropies, the headline subject of this thesis, is covered in Chapter 3 on page 56.

Above $\sim (10^{14} - 10^{15})$ eV, CR information is collected in sets of statistics fitted to hadronic interaction models embedded into shower reconstruction codes, such as CORSIKA (COsmic Ray SImulations for KAscade) and AIRES (AIR-shower Extended Simulations). The data gathered on

EAS and how they are used to estimate energy and direction, plus to infer the nature of primaries, is set out below in the overview on EAS.

2.1 Extensive Air Showers

The blizzard of particles recorded by Earth-based detectors due to atmospheric cascades is known as 'Extensive Air Showers' (EAS) - the total flux of what is assumed to be all particle types. These are rebuilt or 'unfolded' according to shifting interpretations of Quark Gluon String (QGS) laws. As new information becomes available, hadron interaction models are modified in an effort to refine accuracy. The current suite of models include SYBILL, the QGSJet line, Epos and NeXuS.

The EAS properties are such that energy calibration and the direction of the primary particle can be reasonably deduced from detectors. A primary's energy is carried by its EAS, and most of this is discharged into the atmosphere, and the primary's direction is largely conserved along the axis of the EAS.

Questions involving composition of the primary particles remain. The measurements that piece together the composition of a primary particle can vary from detector to detector, and there can be ambiguities in measurements between different primary particles. For instance, measurements of a shower's longitudinal development has sets of fluctuations from shower to shower for the same primary. The X_{max} distributions for proton and iron primaries overlap. EAS development models predict a theoretical X_{max} and elongation rate not globally quantifiable across all models.

The collapse of the CR flux at energies above the knee and the primary's devolution to an EAS 'footprint' of increasing area with increasing energy has standard primary mass profiling responses.

When a CR primary enters the atmosphere, sooner or later, typically within the first 100 g cm^{-2}, it interacts with an atmospheric nucleus. The primary responds by initiating a cascade of secondary particles, each of which may also interact. At groundfall, the EAS particle number has multiplied many times. For example, considering the atmosphere to have a vertical thickness of 26 radiation lengths and 11 interaction lengths – the EAS produced by a vertical proton primary of energy $\sim 10^{19}$ eV, reaching sea level may contain $\sim 3 \times 10^{10}$ particles (99% being photons and

electrons/positrons in a 6:1 ratio in the 10 MeV range). The other particles are muons, pions and neutrinos/baryons. The energy content is, in order, $\sim 85\%/10\%/4\%$ and 1%. The EAS comprises 3 distinct parts - a hadronic component of nuclear fragments comprising the shower core, an electromagnetic (EM) component of sub-showers, and a muonic component.

The longitudinal axis of the shower is generally too remote for direct ground-based detection with a sparse array, and thus the shower's lateral distribution and intensity are the observables used to construct a scaffolding of measurements leading ultimately to the determination of the properties of the primary particles. Each shower part has a characteristic lateral distribution, and intensity, and best results are achieved with a conjunction of assorted detectors.

For the structural hadronic component, a ground-based sampling calorimeter, as used in the KASCADE experiment, may measure the energy and number of associated hadrons. Muons are also registered there at energies > 2 GeV [82].

The EM component may be studied by optical detectors. The electron, e, population emits a narrow cone of Cherenkov light along the shower axis, and, at energies $> 10^{17}$ eV, atmospheric nitrogen atoms light up UV fluorescence in the shower's wake. Fluorescence telescopes effectively track X_{max}. The measurements of $\langle \ln A \rangle$ and $\langle X_{max} \rangle$ involve the depth relationship at shower maximum being $\propto \log(E_0/A)$ (see section 2.4).

The most penetrating component, the muonic fraction with passenger neutrinos, can be studied with buried detectors for the higher energies. Muons at GeV energies are recorded at around 2 metres and muons initially at TeV energies are found at depths > 1 km. Arrays, such as those of the Pierre Auger Observatory, are able to measure both the EM portion and muons, typically with $E_\mu > .23$ GeV. Muon production occurs high in the atmosphere and muon lateral distributions can extend to kilometres from the shower core [82]

Each step in the hadron cascade has $\sim 1/3$ of the pions decaying, bleeding off energy into the fluorescent EM display. The remaining pions may decay to muons, but the more steps taken before muon production begins, the less energy is available for their production.

With regards to the example of the down-going proton shower (see also subsection 2.1.2 on page 13), each square metre of ground receives a few muons. Providing there are enough muon

detectors over a large area, the simpler uncertainties of the N_μ/N_e measure (as opposed to hadronic model measures), has these observables favoured [69, 79, 82].Figure 2.1 illustrates the logitudinal developments of shower components. The EAS modelling is instructive, rather than accurate.

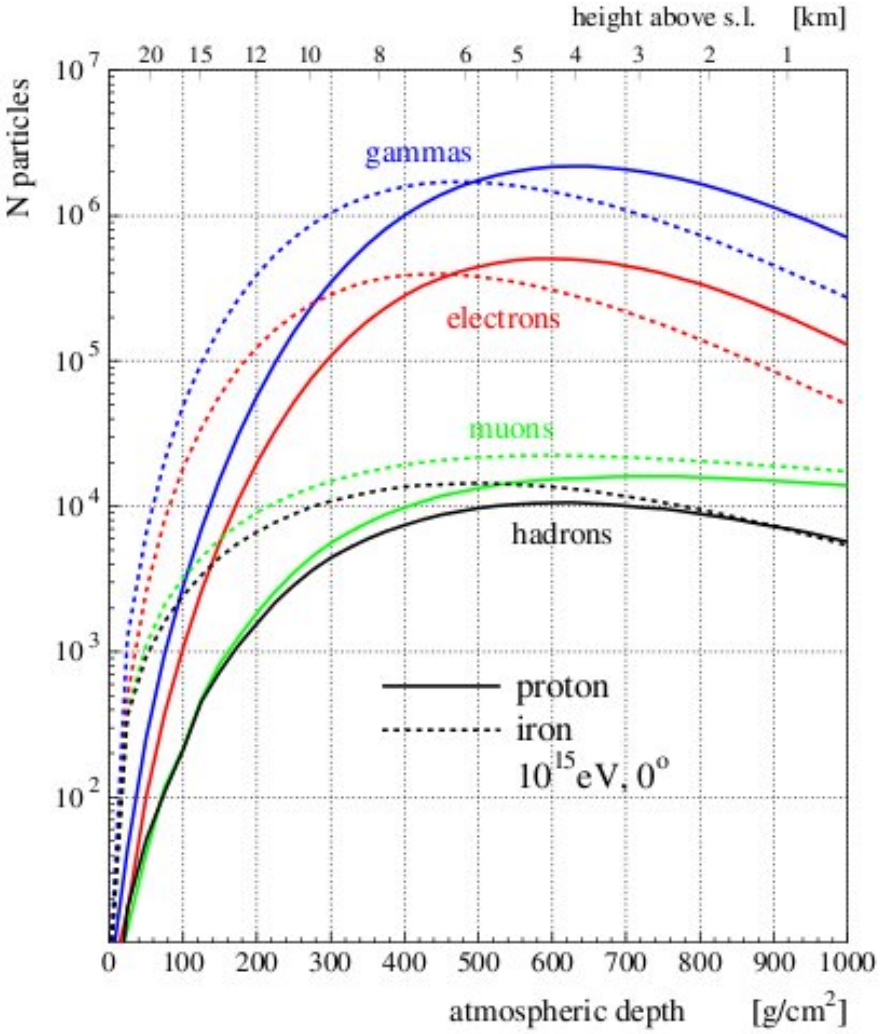

Figure 2.1: **Example of the longitudinal EAS development of the by-products of a vertical primary CR proton and a primary CR iron ion. From [82].**

Models provide a reliable guide to understanding the physical processes underlying their complex structure. EAS are then best understood by modelling their simplest component first, and building on that.

2.1.1 EM Cascades

Heitler fashioned a simple binary tree (see chapter 5 on page 118 on MST theory) to model the physics of EM cascades. EM cascades are considered to describe energy losses through the radiation of interacting electrons/positrons or photons with a medium (air). The number of particles is assumed to double each time there is an interaction, until such time when the radiation interaction length of the particles is larger than the interaction length of other types of energy losses, such as ionization. Heitler's model requires that collisions be inelastic, all energy transferred to the developing cascade, and cross-sections are energy independent. Shower attenuation and the loss of e^-/e^+ into the medium are not considered [128]. Bremsstrahlung radiation is minimized and produces only one photon, and the radiation length of particles, λ_r, in the medium is the same [126].

At high energies ($E > 1$ GeV) the radiation length, λ_r, (for electrons/positrons) is described as the median thickness the particle must travel through before it loses on average of $63.2\% \sim 1/e$ of its energy through bremsstrahlung. For photons, λ_r is $7/9$ of that mean free path length. In the model these values are approximated as equal. The interaction step length or 'splitting length', d, a particle travels, is given by $d \sim \lambda_r \ln 2$. Thus for electrons (electrons and positrons, from now on, are referred to as electrons), we have, for an incident electron, with energy, E_0, at a depth of z [202]

$$E = E_0 . \exp^{-z/\lambda_r} = E_0 . \exp^{-t}$$

where t is the depth measured in units of λ_r.

Reductive although it is, this model reproduces the EM cascade structure and qualifies three useful properties of EM cascades.

Two types of radiation energy loss are considered. The particles are either electrons or positrons or photons. Photons undergo equal energy pair-production where

$$\gamma + N \rightarrow e^- + e^+ + N$$

and each electron produces a photon via bremsstrahlung emission

$$e + N \rightarrow e + N + \gamma$$

where N is a nucleus in the medium the particle crosses.

So an electron, e, with initial energy, E_0, after travelling one interaction step length, $\lambda_r \ln 2$, and radiating a photon, now has $E_\gamma \equiv E_e = 1/2E_0$. Each and every particle continues on producing 2 more, dividing their energy in half. At a point of critical energy, where $\zeta_c^\gamma \equiv \zeta_c^e$, and after n_c^γ steps, the particle multiplication ends, the particle maximum number, N_{max}, has been reached, and the particle energy losses are dominated by collisional and ionization energy losses. We see

$$E_0 \propto \zeta_c^\gamma N_{max} \tag{2.1}$$

The primary energy, E_0, is divided equally between N_{max} particles, or rather an EM shower produces of a maximum of $N_{max} = 2^{n_c^\gamma}$ particles, each having a critical energy, $\zeta_c^\gamma \equiv \zeta_c^e$, after n_c^γ critical steps covering a depth of $n_c^\gamma \lambda_r \ln 2$. Thus, from Eqn. 2.1 and substuting $N_{max} = 2^{n_c^\gamma}$, we have

$$n_c^\gamma = \ln[E_0/\zeta^\gamma]/\ln 2$$

So, given an EM primary with inital energy, E_0, interacting within a medium with a splitting length, $\lambda_r \ln 2$, at a starting position of X_0, and radiating its maximum number of particles, $2^{n_c^\gamma}$, after n_c^γ steps - we arrive at the conclusion that the shower depth at maximum (critical energy$\equiv$ $\zeta_c^\gamma, N_{max} \equiv 2^{n_c^\gamma}$)

$$X_{max}^\gamma \propto X_0 + n_c^\gamma \lambda_r \ln 2 = X_0 + \lambda_r \ln[E_0/\zeta_c^\gamma] \tag{2.2}$$

Thus we see two important EM shower properies are related to the primary energy. The shower has its maximum number of particles proportional to E_0, and the depth of shower maximum increases proportionately to the logarithm of that primary energy.

We also propose an important variable, the rate at which X_{max}^γ changes with energy - the elongation rate, denoted D_{10} or Λ. We find, from Eqn. 2.2, and using the critical energy of an EM

shower in air, $E^{\gamma} \sim 85$ MeV with a radiation length, $\lambda \sim 37$ g/cm^2, that D_{10} is coupled to the radiation length [126, 128, 124].

$$D_{10} \equiv \frac{\mathrm{d}X_{max}}{\mathrm{d}\log_{10} E_0} = 2.3\lambda_r \tag{2.3}$$

The Heitler model inaccurately predicts the ratio of electrons to photons - one reason being Bremmsstrahlung radiation results in several photons, not one. The ratio of electrons to photons is more like 1:6 [128].

Heitler's binary tree model for electromagnetic (EM) cascades was extended by Matthews [128] to explain the structure of a hadronic cascade and qualify similar properties.

Abridgements include a hadron interaction length in air, $\lambda_I \sim 120$ g/cm^2, $\zeta_c^e = 85$ MeV, $\zeta_c^{\pi} = 20$ GeV, $\lambda_r \sim 37$ g/cm^2 and it is assumed each hadron interaction produces N_{ch}=10 charged pions and $1/2N_{ch} = 5$ neutral pions.

The most basic hadronic cascade is proton initiated.

2.1.2 Proton Cascade

The following is based on Matthews [128]. A proton of energy, E_0, travels a typical distance, $d = \lambda_1 \ln 2$, before interacting to produce 3 pion types - the charged π^+, π^- and the neutral π^0.

The neutral pions do not interact and instead decay into 2 photons, thus an EM cascade begins. The logic of a proton cascade follows closely upon the logic of an EM cascade, excepting we must factor in an EM cascade, which, at first interaction, carries $1/3$ of the primary particle's initial energy, E_0, and a hadronic cascade, which at first interaction carries $2/3$ of the primary particle's initial energy. Thereafter, each charged pion decays to a muon $N_{\mu} = N_{ch}$ (actually muons and neutrinos). After experimental corrections, we are potentially able to reconstruct E_0 by just counting EM particles and muons.

The proportions of muons and EM particles, from now on referred to as electrons, depend on the number of interaction steps to reach their *individual* critical energies.

The more steps taken to reach the critical energy of decay of the charged pions, ζ_c^{π}, to muons, the more the proportion of muons to electrons decreases.

We know the primary particle's energy, E_0, is approximately conserved as the sum of the EM cascade energy, E_{em}, plus the hadronic energy, $E_h = N_\mu \zeta_c^\pi$. i.e.

$$E_o = E_{em} + E_h$$

The EM fraction of the primary particle energy, E_0, is then [128]

$$\frac{E_{em}}{E_o} = \frac{E_0 - N_\mu \zeta_c^\pi}{E_0} = 1 - \left(\frac{E_0}{\zeta_c^\pi}\right)^{\beta-1} \tag{2.4}$$

where $\beta \sim 0.85 \rightarrow 0.92$.

2.1.2.1 Proton Depth at Shower Maximum: X_{max}^p

To calculate X_{max}^p, we restrict the X_{max} of EM showers to the first generation of the EM showers. Thus, the proton depth of shower maximum for a first proton interaction depth, X_1, is, given that one third of the primary's initial energy belongs to the EM shower initiated by a photon,

$$X_{max}^p = X_1 + \lambda_r \ln\left(\frac{E_0}{N_{ch}\zeta_c^e}\right) \tag{2.5}$$

which can also be expressed [128] as

$$X_{max}^p = X_{max}^\gamma + X_0 - \lambda_r \ln(3N_{ch}) \tag{2.6}$$

where X_{max}^γ is the EM depth at shower maximum.

The resultant elongation rate is

$$D^p = D_{10} + \frac{d}{d\log_{10} E_0}(X_0 - \lambda_r \ln[3N_{ch}]) \tag{2.7}$$

A common practice when studying nuclear cascades, is then to consider an EAS as a nucleon superposition model.

Regarding a primary nucleus of atomic number, A, as a superpositon of A nucleons of indi-

vidual energy, E_0/A, we see that heavy nuclei showers mature more rapidly and higher than their proton limit, and their shower to shower fluctuations are less. The relationships between nuclei, A, and protons, p, simplifies to

$$X^A_{max} = X^p_{max} - \lambda_r \ln A \tag{2.8}$$

where λ_r is the radiation length of the medium (see Figure 2.2), and

$$N^A_\mu = N^p_\mu A^{1-\beta} \tag{2.9}$$

where $\beta \sim 0.9 \rightarrow 0.95$.

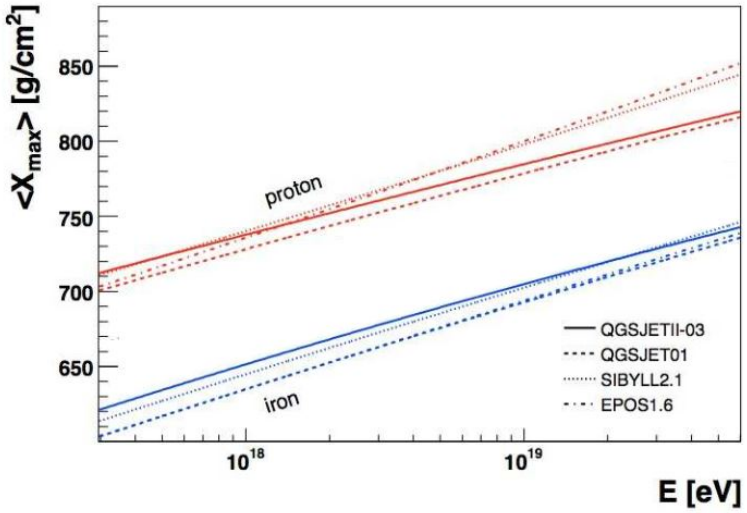

Figure 2.2: $\langle X_{max} \rangle$ **vs Energy for proton and iron modelled EAS (4 model types). The depth of maximum of each ion consistently differs by a constant of ~ 100g/cm^2, the iron $\langle X_{max} \rangle$ being lower. The elongation rates are the slopes of the graphs. From [126]**

The depth of maximum for iron and protons differs by a constant of ~100 g/cm^2 [124, 126]. The cross section and multiplicity change with energy are the same. Iron showers have a larger N_μ count than proton showers of similar energies. The ratios are ~1.8:1 [126]. Showers of any primary at higher energies have a larger X_{max} than they have at lower energies.

2.1.3 Mixed Composition

A set of EAS primaries is not necessarily composed of a single element. Their composition may well be mixed. Therefore their (average) statistics are taken to be combinations of their contributing fractions.

In the superposition model of air showers, the shower development of heavy nuclei with mass, A, and energy, E_0, is described by the sum of A proton showers of energy $E = E_0/A$. The shower maximum, X_{max}, penetrates into the atmosphere as X_{max}~$\ln E$. Hence, most shower observables at ground level scale proportional to $\ln A$. and the mean logarithmic mass becomes

$$\langle \ln A \rangle = \sum_i f_i \ln A_i$$

where f_i represents the fraction of nuclei with mass A_i.

Linsley [62] showed that the elongation rate for protons cannot be greater than than λ_r, and, when this boundary condition is violated, there must be a change in composition from light to heavy. Heavy nuclei shower maxima are smaller (happen sooner) than for lighter nuclei, and their respective shower to shower fluctuations are also less. There is, however, an overlap in the X_{max} fluctuations. There may be times when the elongation rate is too great - when composition is adjudged changing from heavy to light, or there is a smaller elongation rate than expected when changing from light to heavy. These violations are considered to occur when the primary composition is mixed.

Simulation models all share the above properties and all attempt to differentiate between X_{max} and its fluctuations from shower to shower. That is, these 'observables' are subject to their hadron models, and it can be hard to relate them to a particular mass composition. However, the $\sigma(X_{max})$ is less model dependent, and with the present generation of detectors, we can be reasonably sure of the position of X_{max}. So, we assume shower to shower fluctuations of X_{max} are primarily a function of X_0, which has an exponential distribution, and take $\lambda_r \ln 2$ as the average distribution length. For protons at 10^{18} eV, fluctuations in X_0 are 60 g/cm^2 which accords with simulations. For iron, using the superposition model, $\sigma(X_0) \sim 14$ g/cm^2. This gives a lower bound for the fluctuations

in $\sigma(X_{max})$[1]. As has been previously stated, the effects of any fragmentation (the first and all that follow) are ignored in the superposition model. Simulations have $\sigma(X_0) \sim (20 - 25)$ g cm^2. In the gross sense, mass composition of *different* showers can be *statistically* separated. They are light or heavy or mixed and are ultimately constrained by the hadron theory.

2.2 Composition

Many CRs captured by Earth-based detectors are secondary CRs – particles that result from the primary CR particle fracturing through collisions in the ISM and with the planetary atmosphere - 'Cosmic Ray Spallation'. Satellite and balloon borne detectors remain the chief functionaries of primary CR identification at the lower energies. For the energy range $\sim (10^{10} - 10^{15})$ eV, there are results for individual element fluxes from AMS-02 (H and He) and CREAM (H, He and Fe) [7] . For energies $> 10^{15}$ eV the energy spectrum for CRs, albeit secondaries, directs mass composition. For those CR primaries whose energies are low enough to allow reconstruction using data obtained by observing collisions in hadronic colliders, the mass composition is well appreciated. The steep falloff in CR flux and its coeval rise in energy, plus limiting factors in detector size and the power of hadronic collectors makes for an increasingly uncertain primary CR mass composition at high energies - it has no direct measure. Primary mass identifiction becomes problematic for energies above $\sim 10^{14}$ eV [65]. Model-dependent mass tables can be contradictory and even the same models can offer a bewildering choice of variation [76].

Direct observations of CR subsidiaries, such as neutral γ-rays at lower energies can be usefully paired with UHECR events/sources, where EAS are all we have have to work with. For example, the Compton γ-ray Observatory directly observed γ-rays up to $\sim$ 100 GeV [77]. SNR have distinct synchrotron electron radiation at these energies and the high energy X-rays have also been observed by Chandra [79].

For energies $\sim 10^{19}$ eV there are important flux limits for the photon and neutrino secondary CR particles. These flux limits are tied to the composition of their primary particle. In the UHECR energy region there are conflicting compositional results from the Pierre Auger Obsevatory (Ar-

[1] elongation rate for EM showers is the upper limit

gentina, southern hemisphere) and HiRes and the TA (both of the latter arrays are in Utah USA, northern hemisphere). Independent measures of these particle fluxes will clarify whether the primary CR composition in the disputed energy regions is heavy (Pierre Auger Obsevatory) or light (HiRes, TA). All three experiments do agree with one another within systematic errors of 20% [215].

2.2.1 Before the Knee $\sim (10^6 - 10^{14})$ eV

For CRs with kinetic energies of less than 100 GeV per nucleon, abundances are confidently detailed up through the periodic table until Ni [73]. These abundances accord well with the chemical composition of the ISM, an equilibrium of particles formed by stellar and Big Bang nucleosynthesis. The composition consists overwhelmingly of charged particles, approximately 86% protons, 11% helium nuclei, 2% electrons, 1% heavier nuclei and less than 1% secondary anti-particles. There is a small neutral component along with γ-rays that fills out the percentage [72]. These proportions differ slightly from reference to reference [69] [73].

Regarding anisotropies in the flux of CRs and possible composition effects, in the energy range below $\approx$ 1 GeV, solar activity is strongly coupled to the CR flux - **solar modulation** - the magnetic field of the heliosphere suppresses flux ingress, particularly of the heavier elements. In the TeV range this solar modulation anisotropy has to be separated from the anisotropy of the galactic Compton-Getting Effect (CGE) [28]. There also exists a dominant CR temporal anti-correlation which follows after the cycles of solar activity. This anti-correlation points to a non-solar source of CRs at these energies.The 'Forbush decrease' describes the drop in CR flux caused by a Coronal Mass Ejection (CME) or 'solar flare', when a convulsion above the Sun's surface ejects a cloud of plasma. The Sun's CME's magnetic field deflects CRs from its path as it travels towards Earth, the measured CR flux drops, and then recovers once the CME has passed by.[2]

The mass composition at low energies is similar to the composition of the solar system (SS) [89], excepting for some significant abundances. The SS condensed from the collapsing interstellar medium (ISM) some 4.5×10^9 years ago and since then these processes have continued to this day,

[2] A solar storm on the 07/10/2005, dropped levels of CRs on board the International Space Station by 30% [78].

the present ISM composition has probably continued unchanged through the ages [75], although alternative views have merit [74]. The compilation of SS elemental abundances has many sources. They are extracted from terrestial, lunar and meteoritic samples, and complemented by solar spectroscopic measurements which have been compared to like stellar and ISM measurements.

Some of those CR elements whose abundances do not reflect solar system distributions are the 'secondary nuclei' - the light element triplet Li, Be and B and the sub-iron group Sc, Ti, V, Cr and Mn. These overabundances are evidence of a galactic CR component. The overabundant secondaries (1 : 100 for Li, Be, B versus solar abundances; 1 : 3 for Sc, Ti, V, Cr and Mn versus solar abundances) are speculated to be the result of cosmic ray spallation - the fragmentaion of C and O nuclei interacting with the ISM for the former secondaries and the breakup of Ni and Fe in the latter. This departure from the standard solar system abundance is exploited by calculating the spallation cross-section, and thus the ISM column density necessary to have generated the said overabundances.

Other CR abundances inconsistent with the SS/ISM are H and He, and a considerable block of elements separated by a defining factor, their Condensation Temperature, T_c. It is believed that many of the GCRs are ISM accelerated in SN blasts. The ISM consists of gases and dust. The T_c, at which elements condense to become solid or 'grains', divides those elements with $T_c < 400^\circ$ into 'volatile', and those that crystallise at a few thousand degrees celsius, into what are termed 'refractory'. The volatile elements follow a general trend of slight CR enhancement as an increasing function of their mass-to-charge ratio. There is the presumption that elements with higher charge are accelerated or injected more efficiently. The elements that form dust grains more easily, the refractories, are substantially overabundant with respect to the gas-phase, volatile elements and their abundances seem to to be independent of mass. Whether the refractory elements are preferentially abundant in accordance with acceleration efficiency, and if they are indeed ISM dust grains or fresh SN ejecta have their champions [16, 68]. The Ellison, Drury and Meyer acceleration model [16] with ISM dust grains agrees well with source spectra of CRs. Figure 2.3 on the following page illustrates the close correlation between SS refractory abundances and CR refractory abundances, indicating CRs are accelerated out of a well mixed ISM.

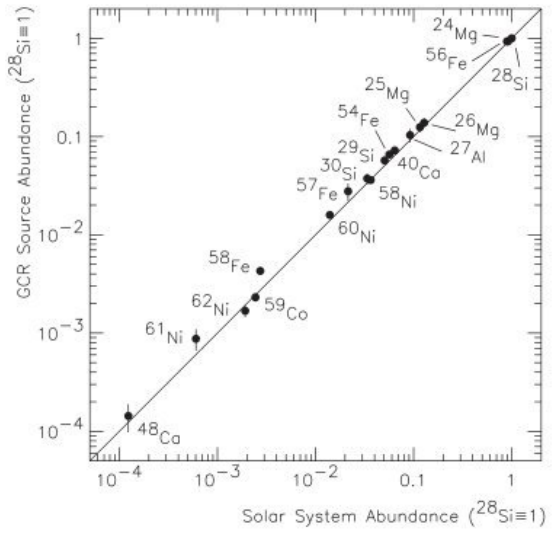

Figure 2.3: **The almost direct 1-to-1 relationship between Solar System abundances of refractory nuclides and derived CR refractory abundances, normalized to Si^{28}. From [89].**

Those elements that do not fit the T_c pattern, e.g. C, O, Ne and Ni, may have their overabundances explained as products of the enormous Wolf-Rayet stars (mass $> 40_{\odot}$), once they have become black holes. Stars with masses $> 20_{\odot}$ are calculated to return 90% of their mass to the ISM compared with 50% of the mass for lesser stars. Stars of masses $< 3_{\odot}$ return material which is largely unprocessed ISM, but for the massive stars nucleosynthesis plays an important role [75].

Most experiments have the composition towards the knee as increasingly heavy, and there is a flattening of the spectral slope just prior to the knee, although the position of the knee itself is subject to experimental uncertainties. The spectral index up to the knee has been measured at $\alpha \approx 2.7 - 2.75$ [126].

2.2.2 Knee to Ankle $\sim (2 \times 10^{15} - 10^{18})$ eV

In this region, the rest-frame of galactic CRs is in sync with galactic rotation and there are no shift-in-reference frame anisotropies [93],and anisotropies for the heavier elements are washed out by the diffusive propagation regime of galactic magnetic field deflections [79].

The steepening that occurs between $3 - 5 \times 10^{15}$ eV is known as the knee of the spectrum. The

feature around $(4-10)\times 10^{18}$ eV is called the ankle of the spectrum [171] [133] [94].

Above E $\sim 10^{14}$ eV, mass composition measurement relies substantially on indirect methods and ascribing composition to unique elements is fraught with inaccuracies - although some EAS observables can now be compared with their counterparts generated in the Large Hadron Collider (LHC) at CERN for energies up to $\sim 10^{17}$ eV (by the end of 2015 CERN aims to reach centre of mass energies of $\sim 14\times 10^{12}$ eV). In the meantime, in the region $(\sim 10^{14}-10^{17})$ eV of the energy spectrum, we are reasonably comfortable in tallying composition in terms of primary particle groups, albeit put together by theoretical models based on a matrix of differing assumptions, and having results irregular in detail [76] [171]. These irregularities become more extreme at energies $> 10^{17}$ eV, where we can only attempt to distinguish between light, medium or heavy primary nuclei.

The KASCADE (Karlsruhe Shower Core and Array Detector) experiment at Karlsruhe, Germany, provided an extensive multi-detector array, and the most thorough theoretical inquiry to date for energies in the $(10^{14}-10^{17})$ eV region [83] [183]. KASCADE began running in 1996 [181].

The KASCADE detectors are independent, and the correlation of their data into single EAS events produces superior data, which is then further subject to qualifying cuts [171].

A Monte Carlo EAS simulation code, developed at Karlsruhe, the CORSIKA project, maintains and updates a scrimmage of different hadronic interaction models that process the KASCADE data. Models at low energies are GEISHA, FLUKA and URQMD, and at high energies DPMJET , NEXUS, QGSJET , SYBILL and VENUS. The results are compared and weighed, each one against the other. Results are often not consistent across model types, each modelling possessing characteristic instabilities. KASCADE reports for high energy interactions the models QGSJET 01, SYBILL 2.1 and DPMJET 2.55 [84] to be the least fractious. What is consistent, apart from the primary mass increase as a function of energy up to 10^{17} eV, is the presence of a knee position on all 3 shower components at around $(3-4)\times 10^{15}$ eV [130]. See [82] for an early publication detailing their approach.

KASCADE reports significant results for the 'all' particle energy spectra and the position of the knee with some ambiguities in the individual elemental groups energy spectra [83]. The knee

is the domain of light elements, $Z < 6$. First, a break in the proton intensity is responsible for the knee, E^{p}_{knee} being $\sim 3-5\times10^{15}$ eV, followed by a 'kink' in the helium intensity, and a continuing fall in flux until $E^{A}_{knee} \sim Z \times E^{p}_{knee}$ [171], i.e. for elements in the region up to $\sim 10^{17}$ eV, there is a $Z \times 3 \times 10^{15}$ eV relationship to be seen in the CR energy density flux, the iron knee being at $\sim 8 \times 10^{16}$ eV. So, the knee in the elemental groups follows the pattern of magnetic rigidity dependence, $E \propto Z$.

We do not know why the post-knee slope does not continue a rigidity implied sharp decline at energies $> 10^{17}$ eV, as is seen when the spectrum is decomposed (or unfolded) elementally (see Figure 2.4). It may be a changeover in source, either galactic, or an incursion of CR nuclei from an extra-galactic source [79]. In both the galactic and extra-galactic regimes, a collection of sources is likely to have produced the energy spectrum given its temporal stability. Neither do we know whether this rigidity is due to acceleration processes at the source or that predicted by the 'leaky box' model [80] [82].

The 'leaky box' model – with the mean escape time of nuclei as a function of particle energy, has a GMF that lets slip the lightest elements at the lowest energies first, relinquishing each element in a discrete, coupled, energy-mass progression up the periodic table – successfully predicts data from their stable primary and secondary nuclei. The widely used Galactic Propagation (GALPROP) numerical code for CR diffusive transport has been reported as having had success with such a model for all the stable nuclei included in its code [132].

The acceleration process predicts a rigidity-based acceleration of nuclei from a source reaching the limit of its power according to the Larmor radius of nuclei.

Figure 2.4 on the next page highlights the acute rigidity downturn for the CR elemental flux, contrasted with the slighter decline of the all-particle CR flux after $\sim 10^{17}$ eV. The dashed line, TG (Total Galactic), is a superposition of the sum of these elemental fluxes. The hatched areas represent the use of the different hadron models, QGSJET and SIBYLL - their $\frac{N_\mu}{N_e}$ ratios differ slightly. This region marks the end of the SNR galactic contribution.

Observables have many dependencies, both the composition and energy spectra are co-dependent and imprecise, and so the KASCADE team developed a series of coupled integral equations which

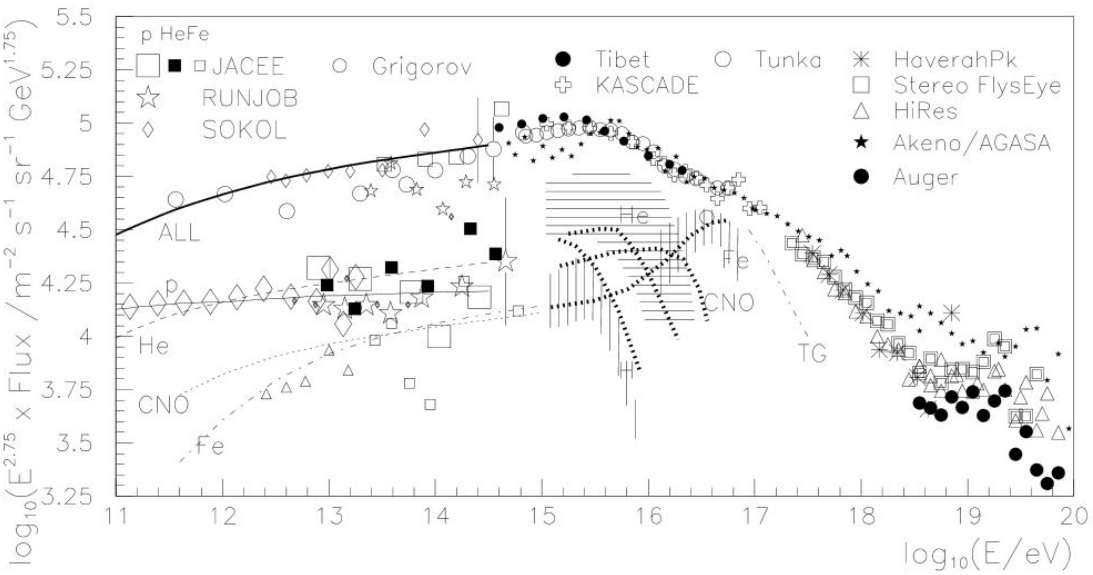

Figure 2.4: **elemental and all-particle energy density flux. From [79].**

would yield the composition of a representative all particle primary mass spectrum - H, He, Ca, Si and Fe at a summation over all possible energies - the entire spectrum being too ambitious to deconstruct. These coupled equations took on a matrix form based on the two dimensional frequency distribution of $\log N_e \geq 4.8$ vs $\log N_\mu^{tr} > 3.6$[3]. The equations were 'unfolded' using mediating algorithms and run through the QGSJet and SIBYLL shower distribution simulations. Literature explaining the method and results can be found in [82, 83, 86] .The results for protons and iron are shown in Figure 2.5.

An extension of KASCADE, KASCADE-Grande is a 7 station array of scintillators [129]. Designed to complement KASCADE, this experiment operated in the $(10^{16} - 10^{18})$ eV range. A second knee in the all-particle spectrum, the heavy analogue of the first, light component knee, should lie $\sim 10^{17}$ eV. Preliminary results for zenith angle $(0 - 42)^\circ$ indicated an increasingly heavy component around here [129]. Later results from KASCADE-Grande for similar zenith angles, had three approaches exploiting correlations of E_0 with the ratio of N_{ch} to N_μ between events [130]. Fine structural features have been found, with a concave section of the energy spec-

[3]the exponent 'tr (truncated)' indicate further quality cuts

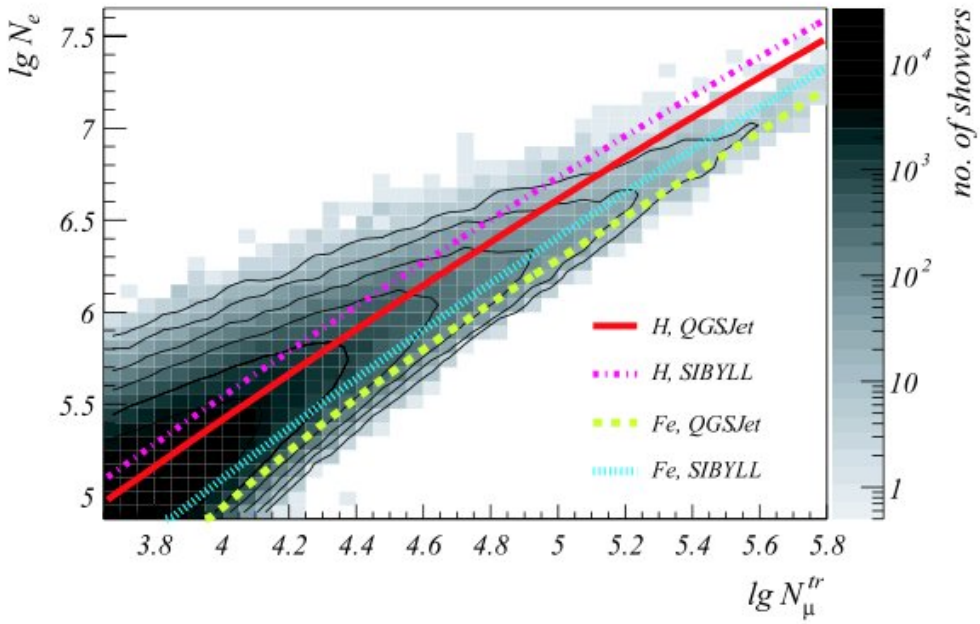

Figure 2.5: **frquency distribution of** $\log N_e$ **vs** $\log N_\mu^{tr}$ **and lines/isolines for most probable values of proton and iron induced showers for QGSJet and SYBILL shower reconstruction simulations.** *tr* **indicates KASCADE quality cuts. From [83].**

trum, a 'dip' just above 10^{16} eV, and a break at $\sim 10^{17}$ eV yielding a slope -3.24 ± 0.08 [130]. These 3 calculations are shown in Figure 2.6 [130]. Several further refinements are then implemented, including the unfolding techniques already mentioned [183].

Tibet III results [131] show a knee at 4×10^{15} eV, with a spectral index $\gamma = -2.68 \pm 0.02$ below 10^{15} eV, and -3.12 ± 0.01 above the determined knee position. Data from KASCADE and EAS-TOP [126] indicate a progression of light to heavy composition with increasing energy between $(10^{15} - 10^{17})$ eV.

Another report from KASCADE-Grande [171] has the all particle energy spectrum with a hardening at $\sim 2 \times 10^{16}$ eV, and a steepening at $\approx 8 \times 10^{16}$ eV. Shower reconstruction at lower energies was with the CORSIKA code, for the FLUKA and QGS-Jet II hadronic interaction models. At high energies, up to 3×10^{18} eV, where the statistics were few, the EPOS hadronic interaction model was used. The QGSJet II model predicts a smaller ratio of muons to charged particles at ground than the EPOS model. Composition derived using EPOS is much lighter than that of either

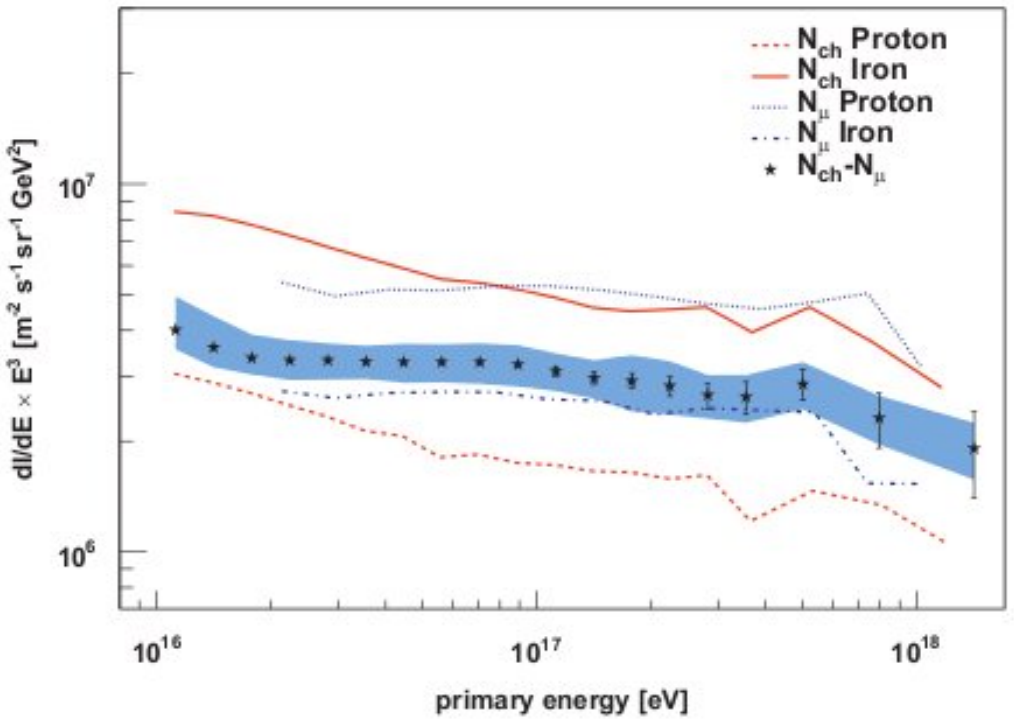

Figure 2.6: **KASCADE-Grande all particle energy spectrum over** $(10^{16} - 10^{18})$ **eV for proton and iron. Results from 3 correlation methods for ratio of** E_0 **to** N_{ch}, N_μ **and** $N_{ch} - N_\mu$**. Systematic errors for** $N_{ch} - N_\mu$ **are shaded. From [130].**

SIBYLL or the QGSJet line. In particular, EPOS model results are incompatible with those of KASCADE where their energy ranges overlap.

The energy spectrum according to the EPOS intepretation has not been finalized due to poor statistics. However the composition can be inferred, whilst noting that the interaction models have inconsistencies. The energy spectrum hardening is expected with rigidity based galactic models, caused by the gap between the light and heavy nuclei abundances, and the steepening is around where an iron knee would be located. However these features could also be explained by the Hillas thesis of a second, galactic based component B. For more, see references [171, 79, 102]. The dip model has a heavy compostion around the second knee and the Hillas model would have a mixed composition between $(10^{16} - 10^{18})$ eV.

The experimental aspect of KASCADE/KASCADE-Grande finished running in March 2009,

but the KASCADE group, with their curation of scrupulously screened data continues to publish results.

One explanation for the knee is that it is due to a second component of accelerated ejecta from young supernova remnants (SNR). Support is from the ATIC experiment which observed a pileup of electron/positron flux at tens of GeV energies, and the ground-based H.E.S.S., and the satellite *Fermi* experiments, have both noted an increase in flux around these energies. The PAMELA experiment records a positron excess at $\sim$ 100 GeV. A young SNR environment, and e^+e^- pair-production by high energy CR nuclei on the photon background predicts such excesses [182]. Indeed, SNRs of many classes, e.g. Wolf-Rayet stars, together with magnetic field generation theory, may satisfy the rigidity constraints of nuclei present in the knee and approaching the ankle [79]. A galactic CR production from primarily SNRs has support.

2.2.3 From the Ankle and Beyond $\sim (10^{18} - 10^{21})$ eV

The discussion of where, and indeed, if, the CR spectrum enjoys an extra-galactic contribution, has concentrated in this region where further changes in its slope are evident and galactic origin models fail.

The ankle is identified as that part of the CR energy spectrum where the flux is observed to flatten at $\sim 3 \times 10^{18}$ eV. This levelling is traditionally interpreted as a diminishing galactic CR remnant overlain by a mounting extra-galactic CR flux [19, 75].

The ankle has also been interpreted as an artifact of an extra-galactic CR component with proton primaries, the result of a 'pair production dip' - the signature of energy losses through e^-e^+ production by protons interacting with the 2.7 K Cosmic Microwave Background (CMB) radiation. This Bethe-Heitler pair production plus adiabatic losses engendered by the expansion of the universe, suggests an extra-galactic transition at energies below 10^{18} eV. A variant ankle model, supported by Hillas [79], replaces the failure in SNR acceleration models at ankle energies with other galactic CR sources, possibly red giants or Magnetars [79]. A predicted feature, the onset of the GZK limit, assuming an isotropic distribution of sources, imposes a horizon of < 200 Mpc from Earth at energies, $E > 60$ EeV, for CRs of all types - interactions with the Cosmic

Microwave Background (CMB) would inhibit nuclei from further distances reaching us [17, 18]. Because standard acceleration models for large-scale galactic sources do not possess the reach of UHECRs and a posited GMF of $\sim \mu$G would not retain them, it is usually agreed these UHECRs originate from sources outside our galaxy.

Being able to differentiate composition before and after spectral breaks is important because composition helps determine origin - galactic, extra-galactic, point or large-scale sources. Both the extra-galactic *introduction* and the *post* extra-galactic dip interpretations of the ankle have predictions on composition, the former admits to a proton dominated early extra-galactic presence at the ankle, and the latter model places the extra-galactic presence well before the ankle at around the second knee. A change in EAS from a heavy composition to light nuclei composition between $(10^{17} - 10^{18})$ eV would support this idea and lend weight to an earlier introduction of extra-galactic CRs (EGCRs).

With regards to the previous section for the second knee, we see the second knee seems to be heavy. Then the spectrum holds steady until another structure, a dip, is observed, then again the spectrum holds steady until where the ankle is found. The data just after the second knee do suggest a lighter component. The dip also seems to be light, with a reduced flux. Composition from the ankle on, once again, seems to become heavier. The highly regarded KASCADE results support the galactic CR contribution ending at $\sim 10^{18}$ eV with a heavier element (iron) dominated composition which would favour an early changeover staging, but a heavy element transition region between $(10^{18} - 10^{19})$ eV would present problems [19].

All of the data, really from the second knee, have serious associated uncertainties because of reasons already mentioned in the section 2.1 on page 8. We just don't know the composition where the dramatic discontinuity at the *predicted* proton GZK is seen.

The GZK cut-off for protons at energies $> 4 \times 10^{19}$ eV, would see a drop in the flux of protons and probable sources closing from the Gpc range to that of Mpc. If UHECRs consist mainly of nuclei, the flux attenuation, due now also to interactions with the optical and IR background radiation, is expected to be translated to between 2×10^{19} eV for the lightest nuclei, to 2×10^{20} eV for the heaviest, with the same contraction in source range as for protons [134].

The tourney over composition, between an ankle introduction of EGCRs, with a flux dominated by the lighter elements, or the scenario with an ankle of mixed composition continues. Debate over whether features in the CR energy spectrum above the knee indicate a changeover between two galactic sources or a transition between galactic and extra-galactic sources also continues [7, 6, 8].

Sidestepping the acceleration problems that come with sources illustrated in the Hillas diagram in Figure 3.1 on page 65, are top-down decay scenarios. The recent decay ($<$ 100 Mpc) of supermassive 'relic' particles, or topological defects collapsing to unstable Grand Unified Theory (GUT) particles, then directly supplies UHECRs with the necessary 'acceleration'. These exotic particles devolve into ultra-high energy photon and neutrino primaries which should dominate this region with fractions of $\sim$ 90% [126]. Searches for photon fluxes in showers incident at large zenith angles in the EeV range have a proportion of photon primaries to baryon primaries of $< 0.4 : 1$ [71].

The Pierre Auger Observatory has actively sought for such primaries. Its hybrid detection arrangement with its measures of X_{max}, sees such measures of photon content even more tightly controlled (see subsection 2.5.1 on page 44). A search for photons would be fluorescence-based measurements on late-developing X_{max}, due to low multiplicity (one must explicitly exclude bias). Zenith angles are only accepted above 35°, and Cherenkov light contamination is voided with angular directions between camera pixels and directions $>$ 10°. At higher energies, SDs alone are seconded in the photon search. The photon deep showers would have large values of the parameters, R_c, the radius of shower curvature; and the shower risetime, $\tau_{1/2}$, the time when between (10−50)% of the shower front has arrived at the detector. Across the energies $\sim (2{\times}10^{18} - 10^{20})$ eV, none of the low photon fraction statistics published by the Pierre Auger Observatory encourage any top-down models. The large flux of photon primaries that is predicted at 100 EeV has not materialized [126].

At PeV energies there is a predicted low photon flux from photo-pion production of low energy protons. It may be that IceCube has measured this flux [216]. A consistent result across CR experiments for this energy range is that the CR flux is proton dominated.

At ultra high energies $\sim 10^{20}$ eV, there is a measurable predicted neutrino flux at $\sim 10^{18}$ eV if the CR composition is proton, as indicated by TA and HiRes. Conversely, if the composition is primarily heavy nuclei, as indicated by the Pierre Auger Observatory, the expected neutrino flux would not be measurable by current detectors. The HiRes, TA and Pierre Auger Observatory results are based on their *respective* calculations of X_{max} [6].

Neutrino primaries could manifest with almost horizontal showers having a 'young' shower waveform (see subsection 2.5.2 on page 46). Neutrino primaries could interact below the surface and SDs would then register the 'upward' early EAS. The Pierre Auger Observatory's search for a sizable flux of tau neutrinos, at ground or atmospheric, and their Monte Carlo derived E^{-2} spectrum has thus far found nothing [6]. HiRes, in its search for tau and electron neutrinos having horizontal or upgoing showers, has also found no neutrino candidates [126, 127].

TA reports that above $\sim 10^{18.2}$ eV, there is a dip in the observed energy spectrum at $\sim 10^{18.7}$ eV and a major flux suppression at $\sim 10^{19.7}$ eV. These features are in accordance with a CR energy spectrum which is predominantly of proton composition.

TA has also reported a $r = 20^{\circ}$ sized anisotropy at energies just beyond 57 EeV, in the direction of Ursa-Major. The anisotropy centre is 19° off the Super-Galactic plane and has a chance probability of 4σ. There are no candidate nearby sources [5]. An important positive result by the Pierre Auger Observatory, was a reported strong angular correlation ($< 3.1^{\circ}$) between CR events of E $>$ 60 EeV and AGN within 100 Mpc. The Pierre Auger collaboration maintains these primaries are protons because their expected magnetic deflections are minor over supposed coherent GMF of scales $\sim$ 1 kpc [106]. Composition measurements, however, are equivocal, and point proton sources do not invalidate heavy composition nuclei from cosmological sources.

2.2.4 GZK Astrophysical Models

A necessary, and hopefully convergent, deconstruction of the question of UHECRs and their composition, is to consider representative astrophysical models of source and propagation. Since a GZK limit or at least a strong flux suppression seems confirmed, we have an annulus, with an inner radius of our galaxy, of horizon distance possibilities. A proton composition around the

discontinuity, recorded by HiRes, the Pierre Auger Observatory and others, would have sources within a ∼30 Mpc radius, which includes isotropic source distributions and at its limit, a section of the supergalactic plane and an associated large-scale anisotropy. Medium to heavy composition, when accounting for path lengths of nuclei, extends this distance to ∼ 75 Mpc. Proton sources could be relatively spread with regards to isotropy. Medium to heavy nuclei, at the minimum set by their path lengths, would not exhibit anisotropies even if the sources were point-like, because of magnetic field deflections. Anisotropic manifestations would only be possible at the outer limits of the heavy nuclei horizon.

A paper [60], on the intergalactic propagation of nuclei, published in 2005, is interesting with regards to forecast extra-galactic composition. They project composition anisotropies from the extra-galactic point of view. The composition of the UHECRs is for protons, mixed and heavy nuclei.

The Hillas criteria-set for the acceleration and containment of protons at $\sim 10^{20}$ eV is stretched, but does solidly accommodate the acceleration and containment of nuclei at these energies. Iron, for example, reaches energies 26 times greater than protons. [60] look at a variety of nuclei primaries propagated through the cosmological infra-red background (CIB), a release of radiant energy in connection with the formation of large-scale-structures (LSS). As nuclei propagate, they fragment on the CIB. They transition down in charge and in their atomic number, Z. The CR spectrum as observed on Earth will then be different from the original source spectrum and dependent on the source distance. Three CIB models were tested - the outcomes not being so terribly different, they present most results using the Malkan-Stecker model of the CIB. They also consider extra-galactic magnetic fields (EGMFs). The composition at the source is contrasted with what we may expect to be recorded by CR detectors here.

Proton electron-positron pair-production ($p + \gamma_{CMB} \rightarrow p + e^- + e^+$) on the CMB is treated as a continuous energy loss, outstripping losses due to the Hubble expansion for energies $\sim (10^{18} - 10^{21})$ eV. It reaches a maximal depletion at $\sim 3\times10^{19}$ eV. This type of energy loss, i.e. $p+\gamma_{CMB} \rightarrow p+\pi^0$ and $p+\gamma_{CMB} \rightarrow n+\pi^+$, pertains to the proton induced cosmogenic neutrino spectrum. The

source injection spectrum is

$$\frac{dN}{dE} \propto E^{-\alpha}, E_P < (E_{max}/26) \tag{2.10}$$

and a constant co-moving density of sources yields

$$\frac{dN}{dV} \propto (1+z)^3 \tag{2.11}$$

where $z \ll 1$, given the propagation distance of UHECRs is typically constrained to (10-100) Mpc. Fig 2.7 is a display of energy loss lengths for protons using the 3 test models of the CIB.

As is evident, at higher energies the 3 models have converging results.

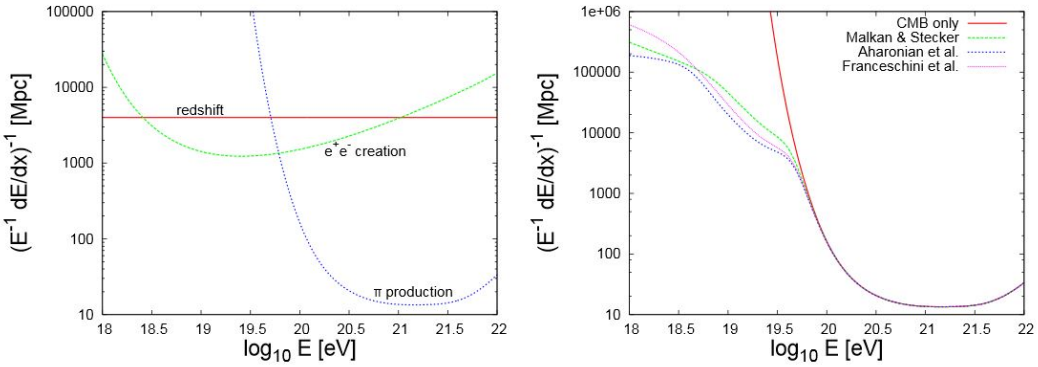

Figure 2.7: **a) All considered energy loss lengths for UHE protons. e^+, e^- production, π^0, π^+ and multi-pion production. Calculations incorporate "redshift" expansion.**
b) Energy-loss lengths for protons on the CMB and CIB by pion production for 3 CIB models. From [60].

The photo-disintegration as seen of heavy nuclei on the CMB and CIB are different from proton photo-disintegration. Heavy nuclei are more likely to interact with the CIB. Photo-disintegration is at its most effective when nuclei interact with photons of energies $\sim (0.01 - 0.1)$ eV. They are found in the CIB with wavelengths $\sim (10 - 100)\,\mu$m. Energy-length loss lengths through propagation of nuclei species in the Malkan-Stecker CIB are displayed below. In the GZK proton limit, $\sim 3 \times (10^{19} - 10^{20})$ eV, intermediate mass nuclei and proton nucleic all have a similar pronounced minimum. According to [60] heavy nuclei still present as heavy, but a recent paper [59] contends the iron and proton and mixed nuclei GZK limits all occur at about the same energy, thus the

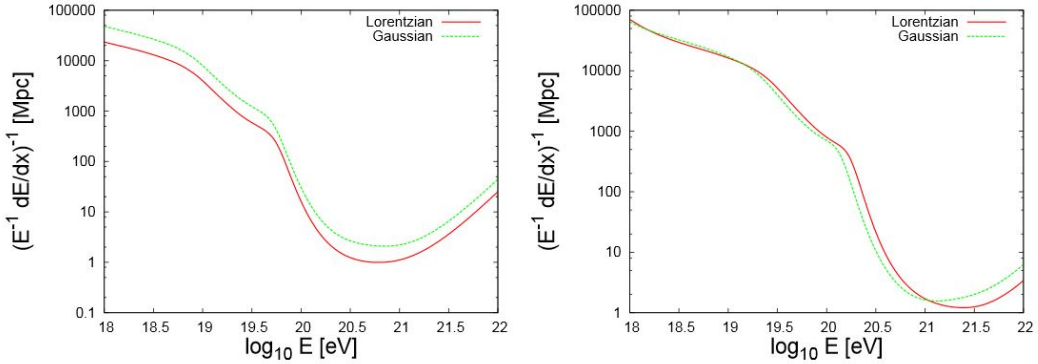

Figure 2.8: **a) All considered energy loss lengths for oxygen, with both Gaussian and Lorentzian cross sections.**
b) energy loss lengths for iron on the CMB and CIB by pion production for 3 CIB models. From [60].

suppression at the very end of the CR energy spectrum, if a GZK propagation effect, cannot yet be further attributed to light, heavy or mixed nuclei. [60] also incorporate the effect of EGMFs on propagation. They do not mention any GMF models. Their calculations demonstrate that such fields have negligible impact on light nuclei at UHE's. Heavy nuclei, however, can have deflections telescoping the source distance from, e.g. 50 Mpc, to distances 30% longer. In the context of energy-loss length, this length is reduced by 30%.

Their calculations take in an angular deflection, θ, such that

$$\theta(E,Z) \approx 0.8^{\circ}\left(\frac{10^{20}eV}{E}\right)\left(\frac{L}{10Mpc}\right)^{0.5}\left(\frac{L_c}{1Mpc}\right)^{0.5}\left(\frac{B}{1nG}\right)Z \tag{2.12}$$

for a particle travelling a distance, L, through random EGMFs with coherence lengths, L_c.
The effective increase in distance is

$$\frac{L_{eff}}{L}(E,Z) \approx 1+\frac{\theta^2}{2} \approx 1+0.065\left(\frac{10^{20}eV}{E}\right)^2\left(\frac{L}{10Mpc}\right)\left(\frac{L_c}{1Mpc}\right)\left(\frac{B}{1nG}\right)^2\left(\frac{Z}{26}\right)^2 \tag{2.13}$$

Figure 2.9 on the following page presents the effect a nano-Gauss EGMF would have on oxygen and iron primaries. We see that these EGMFs have an effect, switching on for oxygen at around 5×10^{19} eV. For iron around these energies, the effect is magnified, and at lower energies Figure 2.9

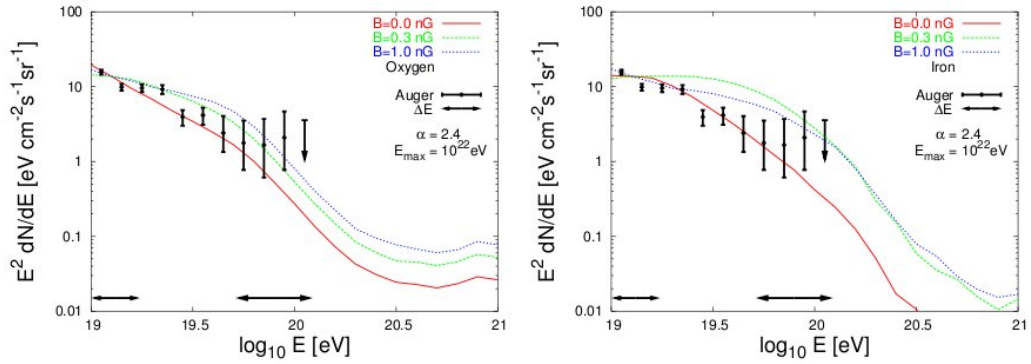

Figure 2.9: **spectra as seen on Earth for oxygen and iron propagated through EGMF** $\sim 10^{-9}$ **G. From [60].**

is unreliable.

[60] also consider mixed nuclei composition at source in the proportions of low energy galactic nuclei compared with a pure proton injection at source. There is little difference in the propagated spectra.

Another measurement [60] model, is $\langle X_{max} \rangle$ for helium, oxygen and iron nuclei at a cutoff energy E $\sim 10^{20.5}$ eV, and 2 different spectral indices. The data are taken from those extant at the time of publication. The relevant plots are shown in Figure 2.10. We can see, for an $E_{max} = 10^{22}$ eV at source, carbon and iron don't disintegrate significantly, and their spectrum at Earth is similar to that at source. There is an expected transition of heavy nuclei at high energies to light at low energies, largely independent of the nuclei proportions at source. At energies below $\sim 10^{19}$ eV, a predominantly proton/helium composition exists and an almost pure iron composition above.

The above is repeated for a higher cutoff energy $\sim 10^{22}$ eV, where heavy nuclei would substantially change their composition.

As is evident, all the source nuclei species end up looking similar at Earth. [60] note a pure oxygen source composition is best fitted to the X_{max} data. The pure iron data are also a reasonable fit for an injection spectrum with $\alpha = 2.0$.

As shown by [60], different CIB models have little effect on results (also different cross section models), but the choice of source spectra and the introduction of EGMFs do count.

[95] mention a pure iron injection spectrum, $\propto E^{-2.4}$, of uniformly distributed sources, with cos-

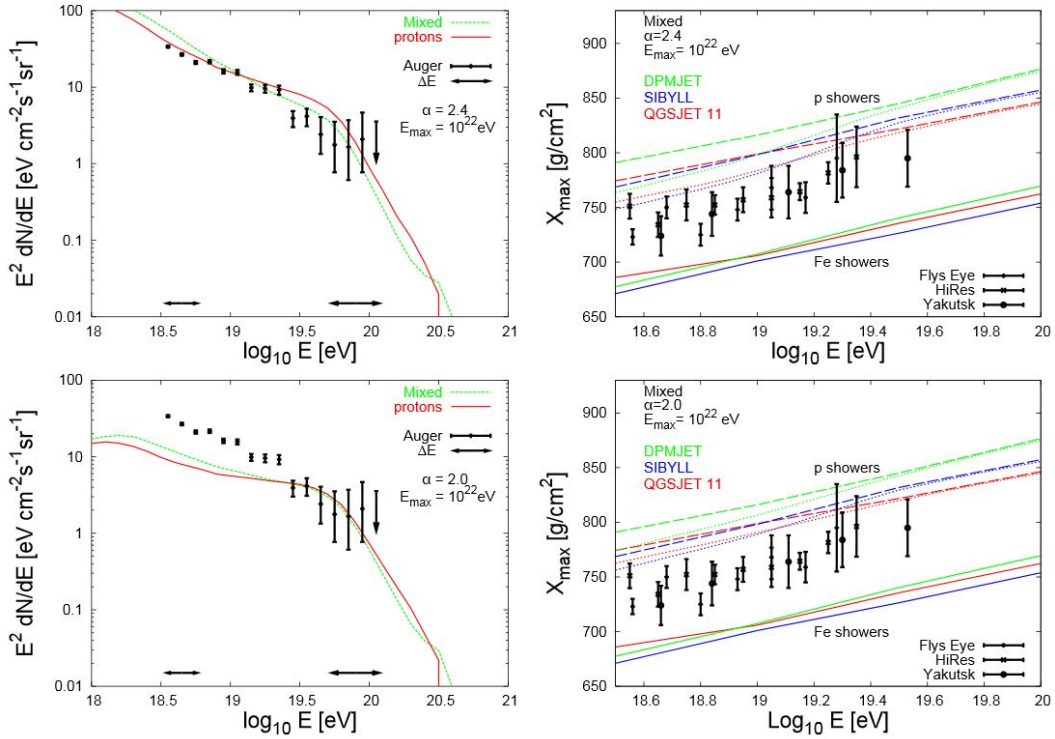

Figure 2.10: **Spectrum and $\langle X_{max} \rangle$ for E = $10^{20.5}$ and mixed nuclei From [60].**

mological evolution, $m = 0$, i.e. $(1 + z)^0$, is able to fit the Pierre Auger Observatory spectrum as seen in Figure 2.11. Heavy nuclei at source are thus viable.

Thus far, what can be said for overall experimental UHECR results is that composition just before the ankle appears light, and the extra-galactic proton pair-production dip fits nicely in this region. There is also a further general agreement regarding a heavier, later developing, higher energy second knee. This accord would seem to displace the presence of extra-galactic CRs to before the ankle.The composition just after the ankle is generally agreed to be light, and thereafter opinions diverge. The conclusions reached by the Pierre Auger Observatory and HiRes plus TA at the highest energies are contradictory. They agree on a GZK limit, but ascribe a different compositional follow-through.

A comprehensive compositional table of CRs from around the second knee onwards, will determine where extra-galactic CRs first make an appearance. Just as with the composition detail up to the first knee, composition will also narrow the choices of possible acceleration sites (see

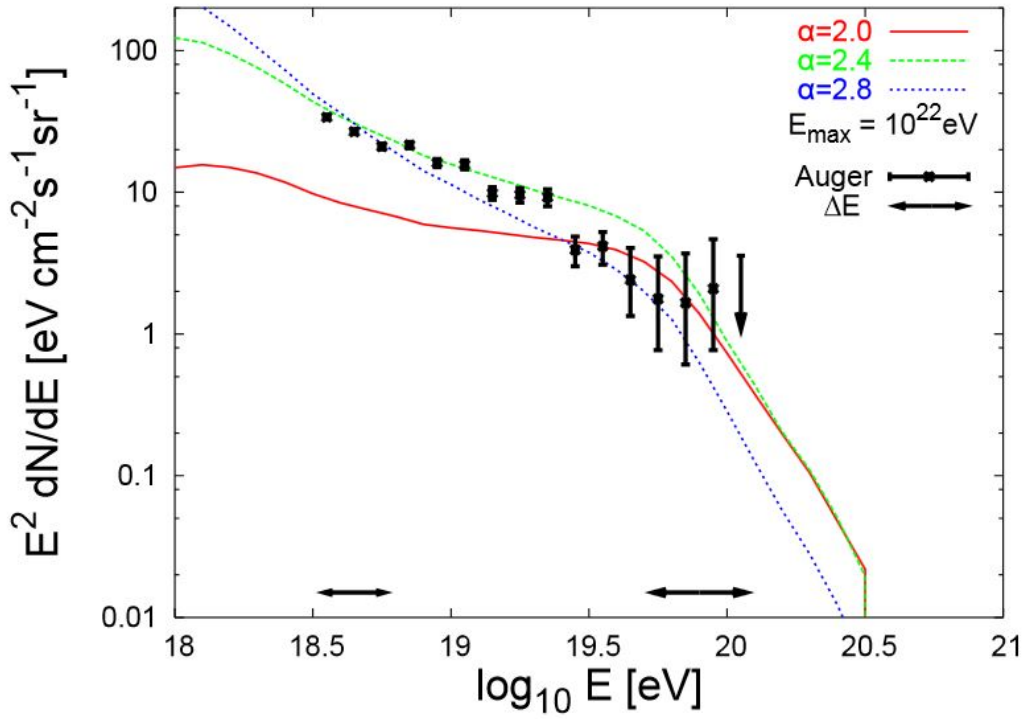

Figure 2.11: **predictions of different cosmological evolution scenarios injection spectrum $\propto E^{\alpha}$, for $(1+z)^{m}$, and measured Pierre Auger Observatory energy spectrum. From [95].**

section 2.2.1) - first of a galactic nature and then extra-galactic.

Galactic sources could be large-scale objects such as the galactic core or disk, or point-sources, such as SN remnants. In the extra-galactic domain, a large-scale source would be the supergalactic plane, and point-sources, for example, AGNs or radio loud galaxies. The spatial scale of sources of light nuclei should also be directly reflected in the scale of anisotropies.

2.3 Energy Spectrum

As has been discussed in the previous section on CR composition (see section 2.2), the thinning flux of primary particles at $E > 1 \times 10^{15}$ eV has redirected focus to the by-products of their refractive interaction within the atmosphere, namely EAS. Composition tables above this energy are patchy and not so well founded, due in part to stochastic uncertainties involving observables.

The simplicity of the all particle energy spectrum, however, is confirmed across experiments that span those cardinal points where its spectral slope becomes steeper (softer), or hardens to a more gradual gradient. Nonetheless, experiments are only able to draw qualitative conclusions about the energy breaks and the slopes that surround them. These features are interpreted differently [159, 85, 93] by theories that are able to reproduce current measurements. They will stand or fall by virtue of their predictions (or probably be modified). Agreements within narrowing details of these break points should also favour one theory above another - for instance the point of transition from galactic to extra-galactic CRs, is it the ankle or the dip?

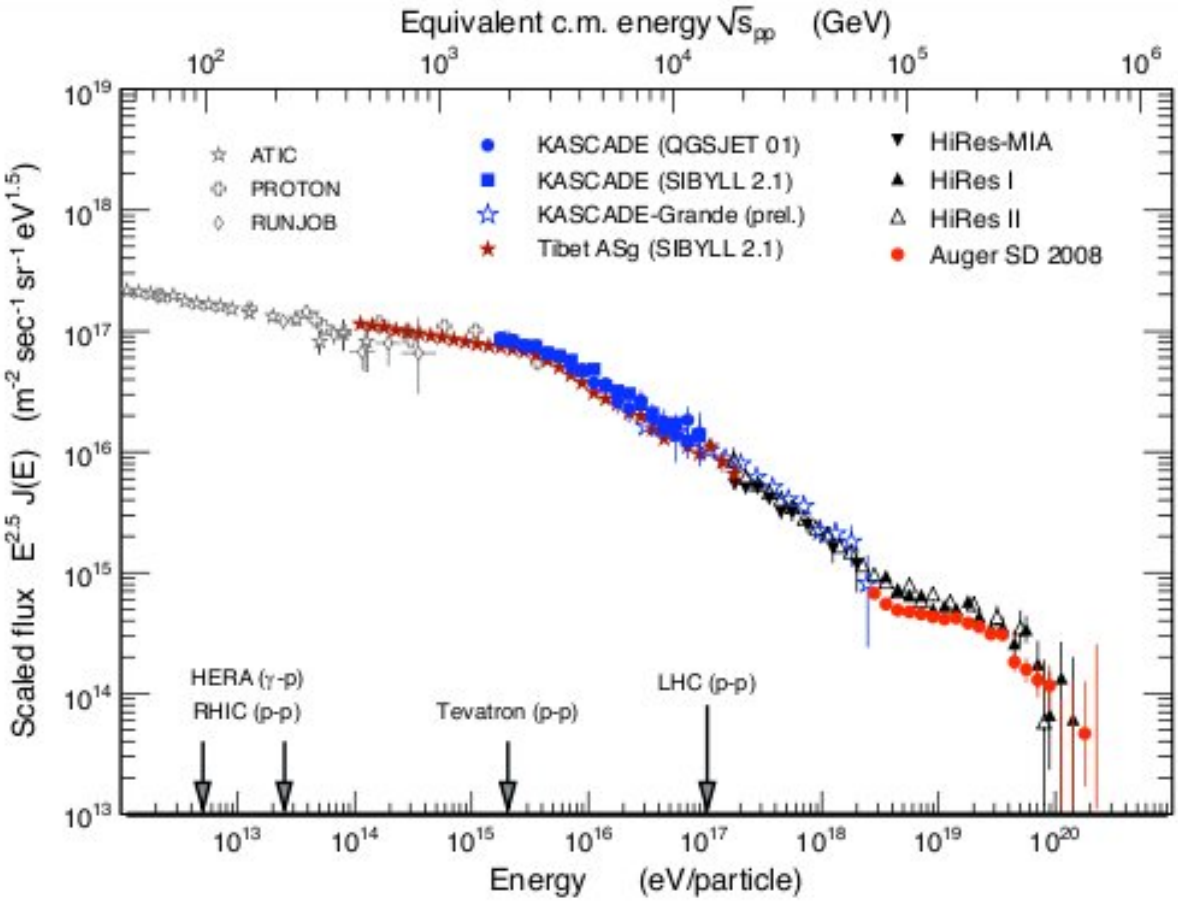

Figure 2.12: **All particle energy density spectrum. Direct measurements are from ATIC, PROTON and RUNJOB. Results obtained by EAS measurements and Hadron interacton models are seen between 10^{14} eV and $\sim 10^{17}$ eV. The UHECR measurements are $\sim 10^{18}$ eV and beyond. Note the projected CERN LHC energy at $\sim 10^{17}$ eV. From [94].**

The all-particle primary energy spectrum is displayed in Figure 2.12 on the previous page. The slight changes in structure are thrown into relief when the spectrum is magnified by factors close to spectral slope values. It is in the fine detail of these structures that we will see a partway resolution between conflicting structure explanations, and before that we must at least have an accurate energy spectrum.

The spectral breaks are roughly to be seen in the decades

the knee: $E_k \approx 10^{15}$ eV

the 2nd knee: $E_{2k} \approx 10^{17}$ eV

the ankle: $E_A \approx 10^{18}$ eV

the GZK limit or flux suppression: E_{GZK} predicted $\approx 10^{19.8}$ eV

An experiment's exposure constrains the energy range within which it is likely to record at statistically significant flux levels. UHECR array exposures (integrated aperture) are displayed in Figure 2.13 on the following page.

Disparities exist in the various spectral experimental normalizations where no satisfactory absolute energy scale can be obtained. Clearly, an experiment spanning as many of the energy spectrum features as possible is highly desirable since an idea of scale can be obtained. Better still, measures of the spectral slopes and the logarithmic difference between features displace reliance on such a scale.

Given the caveat that the energy density breakpoints 'float' according to experimental differences and limitations, *Bergman and Belz* in their 2007 review [86] , fitted their own calculations for energy breakpoints and slopes to data available. Each experiment referenced has an individual fit and then there is a global average for all experiments involved.

Their approach is useful, unifying energy spectra and slopes, particularly discarding reliance on absolute energy values through the use of slope ratios. A global fit to experimental data is predicated on stabilizing an energy point (generally around a discontinuity) to a common scaled value of flux, the original flux preferentially found in experiments with 2 feature exposures. Their work can no longer be considered recent and was incomplete. Significantly, in the fits around the

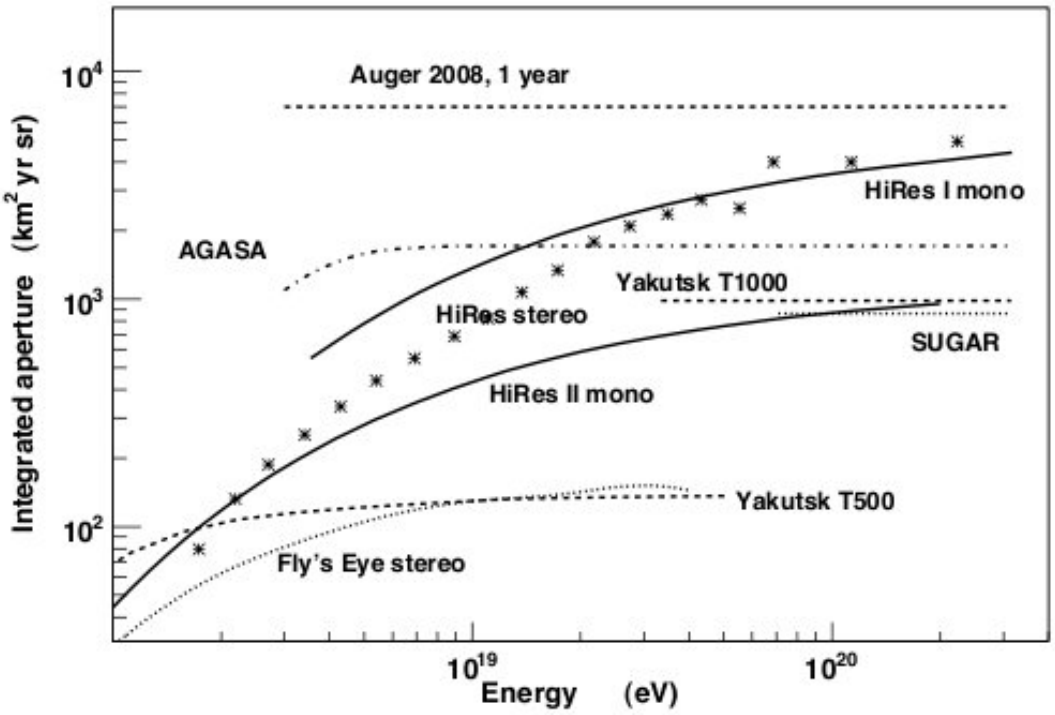

Figure 2.13: **Exposures of UHECR arrays. From [94].**

ankle, data above the GZK limit ($10^{19.8}$ eV) are not used, and so we come to a more current review by Watson [206] on data more recently obtained. By the time of publication Watson had access to post GZK energy results from more than one experiment and so his table listing energy spectral slopes is more complete. Watson's review concludes that the energy spectrum steepens at between $(3-4)\times10^{19}$ eV, which is certainly close to the GZK limit, but maintains that this suppression could also be caused by the physics of the sources themselves. The sources may indeed have reached their maximum acceleration limit and have a heavier composition than proton sources. Spectral shapes from proton or iron injections at source are not believed to be so different and we need more than mass composition to settle the vexed question of whether this flux suppression is because of the GZK limit or not. We need more precise measurements of primary particle directions at around these energies and so we probably need experiments with exposures even larger than the Pierre Auger Observatory exposures. There are plans to place a fluorescent detector on the International

	HiRes Mono	HiRes Stereo	Auger Combined	Telescope Array	Yakutsk Array	AGASA
Power Law before ankle, Region I	3.25 ± 0.01	3.31 ± 0.11	3.27 ± 0.02	3.33 ± 0.04	3.29 ± 0.17	3.21 ± 0.04
Power Law (intermediate), Region II	2.81 ± 0.03	2.74 ± 0.05	2.68 ± 0.01	2.68 ± 0.04	2.74 ± 0.20	2.69 ± 0.09
Power Law above suppression, Region III	5.1 ± 0.7	5.5 ± 1.8	4.2 ± 0.1	4.2 ± 0.7		
log(E/eV) (ankle)	18.65± 0.05	18.56± 0.06	18.61± 0.01	18.69± 0.05	19.01± 0.01	18.95± 0.05
log (E/eV) (suppression)	19.75± 0.04	19.76± 0.11	19.41± 0.03	19.68± 0.1		

Table 2.1: **Final fits to all experiments across energy features. From [206].**

Space Station [220] which will increase exposure and further refine the positions of CR sources, although the enhancement will not be useful for accessing the primary particle composition of CRs. The Pierre Auger Observatory analysis of particle composition using hybrid events will remain superior.

2.4 EAS Measurements

EAS measurements are taken when it becomes no longer feasible to deploy direct spacecraft measurements (see Figure 1.1 on page 5).

- **The Longitudinal Developement Profile (LDP):** An architecture of the shower stages passing through the atmosphere, where its size, the function $N(X)$ - the number of particles in a shower, at depth X g cm^{-2}, is fitted with 4 parameters. The number of particles at shower maximum, N_{max}: the depth at shower maximum, X_{max}: the "depth of the first interaction", X_0: and the elongation rate, D_{10}. The latter 2 parameters are free.

  The inter-relationship between parameters and $N(X)$ is described by the Gaisser-Hillas pro-

file where

$$N(X) = N_{max}\left(\frac{X - X_0}{X_{max} - X_0}\right)^{\frac{X_{max}-X_0}{D_{10}}} e^{-\frac{X-X_{max}}{D_{10}}} \tag{2.14}$$

The EAS's total energy is proportional to the integral of $N(X)$. Showers of different primary composition will have different structures. The most singular feature of an EAS is considered to be the depth at which a cascade reaches a maximum number of charged subsidiaries. There exists a strong correspondence between this depth and the nuclear primary's mass. The elongation rate is the rate of change of the average X_{max} with the logarithm of primary energy. [95] provides one of many good parameters for the interpretation of shower data.

- **Mean logarithimic mass, $\langle \ln A \rangle$, and average $X_{max}, \langle X_{max} \rangle$:** Derived from EAS measurements, consequent upon the stochastic nature of EAS, where uncertainties are important and indeed can dominate.

- **sets of fluctuations in X_{max}:** X_{max} fluctuates from shower to shower primary mass and its standard deviation, $\sigma(X_{max})$, is also used to qualify whether a shower composition is light, heavy or mixed. Fluctuations in X_{max} are thought to be smaller for showers initiated by iron nuclei than for protons (see sub-subsection 2.1.2.1).

- **$\mathbf{N_\mu/N_e}$:** The proportion of muons to electrons in a shower at observational level, the level being concomitant upon the required energy range, this is a further mass parameter. For energies around the knee ($10^{14} - 10^{15}$) eV, a high altitude setting is best, where the shower is closest to its maximum and fluctuations at a minimum.

- **The Lateral Distribution Function (LDF):** A composition sensitive parameter. Particle density is registered by surface detectors as a function of lateral distance from the shower axis. The shower core is thus determined, and E_0 (the primary particle energy) can be estimated, although not as reliably as with the LDP, because of a relationship dependent on extrapolated hadronic interaction models. The Pierre Auger Observatory's LDF is approximated by fitting to a Nishimura-Kamata-Greisen (NKG) structure function. Steep LDF's are from late developing showers.

To establish primary particle composition on a shower to shower footing, rather than averaging across showers, both the measurement of X_{max} and N_μ (or their proxies) are needed.

Because UHECRs are rare ($\sim$ 1 km^2 sr^{-1} century^{-1} for E $>$ 60 EeV), a large collection area is necessary to register the EAS of these primaries. The spacing between detectors is important in establishing an energy threshold and the accessible angular scale of CR directions. The CR detection rate is proportional to the collection area. For energy spectra around the knee, arrays of a few tens of thousand square metres are sufficient. For the GZK point, $\sim 4 \times 10^{19}$ eV, arrays of a few thousand square kms will contain a useful flux of EAS. An energy $> 10^{18}$ eV allows a detector spacing of $\sim$ 1 km[126]. Lower energies need spacings less than this. Parametric limitations placed on the resolution of large arrays (even the Pierre Auger Observatory) also contribute to only the average mass of primary particles or the mass group being derived. Figure 2.14, displays the exposure of different surface array detectors of the Pierre Auger Observatory.

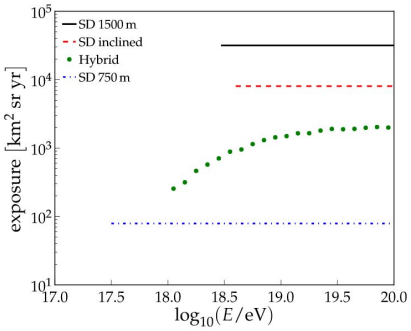

Figure 2.14: **The integrated exposure as a function of energy of the different surface detector arrays of the Pierre Auger Observatory detectors. The various SD array exposures are flat at full trigger efficiency. From [208].**

2.5 Pierre Auger Observatory Design

In the 1930's, Pierre Auger and contemporaries realized that EAS were the result of a primary particle's cascading interaction with the atmosphere. Pierre Auger made the startling calculation that some of these primary particles existed at the hitherto unthinkable energies of $\sim 10^{15}$ eV [194]. Subsequent detection and calculations of even higher primary energies had scientists planning for

experiments that would run longer and cover ever greater areas. In order for an array to be useful, its integrated exposure (in units of km^2 sr yr) must be tied to the energy of the primary and the subsequent spread of its EAS.

The Pierre Auger Observatory was first conceived of as an experiment that would last over 20 years, and be physically large enough to capture a statistically significant number of CRs above the EeV range. Now that this is close to being achieved at the southern site and solid numbers are being continually generated, the Pierre Auger Observatory is augmenting the original design with hardware that subtends its energy range down to $\sim 10^{17}$ eV. Results are mostly yet to be seen. This thesis is concerned with the Pierre Auger Observatory's prime objective to provide CR directional data above 10^{18} eV in statistically significant numbers, and the description here of its design, which was completed in 2008, only mentions in passing its latest additions. Inferences made in the conclusion incorporate the fairly standard observables in section 2.4 on page 39. Whatever can be gleaned from recent data is noted.

The Pierre Auger Observatory was designed to optimize results through its choice of location, hardware and software. Previous experience has been built on from many of the experiments involving CRs over the years [111]. The Observatory is positioned at $(35.1-35.5)^\circ$ **S**, $(69.0-69.6)^\circ$ **W**, in Mendoza province, Argentina. This latitude is deemed preferential in both hemispheres, and would give nearly uniform sky coverage when joined by a similar northern observatory. The site altitude is $\sim 1,400$ m above sea level, and its atmospheric transparency is clear for its UV telescopes.

Different properties of EAS can be registered by different instruments. When the same EAS is caught by more than one detector type, this enables a more accurate shower analysis to be obtained. In particular, an accurate primary particle energy measure, E_0, of these EAS is needed. At the very extreme energies, hadronic interaction models are the only recourse, and their reconstruction of E_0's currently don't converge inside model uncertainties.

The centrepiece of the Pierre Auger Observatory design is the use of 2 detector types, one on the ground and another in the air - the ground or Surface Detector Array (SD), and the Fluorescence Detector, (FD).

The SD is spread over a regularly gridded $3,000$ km^2 array. On each 1.5 km spaced grid point is located a water-Cherenkov station. There are 1,660 [114] of these surface detector stations. They respond to EAS particles - muons, electrons, positrons and photons, and from their ground densities reconstruct the showers LDF (see page 40). The Cherenkov light is tracked by the 3 photomultipliers (PMT) per tank and the signal is timed by GPS.

On the perimeter of the SD and overlooking its 'volume of air' is the FD, 27 [114] optical telescopes at four sites, each with a companion communications tower. Each has a field of view of $\sim 30^\circ$ in both azimuth and elevation. Six of the telescopes per building enable an azimuth coverage of 180°. The FD chases the UV fluorescent photons which atmospheric nitrogen emits as EAS particles pass through the atmosphere. Atmospheric Cherenkov light is also emitted for relativistic particles (chiefly e^-/e^+), except when the shower direction points to a telescope its percentage is less than that of the UV fluorescence. The light signal is treated by the subtraction of this Cherenkov component from the total light [4]. The shower development profile is constructed from this information. The LDF is expressed as a function of energy deposited in the SD stations by the shower particles. Experiments confirm that photon numbers recorded by the PMT's can be seen as related to this energy deposit.

Those EAS which are recorded by both detectors (SD and FD) are termed hybrid events. The virtual model-independence of FD energies allows for a cross-checking of energies, and this refinement can be extrapolated **up** with a degree of confidence, freeing higher energy calculations from interaction models when a shower is only recorded by the SD.

Local atmospheric conditions are factored into data and all detector components in both the FD and SD are checked every 10 minutes. The calibration of each SD station is continually being set by the local muon background and all signals are interpreted in terms of the reference unit, VEM. VEM being the average charge deposited by a downward vertical muon [109]. Such a muon of $\sim$ GeV will deposit ~ 240 MeV into the tank.

The FD's are calibrated such that incident photons on the PMT cameras and their ADC measures are translated to photon counts at the telescope aperture. Regular calibration is conducted twice every night-time run to check the status of every camera pixel.

[4] [121] have developed a method where this subtraction is unnecessary

Data from the FD telescopes are relayed through their communications tower and SD data is sent via a wireless connection to one of the 8 base stations housed in each communications tower. All transmitted data ends up at the Central Data Aquisition System (CDAS). It is CDAS that sorts data to determine what data will be reconstructed into a description of an EAS. CDAS also determines which FD data are co-incident with SD data and hybrid events are identified.

A birds-eye view of the Pierre Auger Observatory is displayed in Fig 2.15 on the following page. The Central Laser Facility (CLF) enables comparisons between SD events and FD events, to refine the geometry of reconstructed hybrid events. Whilst an array's exposure is determined by its surface area and detector arrangement, its aperture is determined by its surface area and zenith angle dependence. The water-Cherenkov stations, being $\sim$ 1.2 m high, enable the array's aperture to take-in nearly horizontal showers [126].

Because data-taking commenced in 2004 when the SD was still being constructed and only half of the FD telescopes deployed, data from the period between 2004 - 2008 have their aperture and exposure calculations adjusted. This process continues as SD stations have a high, but not 100%, on time. The collection of legitimate data is considered to have started in January 2004. The data used in this thesis begin at this date. The recorded EAS meet certain criteria before they are even considered as data. Some conditions are met by the SD stations, some by the FD, and some by both.

2.5.1 Conditions

- Zenith angles of events are $\leq 60°$. This zenith angle cut reduces the number of events by about 30% [13]. EAS with zenith angles of events $> 60°$ are termed inclined EAS (see Figure 2.14 on page 41 for integrated exposures), and are the purview of Pierre Auger Observatory enhancements. Photon primary events are searched for in zenith angles of $> 35°$.

- The estimate of primary energy is made using the SD signal at $1,000$ m from the shower core, $S(1000)$.

- The angular resolution is defined as that angle containing at least 68% of an events recon-

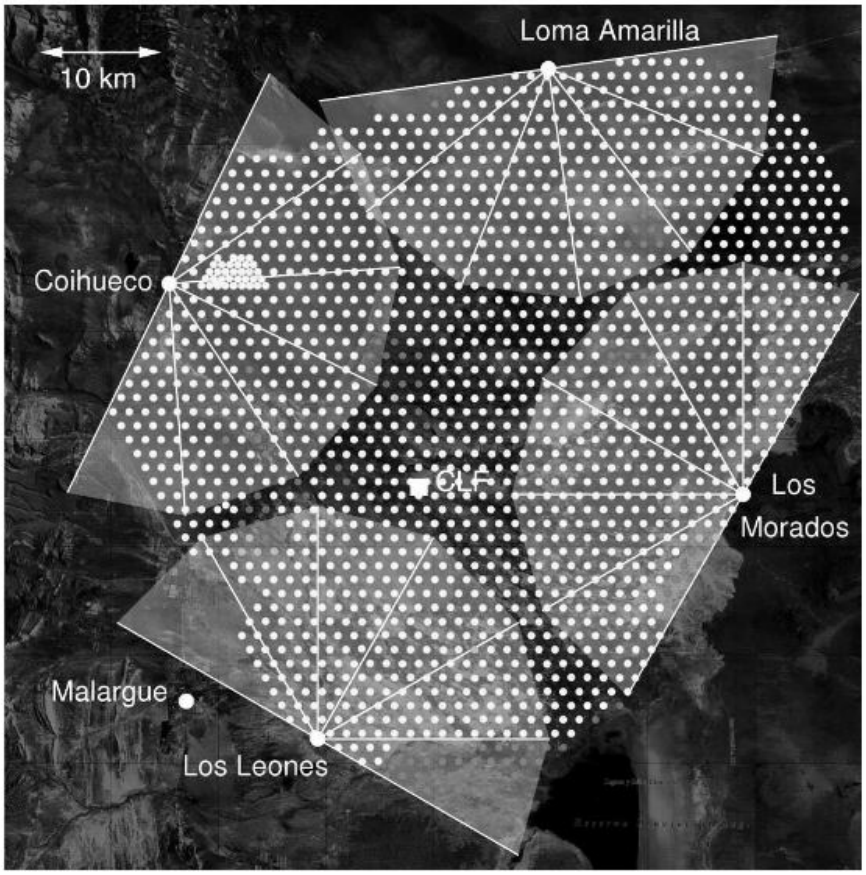

Figure 2.15: **Aerial view of the Pierre Auger Observatory in March 2009. The 4 large light grey segments represent the field of view of 24 fluorescence telescopes and their housing locations are indicated as being at Los Leones, Coihueco, Loma Amarilla and Los Morados. The Central Laser Facility (CLF) used for hybrid timing analysis and to monitor atmospheric properties is indicated by the white square. The SDs are shown as dots. The lower energy detection ($\sim 10^{17-17.5}$ eV) infill array, AMIGA, is positioned at Coihueco. From [108].**

structed shower core.

- The event angular resolution is cross-checked if there is simultaneous fluorescence detection.

- The event angular resolution accuracy is $< 0.9°$, for events that trigger ≥ 6 SD stations [109].

- Hybrid events are established with data from a FD plus at least 1 co-incident SD station.

2.5.2 Some SD Details

The SD runs continually and hence has uniform coverage in **RA**. It is largely unaffected by weather. The detector grid spacing yields a precision of $\sim (1-3)^\circ$ in direction estimates - comparable with the expected GMF deflection [126] for protons, being of the order of one magnitude higher for iron nuclei [123]. A shower's attenuation is zenith dependent. At the vertical and nearly horizontal extremes, showers are termed 'young' and 'old'. SD's find that young shower waveforms cover periods of $\sim \mu$s whilst old showers have short $\sim$ 100 ns waveforms. These old showers are chiefly muonic. Horizontal showers exibiting young waveforms would have the signatures of neutrino primaries [126] Showers with zenith angles of $\leq 60^\circ$ have an axial symmetry of the shower profile in the transverse plane [126] (which is geometrically convenient for a shower reconstruction of the LDF from which the E_0 is to be gauged). Larger angles require an allowance for the inclusion of the Earth's geomagnetic field (muons begin to be appreciably deflected) and a more involved estimate of the shower core position . Also, the shower electromagnetic (EM) (photons, electrons and positrons) component tends to be absorbed into the atmosphere. All data used in this thesis have this zenith cut ($\leq 60^\circ$).

For the Pierre Auger Observatory, a shower particle density measurement at 1000 m, (S(1000)), minimizes the sum of the fluctuations from shower to shower, and fluctuations in particle counts. $S(1000)$ is dependent on the depth in the atmosphere, which changes as $\sec(\theta)$ for a zenith angle θ [95]. A measure of the density of a shower's particle content at this point can be made proportional to the primary's energy. $S(1000)$ is measured in units of VEM and then used with an allowance for the signal attenuation to a standard zenith angle of $38^\circ - S_{38^\circ}$ [141]. A calibration obtained from hybrid events is then used to determine E_0 in terms of S(1000). The systematic uncertainty of this calibration increases with energy to $\sim 15\%$ for energies around 10^{20} eV.

The water-Cherenkov tanks measure the secondary particles' total incident energy, their direction, and how the type of particle numbers vary with time through the shower front. The nature of the shower development is linked to the primary mass. Unlike EM particles, muons arrive at the detectors mostly unchanged from their time of production along the shower axis and this enables X_{max} to be back projected. The detectors are particularly sensitive to muons (hence the

shower signal reference unit, VEM). A particle's Cherenkov light emission is proportional to its path length in water, and the amount of Cherenkov light deposited is primary mass sensitive. The slope of the LDF is also related to X_{max} - the muon LDF is flatter than the electron one and muon numbers dominate at larger distances from the shower core. This distinction in shape allows an X_{max} estimation using the muon core distances. Thus the LDF of the muons estimates composition and E_0.

The SD data are sorted through a series of triggers. The first two triggers, T1 and T2, are local to their detector and must satisfy a certain measure of VEM signal and other conditions. If the T2 requirements are met, the detector signals are sent to their assigned base station which passes data onto CDAS. This T2 data level is then upgraded to T3 data if the triggered stations occupy certain configuration patterns within the entire SD array. A T3 figure is the minimum level for EAS reconstruction. The most desirable trigger level for shower reconstruction is T5. The station configuration must then have the station with the highest signal being surrounded by 5 active stations (T2).

Further SD details being

- On time > 87%.
- Shower direction obtained by fits of a parabolic curvature function to the timing of T1 and T2 triggers that have passed the T3 filter. The angular resolution for 3 neighbouring stations is $\sim 2.2^\circ$, improving as more stations register an event [112].
- When E > 3 EeV, the efficiency for a trigger chain is 100%.
- When E > 10 EeV, the angular resolution is better than 1°.
- Once 100% efficiency is reached, the array's exposure becomes independent of energy and is $\propto \cos(\theta)$ (where θ is the zenith angle), and the exposure uncertainty is $\sim$ 3% [109].
- Energy resolution for E better than 15% [110].

To establish the primary particle's composition on a shower to shower footing, rather than across showers, the measurements of X_{max}, S(1000) and N_μ are needed.

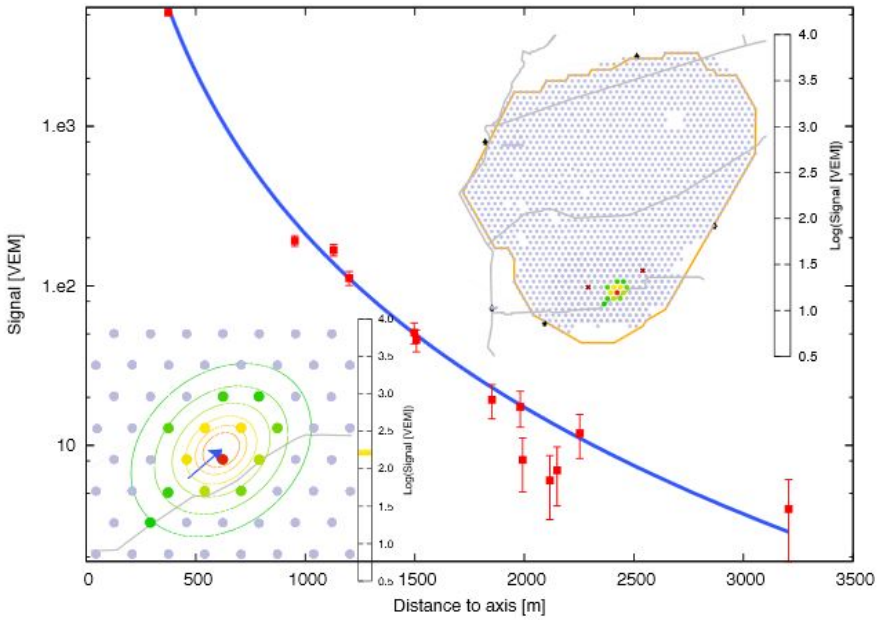

Figure 2.16: **Schematic of the 'footprint' of an EAS. The upper right insert depicts the Pierre Auger Observatory SD array and the proportion covered by an EAS. The lower left shows the EAS footprint as particle density contours. This density is a function of lateral distance from the shower axis and thus determines the shower's core (red dot). The blue curve illustrates the shower's LDF and measurements (red squares) of detector stations distances from the core. The grey dots are the SDs. From [126].**

2.5.3 Some FD Details

Each FD telescope is filtered to admit photons only in the nitrogen fluorescence UV bandpass. It is equipped with circular diaphragm (2.2 m diameter) positioned at the centre of curvature of a spherical mirror. This defines the aperture of the Schmidt optics for imaging. A 440 PMT hexagonal pixel camera is also mounted at the focal point of each 13 m^2 mirror. Light collection is maximized, and the pointing direction of each pixel, which covers 1.5° of night sky, is within 0.1°.

The performance of FDs is constrained by the local weather. Aside from the requisite minimally lit night-time runs, lightning strikes must be excluded from the data set, and a constant monitoring of atmospheric fluctuations which interfere with EAS propogation is maintained. Shower reconstruction is adjusted accordingly.

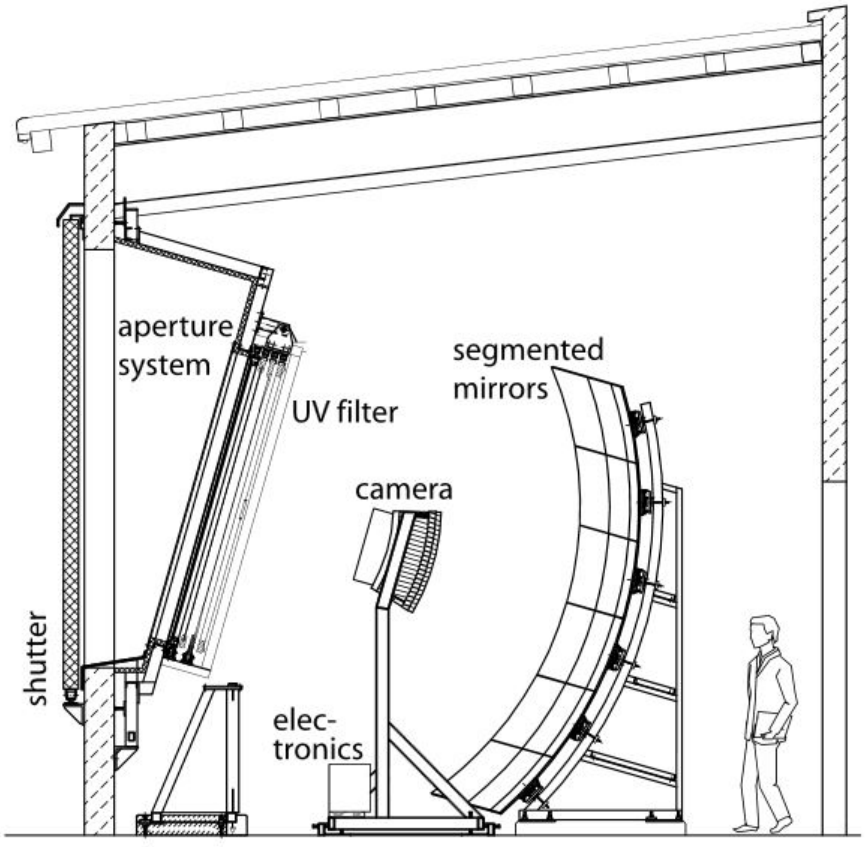

Figure 2.17: **An illustration of a single FD telescope and the arrangement of its main components. From [67].**

The FD's provide a direct longitudinal profile measurement of each shower and provide the vantage point of a visible X_{max} - thus the shower energy estimation of E_0 is relatively straightforward. The UV nitrogen fluorescence emission is isotropic, a decided edge over other EAS observables (excepting microwave radiation), in that it can be detected at 10+ km distances regardless of the shower axis[5] - other shower observables such as Cherenkov light follow the shower axis. The fluorescence strength is proportional to the electron number in the EM cascade and its development timeline also marks the timeline of the EM shower part, which is sensitive to the primary mass composition.

FD data also pass through a triggering protocol. Embedded in the telescopes electronics, a first

[5]has been seen > 20 km from the axis for some UHECRs[126]

level trigger board (FLT) initiates a pixel's trigger when a threshold rate of > 100 Hz is measured. A second level trigger board (SLT) then processes pixel configurations. Those pixels that can be read as straight tracks are passed to an algorithm, the third level trigger (TLT). This discards lightning strike data and chance alignments. Data are then collected from all telescopes at the site and reconstructed to find where, and the time 0, ground impact occurred. When this is complete, a hybrid T3 trigger is sent to CDAS. CDAS responds by initiating data collection from the SD that may correlate with the FD reconstructed parameters. From this we see that data from SD T2 levels can potentially be incorporated into what results as an hybrid event.

Further FD details being

- On time $\sim$ 13% [110].
- Shower geometry –
    1. The shower detector plane (SDP) is fitted to triggerd pixels to within 0.1°.
    2. The shower axis is found from fitting the time sequence of triggers - this can have large uncertainties.
- The X_{max} must be in the field of view - accuracy once geometry is decided is $\sim$ 20 g/cm^2.
- The energy deposit profile calculation is from the slant depth of the shower.
- The absolute energy scale, E_0, has an overall systematic uncertainty of 22% [109]. This is an accumulation of systematic uncertainties that include –
    1. An energy deposit profile calculation $\sim$ 10%.
    2. A fluorescence yield of $\sim$ 14%.
    3. A FD calibration of $\sim$ 9%.
- All events > 10^{19} eV trigger at least 1 FD (monocular), i.e. the array has full detection efficiency at this energy.

- All "GZK" events trigger at least 2 FD (stereo). The intersection of their respective SDP's increases the accuracy of the shower geometry and defines the shower axis orientation in 3-D to within 1°.
- The Cherenkov light component in the shower < 50%.
- The energy estimator resolution is ~ 15% [109].

2.5.4 Some Hybrid Event Details

The derived E_0 from cross-checking of data is superior to that of the estimated E_0 from either detector type alone. This allows for a more reliable E_0 at any energies (see 2.6.1 on the following page), when an event is only registered by one detector or the other. Further hybrid details are:

- A shower is detected by the FD and at least 1 SD station. The coincidence of the event timing between the SD and FD is fixed by the CLF.
- The Cherenkov light from the FD must be < 50% of the total.
- The reconstructed shower core must be within 1,500 m of the station used in reconstruction.
- The fitting of the Gaisser-Hillas (see equation2.14) profile condition to a reconstructed LDP must satisfy $\chi^2/Ndof < 2.5$.
- The uncertainty of the energy reconstruction must be < 20%.
- To establish hybrid events –
    1. From the FD: an estimated time of arrival of shower light front plus the geometry of a Shower Detector Plane.
    2. From the SD: all the space-time information co-incident with the FD event.
    3. The flux composition is considered as mixed [110].

Every admitted FD event with E > 1 EeV has at least 1 SD station triggering. This is regardless of the incoming direction and particle mass [110] and the energy resolution above 10^{18} eV is better than 6%. The statistics for hybrid events are limited by the small FD on-time.

2.6 Pierre Auger Observatory Enhancements:

Now that the Pierre Auger Observatory's main objective is complete and there are reasonable data numbers for CR events $> 10^{18}$ eV, its purview has been broadened. Installations sensitive around knee-to-ankle energies, that region where hadronic models fail, and where the galactic to extra-galactic transition is undecided, are operating or being constructed. Three such lower energy detection arrays are mentioned below. The first two are variations on the Pierre Auger Observatory Hybrid design theme, and the last is the introduction of a new technique for EAS detection.

2.6.1 AMIGA

The acronym AMIGA stands for Pierre Auger Observatory Muons and Infill for the Ground Array. Two sections of the SD array have been/are to be modified to decrease the spacings between water-Cherenkov tanks. One section becomes fully efficient at $\sim 10^{17}$ eV, and the other becomes fully efficient at $\sim 10^{17.5}$ eV. Each tank has a 30 m^2 scintillator buried nearby at a depth of ~ 2.5 m. The muon detection capabilities of these scintillators provide a counterpoint composition measurement to the depth at shower maximum, X_{max}, measurement of the FD's. In Figure 2.15 on page 45 we see the positioning of AMIGA at Coihueco.

2.6.2 HEAT

The acronym HEAT stands for High Elevation Pierre Auger Observatory Telescopes. A three-telescope group directed above AMIGA provides hybrid measurement capabilities at $\sim 10^{17}$ eV. These telecopes are equipped with pivot mechanisms extending the telescope's field of view to elevations of $\sim 60^\circ$ (see section 2.5). The X_{mass} of showers at lower energies is higher, and a higher elevation enables a more likely secure detection of such early developing showers.

2.6.3 AERA

The acronym AERA stands for Pierre Auger Observatory Engineering Radio Array. This radio array complements AMIGA and HEAT and is nearby. There are 21 stations of antennae sensitive to the (27 − 84) MHz range. A magnetic field's acceleration of charged particles results in radio frequency emissions, and so the Earth's geomagnetic field is utilized in studies of radio emission for energies $\gtrsim 10^{17}$ eV. The radio array, LOPES, at KASCADE-Grande is producing results through the correlation of radio pulse height with primary mass groups, and the co-incident EAS of KASCADE-Grande [125].

2.6.4 Pierre Auger Observatory Prime Upgrade

R&D efforts are now focused on mass composition discrimination at the highest energies where the CR flux suppression is evident. Deployment of 2 different ground array detectors placed one above the other is being considered. The separation of the EM, N_e, and muonic, N_μ, shower components so successfully applied by KASCADE and KASCADE-Grande in determining primary mass composition, could be achieved with one detector, e.g. a scintillator, placed over a water-Cherenkov tank. The top detector responding to the EM shower component and the bottom detector to the muonic shower component. Such shower component separations would also enable the detection of photons and neutrinos at the high energy scale of the CR spectrum [8].

2.7 The Pierre Auger Observatory Prescription:

Capitalizing on the previous experience of experiments involved in anisotropy searches, from the very beginning, the Pierre Auger Observatory collaboration set out to avoid the statistical pitfalls of a so called *a posteriori* analysis [13, 14].

When data are gathered and analysed many times, there is the danger, in an increasing number of searches performed on a finite data set, of finding something apparently statistically positive. Searches that are *ad hoc*, not predefined, on the same collected data, potentially over and over again, are deemed *a posteriori*, and statistical penalties are uncontrolled.

To reduce this danger the Pierre Auger Observatory collaboration predetermined some of the parameters of directional searches. Regions of interest were defined, e.g. the supergalactic plane, the galactic centre (GC) - regions that are intrinsically anisotropic. Events with zenith angles greater than 60° were initially excluded[6] , and the minimum energy of events considered was 1 EeV. The searches were in specified areas with a set radial angle around a potential source direction, or in parts of the sky where there were potential point-sources and those regions were assigned a probability distribution value.

An anisotropy is considered as such - when the chance probability of occurence is less than 0.001. That is, more or less, for a significance of 3σ, where we have a probability of 0.0013. The probability of 0.001 is global, and partitioned among the prescibed regions of interest. For the initial data set, collected from 16th May 2003 to the point of 500 working SDs, the Pierre Auger Observatory collaboration hoped for a data set of $\sim 10^4$ events > 1 EeV, with the global probability to be summed over set regions of interest.

Even though we are far beyond this initial data set, any further data sets are treated in much the same manner. A few instances of regions of interest and their assigned probabilities are listed below. This is to be considered a sample prescription.

- The GC showing an excess with a chance probility of less than 0.0035. The entire data set was to be used, the GC at (**RA,dec**) co-ordinates of (**17 h 42 m**,$-$**29.0**°) and a 15° radius.

- The GC having a point-source chance probability less than 0.00025. Data, E < 18.5 EeV; the GC at galactic (**l, b**) co-ordinates of (**0**, **0**) - the approximate site of the AGASA/SUGAR excess [195].

- The Centaurus A (Cen A) centre with a chance probility of less than 0.0005. Data $E < 19.5$ EeV; the (**RA,dec**) co-ordinates of Cen A, (**13 h 25 m 27.6 s**, $-$**43°01′09″**) or [$-50.5°, 19.4°$] [198], and a 5° radius.

  The Cen A cluster is our nearest active galaxy, comprising the nearest part of the Super Galactic Plane to us, at a distance of $\sim$ 3.4 Mpc. The 5° radius is intended to take in the Cen A radio halo and allow for the effect of the EGMF.

[6]The data set used in this work stays within this exclusion.

2.7.1 Shuffled Data Sets

Simulated data sets within the researchers own predetermined energy ranges, were to be assembled from Pierre Auger Observatory data, with the Pierre Auger Observatory prescribed presumption of event isotropy within zenith angle ranges. For example, the zenith angle data used in this research which has been retrieved from zenith angles, $\theta < 60°$, might be 'binned' into $10°$ intervals with an assumption of an isotropic distribution of events within each zenith interval. The generation of simulated data sets is based on real data events - each Pierre Auger Observatory event has its **RA**, **dec**, galactic **l**, galactic **b** co-ordinates listed together with, amongst other associated observables, the event's energy, its local sidereal time (lst), azimuth angle (az), and zenith angle value (zn) and EeV energy. The simulated data events have their co-ordinate directions recovered from the shuffled lst, az and zn values of original Pierre Auger Observatory events that lie within each zenith angle binning for $\theta < 60°$, for each energy range. A prescription suggestion was that for small data sets, < 3000 events, the researcher includes events taken from lower energy ranges to make up the data numbers, to ensure the shuffling of zenith data maintains an isotropic distribution within the preset zenith bin ranges.

However, for this research method, we applied no zenith binning or sampling of event directions from lower energy ranges. The three energy ranges selected were considered to be just below or just above the GZK proton energy limit and the energy range data set numbers *were* small, below 300 events, in fact. Uncertainties surrounding the energy and primary particle types at which an extra-galactic transition of primaries within 200 Mpc might take place led to the decision not to sample from lower galactic primary particle energies which might contaminate the reconstructed primary particle directions for those event directions we *were* sure exceeded the GZK energy limit. Also, the decision not to bin zenith angles is upheld in the work of [112]. A more complete discussion about the shuffling method used is to be found in section 6.5 on page 166.

A posteriori searches are also to be conducted, but without the assurance of significance levels. These types of searches are often the prescriptive basis of future *a priori* efforts. The analysis of the Pierre Auger Observatory data in this thesis is considered *a posteriori*.

Chapter 3

Galactic Magnetic Field Models

The final component in the observational index used to qualify CRs is direction. A preference in the pointing direction of CRs arriving at detectors is linked to anisotropies in the dispersal of a supposedly isotropic population of CRs. That the flux of CRs is broadly isotropic is a consequence of a sequence of interactions with the ISM and G/EGMF's. Whatever their source, CRs lose memory of their beginnings - unless their source and/or acceleration process dwarfs any of the other processes inhibiting or hastening their journey here[1].

The dearth of evidence supporting top-down acceleration models (see page 28) has one returning to the least complicated model of CRs being massively accelerated at source, and their diffusive propagation to CR detectors.

Providing there are no GMF entanglements, galactic anisotropies from point-sources such as supernovae would be fairly localized and related to their galactic source. A sidereal diurnal anisotropy at high energies of 0.03% resulting from the ~21 km sec^{-1} motion of our solar system with respect to the ISM would be observed [64]. Large-scale sources such as the galactic disk and galactic bulge, can be reasonably expected to have a spread of CR directions on a similar scale.

Extra-Galactic anisotropies may be a bit more complicated. A point-source, even a light proton one, within a large-scale-structure (LSS), may exhibit a magnetic lensing effect, with clustered energy ordered CR directions (tails), and a single point-source with a complex composition within

[1] This truism is founded on the assumption of CR acceleration at source. If CRs encounter conditions such that they undergo further very strong acceleration processes or deflections between the source and receiver - by whatever means, one would confront a more complicated CR diffusion scenario.

a LSS may have several different pointing directions with tails. An extra-galactic *Compton-Getting effect* should result from the relative motion of our galaxy with respect to any LSS of diffuse sources (e.g. the super-galactic plane). An anisotropy would be expected to be manifest at such an extra-galactic transition [27, 64][2]. Large-scale regular magnetic fields may result in patterns of directions, possibly across energy ranges.

Diffusive propagation models are inadequate when one supposes an extra-galactic CR introduction at ankle energies. Either another galactic source enables a smoothly declining continuation of supply, or extra-galactic sources begin to feature somewhere in the region after the second knee. Slight spectral structures do interrupt the overall energy spectrum in this region, and so the galactic/extra-galactic debate continues.

In whichever context CRs and anisotropies are discussed, magnetic fields are very important, and only light and heavy nuclei in the extra-galactic regime are expected to survive [207, 93].

On the cosmic scale, gravitation is seen as an overarching, governing principle, but in the astrophysical sense, where one discusses separate large-scale-structure phenomenology, magnetic fields are largely responsible for regulating the dynamical behaviour of systems. The GMF, the ISM and CRs are the prevailing interdependent components establishing the quasi-equilibrium of the galactic system. The GMF distributes CR densities, the high conductivity of the turbulent ISM decouples the Lorentz force on particles, and CR densities define the nature and shapes of the GMF. Extra-galactic propagation scenarios also invoke GMF models. These EGCRs must overcome their source's magnetic fields, travel through intergalactic space and perhaps EGMF's, and then run the gauntlet of our GMFs.

First, the Lorentz force on particles moving through electric and magnetic fields is discussed and its simplification with respect to the ISM, and galactic anisotropies. Recent adaptations of the Hillas criterion for plausible extra-galactic sources then precedes mention of GMF measurement techniques, generic features of representative GMF models, plus a short synopsis about a recent GMF model. The chapter is concluded with mention of the release of results from the S-Band Polarization All Sky Survey (S-PASS).

[2] [64] also predict an extra-galactic *Compton-Getting effect* relative from motion to the CMB.

3.1 Lorentz Force

The Lorentz force on a particle of charge, Z, moving with velocity, **v**, through an electric field, **E**, and a magnetic field, **B**, is defined as

$$\mathbf{F}_L = Ze(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{3.1}$$

where e is the fundamental charge.

Most CR particles are charged and Eqn. 3.1 can be considered as a rudimentary propagation equation, excepting the conductive ISM allows no charge separation, and hence nuclei do not move through any large electric field structure, the large GMF is static[3] and Eqn. 3.1 is reduced to

$$\mathbf{F}_L = Ze(\mathbf{v} \times \mathbf{B}) \tag{3.2}$$

i.e.

$$\mathbf{F}_L = \frac{d\mathbf{p}}{dt} = Ze(\mathbf{v} \times \mathbf{B}) \tag{3.3}$$

The time integral of the momentum of the CR nucleus provides a description of its propagation. In the allowed range of diffusive propagation, there are tens of pc length irregularities in the GMF, random subsets of the field termed 'turbulent fields', and these irregularities with a characteristic 'coherence length', L_c ($\sim$ 100pc [23] and see Eqn. 2.12), of the turbulent field deflect the passage of the cosmic rays, and their propagation is isotropized.

When charged particles interact with a magnetic field, their linear propagation is acted upon by the centripetal force of the magnetic field, such that

$$|\mathbf{F}_L| \propto |\mathbf{F}_c| \tag{3.4}$$

and the radius of curvature, r_g, of the cosmic ray propagation is contained within

[3]Not really, the quasi-equilibrium state $\Rightarrow$ activity, but this activity is on time scales $\sim 3 \times 10^7$ years, or 10^{-1} the galaxy rotational period [187] pg 796.

$$Ze|(\mathbf{v} \times \mathbf{B})| = \frac{\gamma m |\mathbf{v}_\perp|^2}{r_g} \tag{3.5}$$

or rather

$$Ze|\mathbf{v}_\perp||\mathbf{B}| = \frac{\gamma m |\mathbf{v}_\perp|^2}{r_g} \tag{3.6}$$

for CRs of rest mass, m, and Lorentz factor, γ. Thus

$$r_g = Ze|\mathbf{v} \times \mathbf{B}| = \frac{\gamma m |\mathbf{v}_\perp|}{Ze|\mathbf{B}|} \tag{3.7}$$

A nucleus travelling *not* at right angles, $\mathbf{v}_\perp$, to the magnetic field, but at an angle, δ, has the perpendicular component of $|\mathbf{v}_\perp| = \mathbf{v}\sin(\delta)$. Given that the relativistic momentum for the CR, $\gamma m|\mathbf{v}| \approx \frac{E}{c}$, Eqn. 3.7 becomes

$$r_g = \frac{E}{Zec|B|}\sin(\delta) \tag{3.8}$$

r_g is termed the gyroradius of the particle, and when

$$r_g \propto \frac{E}{Zec} \gg L_c \tag{3.9}$$

the diffusive propagation regime is left behind. The term $\frac{E}{Zec}$ is known as a particle's rigidity, and an approximation to rectilinear propagation in the first order is assumed for UHE nuclei when $r_g > L_c$. UHE particles are now deflected by the much larger structure of what is termed the regular GMF . This larger structure's coherence length is on the kpc scale, and only when the UHE nuclei gyroradius, now termed the Larmor radius, r_L, exceeds this coherence length, is the particle able to break free of the regular magnetic field. The deflection of the CR path depends on the particle's charge, Z, such that different particles at different energies may have the same rigidity and hence the same deflections. A knowledge of the energy spectra profiles of different species of CRs in the disputed galactic/extra-galactic region and thereon, continues to be a frustrated goal (see Chapter 2 on page 7 from subsection 2.2.2).

3.2 Magnetic Field Measuring Techniques

Measurements of the GMFs are integrals averaged over the line of sight at large distances. They express quantities in terms of source polarization, whether it be parallel or perpendicular. Among such integrated measures is the optical polarization of starlight by ISM dust grains (first intimation of GMFs, optical arms), Zeeman splitting of spectral lines (magnetic fields in molecular clouds/stars), the thermal polarization of dust grains (from sub-mm frequencies into the infrared frequencies), polarized synchrotron emission (PI) and total synchrotron emission (I) (CRs/external galaxies), and Faraday rotation measures (RM) of linearly polarised radiation (GMFs outside local arm). The most pertinent GMF insights thus far, are offered by PI and RMs. Both measures are in the radio frequency band and are measured according to the density of relativistic CR electrons, n_{cre}, and the total density of free electrons, n_e, along the line of sight. Their polarizations are along mutually orthogonal magnetic field components, and in different ranges of the radio band [23].

The polarisation properties and intensity of synchrotron radiation are proportional to the product of the CR n_{cre} density and the integrated GMF along the line of sight.

Now, the polarization plane of linearly polarized source EM radiation, propagating through a magnetized plasma, rotates by $\Delta\phi = RM\lambda^2$, where λ is the measured wavelength at the observer. We also know that the number of free electrons (mostly thermal) along the integrated line of sight, l, gives RMs $\sim n_e B_{\|} l$. Together,

$$RM \simeq 0.81 \int_0^L \left(\frac{n_e(l)}{\mathrm{cm}^{-3}}\right)\left(\frac{B_{\|}(l)}{\mu\mathrm{G}}\right)\left(\frac{dl}{\mathrm{pc}}\right)\mathrm{rad\ m}^2 \tag{3.10}$$

and

$$\Delta\phi = \phi - \phi_0 = RM\lambda^2 \tag{3.11}$$

where ϕ is the polarization angle at observer and ϕ_0 the angle of polaraizarion of source. L is the distance between the collector and source. Combining Eqns. 3.10 and 3.11 we have

$$\Delta\phi = \lambda^2 0.81 \int n_e \mathbf{B}\cdot d\mathbf{l} = RM\lambda^2 \tag{3.12}$$

Positive/negative RMs point magnetic fields towards our direction/away from our direction. RMs are usually taken from compact radio sources within the galaxy (pulsars) and radio-loud galaxies, these extra-galactic sources being the most abundant.

Both PIs and RMs can only be taken when there is an appreciable density of the relevant electrons, and estimates of these densities can be in error [23].

3.3 Galactic Anisotropies and the Need for a Halo

For the allowed diffusive galactic propagation, anisotropies are found at the lower energies from CRs originating from the Sun (the short-lived Forbush Decrease, see page 18) - a single source. The *Compton-Getting Effect (CGE)* produces a permanent large-scale anisotropy (large in spread albeit small size, in the $10^{14} - 10^{15}$ eV range) due to a difference in reference frames as the Earth rotates around the Sun [28] .

Since our galaxy comprises the ISM, stars and a super-massive central black hole, one may expect associated anisotropies. The majority of galactic CRs are assumed to be from supersonic ejecta of SNRs being accelerated in the non-relativistic shock waves that follow a SN event [19]. If the SN events are recent and close enough, we may expect surges of CRs from their location and a point anisotropy. Whilst this anisotropy is temporary, its time scale is relatively long. A steady large-scale anisotropy may be connected with the galactic core or disk. Neither have been found. The familiar low energy anisotropies *do* disappear around Knee energies [93], so it may be that the isotropy thereon is due to a constant supply of galactic CRs, i.e., the rest frame of supply co-rotates with the galaxy. One must also remember that anisotropies can present when fluxes of different species fall off due to loss of containment - the proton knee being an example. Once the most abundant species can no longer be contained for acceleration, the flux of other CRs must bear some relationship with the percentage composition supposed to reside at acceleration sites. However, heavy galactic nuclei will experience strong GMF distortions and these deflections would not facilitate backtracking to their sources. Because of this, most UHECR source related searches concentrate only on a proton component of the source.

The diffusive galactic CR propagation is supported by two notable ratios of spallation nuclei

fluxes.

- The proportion of spallation products to their parent nuclei wanes as their energy waxes, indicating that the CR escape path length, $\xi_e(E)$, within the galaxy is energy dependent such that, $\xi_e(E) = \xi_0(E/E_0)^{-\alpha}$, where α is positive. Data analysis affords the relation, $\xi_e(E) \propto E^{-0.6}$. This translates, in the GeV range, to a typical path length ~ 5 g cm^{-2} of hydrogen matter for high energy particles. Referencing an interstellar mean particle density of $\bar{\rho} = 10^6$ m^{-3}, and given the relation $\xi = \bar{\rho}c\tau_e$, where $\tau_e \equiv$ escape time we arrive at $\tau_e \sim 3 \times 10^6$ years.

  The above escape time applies to CRs assumed to be confined to the volume of the disk, with radius $\sim (10 - 15)$ kpc and thickness $\sim (300 - 500)$ pc (see [100], page 319).

- the decay of unstable isotopes, measures the total time spent between spallation and observation. It is this figure, together with a calculated path length of around 5 g cm^{-2} [100], that gives up an estimate of mean particle density crossed, which is set at $\bar{\rho} \sim 2 \times 10^5$ m^{-3}. The CR confinement time being $\tau_e = \frac{\xi_e}{\bar{\rho}c}$, one arrives at $\tau_e \sim 10^7$ years, a value many times larger than that obtained by high energy CRs propagating directly through a galaxy of our dimensions. The confinement volume is thus taken to consist of not only the galactic disk but a galactic halo, an oblate spheroid surrounding our galaxy with a semi-major axis of ~ 10 kpc and a semi-minor axis of $\sim (3 - 4)$ kpc [100].[4]

The phenomenon of the galactic halo is now a major component in GMF models.

Parker observed the need for such a halo because our GMF must be unstable. If CRs are continually being generated throughout the galaxy, and the magnetic field weighted by the ISM (intergalactic space being essentially a vacuum), that magnetic field must inflate outwards as the CRs spiral along its lines to a point where the CR pressure is $\geq \frac{B^2}{8\pi}$, the magnetic field pressure. The field is, in effect, a pressure regulator with respect to the generation of CRs. The CR energy density of high energy particles (10^6 eV m^{-3}) is approximately equal to the GMF energy density, 2×10^5 eV m^{-3} [100].

[4]This assumption of 'northern' and 'southern' hemi-spherical symmetries about the galactic disk is now overturned [23], but we maintain the approximation, especially in the cosmological scale of measure.

The observational evidence for a galactic halo phenomenon is that haloes are found surrounding other spiral galaxies[5], a measured radio halo of highly polarized electrons due to synchrotron radiation (see [100] pages 270-271), and broad lines of CIV and SiIV are found for certain stars in the Magellanic clouds. This dragged out structure in the ionization lines is interpreted as a line-of-sight effect of the spectral lines passing through a hot gas closely associated with our galaxy. The galactic disk gravitational potential has been observed as being in hydrostatic equilibrium with the gas [100]. In addition, thermodynamics would require a hot gas halo as compared to the cold matter present in the disk.

In the extra-galactic regime, GMF models are necessarily employed in decriptions of overall propagation. Our GMF is at its most ordered and strongest along the galactic plane, so one would expect this region to exhibit characteristics reflecting these properties. The galactic disk, though, is obscured by interstellar dust, and our position along the outer edge of one of its spiral arms, does not allow for a comprehensive mapping of its features. Still, GMF models satisfying boundary conditions are used to delimit expectations involving extra-galactic propagation.

Whilst extra-galactic anisotropies are discussed in the context of GMF models, a section on source acceleration capabilities, distinguishes which extra-galactic sources are reasonably expected as able to generate the CR accelerations recorded here at Earth.

3.4 Extra-Galactic Sources

With respect to UHECRs, the inability of galactic source models to supply sources of sufficient maximum energy indicates a need for extra-galactic sources of CRs at the highest energies.

The magnitude of CR energies span such an enormous range that it is not unreasonable to posit their sources are diverse.

There is also no reason not to expect anisotropic phenomena at these UHECRs. Point-sources could be due to extraordinarily powerful extra-galactic sources of a certain longevity and possibly involving some type of lensing effect [185]. Large-scale anisotropies may be due to a collective diffuse spread of sources more powerful than our galactic core, in a rest frame distinct from our

[5]There are now definitive observations of our local halo, see [26, 199], which will be covered later in this chapter.

galaxy [186].

The EGMF is proposed to be even weaker than our galactic magnetic field ($\leq 10^{-8}$ G [161]), and so we retain the concept that acceleration and propagation of UHE extra-galactic CRs are based on collisionless processes, and their spectra depend on magnetic rigidity [69, pg 105].

Given the likely necessity that CRs must be accelerated within size limits, R, at their source, and be contained for a time by the magnetic field, $\mathbf{B}$, of the source, their Larmor Radius, R_L, must be within R - or rather, when qualifying composition, a particle's R_L defines the accelerator size. This is a Hillas criterion as the *least* necessary, but not sufficient condition, for a particle of charge, q, to gain a maximal energy, E_{max}, such that

$$E_{max}(B, R) \leq E_H(B, R) = qBR \tag{3.13}$$

Inclusion of an acceleration term, diffusive or one shot, for $E >$ EeV, gives the maximum energy, E_{max}, of a particle, size Z,

$$E_{max} \approx 10^{18}\text{eV}Z\beta_s(\frac{R}{\text{kpc}})(\frac{B}{\mu\text{G}})$$

$$\text{where } \beta_s \equiv \text{ shock velocity}$$

$$B \equiv \text{ magnetic field strength}$$

$$\text{and } R \equiv \text{ acceleration region size}$$

This relationship is illustrated in the Hillas plot of required magnetic field strengths with respect to source size in Figure 3.1 on the following page.

The Hillas limit is not a sufficient condition for the acceleration of particles. Limits are imposed by radiative losses within the acceleration region, and external losses due to interaction processes are a consideration. Sources that are collective (large-scale) need a density distribution that will satisfy the overall observed flux, and radiation byproducts must satisfy the observed spectrum.

All these provisos make for an elaborate inequality comprising several terms. A transport equation between source and receiver will be equally complicated. CR propagation models pro-

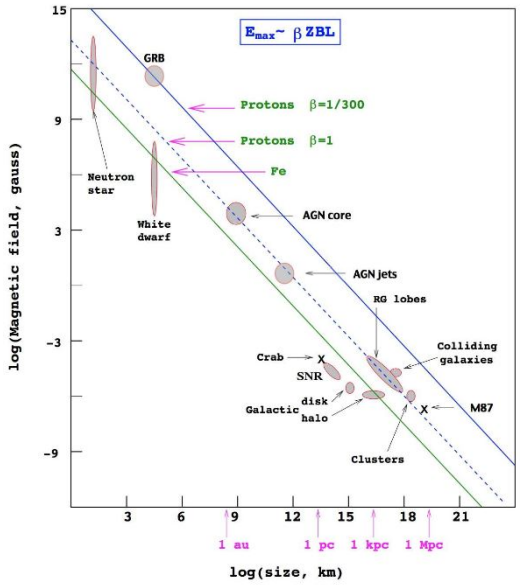

Figure 3.1: **Hillas candidate acceleration sites for protons and iron above 100 EeV. Sites must lie above relevant lines. Green for iron. Blue for protons. From [126] .**

vide further subsets of conditions. Imposition of another necessary, but not sufficient, condition that does not involve introduction of further variables, is achieved by a radiation loss term *within* the acceleration region, such losses are also contingent on the sources magnetic field. An extended Hillas plot including radiation losses is possible.

The Hillas plot has been recently revised by Ptitsyna and Troitsky [11] into 2 diagrams, dividing proton and iron sources. Hillas' original plot was model independent and geometric (size R), these also incorporate radiation induced contraints within a geometry. Their review article, [11], also seperates astrophysical source cases - the source defining acceleration regimes.

Acceleration source scenarios are diffusive or inductive. Diffusion is time-weighted and stochastic. Repeated Fermi-style interactions, order 1 or order 2, are mostly invoked. An inductive, or one shot, acceleration has a massive surge of the source electric field, **E**, overwhelming the source magnetic field. Extreme spatial curvatures (black holes, neutron stars) or synchrotron radiation (large-scale jets in powerul galaxies, γ-ray bursts (GRB)) in highly ordered fields would provide these conditions [11]. This acceleration scenario may well be that required by the latest GMF model (see section 3.4.3 on page 77, in particular the subsection on an Halo X-field found on page

80).

Because we are dealing with extreme conditions, equations governing the processes are taken to their limits. I reproduce/follow the logic of [11] in the discussion below. Light speed, $c = 1$, and efficiency coefficient $\eta \cong 1$. A particle's charge is q, and its mass, m. An electric field,

$$\mathbf{E} = \eta\mathbf{B}$$

has a particle energy gain rate

$$\frac{d\mathcal{E}^{+}}{dt} = q\eta B \tag{3.14}$$

and a particle moving with velocity, $\mathbf{v}$, in arbitary fields $\mathbf{E}$ and $\mathbf{B}$ is subject to radiation losses such that

$$-\frac{d\mathcal{E}^{-}}{dt} = \frac{2}{3}\frac{q^4}{m^4}\mathcal{E}^2((\mathbf{E} + [\mathbf{v} \times \mathbf{B}])^2 - (\mathbf{E}.\mathbf{v})^2) \tag{3.15}$$

the relativistic equations of motion give

$$= \frac{2}{3}\frac{q^2}{m^2(1-v^2)}(\mathbf{F^2} - (\mathbf{F}.\mathbf{v})^2) \tag{3.16}$$

and since $\mathbf{F} = \mathbf{F}_{\perp} + \mathbf{F}_{\parallel}$

$$= \frac{2}{3}\frac{q^2}{m^2(1-v^2)}(F_{\perp}^2 + F_{\parallel}^2(1-v^2)) \tag{3.17}$$

The first term in the R.H.S. of Eqn. 3.17 is the synchrotron term and the second is the curvature term.

Within an accelerator, size $R \mapsto \infty$, a particle's maximum energy gain, $\mathcal{E}_{max}$, is limited by its Hillas constraint energy, $\mathcal{E}_H$, and its radiation loss energy, $\mathcal{E}_{loss}$. At the very least,

$$\mathcal{E}_{max}(B,R) = \min[\mathcal{E}_H(B,R), \mathcal{E}_{loss}(B,R)] \tag{3.18}$$

Now, in equilibrium, assuming the rate of energy gain equals the rate of energy loss, we have

$$\frac{d\mathcal{E}^{+}}{dt} = -\frac{d\mathcal{E}^{-}}{dt} \tag{3.19}$$

a critical value of the magnetic field, $B_0(R)$, occurs when

$$\mathcal{E}_H(B,R) = \mathcal{E}_{loss}(B,R)$$

Thus

$$\mathcal{E}_{max}(B,R) = \begin{cases} \mathcal{E}_H(B,R) & , \quad B \leq B_0(R) \\ \mathcal{E}_{loss}(B,R) & , \quad B > B_0(R) \end{cases}$$

In all accelerator regimes, $R \equiv R_L$, or $R \equiv$ curvature radius, r, for field lines. When a curvature radius, $r > R$, r is then used in estimates. Scenarios of violent fluctuations in magnetic field configurations are not considered.

Ptitsyna and Troitsky [11] looked at experimental data of source magnetic fields. These fields can vary by orders of magnitudes for different sources, and they emphasize that all field sizes quoted are in orders of magnitude and energies are *limits* on maximal energy, $\mathcal{E}_{max}$. The sources they consider are

- Neutron stars: NS
- anomolous pulsars and magnetars: AXP
- Active galaxies - including

  Seyfert galaxies: Sy

  Radio galaxies: RG

  Blazars: BL

  Supermassive black holes and their environment

  Jets and outflows of active galaxies

  Jet knots: K, hot spots: HS and lobes: L, of powerful galaxies

- Star formation regions and starburst galaxies
- γ-ray bursts: GRB
- Galaxy clusters, superclusters and voids

An interesting result is the emergence of a viable Seyfert galaxy scenario. These are probably the least powerful of the AGN, they are radio weak, their jets are found to be non-relativistic and are often non-collimated. They have prominant nuclei emission lines and are relativley abundant. Their X-Ray emission is thermal. Ptitsyna and Troitsky [11] conclude while they would not be able to accelerate protons to energies $> 5\times10^{19}$ eV, within a few Schwarzschield radii, R_S, of their black holes, they should be able to accelerate nuclei to energies $\sim 10^{20}$ eV if their interaction with the ambient photon field does not dominate. Given the heavier composition of 'post' GZK Pierre Auger Observatory data - both $\langle X_{max}\rangle$ and the less model dependent, RMS values, it appears that Seyfert galaxies are possible candidates. The proton and iron extended Hillas plots are seen in Figure 3.2 on the next page and Figure 3.3 on page 70.

The large-scale isotropy of the universe admits to intricately worked nettings of galaxies. These inhomogeneities can be smoothed out in the very large units of cosmological measure, but within the discipline of the GZK horizon, those parts of the supergalactic plane and its resident galactic superclusters satisfying this limit, are likely sites not only of large-scale anisotropies but point-sources. A large-scale CR dipole would indicate evidence of a *Cosmological Compton Getting* (CCG) Effect and/or a collective diffuse spread of sources with nuclei of a heavy species.

3.4.1 Generic Galactic Magnetic Field Models

Before opening a chapter about UHE anisotropies, a foreword is now supplied describing common features of GMF models and a brief description of a recent (at the time of writing) GMF model. This model is more generalized than other GMF models and better fits observational results. The GMF models in the Energy-Energy ordering section of 4.3.3 on page 97 are precursors to this model.

The genesis of GMFs is unknown, but measurements of other galaxies and their optical ap-

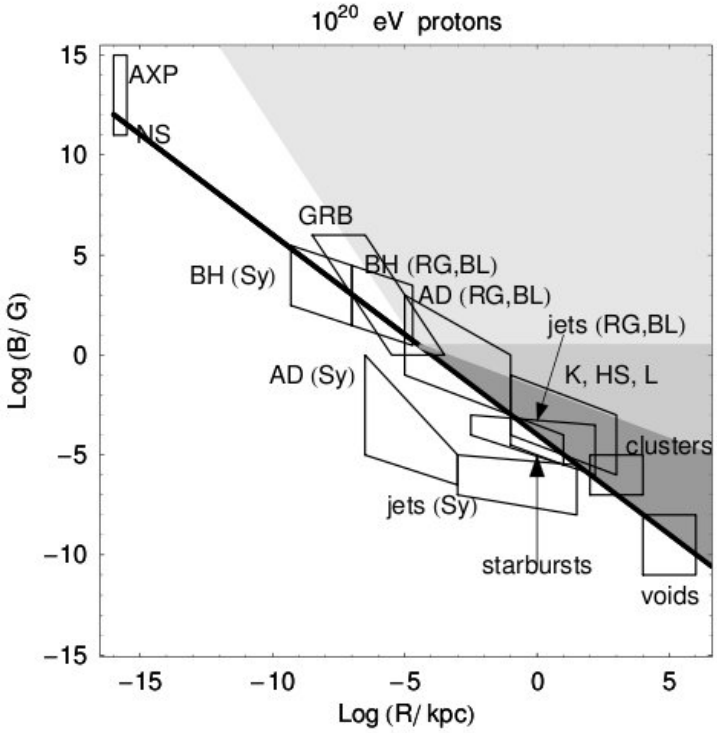

Figure 3.2: **proton sites at $\sim 10^{20}$ eV incorporating geometrical and radiative losses. Shaded areas are possible sites bounded by a lower limit. light grey: one shot acceleration in curvature regime. middle grey: one shot acceleration includes curvature and synchrotron regimes. dark grey: includes one shot acceleration for curvature and synchrotron regimes as well as shock acceleration. Neighbourhoods are central parsecs (AD) for low-power Seyfert galaxies (Sy) to powerful radio galaxies (RG) and blazars (BL) neutron stars (NS), anomalous X-ray pulsars and magnetars (AXP) and of supermassive central black holes (BH) of active galactic nuclei, from low-power Seyfert galaxies (Sy) to powerful radio galaxies (RG) and blazars (BL). Relativistic jets, knots (K), hot spots (HS) and lobes (L) of powerful active galaxies (RG and BL); non-relativistic jets of low-power galaxies (Sy); starburst galaxies; γ-ray bursts (GRB); galaxy clusters and intercluster voids. From [11].**

peareance, combined with measurements of our own, are thought to describe a present day GMF structured like a magnetic dynamo field (see [187] section 19.3.3). These dynamo models allow for bi- and axi-symmetric configurations. Some of the models also predict not only an azimuthal field but a poloidal field [23]. In accordance with, but not necessarily because of these predictions (the visible spectral shape lends itself to the maths), GMF models describe different magnetic field configurations thought possible with azimuthal and poloidal fields. We find in a GMF model superpositions of dominant magnetic fields in distinct regions and over a wide scale in size.

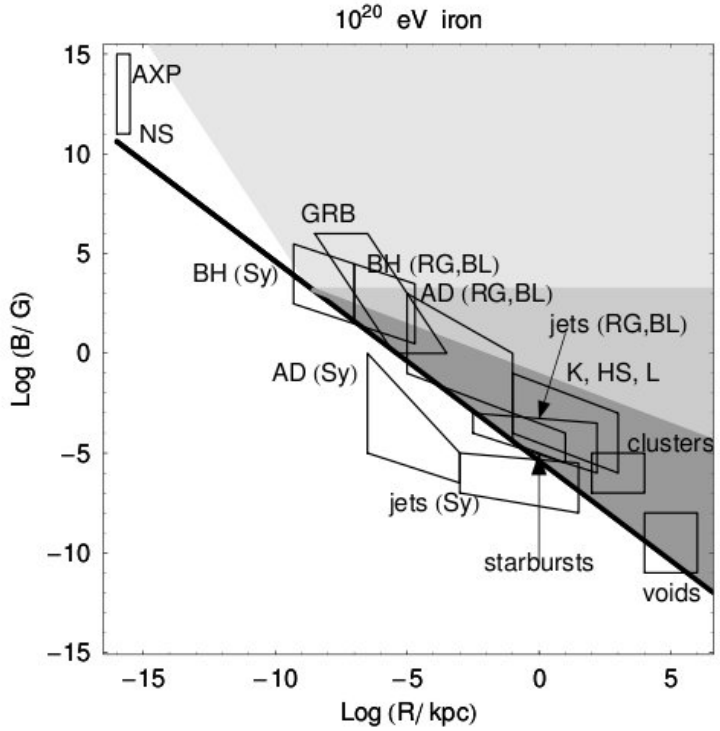

Figure 3.3: **iron sites at E $\sim 10^{20}$ eV. Shading and neigbourhood legends same as in Figure3.2. As is apparent, acceleration from Seyfert galaxies is possible for iron. From [11].**

Some models only consider the galactic disk, others the disk plus a halo. The disk is generally subdivided into a central bulge plus a thin disk finishing some few kpc from the Earth's location in the x-coordinate. The axes of symmetry of the disk and halo are the same. The GMFs are believed to comprise regular, large component fields, and a pervasive, small-scale turbulent component field. The difference in scales lends itself to distinct analyses of the regular large GMFs and the random magnetic fields. When GMFs are invoked, it is in reference to the large-scale regular fields, $|B|$, describing large-scale galactic dynamos and deflections, $\Delta\Theta \sim 13^\circ$ in the EeV energy range over a kpc length range, i.e. the random component can be viewed as isotropic background within the limits of kpcs.

The random component, describing small-scale turbulent processes mainly in the ISM, is referenced by its RMS strength, B_{rms}, and small random angular deflections, δ_{rms}, ($\sim 1.5^\circ$) over the tens of pc length ranges, i.e. in the limits of tens of pcs, the regular component can be viewed as

isotropic background. The regular galactic magnetic field (GMF), $\sim 6\,\mu$G, can be inferred from a plurality of line of sight integral measures over several coherence lengths (see section 3.2 on page 60). A dominant regular field figures largely in extra-galactic propagation models, responsible for the reception of rigidity-based CR nuclei deflections.

The smaller range, greater strength, random fields, which are one half to twice as strong as the regular, coherent GMF [21], having no coherence, until recently had not been considered serious players in galactic propagation scenarios.

Because of the large-scale symmetry in the appearance of the galaxy, it is necessary to obtain reliable values for the strength and direction of the GMF in its radial profile and its (x,y,z) planar co-ordinates. The galactic shape itself is evidence that the GMFs would be expected to change character depending on their position in 3-D space.

GMF models are diverse. Recently, though, some agreement over certain large-scale features has settled. What follows below is an account of some of those agreements and which representative GMF models survive, or at least contribute building blocks the a more generalized GMF model. Given our position within the galaxy, most models refer also to the GMF local to our Sun, $B_0 \sim 2\,\mu$G in a radius ~ 3 kpc, and our Sun's radial distance from the GC, $R_\odot \sim 8.5$ kpc.

3.4.2 Galaxy Volume

The sum volume of the GMFs is taken over a ~ 20 kpc radius. The radial GMF profile of the galaxy seems to be at its strongest around the galactic centre (GC). This profile is loosely described by a decaying exponential function with a scale length, R_B, such that

$$B_{reg}(R) = B_0(R) \exp\left[-\frac{R - R_{GC}}{R_B} \right] \tag{3.20}$$

which could, when divided into z-axis slices, describe a regular azimuthal GMF in the (x, y) plane e.g. see Figure 3.4 on the next page (a regular poloidal GMF is most apparent in the (x, z) plane as seen in Figure 3.5 on page 77). The exponential decline is over the radius of ~ 20 kpc from the galactic centre (GC). The GMF is commonly taken to be zero outside this radius and also within a central radius containing the supermassive galactic centre black hole.

Radio galaxy RMs manifest opposite directions in an azimuthal field above and below the disk mid-plane, $z = 0$ [21], but not past a height $z \sim \pm 1.5$ kpc. This is the height a disk field is considered to subtend. Immediately, we see a separation between the disk field and halo. RMs within the disk indicate at least one confirmed azimuthal field reversal [22], with the possible detection of others [23]. The GMFs of disk and halo are even further differentiated. Separating the GMF into, at the very least, a disk region and a large halo is now supported.

In Figure 3.4 we see a reductive illustration of the separations of the GMFs between disk and halo [24]. A large, regular azimuthal field is present in the galactic disk of $\sim 6\,\mu$G. This field is

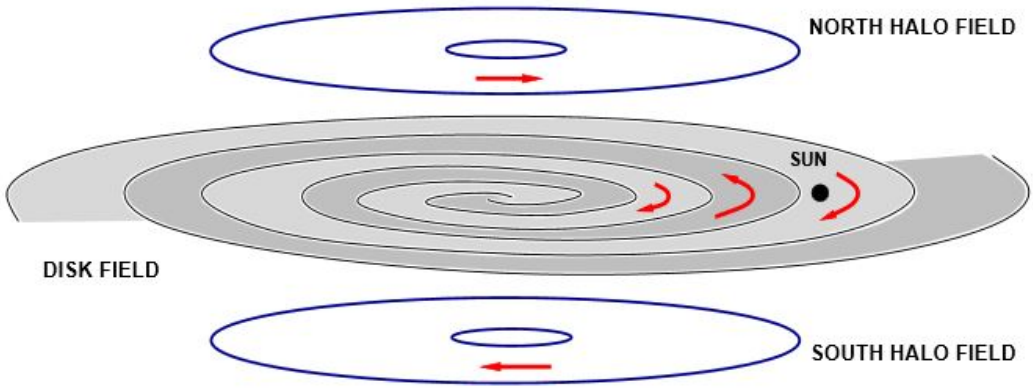

Figure 3.4: **Structure of GMF as considered divided into halo and disk components. RM's reveal a disk field symmetric about the galactic plane and a halo field antisymmetric about the galactic plane. This is seen through correlated anti-symmetric RMs between the disk field with the halo field above the galactic plane, and with correlated symmetric RMs between the disk field with the halo field below the galactic plane. Field reversals within the spiral arms of the disk are also shown. From [24].**

generally represented as a logarithmic spiral or as toroidal. The height of the disk is usually taken as ~ 1.5 kpc. How and where the disk field transitions into the halo invites further modelling.

Poloidal fields (dipole) perpendicular ($\perp$) to the galactic plane are predicted by magnetic dynamo field models [187, 162], and a small dipolar field of $\sim 0.2\,\mu$G within the solar system locus exists, [21]. Large vertical 'striated filaments' of a GMF around the bulge have been measured with strengths of $\sim$1 mG [190, 25]. A strong $\perp$ halo field has also been proposed, (see [23] for a

working model) and recently is verified in polarized synchrotron emission (PI) measures published by [26].

The disk azimuthal field of $\sim 6\ \mu$G is strongest in the inter-spiral arms (magnetic arms) [193], which may number more than two, and the arms themselves are dominated by a random field.

So, we have garnered GMF models with a disk being composed of a likely azimuthal field (whether spiral or toroidal) which has at least one measured field reversal, and perhaps a lesser poloidal field. The halo can now be approached as composed of an azimuthal field distinct from its disk sister's, which swaps field direction with respect to the galactic disk plane. A significant vertical/poloidal field may have asserted its presence. It may also indicate a massive galactic wind.

Excluded from the general GMF properties in the radial and (x,y,z = 0) plane is a region around the galaxy's central black hole, a central bulge - the disk centre, radius denoted R_c. Non-thermal vertical filaments hint at a strong poloidal field here $\sim$ 10 mG, and there also seems to exist a smaller toroidal field. The optimized radius for the disk centre extends to $\sim (1-6)$ kpc , according to the particular model used [23]. Disk models are considered to begin around this perimeter. The latest modelling/observations from e.g. [23, 26, 199], indicates that the bulge is very active and there are huge radio lobes extending well into the halo. This will be further discussed in the context of the generalized GMF model of [23], and in the Chapters 6 and 7.

All GMF modelling is likely only approximate. Any line of sight integral is unlikely to be clear and GMF models are primarily constrained by RMs. Over astronomical distances, the line of sight measure in a direction to a candidate feature can be distorted by RMs of intervening polarizations or GMF reversals. 'Masks' are used to subtract those region's RMs that are not pertinent to the desired RM. As example, for GMFs in the halo, one also measures RMs in the disk. Because of the azimuthal field reversal in the halo, RMs in the northern galaxy should be augmented and in turn reduced in the southern galaxy [193]. By how much, is model dependent. The electron densities (n_e/n_{cre}) of both PI and RMs must be large enough for these meaures to register, and these electron populations are not isotropic.

GMF models are decidedly in transition. The latest halo observations highlight the need for a more consistent GMF model wherein the halo is given prominence, and simplified GMF models

are overturned. For example, because of no definitive global measurements, previously, poloidal fields were often taken to zero both for the disk and halo (see [191] for a table of GMF models). In the large-scale domain, the isotropic random field component is often excluded from modelling [23], but now is utilized in UHECR detection Energy-Energy ordering techniques (see section 4.3.3 on page 97). Small, selective data sets used can yield conflicting results and/or results that do not approach observations, no matter how GMF parameters in various models are optimized.

3.4.2.1 Disk Models

Disk Models are arrangements of magnetic field symmetries. There are logarithmic spiral models and concentric ring (toroidal) models with field reversals [21, 22]. Some models star a purely azimuthal field, for others the GMF is analysed as spiral and the field reversals as circular [22]. The ring models have fallen into disfavour [24, 25], and the symmetries of the remaining spiral models have been reduced [24, 23]. The Bisymmetric Field model (BSS) has the GMF swapping direction in its arms ($\text{arm}_{min} = 2$), and the Axisymmetric Field model (ASS) has the GMF retaining its direction in both arms. Associated with each model are further symmetries across the galactic plane. The BSS model is symmetric/anti-symmetric, BSS-S/A, across the galactic plane, and the ASS model is symmetric/antisymmetric, ASS-S/A, across the plane. No vertical GMF is assumed to be present across the galactic plane, i.e. the lines of the GMF are azimuthal. Spiral models are iterations on the models referenced in [24], as given below.

The GMF co-ordinate systems used in descriptions are cylindrical such that $B = B(r, \theta, z)$. with $B_\theta = B\cos(p)$, $B_r = B\sin(p)$ The rectangular co-ordinates of our location being, $(R_\odot, y, z)$ kpc, $R_\odot = 8.5$ pc, Earth's distance from galactic centre, gives $\theta = tan^{-1}(\frac{-y}{x})$.

ASS model:

$$B(r) = B(r)\Big|\cos\left(\theta - b\ln(\frac{r}{R_\odot}) + \phi\right)\Big|\exp\left(\frac{-|z|}{z_0}\right) \tag{3.21}$$

BSS model

$$B(r) = B(r)\cos\left(\theta - b\ln(\frac{r}{R_\odot}) + \phi\right)\exp\left(\frac{-|z|}{z_0}\right) \tag{3.22}$$

where $\phi = b \ln(1 + \frac{d}{R_\odot}) - \frac{\pi}{2}$

and $b \equiv \frac{1}{\tan(p)}$

the pitch angle, p, is the deviation of the field lines from azimuthal in the (x, y) plane. i.e. $p = -10°$, yields a spiral GMF at $l = 80°$, and $p = -5°$, yields a spiral GMF at $l = 85°$. d is the distance to the first field reversal.

$\mp d \Rightarrow$ reversal towards/away from galactic centre.

Thus the BSS-S model satisfies $B_{BSS-S}(r, \theta, z) = B_{BSS-S}(r, \theta, -z)$, i.e. the disk fields above and below the galactic plane, $z = 0$, look the same.

The ASS-A model satisfies $B_{ASS-A}(r, \theta, -z) = -B_{ASS-S}(r, \theta, z)$, i.e. there is a field reversal above and below the galactic plane. This case does not accord with observations for the disk.

The GMF amplitude is defined as [24]

$$B(r) = \begin{cases} B_0 \frac{R_\odot}{r \cos(\phi)} & , \quad R > R_C \\ B_0 \frac{R_\odot}{R_c \cos(\phi)} & , \quad R < R_c \end{cases}$$

where $R_c \equiv$ radius of a large constant GMF around the galactic centre. Because this region is frequently not discussed, the GMF within this radius can also be taken to zero.

The conservation of magnetic flux, $\nabla \cdot \mathbf{B} = 0$, is not incorporated into many spiral models [23].

There are assumed to be no significant vertical fields.

Pshirkov *et al.* [24] recognize two spiral field benchmark models, with BSS and ASS symmetries in the disk. The halos of these models do not differ in extent and we now know that the southern halo is much bigger than the northen halo [26]. Following observations, we see that the spiral fields now most closely following observations, are for BSS_s and ASS_A [24].

3.4.2.2 Halo Fields

The direction of the halo fields reverse across the galactic plane. Some halo models consider only a diluted vertical extension of the disk field, a toroidal field, whilst others add a poloidal field. The

general form of a toroidal field, which is purely azimuthal, as taken by [24] is such that

$$B_{\theta}^{H}(r, z \geq 0) = B_{0}^{H}[1 + (\frac{|z| - z_{0}^{H}}{z_{1}^{H}})^{2}]^{-1} \times \frac{r}{R_{0}^{H}} \exp(1 - \frac{r}{R_{0}^{H}}) \tag{3.23}$$

when $z < 0$ the field direction reverses.

where $B_{0}^{H} \equiv$ Halo MF strength.

$R_{0}^{H} \equiv$ Halo radius scale.

$z_{0}^{H} \equiv$ Halo vertical position.

$z_{1}^{H} \equiv$ Halo vertical scale.

where $|z| < z_{0}^{H}$ and $|z| > z_{0}^{H}$ and the form of the poloidal/dipole field, which has no azimuthal component, as taken by [194] is such that, in spherical polar co-ordinates (ρ, θ, z)

$$\mathbf{B}_{D}(r, \theta) = \frac{2D\cos(\theta)}{r^{3}}\hat{\mathbf{r}} + \frac{2D\sin(\theta)}{r^{3}}\hat{\theta} \tag{3.24}$$

where D $\equiv$ dipole moment. the conversion between cylindrical to spherical co-ordinates being $\rho = (r^{2} + z^{2})^{1/2}$, $\theta = \theta$, $\tan(\theta) = r/z$.

The strength of the halo fields is much in doubt. The electron densities, n_{e} and n_{cre} are uncertain as predicted by diffusion models, and the line of sight integrals, RM and PI, which are weighted by n_{e} and n_{cre} have associated layers of polarizations belonging to features interfering in the line of sight.

A recent form of the poloidal field, as proposed by Jannson and Farrar [23], is axisymmetric with a divergence satisfying the condition $\nabla \cdot \mathbf{B} = 0$. This field has been dubbed the 'X-field', in part because of structures observed edge-on in extra-galactic galaxies.

The field position, in the (r, z) plane, is defined in terms of a radius, r_{p}, taken at the mid-plane when $z = 0$. A radial elevation angle, Θ, is a constant function outside the boundary value set when $r_{p} = r_{X}^{c}$, the galactocentric radius, and within the boundary, Θ begins to climb, resolving at the galactic centre, $r = 0$, where it reaches 90°. The field is seen in Figure 3.5 on the following page.

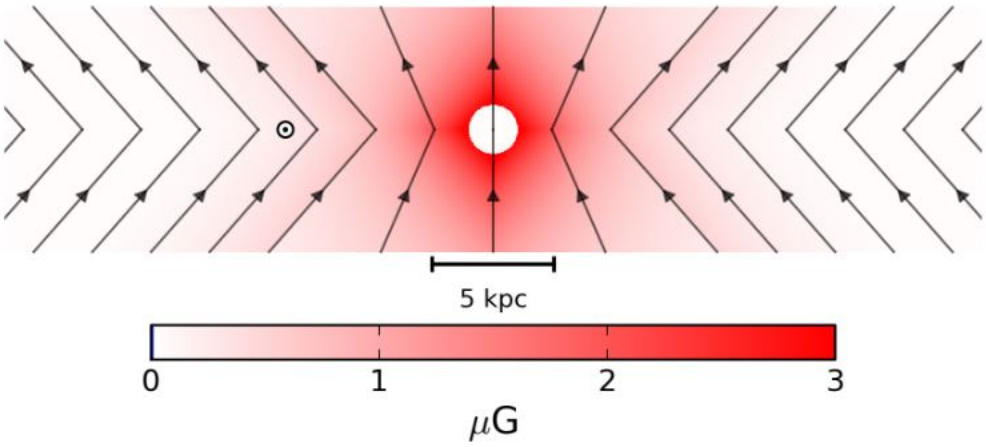

Figure 3.5: **The poloidal X-field in the x-z plane of the galaxy. The inner region, reflected about the z-axis is bounded by the X-field magnetic lines which show the field direction. The inner region is where a varying internal elevation angle, Θ, is posited, such that$|\frac{\partial\Theta}{\partial r}| > 0$. Outside this region, Θ is considered an external constant, such that$|\pm\frac{\partial\Theta}{\partial r}| = 0$. The same field with a more complete explanation is seen on page 82. From [23].**

3.4.2.3 Turbulent Fields

Density fluctuations in the ISM are the template for turbulent fields. They can scale up to $\sim$ 100 pc in the vicinity of SNRs. Turbulent fields are considered to occupy the entire volume of the galaxy, although in some galactic modelling they are taken to be zero, but not by all (see [23] and the reference therein). Their scale of turbulence is considered to have close to a Kolomogorov spectrum, although alternative power law spectra have also been used (see [25]). An example of turbulent field structure is found in [194].

3.4.3 Generalized GMF Model of Jannson and Farrar

The preceding sections concerned with GMF models, contribute to the language regularly used by Jannson and Farrar [23] in their 21 free-parameter model of GMFs.

This model has given the best fits to observations up until now. The RMs used come from multiple sources and are extra-galactic (galactic pulsar distances are deceptive), some 40,043 RMs in total. The PI measures are from the WMAP7 release of its K-band data - galactic synchrotron

emission is primarily in this 22 GHz band.

The model takes on the character of 2 coordinate systems in 3 dimensions, the cartesian (x, y, z), and cylindrical (r, ϕ, z) [6]. The galactic centre is at $= \vec{0}$, Galactic north is the +ve z axis, and the Earth's location is at $x = -8.5$ kpc.

Briefly, their model is composed of a disk component, a striated field component, and a halo component. **Each** magnetic field component observes flux conservation, $\nabla \cdot \mathbf{B} = 0$. The model enjoys the simplification of no total synchrotron emission factor, I, ($\sim$ 75% is estimated to be polarized), and the omission of the galactic-wide random fields themselves. The total field is set to 0 for $r < 1$ and $r > 20$ kpc. A transition expression, $L(z, h, w)$ of parameter width, w_{disk}, and disk height, h_{disk}, is defined for that region where the disk field morphs into the toroidal halo field component. To achieve this, the disk field component is multiplied by $(1 - L(z, h_{disk}, w_{disk}))$, and the halo by $L(z, h_{disk}, w_{disk})$.

The striated field component, wherein there is an overall orientation over large-scales, has small random field units which change direction and strength. In this model it is assumed to be orientated with the local GMF and to be of a constant strength. Jansson and Farrar [23] speculate that these fields could be within detached bubbles of hot plasma or produced by the differential rotation of random fields. The striated fields contribute to the PI measures but not the RMs, which would be obfuscated by the field's variable sign.

3.4.3.1 Disk

The disk field is defined in the $x-y$ plane. A sole azimuthal aspect, B_{ring}, represents the molecular ring[7], between (3 - 5) kpc of the galactic plane radius, and 8 logarithmic spiral arms, r_i, extend to between (5 - 20) kpc. Their strength, b_i, is at a *maximum* at the 5 kpc ring boundary and declines as r^{-1}. The spiral fields are given direction, $\hat{b} = \sin(i)\hat{r} + \cos(i)\hat{\phi}$. Magnetic flux conservation leaves 7 of the arm field strengths as free parameters and the 8th arm is proportional to the -ve sum of the 7 free arms $\times$ their cross sectional areas divided by the cross sectional area of the 8th arm. Jannson and Farrar's [23] work is in refiguring the halo fields. They make the point that the disk

[6] ϕ replaces θ in the previous cylindrical co-ordinates to avoid confusion with the parameter assigned Θ by [23].
[7] a resevoir of $(5 - 10)$% of our galaxy's molecules and a vigorous star formation area [189]

field is of unfolding complexity.

Some parts of the disk field are seen as asymmetric with z because of the superposition of the halo fields (both toroidal and poloidal). The azimuthal molecular ring field and the first spiral arm are such, as illustrated by the arrows in Figure 3.6 on the next page. The strength and orientation of the azimuthal component of all of the field is seen in the arrows for the molecular ring and the eighth spirals, and can be read in the colour bar, a negative field strength being considered azimuthly clockwise. A field reversal can be seen in the second arm which is anticlockwise, and the sixth arm which is clockwise. The disk field above and below the plane is similar but not twinned. There may be yet more field reversals found in the spiral arms - [23] may have found another in regions six and eight. The pitch angle of the disk field changes with r under the enforcement of the halo poloidal component field.

3.4.3.2 Halo Toroid

The disk field transitions to the halo toroid at parameter height, h_{disk}.
The halo's azimuthal character's best fit, is a toroidal field, defined such that

$$B_{\phi}^{tor}(r,z) = e^{|z|/z_0} L(z, h_{disk}, w_{disk}) \times \begin{cases} B_n(1 - L(r, r_n, w_h)) & \text{if} \quad z > 0 \\ B_s(1 - L(r, r_s, w_h)) & \text{if} \quad z < 0 \end{cases} \tag{3.25}$$

Note the bipolar identities of the north and south field amplitudes, B_n/B_s, and their radial extent, r_n/r_s. An exponential scale height description is retained (see 3.20). The asymmetrical aspects of the halo can be referenced the second panel of figures in Figure 3.6 on the following page just above the magnetic field strength colour bar.

The pitch angle of the toroidal component is zero. The vertical extent of the GMF is not constant across the disk, and is primarily dependent on the value of the toroidal n_{cre}. Electron densities are not considered well 'locked in' [189]. The constraint chosen was supplied by the GALPROP numerical code for CR diffusive transport. A plot of 2 contours of the optimized GMF field strength for 2 locations in different spiral arms is seen in Figure 3.7 on page 81.

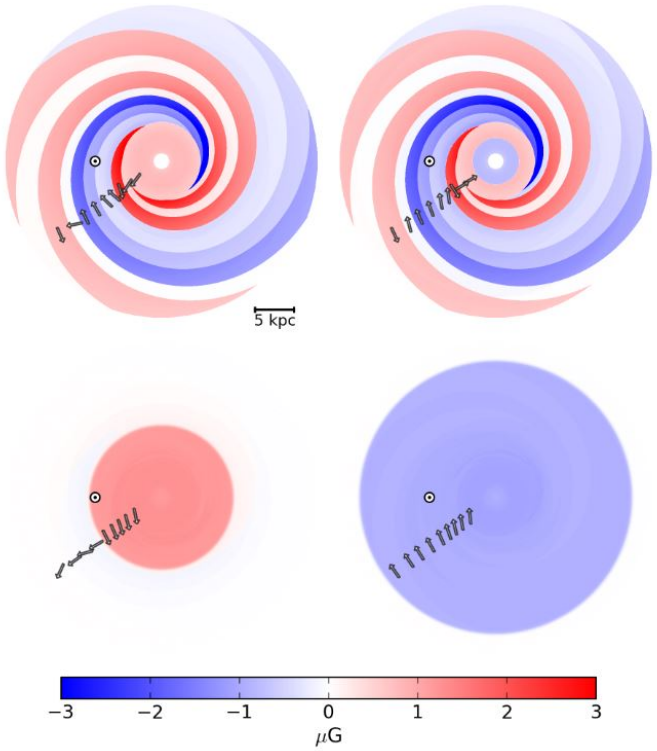

Figure 3.6: **GMF as seen in x-y slices: top left:** $z = 10$ **pc. top right:** $z = -10$ **pc. The asymmetry of the molecular ring and first spiral arm about the** $z = 0$ **plane show the impact of a halo X-field causing actual asymmetry in this region and a varying pitch angle (changes with r) in the disk toroidal field. Both 'north' and 'south' disk fields are similar and roughly symmetric about the** $z = 0$ **plane. The position of the Sun is shown as a bullseye.**
bottom left: $z = 1$ **kpc. bottom right:** $z = -1$ **kpc. Views of the disk field, which is now dominated by the halo toroidal field. At these distances above/below the** $z = 0$ **plane the halo field's antisymmetry is clear. The radius of the top halo field is** ~ 9.2 **kpc and the radius of the bottom halo field is** $\gtrsim 16$ **kpc. The strength of the magnetic field is shown by the colour bar. The X-field superposition at this point is also made apparent. From [23].**

3.4.3.3 Halo X Field

As already mentioned, the significant departure from previous GMF models, is a substantial, functional form of a poloidal, axisymmetric field, reproducing structures we know exist in halos of other galaxies. This field is parameterized by the radius, r_p, the midplane crossing point ($z = 0$), for a field line located at (r, z), with the free parameters, B_X, Θ_X^0, r_X^c and r_X. Field strength in the

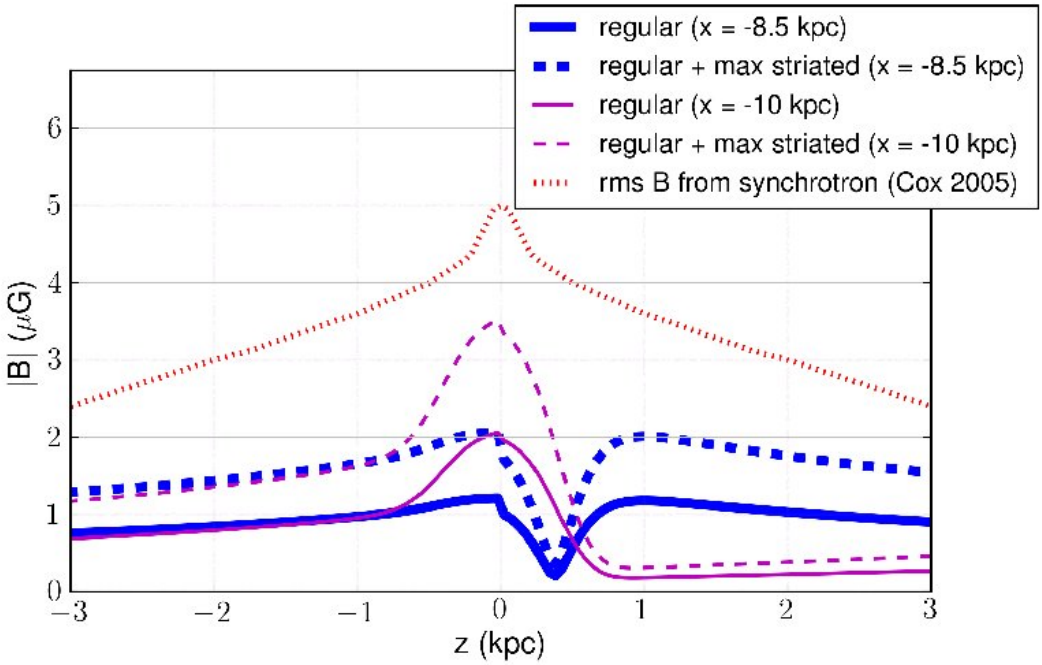

Figure 3.7: **The predicted field strength of the optimized GMF model as a function of z, for $(x, y) = (8.5, 0)$ kpc, and $(x, y) = (10.0, 0)$ kpc. Note the difference in field strength $\sim |z| = 0$. The location of the 2 x-values are in different spiral arms. The dotted line is from Cox (2005) and is an estimate of the total field strength. From [23].**

midplane is given by the amplitude, B_X, of the field and r_p in Eqn. 3.26.

$$b_X(r_p) = B_X \exp^{\frac{-r_p}{r_X}} \tag{3.26}$$

In the region of constant elevation angle, the field strength is an expression of $b_X(r_p)r_p/r$ translating as

$$r_p = \frac{r - |z|}{\tan(\Theta_X^0)} \tag{3.27}$$

and when $r_p < r_X^c$, the field strength varies linearly with r, an expression of $b_X(r_p)(r_p/r)^2$ translating as

$$r_p = \frac{r r_X^c}{r_X^c + |z|/\tan(\Theta_X^0)} \tag{3.28}$$

and the elevation angle is taken as

$$\Theta_X(r,z) = \tan^{-1}\left(\frac{|z|}{r-r_p}\right) \tag{3.29}$$

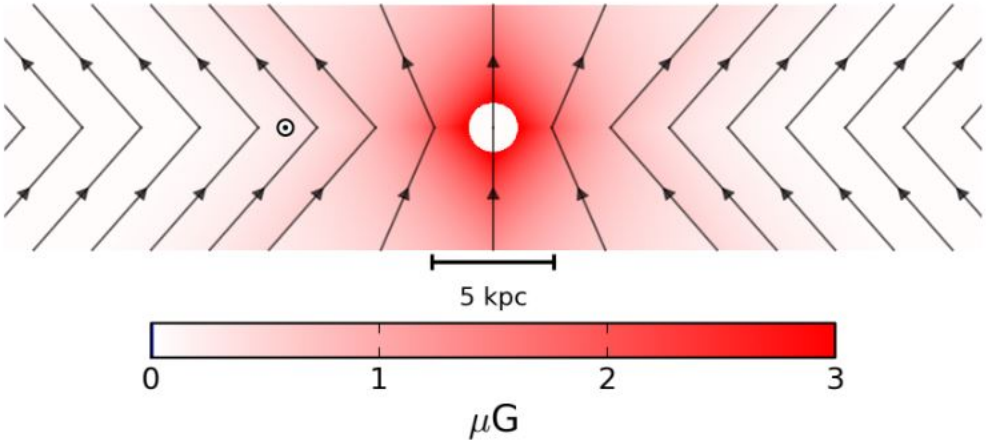

Figure 3.8: **The X-field in the x-z plane. The elevation angle, Θ_X, varies in the area where $x \sim \pm(3-5)$ kpc, and is constant, $X_X^0 \sim 50°$ for $x > |4.8|$ kpc. The black lines show the direction of the field. The varying Θ_X increases linearly from $\sim 50°$ at the boundary to $\Theta_X = 90°$ at the inner boundary for $x \sim \pm 3$ kpc where the X-field becomes vertical. From [23].**

The 'out-of-plane' field is seen in Fig 3.5 on page 77.

3.4.3.4 Striated Random fields

These fields are considered global. They are small-scale, and they have small-scale variations in strength and direction, but their **overall** orientation is the same and considered to be the same as their local regular field. They are also called "ordered random" fields [191]. These fields may be produced by the phenonomen of bubbles of hot plasma gases disconnecting from the central bulge. The bubbles which contain random fields may see these fields become "ordered" by a differential rotation which may align the overall random fields with the prevailing large-scale coherant magnetic field. The PI measures of striated random fields contribute to both the total and polarized emissitiy (thereby engendering a degeneracy with the relativistic CR electrons), and are parameterized such that $B^2_{stri} = \beta B^2_{reg}$, $\beta \geq 0$. The fields lose their striations when $\beta = 1$.

Allowing $\gamma = \alpha(1 + \beta)$, where α is a scaling factor for n_{cre}, results in an underestimation of n_{cre}, when $\gamma < 1$. Doubts over the measurements of n_{cre} have previously been touted in, for example, [24]. Underestimation of n_{cre} or an absence of striated fields achieve similar effects.

Whilst acknowledging the presence of small-scale random fields, [23] do not include them in their model. The overall large-scale coherent field contributes to the RMs, the PI and I. The striated fields because of changing signs, to the first order, only contribute to I and PI and the random fields contribute to I. So, considerations of I are excluded.

3.4.3.5 Optimizations and Implications

GMF optimization values as supplied by [23] are seen in Table 3.1.

Field	Best fit Parameters	Description
Disk	$b_1 = 0.1 \pm 1.8\,\mu$G	field strengths at $r = 5$ kpc
	$b_2 = 3.0 \pm 0.6\,\mu$G	
	$b_3 = -0.9 \pm 0.8\,\mu$G	
	$b_4 = -0.8 \pm 0.3\,\mu$G	
	$b_5 = -2.0 \pm 0.1\,\mu$G	
	$b_6 = -4.2 \pm 0.5\,\mu$G	
	$b_7 = 0.0 \pm 1.8\,\mu$G	
	$b_8 = 2.7 \pm 1.8\,\mu$G	inferred from $b_1, ..., b_7$
	$b_{\rm ring} = 0.1 \pm 0.1\,\mu$G	ring at 3 kpc $< r <$ 5 kpc
	$h_{\rm disk} = 0.40 \pm 0.03$ kpc	disk/halo transition
	$w_{\rm disk} = 0.27 \pm 0.08$ kpc	transition width
Toroidal halo	$B_{\rm n} = 1.4 \pm 0.1\,\mu$G	northern halo
	$B_{\rm s} = -1.1 \pm 0.1\,\mu$G	southern halo
	$r_{\rm n} = 9.22 \pm 0.08$ kpc	transition radius, north
	$r_{\rm s} > 16.7$ kpc	transition radius, south
	$w_{\rm h} = 0.20 \pm 0.12$ kpc	transition width
	$z_0 = 5.3 \pm 1.6$ kpc	vertical scale height
X halo	$B_{\rm X} = 4.6 \pm 0.3\,\mu$G	field strength at origin
	$\Theta_{\rm X}^0 = 49 \pm 1^\circ$	elev. angle at $z = 0, r > r_{\rm X}^{\rm c}$
	$r_{\rm X}^{\rm c} = 4.8 \pm 0.2$ kpc	radius where $\Theta_{\rm X} = \Theta_{\rm X}^0$
	$r_{\rm X} = 2.9 \pm 0.1$ kpc	exponential scale length
striation	$\gamma = 2.92 \pm 0.14$	striation and/or $n_{\rm cre}$ rescaling

Table 3.1: **Optimization Table of GMF with $1 - \sigma$ intervals. From [23].**

This table, with its 1σ values, together with Figure 3.6 on page 80 gives an appreciation of the optimized GMF. The arrows in Figure 3.6 are for the inner molecular ring values and the 8 spiral arms. Their orientations are shown. Field reversals are explicitly stated. One sees the

disk/halo transition heights and widths are variable even down to the disk midplane. Jannson and Farrar qualify the extent that these optimizations can be considered. They leave open the nature of the other side (l.h.s) of the galaxy and the disk. This may not enjoy symmetry with the r.h.s. chiefly because of the nature of the logarithmic functions for the spiral arms. The observables in the logarithmic functions, and in fact all observables taken, are only considered reasonable within several kpc of our location. The northern halo radius at $r_n = 9.22 \pm 0.08$ kpc and field strength $B_n = 1.4 \pm 0.1$ μG is in strong contradistinction to its southern counterpart, with an unconstrained $r_s >\sim 16.7$ kpc and field strength $B_s = 1.1 \pm 0.1$ μG .The elevation angle, Θ is $\sim 50^\circ$ for $r_p < r_X^c$ and ends $\perp$ to the $(x, y, z = 0)$ plane. The scale height of the toroidal field is $z_0 = 5.3 \pm 1.6$ kpc, and one is reminded that the large $\gamma = 2.92 \pm 0.14$ will accommodate either a striated field contribution or a rescaling to a larger n_{cre}. To check how robust the GALPROP value of n_{cre} is, [23] multiplied n_{cre} by $\exp|z|/z_{cre}$ with $z_{cre} = 10$ kpc. The scale height of the relativistic electrons increased $\sim 10\%$ for $z = 1$ kpc and $\sim 20\%$ for $z = 2$ kpc. The optimized parameters, r_s, r_n, w_h and Θ_X changed only by an $\sim 0.4\sigma$ extent and the greatest change was in the value for α which decreased to 2.65 from 2.92. It is noted that the strong vertical field does predict a larger scale height for the diffusion of relativistic electrons, and the GALPROP n_{cre} may be too small, given they do not include an anisotropic, spatially varying diffusion term.

A strong vertical field would disperse UHECRs further than a small one. The deflections are proportional to the integrated transverse magnetic field along the trajectory of the primary. A prediction of a widely asymmetric dispersal of protons is illustrated in Figure 3.9.

The halo GMF components are seen as predominantly responsible for UHECR deflections. Jansson and Farrer make a distinction between halo deflections, which map out of the radio lobes and the deflections of UHECRs along the galactic plane which are very small. The proton deflections are at their largest extent of 15° in the inner galaxy and 'southern' galaxy proton deflections are wider than the 'northern' proton deflections. The average 60 EeV proton deflection is $\sim 5.2^\circ$. We should expect from their predictions a large popuation of UHECRs present in the radio lobes, medium-heavy nuclei would presumably have a further spread. The point is made that the rescaling of of the relativistic electron density only shifts the projection by $\sim 0.3^\circ$. There is no mention

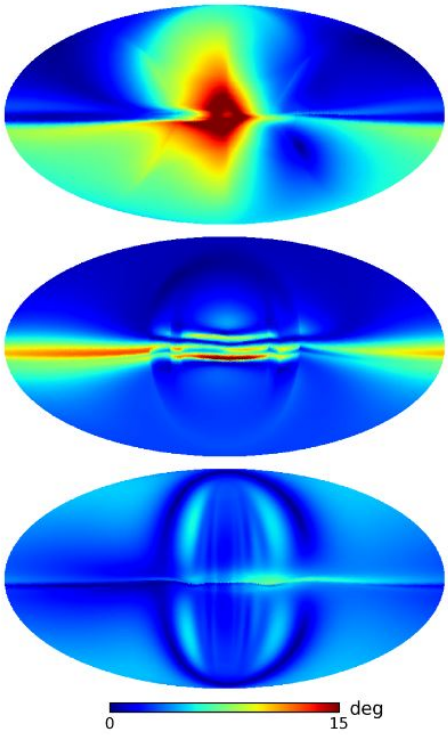

Figure 3.9: **Top: Predicted 60 EeV proton deflections for optimized BSS GMF model. The inclusion of the strong halo X-field deflects protons far more widely than other GMF models, particularly for the inner galaxy. The predicted deflections are proportional to the integrated transverse magnetic field along the trajectories. The deflections are asymmetric about the $z = 0$ plane. Middle: Sun *et al.* 2008 model. Bottom: Stanev 1997 model. From [23].**

of a UHECR reacceleration process.

Observations recently published by Carretti *et al.* [26] in Nature, report two large bi-conical, linearly polarized radio lobes with opening angles of $\sim 60°$, in a north and south outflow from the galactic bulge, transporting magnetic energy of $\sim 10^{55}$ erg. The lobes magnetic field strengths are $\sim 15\ \mu$G. Three major ridged structures appear to be winding around the lobes to the galactic west. A northern ridge, a southern ridge and a younger central 'galactic spur' that appears connected with the galactic bulge. Their magnetic fields are from $\sim (11 - 18)\ \mu$G. These ridges are similar in width, being 300 pc thick and have polarization fractions between $(25 - 31)\%$, the galactic spur which appears still attached to the central bulge having the lowest polarization fraction. All these structures are within the angle subtended by the molecular ring. A representation of these

structures is seen in Figure 3.10.

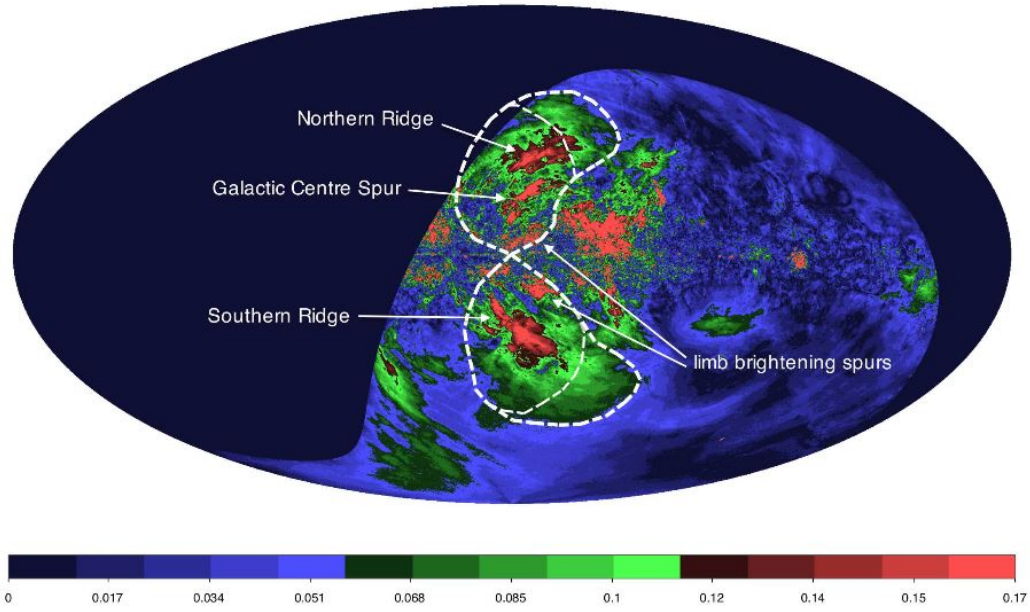

Figure 3.10: **S-PASS, $PI \equiv \sqrt{Q^2 + U^2}$. Linearly polarized intensity, PI at 2.3 GHz. Galactic projection. The map is in Galactic co-ordinates, with Galactic north $+z$, Galactic east on the left and Galactic west on the right. The polarized flux intensity is in units of $Jy/beam$ given by the colour bar, where 1Jy$\equiv 10^{-26}$ W m^2 Hz^{-1} and beam size 10.75'. Thick dashed lines outline radio lobes. Thin dashed lines outline γ-ray Fermi bubbles. Also indicated are the 3 polarized ridges and two limb brighting spurs. From [26].**

The ridges are highly collimated and Carretti *et al.* [26] speculate that they satisfy a topography wherein $B_\phi > B_\parallel$, i.e.'the magnetic structure is self confined'. On either side of the galactic bulge along the disk, are two distinct regions of depolarization extending $b \leq 10^\circ$ and limned by small-scale polar angle modulation. Carretti *et al.* leave aside the implications for transport/acceleration of UHECRs. The data images are from the S-PASS project run by Parkes. The telescopes resolution is $9'$, the frequency bandwidth 184 MHz, at a frequency of 2.3 GHz. These structures compare with the 23 GHz structures found in the WMAP7 data and γ-ray emission recorded by the FERMI telescope.

Chapter 4

Anisotropies for E $\geq 10^{17}$ eV

This chapter discusses anisotropies in the energy range, E $\geq 10^{17}$ eV, which is low enough for one to expect cosmic rays to be predominantly galactic in origin, and high enough to expect at some energy beyond, a transition to an extra-galactic regime. Whether this transition occurs around the second knee, and the ankle is then the signature distortion of proton pion-pion production on the CMB photons, or whether the ankle itself occasions the transition, is still in question. A custom for extra-galactic scenarios has been to take a minimum energy at just over ankle energies, but the latest research refining structures on the CR energy spectrum and composition studies at lower energies may reset that minimum [171, 130, 129, 84].

4.1 Signal Data Sets

Statistical analyses of data, E $\geq 10^{17}$ eV, tend to concentrate on generating mock isotropic backgrounds to compare with data, even at post-GZK energies where the data set is small and there exists extra-galactic LSS local to the GZK limit. The initial stage of analysis involves estimating a method's ability to reconstruct the selected signal, and its maximization. Signal data sets are used. These are then usually many Monte Carlo simulations of a background plus a signal set. The signal set is usually derived from the data set to be analysed. The signals are propagated within environmental settings or models that also require parameterization, or assigning of constants where we may only have a general idea of the actual values. These models entail informed speculation and

none consistently satisfy known facts and limits [61].The parameterization associated with models, such as GMF models, and enviromental settings, for example - the assignment of constants such as spectral index values, is to be regarded as external. Optimization is for parameters intrinsic to the method itself.

Where analysis methods are established, such as the Maximum Likelihood Method (MLM), the method, model setting parameters and constants, the data set and the statistical outcomes are briefly mentioned (see subsection 4.3.1 on page 90).

Where an analysis method has only been recently developed, as in several energy ordering methods (see subsection 4.3.3 on page 97), two studies [61, 194] are coveredin more detail.

The chapter is completed with an overview of a recent paper [61] (see section 4.4 on page 111). Although this paper concerns the energy range between $(10^{16} - 10^{18})$ eV, it forecasts expected anisotropies during the transition between largely GMF sources and purely EGMF sources. Because this transition connects us in a larger sense to the cosmos, anisotropies before and after are especially important.

4.2 Optimizations and Signal Data Sets

Departures from isotropy are routinely sought. This presupposes an isotropic background. When developing a diagnostic tool, Signal Data Sets are commonly constructed starting with this assumption. Optimizations represent a compromise between a maximized signal acceptance and reduction of background effects [168]. Data can be placed in a variety of environments/models. Isotropic data sets are generated and compared with the required signal overlaying the isotropic data within these environments. There is a type of co-dependence between models and signal optimization.

When we have different models (mainly GMF models), explaining similar outcomes, models are often tested to see by what degree the signal measured is affected by the model type. The impact of model types are evaluated, as are the parameter limits. For example, there are times when the galactic model type used does not greatly affect outcomes [114, 194, 168], or when only certain parameters of the turbulent GMF component, such as its amplitude, B_0, and coherence

length, L_c, play a significant role in affecting anisotropies [61]. Deciding to include unknowns as constants, or as threshold parameters, may or may not matter greatly in search outcomes.

Once a signal maximizing procedure establishes intrinsic parameters, the background data sets proper are generated, the method implemented, and the statistics of the signal parameters measured. The same procedure on data sets is also run and the statistical comparisons recorded. Naturally, the intention is to quantify statistically apparent departures from isotropy. Also the efficacy of an analysis may be quantified and improvements suggested.

Wherever possible, parameter numbers are restricted as we are interested, at this stage, in systems that are not too chaotic, and mainly in broad outcomes.

Another approach at GKZ energies, where it is clear that anisotropic large-scale extra-galactic structures exist, coupled with probable point-sources within these structures, is cross correlations with cosmological astronomical catalogues.

A priori analysis is desirable, because we are so limited in our understanding and have a tradition of reconstruction, *a posteriori* analysis is also common. While we can then quantify probabilities we are often unable to quote statistical significances, since *a posteriori* results contain unknown statistical penalties.

Concerning experiments published by Pierre Auger Observatory, the data conform with the requirements as set out in subsection 2.5.1 on page 44, although exposures quoted differ for data collected throughout the construction stage. When publications adjust their background data sets to an experiment's exposure, they are ensuring the background statistics are accurate [169].

4.3 Anisotropy Results for E $\geq 10^{17}$ eV

Because of the current conflicting arguments regarding just where in the CR energy spectrum (the second knee or the ankle) the galactic/extra-galactic transitions of CRs takes place, we take special interest in the energies just beyond the second knee in our search for anisotropies. We note there is an argument for purely galactic CRs and that in this case, the composition of an isotropic flux of UHECR primaries would need to be heavy [27] to ensure containment in the GMF.

4.3.1 Familiar Methods

A principal target amongst anisotropy searches has been our galaxy's central region (GC) which contains a supermassive black hole.

(**1**) : In 1999, reports came from AGASA [175] of an excess of events in the direction of the GC and a deficit towards its anti-centre. These events were at energies $\sim 10^{18}$ eV. They fitted the idea of diffusive propagation, and contributed to the supposition of the ankle as a transition point between GCRs and EGCRs. Later, the data set was enlarged by 25% and more comprehensive statistics published [172]. The energy range was roughly from $(10^{17} - 10^{18.3})$ eV. Harmonic Analysis in the **RA** distribution of events revealed an amplitude of $4.0 \pm 0.8\%$ in the direction **RA** $\sim 280^{\circ}$ in the first harmonic. This result was for a small region in the galactic plane about 10° from the GC [63] and the probability of this amplitude against an isotropic background was 10^{-5} [172]. The likely composition of the source primaries at these energies was declared to be probably light.

The data set events were for zenith angles, $\Theta < 60^{\circ}$, and, given AGASA'S northerly location, $(138^{\circ}30' \text{ E}, 35^{\circ}47' \text{ N})$, declinations south of -25° were invisible, i.e. the GC with a southerly declination $\sim -29^{\circ}$ was not well positioned. The GC excess significance was $\geq 4\sigma$ as was the galactic anti-centre deficit. The event density was in accord with that of a wide angle dipole distribution flow expected in a diffusive regime. The excess was not far from the Cygnus point-source of Bellido *et al* [195]. [172] do not rule out the possibility of a Cygnus point-source rather than a GC anisotropy. They also note the cluster labelled BC1 by AGASA is close by, where a point-source was reported at E $\sim 10^{19}$ eV. Their own point-source seaches in a 6° window found nothing.

(**2**) : Data collected from SUGAR, with the GC better positioned, saw for events in the energy range $\sim (17.9 - 18.5)$ EeV, an excess of about 20% for the GC at declinations -19° and -9° [63] just 6° off the AGASA data. The probability of this occuring by chance was 0.0005 [112]. [63] calculate any directionality from the GC would only be apparent in a small energy range because of competing diffusion from heavier primaries and a possible extra-galactic flux. The GC would be able to accelarate protons at energies $\sim 10^{18}$ eV as an upper limit.

(**3**) : Experiments also revealed the presence of TeV γ-rays in the region and buoyed hopes of detecting predicted GC neutrons in the EeV range that might account for the anisotropies of AGASA and SUGAR. A pp neutral pion decay at GeV energies would produce the γ-rays and neutrons. Neutrons undergo no magnetic field deflections and their directions would point back to the GC, where their rest half-lives of $\approx$ 12 minutes at EeV energies would have us within reach. [163] reported that their modelling did produce neutron fluxes consistent with the AGASA/SUGAR anisotropies.

(**4**) : The Pierre Auger Observatory, which also has the GC favourably positioned, performed an analysis of their own data with more events in 2007 [112]. The isotropic random data was generated in 2 ways, a shuffling method and a semi-analytic technique. Both approaches treating the same variable, the 'time dependence of Pierre Auger Observatory's acceptance', in different fashions. The backgrounds made had only a 0.5% divergence in size, with the shuffled data being larger. Nothing at a statistically significant level was detected. A neutron source search was also negative. Data analysed were from a construction phase of the array, 1st January 2004 - 30th March 2006. With regards to previously mentioned zenith binnings as a part of the Pierre Auger Observatory prescription's shuffling technique (see section 2.7 on page 53), a comparison of binning with no binning showed no significant difference [112].

Large-scale anisotropy searches other than the GC for available data continue to be conducted. At UHEs, data gathering takes years, and searches over wide areas need large data sets –

5) : A large-scale dipole anisotropy analysis for event energies, $E > 2.5 \times 10^{17}$ eV was performed by the Pierre Auger Observatory in 2011 [115]. An analysis of the flux distribution in **RA** aimed to find the first harmonic amplitude, r, and its phase, ϕ. The large data set at the Pierre Auger Observatory's original lower energies, E $\sim$ 1 EeV, allowed for detection of amplitudes as low as 1%. There was found to be a significant but very low anisotropy under 2% with a 99% Confidence Level (CL) at EeV energies. Figure 4.1 on the following page is the plot as published in [115]. An assumption of galactic loss of heavy nuclei at these energies depends on the size of the GMF and the distribution of sources. Very small dipole amplitudes have been predicted as cited in [115], in the range of a few percent. Extra-Galactic anisotropies from isotropically-distributed cosmologi-

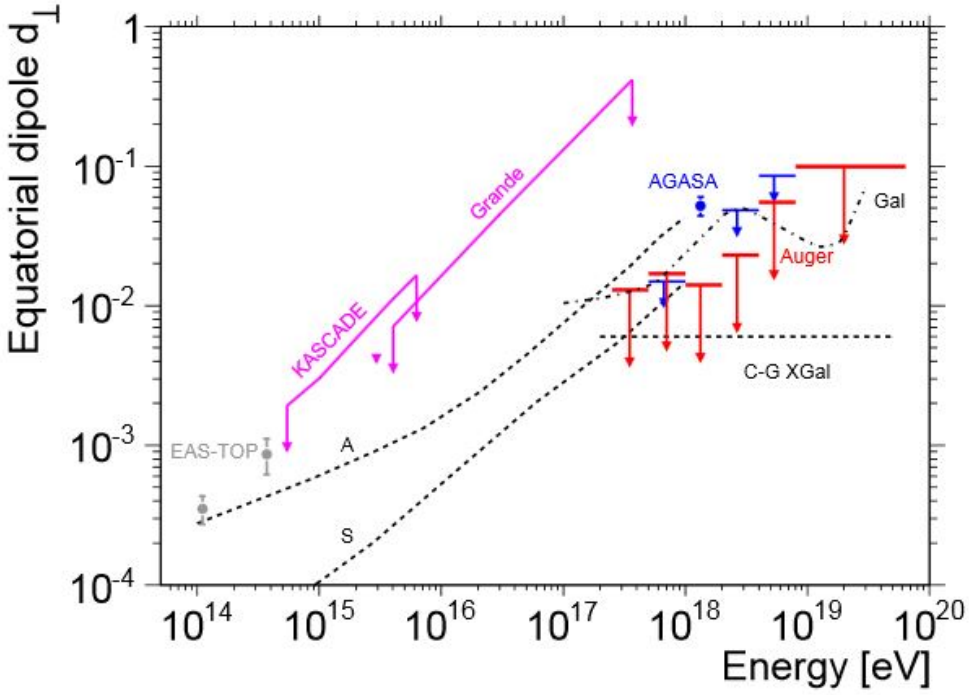

Figure 4.1: **Anisotropy amplitude results of the first harmonic as a function of energy for the Pierre Auger Observatory and 4 other experiments. Likewise predictions up to 1 EeV from an asymmetric (A) galactic model and a symmetric (S) galactic model. Shown are also predictions up to several 10^{19} eV for galactic sources (gal) and extra-galactic sources. The extra-galactic sources factor in the Compton-Getting effect on an isotropic background in a CMB rest frame (C-G Xgal). From [115].**

cal sources are also feasible in a *Cosmological Compton-Getting* effect (CCG) effect. In this case, a dipole anisotropy in the direction of **RA** = 168° with an amplitude of ~ 0.6% is predicted, when ignoring the GMF. It has been noted that anisotropies may be below detector sensitivity. There are not yet enough data, and [27] have calculated that as many as 10^6 events may be necessary, i.e. ~ 3× the data set as used by [115] could be necessary to establish such an anisotropy with a 99% CL. [61] describes predictions for anisotropies of nuclei at energies between $(10^{16} - 10^{18})$ eV with these results as a constraint (see subsection 4.4.1 on page 113). If this is so, the Pierre Auger Observatory's enhancements towards lower energies ~ 10^{17} eV (see section 2.6 on page 52) will be capable of supplying the necessary event direction numbers as will the Telescope Array (TA). A large dipole at these energies would indicate a cosmological contribution of CRs within an energy

loss horizon of Gpc [27].

Sometimes searches for one thing uncover something else –

(**6**) : AGASA found no large-scale anisotropy for energies $E > 10^{19}$ eV. However, a number of events were recorded as being within close proximity to one another (within a 2.5° radius) and above energies $E > 4 \times 10^{19}$ eV. A triplet of events and 3 sets of doublets have been recorded. The chance of detecting a triplet is against an expectation of 0.05 for such an event. These results were *a posteriori* with a 1% chance with respect to isotropy [154].

(**7**) : Similar regions were analysed by HiRes, with their own monocular data above $10^{19.5}$ eV using an antisymmetric autocorrelation technique [145]. They reported no correlation with a sensitivity on the scale of AGASA. Only 13% of HiRes events could be associated with clustering with a 90% CL.

4.3.2 Catalogue Searches

Catalogue searches directly incorporate anisotropies into analyses. Maps of the known sky are compared against CR directions. The higher the data energies, the less the particles undergo deflection in G/EGMF's and the closer their directions will be to their original source directions.

Catalogue maps are limited by our location in the galaxy (we don't get to see the GC and regions behind it), in the regions of sky mapped (northern or southern hemispheres), and often only include found objects in a certain wavelength (EGRET catalogue), or objects of a certain type (Véron-Cetty & Véron (VCV) catalogue - quasars and active nuclei) [173]. Astronomical mapping is restricted to the physics of technologies used, plus each device carries with it systematic errors and catalogue biases. Despite drawbacks, which in any case hold for any method, pattern detection in data, with cross correlations against astronomical catalogues is a valuable search tool. Some of the attempts to match data against catalogues are listed below.

(**1**) : Using two catalogues, the third EGRET γ-ray source catalogue, and the VCV 2001 catalogue, 8 out of 12 BL Lacertae (B Lac) objects that were loud γ-ray emitters were found to correlate with data from AGASA ($E > 4.8 \times 10^{19}$ eV) and for Yakutsk ($E > 2.4 \times 10^{19}$ eV). Again, the analysis was *a posteriori* [173]. Subsequently, the GMF models and CR composition

employed by AGASA were demonstrated to exhibit magnetic lensing effects [21], predisposing data to be backtracked to specific regions of the sky. Positive correlations with sources were thus likely to result [192].

(**2**) : HiRes data from similar regions and energies, E > 10 EeV, were analysed using the Maximum Likelihood Method (MLM). The correlations found for AGASA and Yakutsk were not found, but a different set of BL Lac objects did seem to correlate. The analysis was *a posteriori* [146].

(**3**) : For energies E > 10 EeV, Pierre Auger Observatory data from its surface detector, in the period 1st January 2004 - 31 December 2007, was searched for BL Lac correlations. This search was unsuccessful, correlations found were consistent with isotropic data, but comparisons between results for north and southern sites are not necessarily useful, given their distributions of astronomical objects are unrelated to one another [174].

(**4**) : A Pierre Auger Observatory campaign of searches using the 12th edition of the VCV catalogue which maps the local distribution of large-scale matter, turned up a number of AGN source correlations [106, 109, 120].

A prescription to minimize the possibility that correlations between CR directions and AGN in the VCV catalogue could occur when the CR flux were isotropic was devised from Pierre Auger Observatory data. 13 events taken between 1 January 2004 to 26th May 2006 were used, and the parameters of angular resolution, energy threshold, and maximum redshift were decided. The angular resolution was set at a maximum of 3.1°, the energy threshold energy, E ≥ 55 eV, and redshift $z \leq 0.018$. This redshift was set for 75 Mpc as the maximum radius of VCV objects considered. An *a priori* search was then performed on data taken between 27 May 2006 to 31 August 2007. 9 of these 13 events correlated, such that there was a 69^{+11}_{-13}% correlating fraction compared with 21% or 2.7 isotropic events, and a CL. of 99% of rejection of an isotropic hypothesis.

The data set was widened.

By this time, from data taken between 27 May 2006 to 31 December 2009 the Pierre Auger Observatory had recorded 69 events above 55 $\times 10^{18}$EeV [109]. The 13 events used in the *a priori* search, being excluded, the data set was 56 events and an *a posteriori* search was conducted.

Within the decided parameters, a correlating fraction $38^{+7}_{-6}\%$ compared with 21% for **assumed** isotropic distributions. 21 of the 56 events correlated, the isotropic correlation being 11.6. The probability of such a random correlation is 0.003%. The energy resolution of these events was $\sim$ 15% and the absolute energy scale has a systematic uncertainty of 22%. The angular resolution being better than 0.9° for events triggering at least 6 stations (E $\gtrsim$ 10 EeV). The integrated exposure for events was 20,370 km^2 sr yr. Of the 56 data set events, 9 lay within ±10° of the galactic plane, none of which correlated with objects in VCV catalogue. All catalogues are incomplete here because the Milky Way gets in the way. Also, it may be difficult to correlate directions because of large magnetic field deflections of CR trajectories along the galactic disk. The 5 most energetic events do not correlate with objects in VCV catalogue. There were found to be 18.8% of arrival directions within 18° of Centaurus A (Cen A) compared with 4.7% when flux is assumed isotropic. The galactic co-ordinates of Cen A are $(-50.5°, 19.4°)$. The search being *a posteriori*, no CL's were established.

(**5**) : The search for correlations with large-scale candidates was again widened to include not only AGN's but radiogalaxies and γ-ray bursts. Hence the 2MASS Redshift Survey (2MRS) and Swift Burst Alert Telescope (BAT) catalogues were referenced.

Once again the search was *a posteriori*. The data set now *did* include the 13 events taken between 1 January 2004 to 26th May 2006 previously used in the *a priori* search. These 69 events were cross correlated with galaxies within 200 Mpc in the 2MASS Redshift Survey(2MRS), excepting all the CR events within ±10° of the galactic plane. Flux models based on the 2MRS catalogue predict 13% of arrival directions to lie within 18° of Cen A [109]. This time, the 69 events were cross correlated with AGNs within 200 Mpc found at X-Ray frequencies by the Swift Burst Alert Telescope(BAT) [109]. Flux models based on the Swift BAT catalogue predict 29% of arrival directions within 18° of Cen A [109]. The CR background was assumed isotropic, even though there were small modulations of the exposure in **RA** induced by the growing array size and dead time of detectors [112]. The 2MRS and Swift BAT results were not tightly constrained. The correlations were within the GZK limit, supporting the GZK flux suppression.

Figure 4.2 on the next page is the plot of the 69 events along with the AGN correlations as

published in [109].

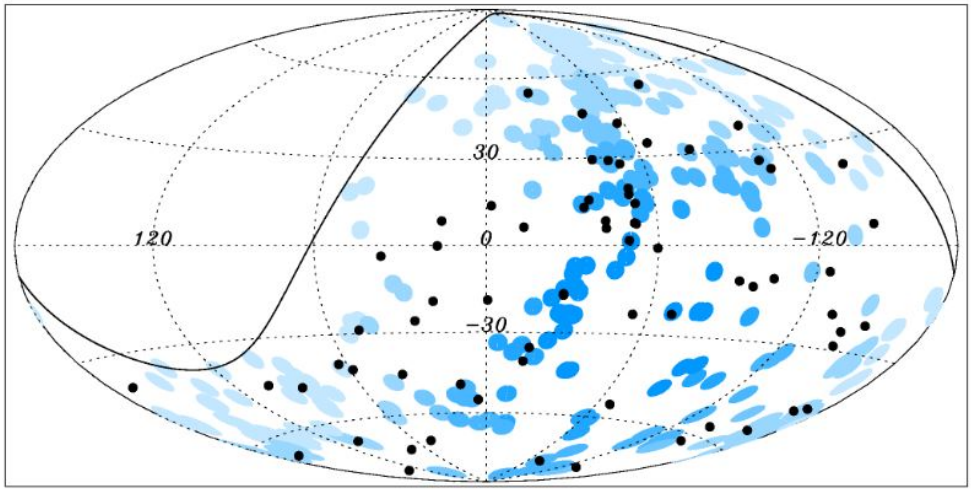

Figure 4.2: **Pierre Auger Observatory arrival directions for E > 55 EeV. 69 events were recorded until 1st Dec 2009. Black line shows the edge of Pierre Auger Observatory field of view for events with** $zn \leq 60°$ **Filled blue circles are** $3.1°$ **centred around 318 active galactic nuclei (AGN) in the VCV catalogue. All these AGN are within 75 pc of the field of view. The darker the blue, the greater the Pierre Auger Observatory's relative exposure. From [109].**

(**6**) : HiRes conducted a similar search for correlations with AGN and the VCV catalogue for its own data in the Northern hemisphere [147]. Parameters as set out in the Pierre Auger Observatory's *a priori* search above were repeated. Only 2 events out of 13 correlated, with an expectation of 3.2. The chance of this was 82%. When HiRes used its own parameters on one half of its data (*a posteriori*), established from the other half of its data to maximize its signal, a small signal with a chance probability of 15% was found. Another scan on HiRes's complete data set found an optimal correlation of 36 out of the total 198 events for parameters.The angular resolution was set at a maximum of $\Theta_{max} = 2°$, the energy threshold, E = 15.8 EeV and redshift $z = 0.016$. The chance probability was 24% and the results were consistent with isotropy.

The VCV catalogue is incomplete in both hemispheres, and the northern HiRes results may not really shed any light on Pierre Auger Observatory's positive correlations.

Cosmological catalogue searches are suitable for long-lived sources and/or small magnetic deflections [113]. Sources that are short-lived are difficult to correlate with recorded events whose

arrival time delay may number many thousands of years. Magnetic field deflections for light CRs from the same source, but at different energies have corresponding time lags. Magnetic field deflections of heavy CRs may be too broad to confidently identify single sources within a region populated by a number of putative sources, also, the heavy nuclei deflections could be large-scale and lack coherence in the inverse energy correlation needed. Signal dilution due to the incompleteness of catalogues is another drawback for this anisotropy search method.

4.3.3 Energy-Energy-Ordering

Establishing structures due to energy ordering in CR arrival directions is another search tool. It is expected that CRs of a common origin and with different energies, propagating through the same magnetic field, will be ordered in terms of their angular deflections. In the same magnetic fields, in a first approximation, CR angular deflections being inversely proportional with their CR energies for a composition species.

For standard source injection spectra, $\alpha > 2$, high energy events should be accompanied by some low energy events [168].This scenario becomes more complicated when the magnetic fields are random and when nuclei are heavy. Heavy nuclei detection has a greater dependence on GMF models than light [169].

In general, Energy-Energy-Ordering (EEO) is restricted to light nuclei and a dominant regular magnetic field. The source density distribution should not be too large, so as to provide separation between sources and avoid confusion.

Fundamental assumptions for EEO are

- there will be a subset of events or multiplet sets, related to the same source.
- the dominant component of the GMF is coherent and its turbulent component plays a lesser role.
- in extra-galactic scenarios, EGMF's are considered negligible.
- angular energy ordering, for δ the angular deflection, is $\delta \propto \frac{Z}{E}$.

For high energies, and small deflections, multiplet sets would satisfy certain geometries, and have a high correlation between the source direction and the inverse CR energy. Low energy particles' gyroradii, r_g, can scale with a magnetic field's irregularities, and the particles undergo diffusive propagation where their directions are scattered. At high energies, diffusive propagation is not expected and sources could thus be detected through their predicted rigidity based deflections.

In some instances, there is also the hope that the integral of the magnetic field orthogonal to the CR paths could then also be measured. This integral gives us an idea of the size of the GMF which, as yet, is poorly constrained.

Energy ordering methods have a number of variants, Some are mentioned below, along with any applied experimental results of such. Thus far, few of the applied methods have yielded significant statistics.

(**1**) : The Pierre Auger Observatory searched for linearity of directions against an isotropic background for data with threshold energies, $E_{th} > 10^{20}$ eV. For this method, the sources should outlast the delays wrought by propagation in magnetic fields that would occur in energy ordering, and the magnetic fields travelled be mainly uniform. The composition of the CRs would be proton or intermediate, as heavy nuclei deflections could be large-scale and lack coherence in the inverse energy correlation needed. The energy threshold has the majority of events occurring around the lower energies, where data sets are not so sparse and where isotropy of the sources is still expected - the known anisotropic large-scale extra-galactic structures requiring larger distances. This also obviates any reliance on astronomical catalogues.

The method relied on CR directions received having aspects characteristic of the coherence length, L_c, of the field's turbulent component. Assume a deflection, δ, produced by the fields regular component, where $r_g \gg L_c$. The displacement, $\delta_{rms}L(E)$, experienced by the particle having travelled distance, $L(E)$, in the turbulent component, is differently configured/arranged according to whether or not $\delta_{rms}L(E) < L_c$ or $\delta_{rms}L(E) > L_c$.

Considering $\delta_{rms}L(E) < L_c$, the displacement is radial around the regular components' deflected source position. This is the scenario for CR energies $\gg$ 20 EeV, with distance $\sim$ 2 kpc.

Considering $\delta_{rms}L(E) > L_c$, all deflections can be expected to be shifted by amounts increasing with the displacement of the regular component's source position. A roughly linear alignment of CR directions would result. The small angular deflections of UHECR primaries caused by the turbulent component yield a linear approximation between the perceived source direction, $\vec{\theta}$, the actual source direction, $\vec{\theta_s}$, and the integral of the magnetic field's orthogonal component to the CR path.

The data set taken was for 1,509 events from 1 January 2004 to 31 December 2010 [114]. One 12-plet was located as were 2 decuplets. Comparisons with the isotropically distributed mock data sets were disappointing - 6% of these had at least one 12-plet or greater, and 20% had a minimum of 3 decuplets.

(**2**) : Another magnetic field search tool is 'Energy-Energy-Correlation' (EEC). An energy-energy correlation method was developed by [166] and applied to Pierre Auger Observatory data > 5 EeV [167]. This correlation does not assume a geometry in arrival directions, only that source events above a threshold energy will possess a close angular relationship with the original source.

Random events, i, j, are not expected to display strong energy correlations. Given that we are seeking point-sources, the sky has appointed Regions Of Interest (ROI's) - in assigned solid angle projections of 11.5° on a galactic $(\mathbf{l}, \mathbf{b})$ co-ordinate system. Events belonging to a ROI are identified by their energies and their angular distance to the centre of the ROI, which is an initial event E > 60 EeV. The energy-energy correlation, Ω_{ij}, between events i and j, with angular distances, $\alpha_{i/j}$ to the centre of the ROI, and with average energy, $\langle E_{i/j}(\alpha_{i/j})\rangle$ for all events is defined as

$$\Omega_{ij} = \frac{(E_i(\alpha_i) - \langle E(\alpha_i)\rangle)(E_j(\alpha_j) - \langle E(\alpha_j)\rangle)}{E_i(\alpha_i).E_j(\alpha_j)} \tag{4.1}$$

and is based on the expectation that when any two energy events in a ROI, one $E_i > \langle E_{ROI}\rangle$ and the other $E_j > \langle E_{ROI}\rangle$, result in $\Omega_{ij} < 0$ - this is attributed to typical background events. When both $E_i/E_j > \langle E_{ROI}\rangle$, or both $E_i/E_j < \langle E_{ROI}\rangle$, result in $\Omega_{ij} > 0$ - this is attributed to source related events.

Signals from the same source would display an average Ω_{ij} that is greater than an average Ω_{ij} from a random background.

This EEC method was applied to Pierre Auger Observatory data, $E > 5$ EeV, over the period 1 January 2004 to 31 December 2010, which consisted of 18,744 SD events. No anisotropies were found [167]. This could be attributed to a higher source density than assumed, or larger magnetic field deflections.

(**3**) : An energy-energy ordering method, developed by [168], searched for bright, high energy proton/light nuclei point-sources in sectors of the celestial sphere. If the GMF is largely responsible for angular deflections of nuclei, near positions of UHECR signals - to one side of a source, there may appear a shifted sector shaped distribution of lower energy events or 'tail'. The vertex of the sector would be the source and the opening angle size a ratio of regular and turbulent magnetic field deflections.

[168] searched for doublet/clusters of events for energies EeV > 19.8, and the approximately linear structure of such a tail. Searching for multiple events lowers the background density requirements. The method and assumptions, e.g. a negligible EGMF and a GMF with uniform and tubulent components, was similar to that outlined in [114] in their search for multiplets. The GMF model was bi-spiral with a toroidal and poloidal halo.

Magnetic field deflections were considered with the regular component deflecting protons along a curve as $\frac{1}{E}$, and the turbulent component deflecting protons around their original uniform deflection. The proton is, in effect, deflected as $1/E$ times a proportionality factor, nominated the deflection power of the regular GMF.

For high enough energies, outside the galactic disk, the curve is almost straight. The disk itself is not considered in the point search for the usual reason of the uncertainty surrounding the value of its magnetic field. Sectors in the sky were chosen that had 2 or more UHECR events > 60 EeV. Low energy events within the sector were sought that maximized the correlation between each UHECR event direction and its energy inverse. Thus sectors of the sky were searched for evidence of EEO.

Background event simulations followed the TA exposure, declination $\delta = 40.1°$, and $zn\ \theta < 45°$ [149], and their energy distribution followed that of HiRes results [218]. For energies > 10^{19} eV, data set numbers were from 2,500 to 100,000. For sources, they used 2 power law spectral

indices, $\alpha = 2.2$ and 2.7, and set the proton source distances $\sim (50 - 100)$ Mpc. The usual propagation energy losses were incorporated. The source maximum energy was $E_{max} = 10^{21}$ eV and the signal modified in accordance with TA exposure.

The source position was reconstructed to within 1°, 68% of the time, though dependent on the parameters of source luminosity, magnetic field strengths and deflection strengths. This bright point search method, now optimized and tested, was then extended to nuclei sources [169].

(**4**) : [169] used the Pierre Auger Observatory data, $E > 55$ EeV, previously published in [109]. While there was evidence on some of the data being ordered in the Virgo cluster direction, this analysis was *a posteriori*.

The problem of multiple and irregular source images posed by nuclei and their large angle deflections can be reduced under certain conditions to a scenario of a roughly Z/E angular deflection. - an 'enlarged proton-like' image. This scenario is dependent on simplified cases of the GMF and source position in the sky. The nuclei are typically taken as iron. The method is a modification of that in [168]. The data set used was the 69 events with $E > 55$ EeV, previously published in [109]. [169] makes special mention of a region they name the 'Cen A' region - a region $(-60 \lesssim l \lesssim -30)^\circ$ and $0 \lesssim \mathbf{b} \lesssim 30^\circ$. This is the location of the Centaurus cluster plus a possible independent source, Cen A, which was formerly promoted by Pierre Auger Observatory [109]. [169] feature Cen A as a region of anisotropy. As a large-scale anisotropy, this region has been previously scanned by [109] with a 4% significance upon application of the KS test, and by [170] with a 3%/2% significance upon application of the 3-pt/4-pt autocorrelation functions (see 4.3.4 on page 109).

The GMF model used is by Smida and Prouza (PS). For The Cen A region of events, the reconstructed source was close to the Virgo cluster of galaxies. The powerful active galaxy M87 is embedded within this cluster. [169] view the Virgo region as an extended source, with a radius set at 5°, and its GMF's all combining in a spread of CR deflections such that even one source would illuminate the region. The probability of such a reconstruction anywhere in the sky is $\sim 7 \times 10^{-3}$, and the probability of such a reconstruction of a source within 10° of M87 is $\sim 4 \times 10^{-3}$. The combined probability of such a reconstruction with a cluster of events, plus a reconstruction of a source within 10° of M87, is $\sim 3 \times 10^{-5}$.

The analyses was *a postiori*. See Figure 4.3 for an image of deflected iron nuclei from Virgo.

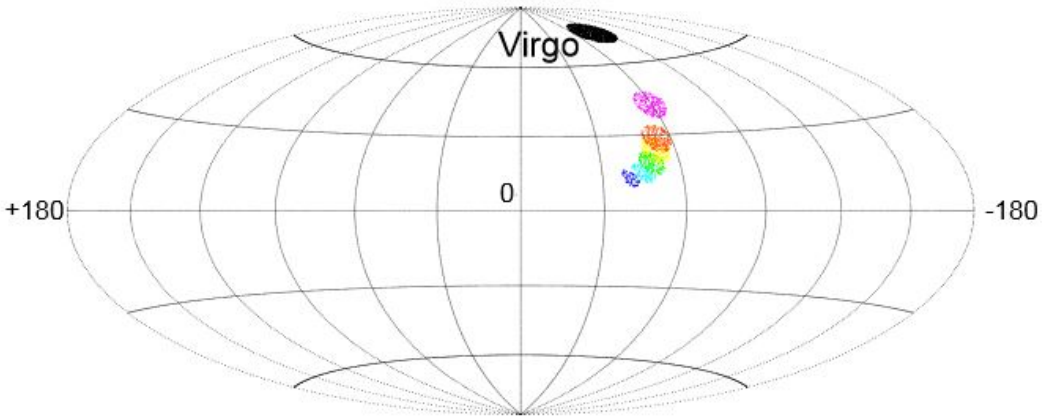

Figure 4.3: **Deflected arrival directions in a regular GMF of iron nuclei from the Virgo cluster. The addition of a turbulent GMF would spread out deflections in the Cen A region for the lower energies, and arrival directions would be similar to those featured in the Pierre Auger Observatory data. Colour code: dark blue for $E > 60$ EeV; light blue for $E > 70$ EeV; green for $E > 80$ EeV; yellow for $E > 90$ EeV; orange for $E > 100$ EeV; red for $E > 120$ EeV; and magenta for $E > 140$ EeV. From [169].**

[169] consider alternative explanations for the composition of the outcome. They cite a possibility that the Cen A regions source (Virgo) could also emit protons with a wider angle of deflection if the EGMF is larger than the value considered.

The heavy nuclei scenario is appropriate for their values of a small EGMF and a large, dominant GMF, with a smaller turbulent component. Outside the Virgo radius, deflections would be minimal until the CR nuclei enter the GMF. At the high latitudes of such a source (away from the galactic disk), where the GMF is parallel to the disk, the source position would be shifted steadily in terms of galactic longitude, independent of galactic latitude. The Cen A region, the 142 EeV event, and Virgo are all located at similar galactic longitudes. Events near the galactic bulge are expected to be dominated by a perpendicular GMF.

Upon modelling deflections of iron nuclei, such that E > (60 – 140) EeV in GMF models with a turbulent component, [169] demonstrated these deflections angular spread could account for the

Auger events in the Cen A region. This actually relocates the previous identity of the source region as Cen A or the object Cen A [109], to a source region of Virgo, capable of proton or heavy CR production. If the source region is Cen A the composition of nuclei would be proton in small EGMF's and if the source is the more immediate Cen A itself, the composition would also be proton.

[169] included large halo pitch angles from -40° to -60° in their calculations and noted limits introduced by heavy nuclei with respect to the regular disk field. The field has maximum impact on these nuclei in the region $30^\circ \lesssim l \lesssim 40^\circ$ for a disk height of $\gtrsim 1$ pc. They also noted that sources near the galactic poles would be hard to identify/invisible because of the purported toroidal and dipolar components of the GMF which, if they exist, are unknown.

[169], in commenting of the anisotropy of region Cen A, note that whatever the GMF and EGMF configurations, heavy nuclei with energies of $(60 - 80)$ EeV act as $(2 - 3)$ EeV deflected protons. If Cen A produces protons and heavier nuclei there should be an anisotropy at these lower energies, which according to the Pierre Auger Observatory, there is not. In turn, the lack of anisotropy could be attributed to a harder source spectrum, or the ratio of nuclei to protons may be around one-to-one. The mean free path of nuclei at Cen A region event energies is > 100 Mpc and a Virgo source would not have many distintegrate into protons. Nuclei mean free paths at energies a few times larger than 10^{20} eV fall short of Virgo. These two competing factors constrain CR acceleration models. An additional factor in support of the Virgo heavy nuclei scenario is a comparable flux of γ-rays seen by HEGRA around M87.

(**5**) : HiRes searched for arcs of directions in energies $> 10^{19}$ eV, [144]. This method was based on the use of Hough transforms, which involve image retrieval from a background of local measurements - 'noise' [210, 211]. Boundary features of these images must be available in a parametrized form, such as the conic sections, which all can be retrieved through the geometries of the conic shapes [212]. The HiRes interpretion of a Hough technique relied on each individual event direction in a data image space having a corresponding great circle set at right angles to that circle in a parameter space. We see an illustration of the *general* idea in Figure 4.4 - data set events lying in an arc in the (**x,y,z**) co-ordinate space undergo a transformation to a parameter space that

sees their corresponding circles intersecting at two points. Every event in a data set of the (**x,y,z**) co-ordinate system is mapped to a great circle in a Hammer-Aitoff projection of the **(l,b)** 'sky'. The **(l,b)** Hammer-Aitoff projection was binned into 4,584 areas of $6^\circ \times 6^\circ$ and the great circles were projected onto this binned space. Each great circle will pass through an individual set of bins. Intersecting great circles will pass through the same pair of $6^\circ \times 6^\circ$ bins. The number of great circles passing through these bins is tallied. The number *in* each bin tally is the number of great circles intersecting. The size of the bins are an indication of an allowed intersection error.

Data were taken from 271 HiRes stereo events collected from December 1999 and January 2004. Full detector geometric acceptance and exposure were adjusted according to HiRes on-time. The data set and the MC simulations were compared for anisotropies.

No departures from anisotropy were found. Illustration of the Hammer-Aitoff binning is seen

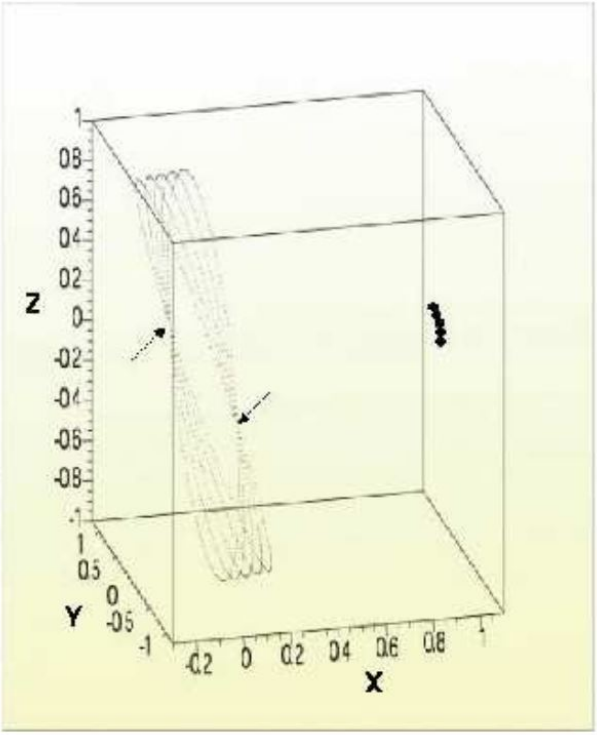

Figure 4.4: Generalized sketch of the great circle hough transforms of CR directions lying on an arc. The CR directions lying on an arc are seen as dots and each corresponding great circle intersection is illustrated, but not very well. From [144].

in Figure 4.5.

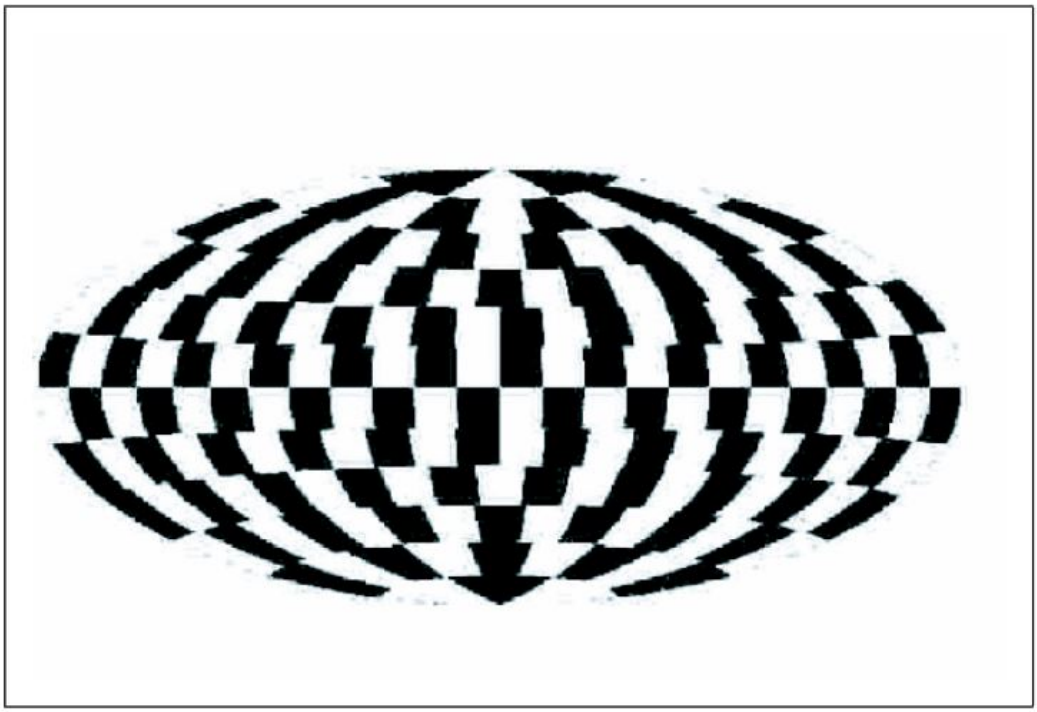

Figure 4.5: Counting bins for great circles mapped to an Hammer-Aitoff projection: From [144].

(**6**) : Energy ordering using a magnetic spectrometer method (MSM) has been developed by [194] and applied to Pierre Auger Observatory data with energies > 55 EeV. The advantages of this method, are, it is virtually independent of magnetic field models in its source reconstruction, and experimental uncertainties are directly factored into the source reconstruction. An estimation of the magnetic field at right angles to the line of sight, $|\mathbf{B} \times d\mathbf{l}|$, should also be possible.

This work was based on the assumption of a dominant single species, the detection of light nuclei, to avoid confusion, for instance, of the possible occurance of several source images that may obsure heavy nuclei source detection.

Fig 4.6 on the following page illustrates an idealized magnetic spectrometer.

CRs travel in a plane at right angles to a uniform magnetic field, **B**. So a CR, from source to detector, placed at a distance, D, is at an angle, θ, to the source such that

$$\sin(\theta) = \left(\frac{ceD|\mathbf{B}|}{2}\right)\frac{Z}{E} \tag{4.2}$$

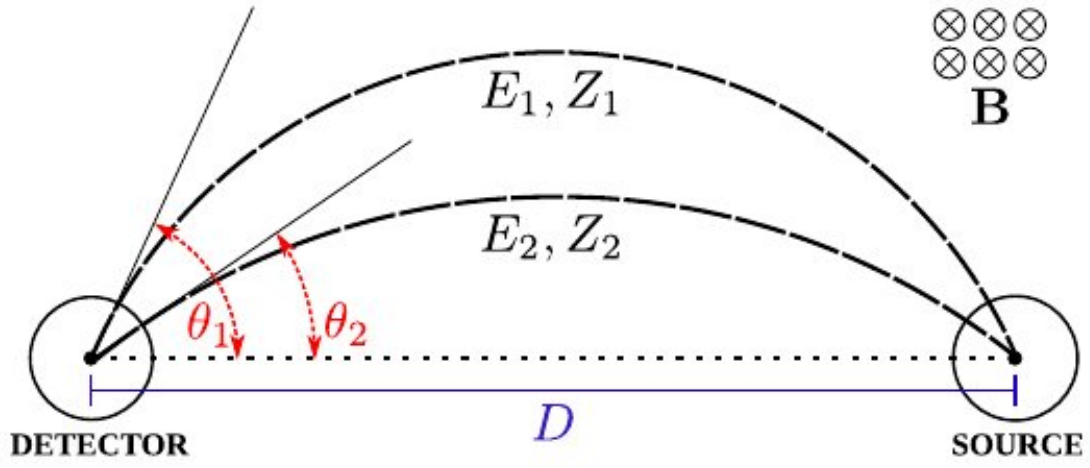

Figure 4.6: **Illustration of the basic geometry involved in a magnetic spectrometer. CRs from a common source, propagating at right angles through a uniform magnetic field, B, arrive at a detector, a distance D away. The angular separation between a source direction and the arrival direction of the CR at the detector is a function of the energy of the CR. The energies are such that, $E_2/Z_2 > E_1/Z_1$. From [194].**

the proportionality constant, $\frac{ceD|\mathbf{B}|}{2}$, is geometric, and the factor 2 is due to the angle, θ, being between the CR arrival direction at the detector and the CR source, NOT the CR arrival direction and its direction upon leaving its source. Consider two related CR events, with species and energy number Z_1, E_1 and Z_2, E_2 and $Z_1 = Z_2$ when $E_2 > E_1$ it follows $\theta_2 < \theta_1$. Thus, from Eqn. 4.2 we have, subject to small deflections

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{E_2}{E_1} \tag{4.3}$$

which results in

$$\theta_1 = \sin^{-1}\left(\frac{E_2}{E_1}\right)\sin(\theta_2) \tag{4.4}$$

Introducing energy ratios frees dependence on detector energy calibrations. The difference in angular arrival between the two events being known, i.e.

$$\Delta\theta = \theta_1 - \theta_2 \tag{4.5}$$

or $\theta_2 = \theta_1 - \Delta\theta$, we can talk in terms of one unknown, θ_2

$$\theta_2 = \tan^{-1}\left(\frac{\sin(\Delta\theta)}{\frac{E_2}{E_1} - cos(\Delta\theta)}\right) \tag{4.6}$$

Each event in a data set is paired with every other event, and the spatial environment is a cartesian co-ordinate system, $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$.

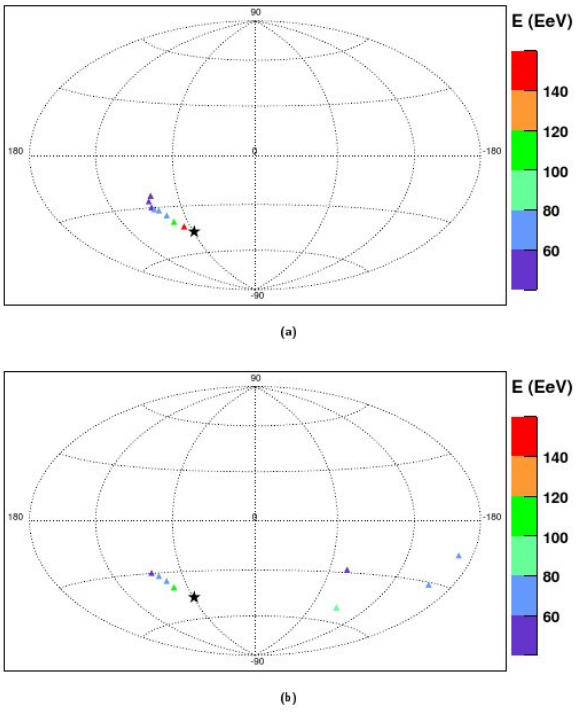

Figure 4.7: **a) Proton flux b) half proton and iron flux.**
Both fluxes are from the same source, ⋆, and from b) it is clear the arrival directions of the mixed composition CR nuclei are significantly deflected. From [194].

θ_1 and θ_2 are found from substituting Eqn. 4.6 into Eqn. 4.5 on the preceding page.

Significant clustering of source directions indicate a possible common source. This concept of source clustering, rather than certain geometries of event clustering being satisfied, e.g. linearity, frees the analysis from the introduction of multiple geometric parameters and their optimizations.

Factoring in the Pierre Auger Observatory experimental uncertainties (see subsection 2.5.1) results in distributions of reconstructed source positions.

Before optimizations of the signal data set, extrinsic parameter prescriptions are decided and representative features of the skymap, S_{ij}, that are compiled from application of the MSM to data are quantified. For example, the energy threshold is $E > 55$ EeV and pairs are separated by no less than a parameter angle $\Delta\theta_{cut}$ (initially set as 90°). Paired events are rejected if $\frac{E_2}{E_1} < 1.02$ or/and if $\Delta\theta < 2°$, this being the considered scale of turbulent magnetic field deflections. A skymap, S_{ij}, of directions is compiled, i.e., each pixel represents the integral of the number of sources at that position (i, j), within the radius angle ζ_{smth}, and peaks in S_{ij} are sought.

Source positions are placed within 100 Mpc of the GC. Signal events are propagated through a number of symmetric and antisymmetric spiral GMF models. The GMF models used do not describe the GMF near the GC and particles entering within 4 kpc of this region are discarded, as are particles which are not collected by the particle detector, a 100 pc sphere centred on the Earth. These particles exit the galaxy. The signal data is initially analysed separately to get an idea of parameter optimizations.

Once the signal data is analysed it is combined with background data as the Signal Data Set.

Sources of mixed composition were considered as part of the Signal Data Set. They were generated in a similar fashion to those original to the Signal Data Set. Source fluxes of pure proton were compared with sources of 50% proton and 50% iron.

The data set used was the 69 events with $E > 55$ EeV, detected between 1st January 2004 and 31st December, 2009. A maximum correlation at this energy threshold was found between arrival directions and AGN in the VCV catalogue (see subsection 4.3.2 on page 94).

$\Delta\theta_{cut}$ was varied in increments of 5° between $15° - 45°$. ζ_{smth} was $5°, 7.5°$ or $10°$.

See Figure 4.8 on the following page for an illustration of the MSM on Pierre Auger Observatory data.

The data peak value was found at the location $(\mathbf{l}, \mathbf{b}) = (-66.4°, -58.8°)$ with 4 events. The chances of such a signal out of an isotropic data set was $< 0.2\%$. The Virgo cluster is $\sim 16°$ from the peak location. A $\Delta\Theta_{cut} = 40°$ is twice as large as a cut considered for proton sources, this cut fits with a larger scale deflection for intermediate particles in the GMF models considered. The largest energy ordering was found to be in the first 27 CR event arrivals. Later epochs of

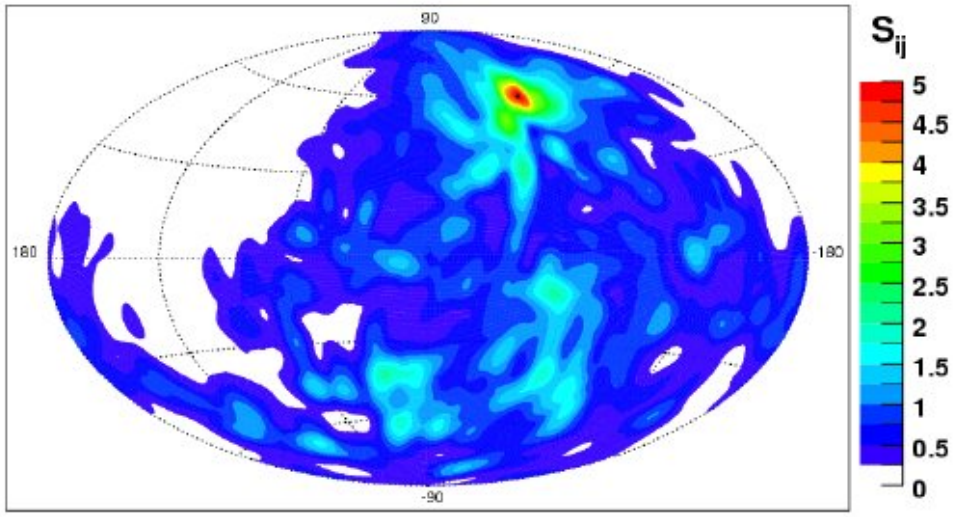

Figure 4.8: **MSM $\zeta_{i,j}$ skymap of 69 arrival directions with the optimal parameter cut of $\Delta_{cut} = 40°, \zeta smth = 5°$. This yields a peak direction (black dot) of $(\mathbf{l}, \mathbf{b}) = (\mathbf{-66.4°}, \mathbf{-58.8°})$. From [194].**

CR event arrivals were consistent with isotropy. [194] noted the large contribution of the highest energy event, $E > 142$ EeV to the consideration of anisotropy.

4.3.4 Self Correlation Methods

In the wake of the dilution of correlations in arrival directions of Pierre Auger Observatory results with AGNs in the 12*th* edition of the VCV catalogue (see [106] and [109]), another approach to analysis of Pierre Auger Observatory event directions was tried by [113] in 2012. Our local universe has inhomogeneities in the form of clusters and superclusters and it is reasonable to expect anisotropies in arrival directions to manifest if these source densities are able to be separated, i.e. source densities are low. Three search methods, variations on self-correlation methods, were developed: the 2pt-L analysis, the 2pt+ analysis and the 3pt+ analysis. These self correlation methods were first applied to mock data sets assembled as set out below, and then, their ultimate analysis was performed on data taken from January 1, 2004 to July 31, 2010 for 110 events in the energy range E > 49.3 EeV, where the composition of CRs is uncertain and G/EGMF's are also unknown. Results for these methods on signal data sets were promising but yielded no positive

results when applied to the Pierre Auger Observatory directions.

- SIGNAL DATA SETS: The work of [113] was directed at testing the efficiency of chosen methods on different models, characterized by both large-scale and small-scale anisotropies. Mock data sets of 60 events following Pierre Auger Observatory's exposure after 2 years, and for energies, E $\approx$ 50 EeV, were generated for each model. Monte Carlo sets of simulated isotopic data following Auger exposure constituted the background of the Signal Data Sets.

- MODEL CONSTRAINTS: Mock data sets were built where the number of extrinsic uncertainties involved were condensed to constraints such that:

  - Volume densities of sources were set within ($10^{-3.5}$, $10^{-4.0}$ and $10^{-4.5}$) Mpc^{-3}.
  - the sources spectral distributions were set at $\mathrm{E}^{-2.3}$
  - energy losses that determined the horizon were
    - redshift
    - photo-pion
    - pair-production

- OPTIMIZATIONS: Each data set arrival direction had a number of fixed Gaussian density distributions (smearing) in the range of $1^\circ - 8^\circ$. The Pierre Auger Observatory's angular resolution at these energies is $\sim 0.8^\circ$. The CR composition was considered to be proton. Heavier nuclei would also involve photo-distintegration. The larger angle deflections of 8° are appropriate for heavier composions. Each data set energy resolution was smeared with a Gaussian density distribution around the original energy, E, with R.M.S., σ_E, such that $\sigma/\mathrm{E} = 10\%$. This was to reflect experimental and systematic uncertainties.

  Smearing angles $> 8^\circ$ yield isotropy as does the setting of lower energy thresholds which increase the CR horizon and dilute signals.

- METHOD: The effect of magnetic smearing and data size, with regards to angular and energy resolutions on self-clustering (doublets and triplets) in arrival directions, was sought. Comparisons with many random data sets of expected results at Pierre Auger exposure were

made. This procedure was to establish whether searches for anisotropies within given resolutions were viable for so few data. Angular and energy resolution affect results. Angular resolutions can smooth out evidence of signal clustering and energy resolutions can permit lower energy events being admitted into data sets.

The effects of the Pierre Auger Observatory's angular resolution and energy resolution at these energies were modelled and smearings applied.

The three methods when applied to mock data sets with high source density and large smearing angles are expected to produce poor anisotropy results, i.e. have small powers. Those with small source density and small smearing angles are expected to yield larger power results. Each method was used on set data numbers over magnetic smearing ranging from $1^\circ - 8^\circ$ and source densities of $10^{-4.5}$, $10^{-4.0}$ and $10^{-3.5}$.

- CONCLUSION: The methods when applied to the mock data sets were sensitive to large-scale anisotropies in the local universe and GZK energy losses for ranges in the parameters of magnetic deflection and source density. However, the methods, when applied to the Pierre Auger Observatory directions, yielded no positive results for clustering.

4.4 Predicted Anisotropies Around 10^{18} eV

KASCADE and KASCADE-Grande have published the most complete statistics thus far on data between ($10^{16} < E < 10^{18}$) eV [171, 82, 83]. They have analysed CR composition within the uncertainties of hadron interaction models, and this, plus the consequent energy density spectrum, is relevant to the location of sources. The all particle energy density spectrum exhibits a hardening at $\sim 2 \times 10^{16}$ eV and a steepening at $\sim 8 \times 10^{16}$ eV [171]. That the energy density spectrum exhibits contrasting structures within this energy range, where hitherto, we had assumed a simple rigidity based decline in CR flux, yet again shifts attention from the ankle as the galactic/extra-galactic crossover. Galactic and extra-galactic sources should be different. Their energy based composition should also be different due to propagation effects.

Extra-galactic propagation would require steady point-sources of enormous power or large-

scale clustering effects, both extrinsic to the galactic regime. Considering they had gone as far as they could with these two key CR indicators, [61] modelledthe expected anisotropies for different galactic primaries in this energy range against the dipole anisotropy as measured by the Pierre Auger Observatory [115].

Their actual data leaves off where the Pierre Auger Observatory's begins, but a discussion of anisotropies with respect to to the galactic/extra-galactic transition period includes their work and informs interpretations of data in higher energy ranges. The subsection 4.4.1 starting on page 113 briefly discusses some of their results.

Wherever the transition point, extra-galactic CRs would start off proton rich [17, 18], and eventually give way to heavier nuclei. The extra-galactic spectrum would continually steepen until all putative nuclei reach their GZK limits, and thereafter, no more nuclei would be detected, unless some new physics is introduced, or 'top-down' decay scenarios are rehabilitated (see page 28). The galactic diffusive regime of CR propagation by frequent SN sources peters out at around E/Z ≥ ($10^{16} - 10^{17}$) eV. The iron knee as quantified by KASCADE is ∼ 1×10^{17} eV [87]. The end of the actinides, i.e. Z = 92, occurs ∼ 4×10^{17} eV, which is the location of the second knee [88]. There is a predicted dip in the CR energy spectrum where EGCR protons scatter inelastically on the CMB radiation. The proton dip for a uniform density of sources appears at ≈ 1×10^{19} eV. The attenuation length of EGCR protons at this energy is ∼ 30 Mpc[1]. The predicted proton GZK cutoff is at ≈ 6×10^{19} eV. For a uniform distribution of sources, $E_{1/2} = 5.3 \times 10^{19}$ eV [98]. For the interval between ($10^{17} - 10^{18}$) eV, the EGCR protons may randomize on the EGMF. There should be other 'dips' in the energy spectrum caused by nuclei interacting with the CMB and CIB, before nuclei reach their respective GZK cutoffs.

If the transition is the ankle, at around these energies, the CR composition should be light. A light composition is supported by results from HiRes [176] and TA results [177, 8]. In contradistinction, results from the Pierre Auger Observatory indicate a heavy composition around ankle energies [117], and the muon count from Yakutsk [178, 169] reinforces this report of an appreciable fraction of heavy nuclei at ≃ 10^{19} eV. KASCADE-Grande results point to a fast rising proton fraction for E > 10^{17} eV, and at the same time, heavy nuclei for E ∼ 10^{18} eV.

[1] i.e. if protons form the composition of the dip, their sources should lie within 30 Mpc

Interpretations vary. Could a non-random proton extra-galactic portion of flux start around $E > 10^{17}$ eV, but by an energy $\sim 10^{18}$ eV be overlain by a non-random heavy extra-galactic component of flux, or is this heavy nuclei flux still galactic and the proton flux happens at an energy where there is no rigidity based galactic contribution of nuclei? Or possibly KASCADE-Grande results are in error for one or both these energy positions. Or perhaps the extra-galactic proton contribution starts at the hardening at $\sim 2 \times 10^{16}$ eV, the steepening at $\sim 8 \times 10^{16}$ eV is galactic, as is the knee and second knee at at $\sim 4 \times 10^{17}$ eV - these galactic contributions washing out the extra-galactic portions until the proton dip at $\sim 1 \times 10^{19}$ eV.

4.4.1 Galactic/Extra-Galactic Energy Constraints

Before fitting energy constraints into a galactic/extra-galactic anisotropy scenario, [61] established lower and upper limits of anisotropy for light, intermediate and heavy CR primaries. The particle primaries are p, He, C, Si and Fe, with C representing the intermediate CNO group. [61] propagate UHECR nuclei p, He, C, Si and Fe through 2 different GMF benchmark models (see subsection 3.4.2.1 on page 74 [24]) that fit with extra-galactic RMs at the time. The GMF is proposed to consist of 2 components: The large regular magnetic disk field, and a turbulent halo field which is taken to be either exponentially decaying with distance from the plane. or constant. Their upper limits are constrained by the published Pierre Auger Observatory's dipole anisotropy amplitude as a function of energy [115]. These upper limits on the dipolar amplitudes obtained for EeV energies are below 2% at a 99% CL. A range of energies above 2.5×10^{17} eV were used. The upper limits of the anisotropy amplitude of first harmonic dipole as a function of energy as published by the Pierre Auger Observatory and other experiments is displayed in Figure 4.1, and the results are briefly mentioned in subsection 4.3.1 on page 91.

An ankle galactic/extra-galactic transition still allows a diffusive/drift motion escape of UHECRs at $\sim$ 1 EeV provided the composition is heavy - departing light composition particles at ankle energies exceed the Pierre Auger Observatory's anisotropies.

A second knee galactic/extra-galactic transition ($\sim 5 \times 10^{17}$ eV) will have predominantly extra-galactic CRs at UHEs of 1 EeV. If the rest frame of these extra-galactic CRs is also the CMB

rest frame, a putative extra-galactic isotropic contribution would exhibit small *Compton-Getting* anisotropies. The GMF being factored out, this anisotropy would be dipolar in the direction of **RA** = 168° with a 0.6% anisotropy. The presence of a GMF would involve higher order harmonics and statistics three times the ones used.

[61] assume a continuous, isotropic source distribution (mainly transient SN's)in the the galactic region -200 pc $\leq z \leq 200$ pc and $r \leq 20$ pc. For E $\geq$ 1 EeV (Pierre Auger Observatory anisotropy measures), they backtrack the random UHECR anti-particle trajectories from the Earth to sources and count their length. They then calculate the dipole amplitude of CR anisotropies.

They find an intermediate galactic contribution is limited to $\sim 1\times10^{18}$ eV and a galactic heavy iron contribution to $\sim 3\times10^{18}$ eV.

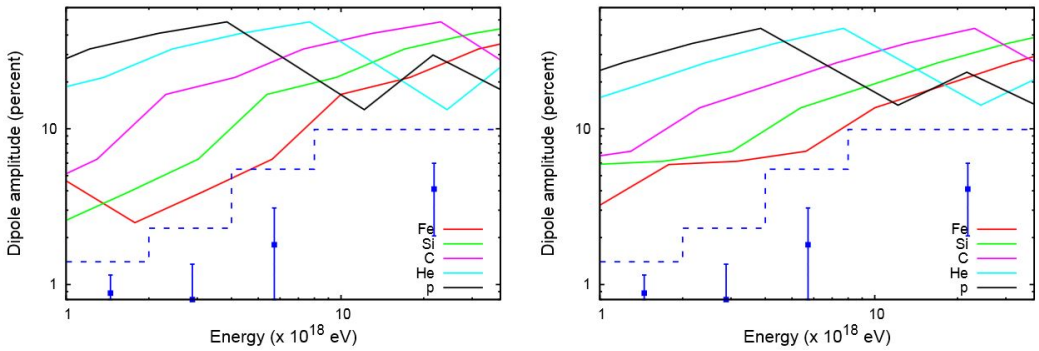

Figure 4.9: **left: Profile 1, (exponentially decaying field), and predicted dipole amplitudes in RA for 5 primaries. Sources are galactic and defined in regions** $(-200 \geq z \leq 200)$ **pc with** $r \leq$ 20 **kpc. Dashed blue line is** 99% **upper limit CL. Blue points are Pierre Auger Observatory dipole measurements with the 'east-west' method for E**$\in [1:2]$ **EeV, and Rayleigh analysis for the rest. The GMF model is PTKN-BSS (bisymmetric). The turbulent component has an assumed** $B_0 = 4\,\mu$**G and** $z_0 = 2$ **kpc. Limiting scale lengths are from** $L_{min} = 1$ **AU and** $L_{max} = 200$ **pc.**
right: profile 2, (uniform field), predicted dipole amplitudes in RA for 5 primaries, as in profile 1. From [61].

The propagation of CR iron nuclei at energies $\sim 1/4/8$ EeV is still diffusive, but by energies $\sim 8\times10^{19}$ eV resembles less of a random walk, and becomes predominately ballistic. The propagation is seen to follow the regular GMF lines in the disk (x-axis). By E $\sim 10^{20}$ eV iron nuclei are deflected $\sim 20° - 40°$. The energy region where both behaviours may be said to operate can

confuse galactic anisotropy readings - thus the individual propagation of sets of 10^4 nuclei in GMF models in these energy ranges for this work. The dipole amplitude error for 10^4 is $\leq \pm 3\%$.

Figure 4.9 on the previous page shows their [61] predicted dipole amplitudes with respect to the Pierre Auger Observatory measurements. Within the 3% error margin, in the case of low energies, it can be seen both silicon and iron nuclei are compatible with the Pierre Auger Observatory upper limits, i.e., at low EeV energies, silicon and iron nuclei may satisfy the measured Pierre Auger Observatory galactic dipole anisotropy. Iron would have a dipole amplitude of $\approx 0.6\%$. Galactic iron nuclei may still be present at energies of $\simeq 10$ EeV.

Having established this important constraint, [61] look for further parameter constraints. Remember, [61] are trying to see how likely an extra-galactic transition is at the ankle. Predicting anisotropies for particular nuclei or nuclei groups, and gauging whether these predictions violate the measured Pierre Auger Observatory upper limits, one can decide which galactic composition scenario is most appropriate at which energies. For example, as seen in Figure 4.9, a predicted light galactic composition at $\sim 10^{18}$ eV violates the upper limits of the measured anisotropy and the ankle as the galactic/extra-galactic transition is unlikely. Also, it is evident, the predicted anisotropies for either an exponential or a constant profile of the turbulent field are similar.

Predicted amplitudes for certain models parameters further allow galactic composition constraints.

For higher energies/rigidities, dipole amplitudes may decrease. There no longer is a rigidity based dipole amplitude after > 4 EeV. This is actually because other multipoles contribute and anisotropy measures must be widened to include them. The spectral indices of turbulent magnetic fields (Kolmogorov, $\mathbf{E_k} \propto k^{-5/3}$ or Kraichnan $\mathbf{E_k} \propto k^{-3/2}$) [219], don't much matter with dipole amplitudes, but certainly amplitudes are affected by the maximum scale variation, L_{max}, taken as (100-300) pc, of the turbulent field. For larger L_{max}, CRs remain diffusive longer, and so less galactic anisotropy is seen [61].

Results are condensed into Figure 4.10 for E/Z = 10^{18} eV. E/Z implies light nuclei, and it is consistently the case that their predicted amplitudes *exceed* the Pierre Auger Observatory observational limits. A contribution of $> 20\%$ of light nuclei at E $\sim$ 1 EeV cannot be galactic *unless*

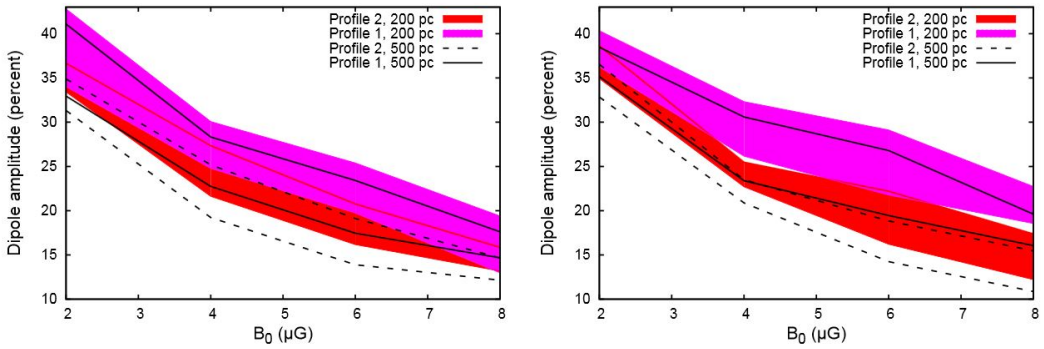

Figure 4.10: **Predicted dipole amplitudes versus turbulent Galactic magnetic field strength for E/Z = 10^{18} eV and Galactic sources defined in region** $(-200 \leq z \leq 200)$ **pc and in region** $(-500 \leq z \leq 500)$ **pc respectively with** $r \leq 20$ **kpc and** $z_0 \in [1 : 8]$ **kpc. PTKN-BSS Galactic magnetic field model.**
left: Kologmorov spectrum with $L_{Min} = 1$ **AU,** $L_{Max} = 200$ **pc.**
right: Kraichnan spectrum with $L_{Min} = 1$ **AU,** $L_{Max} = 200$ **pc. From [61].**

$B_0 \gg 10\mu$G, where B_0 is taken as the B_{rms} of the turbulent field close to the Sun.

The conclusions drawn are that anisotropy is dependent on B_0. Current limits $\sim 10^{18}$ eV do not permit dominant galactic contributions from light or intermediate nuclei. Current limits up to $(10 - 20)$ EeV do permit a dominant galactic contributions from iron nuclei, for turbulent magnetic field strengths of 8 μG. [61] consider composition $\sim 10^{18}$ eV crucial to the determination of thegalactic/extra-galactic transition energy (where fluxes are approximately equal). They also posit the fast rising proton composition of KASCADE-Grande marks the beginning of an extra-galactic flux between $(10^{17} - 10^{18})$ eV.

A study was also made of source density in the context of what sources would be able to accelerate iron up to $\sim$ 10 EeV, if the ankle composition is confirmed as heavy, and the ankle is the location for a galactic/extra-galactic transition. The escape time for galactic iron nuclei at such energies is $\sim 10^5$ years. The minimum rate of flux, $\mathcal{R}$, to maintain iron fluctuation levels, $\sigma F/\langle F \rangle$, under thresholds of $(5, 10, 25, 50, 100)\%$ for timescales $\gg \mathcal{R}^{-1}$ was estimated. Magnetars, with a density of 100/galaxy, and a rate of 10^{-3}/yr would be suitable candidates to satisfy the Pierre Auger Observatory anisotropic constraints.

Thus, overall, if the composition at $\sim 10^{18}$ eV is light, the ankle is probably not the transition

region. Whereas, if it is heavy around there, it probably is.

4.5 Discussion

There has emerged one point-source possibility. Cen A as a source has been 'found' from many analyses e.g. [109, 194], among others. As a source possibility, Cen A has also been linked into the LSS of the Virgo cluster or as a translated large point-source within. Large-scale anisotropies at lower energies are barely discernable [115], and if they exist, may be connected to the galactic/extra-galactic transition [61] of sources. No large-scale anisotropies at the highest energies have been detected. The size of the GMF, density of the source distribution, source distances considered, spectral indices, and other conjectures make for only broad generalizations about the physics in play. As more data become available, these generalizations may become more specific. The difference in CR composition between HiRes, TA and Pierre Auger Observatory at UHEs remains a sticking point.

Chapter 5

Minimum Spanning Tree Theory

5.1 Introduction

This chapter on the theory underlying minimum spanning trees (MSTs) begins by describing the components of MSTs in general terms, and why MSTs are useful.

A graph may be thought of as a collection of vertices, or nodes, or points in the plane, some of which are connected by edges. Thus a graph, $G = \langle V, E \rangle$, is defined by two sets of elements, a finite, non-empty set, V, of single vertices, and a set of edges, E, which are pairs of these vertices. Each edge **connects** these pairs. If the edge pair of vertices, $(u, v) = (v, u)$, we say the vertices are **adjacent** to one another, the vertices are the **endpoints** of the edge, and the edge connection is **undirected**. Thus the graph, G, is **undirected**. If the edge pair of vertices, $(u, v) \neq (v, u)$, we say the vertices are adjacent to one another, the vertices are the endpoints of the edge, the edge connection is **directed**, and the vertices are **ordered**. Thus the graph, G, is **directed**. The vertex set of G is $V(G)$, and the edge set of G is $E(G)$. The order of a graph is its number of vertices, $|V(G)|$, and the size of a graph is its number of edges, $|E(G)|$. A **loop** is an edge where the edge endpoints are the same vertex. A **path** in graph G, from vertex u to vertex v, is a set of vertices adjacent to each other, connected by edges, which starts at u and ends at v. A **simple** path is when all the vertices of the path are distinct.

A tree is a connected, acyclic (no loops), simple graph [45, 43].

A graph often contains redundancy, in that there can be multiple paths between two vertices. This redundancy may be desirable, for example, to offer alternative routes in the case of breakdown, or overloading of an edge (road, connection, phone line) in a network. However, we often require the cheapest sub-network that connects the vertices of a given graph. This must in fact be an **unrooted tree** (see page 122), because there is only one path between any two vertices in a tree; if there is a cycle, then at least one edge can be removed. The total cost, or "weight" of a tree is the sum of the weights (a property or measure of some kind, of an edge) of the edges in the tree. We assume that the weight of every edge is greater than zero. Given a connected, undirected graph $G = \langle V, E \rangle$, the minimum spanning tree problem is to find a tree $T = \langle V, E' \rangle$ such that $E' \supseteq E$ and the cost of T is minimal.

Note that a minimum spanning tree is not necessarily unique. A tree over $|V|$ vertices contains $|V| - 1$ edges. A tree can be represented by an array of this many edges. The Minimum Spanning Tree (MST) is a graph theoretical algorithm[1], used for solving optimization problems, and has numerous applications in a range of fields. It is manifest in its practical aspect for the minimization of costs in network design, virological epidemiology [48], and, in its purely mathematical presentation, it is part of a 'combinatorial' (see subsection 5.4.1 on page 127) class of problems, including the Steiner tree and the Travelling Salesman Problem. In an augmented state, it is to be found in the detection of maximum bottleneck paths, ethernet bridging to avoid network cycles, and is a mainstay in pattern recognition and cluster analysis. A side development in the pursuit of MST solutions is the use of DNA encoding methods [49].

The MST is a particular type of tree, and the concept of a tree was later recognized to be an application of graph theory. A mathematical formalism of the tree was given shape and defined in the work of Gustave Kirchhoff, who introduced trees to determine the fundamental cycles of electrical networks that were a consequence of his conservation law[2]. Arthur Cayley[3] coined the term 'tree', and developed many of its properties over a span of 32 years. D. Huffman was responsible for the principle of minimizing weighted path lengths in his research on effective

[1] 'An algorithm is a description of a pattern of behaviour expressed in terms of a well-understood, finite repertoire of named actions of which it is assumed *a priori* it can be done' [40].

[2] flow in = flow out

[3] developed the language of algebraic invariance used in special relativity

data compression methods, and the 1960s introduced an era of tree based procedures. The 1970s saw the beginnings of computational complexity theory. Computational complexity theory was a nascent response to the questions surrounding the solvability of combinatorial problems in the context of Turing machines and determinancy [3, 4].

The language of MSTs is mostly in terms of graph theory, and so a brief overview of the theory of graphs relevant to trees and then MSTs, will be presented in section 5.2.

The recursive nature of tree algorithms will be discussed in section 5.3 on page 125, as will ways of analysing the algorithm's efficiency, or order of growth, in subsection 5.4.2 on page 127. Two classic examples of MST algorithms will be given, where, in the interest of consistency, both algorithms pseudocodes are taken from [45]. The algorithms are covered in section 5.6 and consist of Prim's algorithm and Kruskal's algorithm plus a mention of Dijkstra's algorithm, known as 'the shortest path problem'.

Finally, in section 5.7 on page 140, there is a discussion on the Fast MST of Bentley and Friedman [31], and its adaptation as a model to be used for the detection of structure in the cosmic ray (CR) directional data of the Pierre Auger Observatory.

5.2 Theory of Graphs

Most of the material in this section is from Knuth [35], and Aho, Hopcroft and Ullman [33].

DEFINITION: A graph, $G = \langle V, E \rangle$, comprises a finite, non-empty set of vertices, $V = (v_1, v_2, v_3, \ldots, v_i, v_{i+1}, \ldots, v_{n-1}, v_n)$, coupled with a set of vertex pairs called edges, $E = ((v_1, v_2), (v_2, v_3), \ldots, (v_i, v_{i+1}), \ldots, (v_{n-1}, v_n))$, where $1 < i + 1 \leq n, i \in \mathcal{I}$

1. 2 vertices that share an edge are said to be **adjacent**, e.g. for (v_1, v_2), v_1 is adjacent to v_2 and v_2 is adjacent to v_1. The number of vertices adjacent to v_1 is called the **degree** of v_1.

2. In a **directed** graph, each edge has a sense of direction and consists of ordered vertex pairs, e.g. $e_1 = (v_1, v_2)$ such that v_1 is called the **head** of the edge leaving v_1, and v_2 the **tail** of the edge entering v_2. There are at most $|V|^2$ edges in E.

3. In an **undirected** graph, no edge has a sense of direction and consists of unordered vertex pairs, e.g. $e_1 = (v_1, v_2) = (v_2, v_1)$. There are at most $|V|(|V| + 1)/2$ edges in E.

4. A graph with every pair of vertices connected by an edge is called **complete**.

5. A graph with a **weight** or **cost** associated with each edge is called a **weighted** graph.

6. If $v = v_1$, and $v' = v_n$, $V = (v_1, \ldots, v_i, v_{i+1}, \ldots, v_{n-1}, v_n)$,
   where $1 \leq i \leq n,\ n \geq 1$ we say $\langle v_i, v_{i+1} \rangle$ is an adjacent edge in G and V is a **path** of length $n - 1$ from v to v'.

7. A single vertex, v_i, is of length 0 from itself to itself, $e = \langle v_i, v_i \rangle$.

8. A **simple** path is when all edges and all vertices, except possibly the first and last are distinct.

9. A **cycle** is a simple, closed path of length at least 1, that begins and ends at the original vertex, i.e., $v_1 = v_n$. In a directed graph a cycle must be of length at least 3.

10. A graph with no cycles is called **acyclic**.

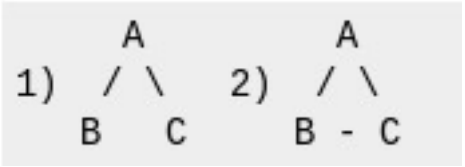

Figure 5.1: The graph on the left of the figure, 1), is an example of an acyclic graph. The graph on the right of the figure, 2), is an example of a cyclic graph. Both graphs are undirected and simple. From [50].

11. A graph is **connected** if there is a path between any two vertices on the graph, i.e. for any vertex v and any vertex v' of a graph, there is a sequence of connecting edges between them.

12. A **subgraph** of a given graph, $G = \langle V, E \rangle$, is a graph, $G' = (V', E')$ such that $V' \supseteq V$ and $E' \supseteq E$.

13. If a graph is not connected, it will consist of **connected components**. A connected component is a maximal (not expandable via an extra inclusion of a vertex), connected, subgraph of a given graph.

5.2.1 Trees

Let us define a **tree**, formally, as a finite set, T, of one or more vertices such that

(a) there is one specially designated node called the **root** of the tree, root(t), and

(b) the remaining nodes, excluding the root, are partitioned into $m \geq 1$ disjoint sets, $T_1, T_2, T_3, \ldots, T_m$, and each of these sets in turn is a tree. The trees, $T_1, T_2, T_3, \ldots, T_m$, are called the **sub-trees** of the root[4].

5.2.1.1 Free Tree

A **free**, or **unrooted** tree, is defined to be a connected, undirected, acyclic graph, G.

The following theorem is a set of equivalency statements[5].

THEOREM *A*: If G is a graph

(a1) G is a free tree.

(a2) G is connected, but if any edge is deleted the resulting graph is no longer connected.

(a3) If v and v' are distinct vertices of G, there is exactly 1 simple path from v to v'.

(a4) G has no cycles, and has $n - 1$ edges.

(a5) G is connected, and has $n - 1$ edges[6].

The idea of a free tree is not only a key property of the MST it also features in the analysis of computer algorithms and will appear later in this chapter.

A **spanning tree** is a free tree, $G =< V, E >$, which connects all vertices in V such that there is a subset of E, and such that there is a unique path between any 2 vertices in V.

If we define the weight or cost of a spanning tree to be the sum of the weights of its constituent edges, then a **minimum spanning tree** of graph, G, is a spanning tree whose weight or cost is minimal among all spanning trees of G.

[4] recursive definition – have defined a tree in terms of trees

[5] proof is in Knuth [35]

[6] Kruskal [34] assembled a set of equivalency statements akin to these, and from those statements and the definition of a MST he proceeded to build one.

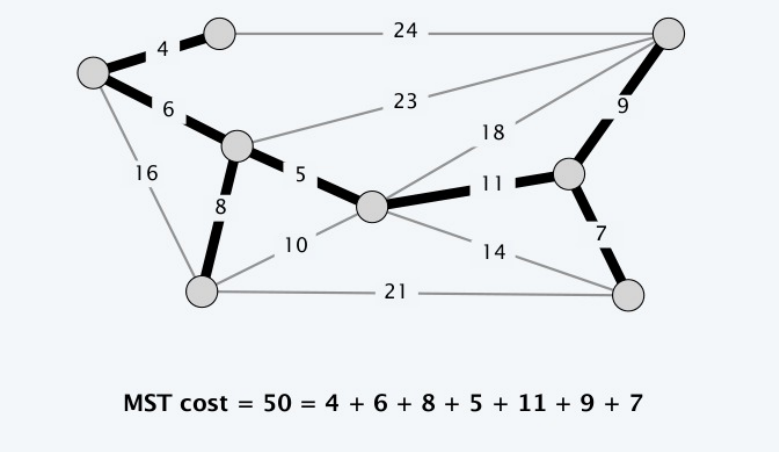

Figure 5.2: Given a connected graph $G = \langle V, E \rangle$ with edge costs, $c(e)$, the MST is a subset of the edges $T \supseteq E$, such that T is a spanning tree whose sum of edge costs is minimized. From [51].

Many MST algorithms have been developed. Two classic such algorithms will later be set down in section 5.6 on page 135.

5.2.1.2 Rooted Tree

A directed graph with no cycles is called a directed, acyclic graph and a directed tree or **rooted tree** is a directed, acyclic graph satisfying the following properties:

(a) There is exactly 1 vertex, the **root**, where no edges **enter**.

(b) Every vertex except the root has exactly 1 entering edge.

(c) There is a unique path from the root to each vertex.

5.2.1.3 Forest

Given that a tree (free tree) is a connected, acyclic graph - a **forest** is a graph that has no cycles, but is not necessarily connected. Each connected component is a tree.

The text below is largely transcribed from [33]. Let $F = \langle V, E \rangle$ be a directed graph which is a forest. If (v, u) is in E then

1. v is called the **parent** of u, and u is the **child** of v.

2. If there is a path from v to u, then v is an **ancestor** of u, and u is a **descendant** of v.

3. If $v \neq u$ then v is a **proper ancestor** of u, and u is a **proper descendant** of v.

4. A vertex with no proper descendants is called a **leaf**.

5. A vertex v and all its descendants are called a **sub-tree** of F. A vertex, v, is called the **root** of the sub-tree.

6. The **depth** of a vertex v in a tree is the length of the path from the root to v. The **height of a vertex** v in a tree, is the length of the longest path from v to a leaf.

7. the **height of a tree** is the height of the root (trees are drawn with the root at the apex a.k.a. the family tree).

8. the **level** of a vertex v in a tree, is the height of the tree minus the depth of v.

An **ordered tree** is a tree in which the children of each vertex are ordered, and we assume the children are ordered from left to right.

A **binary tree**, or decision tree, is an ordered tree such that

1. Each child of a vertex is distinguished either as a **left child** or **right child** and -

2. no vertex has more than one 1 left child or right child.

A sub-tree, T_l (if it exists), whose root is the left child of a vertex v is called the **left sub-tree** of v and similarly, the sub-tree, T_r (if it exists), whose root is the right child of v is called the **right sub-tree** of v. All vertices in T_l, are said to be **to the left** of all vertices in T_r.

A **priority queue** [56] is a data structure (e.g. an array, a binary search tree) where the first value out, is the value with the highest priority.

A **heap** [56] is a binary tree, with two special properties: The ORDER property: For every node, n, the value in n is greater than or equal to the values in its children (and thus is also greater than or equal to all of the values in its sub-trees). The SHAPE property:

1. All leaves are either at depth, d, or d-1 (for some value d).

2. All of the leaves at depth d-1 are to the right of the leaves at depth d.

   (a) There is at most 1 node with just 1 child.

   (b) That child is the left child of its parent, and

   (c) it is the rightmost leaf at depth d.

5.3 Recursion

Recursion exists when an object, or class, is defined empirically in the language of itself, i.e., it exists in terms of itself and so continues in a species of self replication. Some of the more well known examples of recursion are the natural numbers, the factorial function and tree structures [41].

example 1: ***Natural numbers***

(a) *1 is a natural number*

(b) *the next natural number is a natural number*

example 2: ***Factorial function*** $n!$ $(n \in \mathcal{I} > 0)$

(a) $0! = 1$

(b) *If* $n > 0$, *then* $n! = n(n-1)(n-2)...1$

example 3: ***Tree structures***

(a) *0 is a tree (empty)*

(b) *If* t_1 *and* t_2 *are trees, then* $t_1 + t_2$ *is a tree*

Recursive algorithms reproduce the behaviour of the problem to be solved, or the function to be computed, and fold in on themselves in a series of steps that must be subject to a halting condition lest the algorithm never terminates. There are problems of a recursive nature which are believed to have no solution, the Turing Machine being a notable example - where the input into an algorithm is in terms of the algorithm, the problem being - will the algorithm cease?

Minsky's proof that the machine (or program) would indeed go on, is recognized by his peers [40] and Minsky considered the problem and its elucidation the most important section of his book [44].

Gödel [46], in his study of systems of metamathematics [7] arrived at the generalized solution for essentially the same problem. **Axiom - systems of theorems that prove their own consistency, are inconsistent**[8].

All tree structures are recursive, irrespective of the branching decisions that direct their growth. The halting condition is avoided, the data-sets being finite and the recursion paradox is irrelevant, since we work **within** the framework of the boundary constraints or rather initial conditions. Tree structures are engineered applications of recursion.

5.4 Analysis of Algorithms

The analysis of algorithms usually covers

(a) proof of correctness

(b) order of growth

(c) optimization

(d) generality

(e) simplicity

Here we will be concerned chiefly with individual algorithm's proof of correctness and order of growth. Optimizing an algorithm is tied up with its order of growth, and the Fast MST (see section 5.7 on page 140) can be regarded as an optimized MST. Given their recursive identity, the generality and simplicity of MST algorithms is what tends to render them unwieldy in the face of large inputs and complex data structures, their simplicity being best exploited for small data-sets and their generality customized to conform to the geometry of the data spaces. As will be

[7] A mathematical theory about mathematics.

[8] Gödel's theorem is actually stated using recurrence.

evident in Table 5.1 on page 131, no matter how fast computer hardware enables the performance of calculations, the order of growth of a tree algorithm must somehow be constrained if many tree problems with meaningful sized data-sets (data sizes that will yield useful statistics), are to be resolved. The asymptotic order of growth of the algorithms of Prim and Kruskal are discussed, as is the asymptotic order of growth of our fast MST, the Yggdrasil. The Yggdrasil does not employ the optimized priority queue strategy of [31], but does fragment a property of the Pierre Auger Observatory data metric space as they [31] suggest, that yields a useful order of growth.

5.4.1 Proof of Correctness

Tree structures are recursive from root to leaf and their algorithms duplicate this architecture. Mathematical induction as a proof of correctness follows naturally upon the sequence of iterations which provide the inductive body of the proof. Inductive proof by contradiction also features extensively in algorithms of this type.

The theory of an algorithm's proof of correctness does not necessarily mean it possesses its practical counterpart as a computer program. Graph algorithms of the type known as combinatorial, problems that look for solutions that are permutations, combinations, or a subset satisfying certain constraints and optimizing a desired characteristic, often prove irreducible in terms of space (memory) and time complexity, and have approximate solutions only. The MST is a combinatorial problem which has a solution, unlike the Travelling Salesman Problem or Knapsack problem, albeit for data sizes which remain small in terms of computational complexity theory (see sub-subsection 5.4.2.1 on page 130).

5.4.2 Order of Growth

An algorithm's space and time complexity or efficiency, is implicit in its order of growth. The analysis of an algorithm's time complexity is analogous to its space complexity, and it is now customary to address the question of time complexity, rather than space complexity, or both. The time complexity can be assessed experimentally - where the program is run on a computer and its time for various inputs benchmarked, but this is dependent on the computer hardware, or analytically,

where we deduce certain time-related quantities with mathematical dialectic.

These time related quantities are the algorithm's ⟨max, min, ave, st dev⟩ and D. Knuth pioneered this area of computer science [35]. An algorithm's time complexity is a measure of the size, n, of its input data. We are almost always interested in time complexity in terms of large sizes, thus we look to the limiting behaviour, or asymptotic time complexity as $n \mapsto \infty$, of an algorithm. The time units assigned to time complexity are relative, it is the description of the function's time evolution that is important in determining whether an algorithm is effective or not.

Three different asymptotic orders of growth, which correlate to the max, min and ave [45], are called into service by computer scientists, $O(f(n))$ - big Omicron, $\Omega(f(n))$ - big Omega, and $\Theta(f(n))$ - big Theta [37].

i) The notation $O(f(n))$ is read as 'order at most $f(n)$.'

ii) The notation $\Omega(f(n))$ is read as 'order at least $f(n)$.'

iii) The notation $\Theta(f(n))$ is read as 'order exactly $f(n)$.'

Formally stated [36]

1. $O(f(n))$ denotes the set of all $g(n)$ such that $\exists + ve$ constants c and n_0 with $|g(n)| \leq cf(n)$, $\forall n \geq n_0$.

2. $\Omega(f(n))$ denotes the set of all $g(n)$ such that $\exists + ve$ constants c and n_0 with $g(n) \geq cf(n)$, $\forall n \geq n_0$.

3. $\Theta(f(n))$ denotes the set of all $g(n)$ such that $\exists + ve$ constants $c,\ c'$ and n_0 with $cf(n) \leq g(n) \leq c'f(n)$, $\forall n \geq n_0$.

This analysis can be performed micro-analytically or macro-analytically [42]. Knuth devised a micro-analytic system, a frequency analysis, where the four quantities, ⟨max, min, ave, st dev⟩, were determined as functions of the number of times each 'primitive' or 'elementary' operation (e.g. addition, subtraction) was performed. He first organized a program in its flow chart form, and then utilized Kirchhoff's Law to establish the fundamental cycles of the flow chart, and thus count the number of iterations for various instructions. He then instituted a more formal analysis

by incarnating a hypothetical computer, MIX, and its score of MIX instructions to quantify all four values.

Program flow charts are no longer extant and although micro-analysis still exists, it is considered time consuming and unduly detailed. The analytic standard of time complexity is macro-analysis, and the order of growth most commonly used is the $O(f(n))$ notation [33, 39, 40, 37], although [45] favours the $\Theta(f(n))$ notation[9].

In macro-analysis, the one or two dominant basic operations of an algorithm are isolated, and we count the times that these operations are executed to arrive at the algorithm's order of growth. Identifying the dominant operations frees us from quirks of computer hardware, programming languages and such. Macro-analysis involves locating a **bottleneck** in the code - that principle section of code that consumes a major portion of running time. In software engineering, it has been observed that $80\% - 90\%$ of the time spent [43] is produced by $10\% - 20\%$ of the code - an instance of the Pareto principle[10]. It is also these parts of the code that the later software design process of 'optimization' seeks to better reduce.

The problem to be solved, or the calculations to be computed, fall into a number of clear categories of a representative design type, with just a few common basic operations. Some of the important categories are graph problems and geometric problems, and some basic operations are searching, sorting and matrix multiplication. If the basic operations are properly identified, and there is an accurate assessment of the sum total of operations performed, we possess a good independent measure of the algorithm's efficiency, or order of growth.

These basic operations have already quantified asymptotic efficiency classes, and so when an algorithm's order of growth is reckoned, we need only to nominate those central basic operations, which are often self-evidently the algorithm's inner loops, and refer to their efficiency classes.

Decision trees represent the basic operation of sorting, and we illustrate, by way of example, the derivation of the sorting efficiency class, a structural component of the MST.

[9] Indeed, while the order of growth **functions** are the same for identical algorithms, their notations differs [33, 39, 38, 45, 37].

[10] States that for many phenomena 80% of the consequences stem from 20% of the causes. Pareto observed 80% of Italy's income went to 20% of its people.

5.4.2.1 Decision Trees

Sorting is a basic operation predicated on the notion of a **decision tree** - a binary tree (can be ternary or greater) where every vertex represents a decision, sorting being a re-arrangement of input, where there are a number of comparisons made in a branching series of instructions or decisions.

At every parent vertex, a test is given, and depending on the outcome, control passes to one of its children, and so on down the tree until a leaf is reached. The time complexity is the number of decisions taken, or comparisons made, to traverse the tree from root to leaf. This is the height of the tree as a function of the size of the problem.

A binary tree of height, h , and leaves, l (see sub-subsection 5.2.1.3 on page 123), has a maximum of 2^h leaves [45] and hence

$$2^h \geq l$$

and

$$h \geq \log_2 l$$

the logarithmic base being the number of comparisons made per vertex.

If we wish to sort n numbers, for example, there are $n!$ possible outcomes, and to sort them using comparisons, we come to, employing Stirling's formula

$$\lceil \log_2 n! \rceil \sim \log_2 \sqrt{2\pi n}(e/n)^2 \approx n \log_2 n$$

i.e., we need to make $n \log_2 n$ comparisons to sort a list of n numbers and, in this case, the maximum complexity is $O(n \log_2 n)$. For all efficiency classes, log is assumed to be $\log_2$ (binary) unless stated otherwise.

The number of trees needed to be calculated in order to find a MST was derived by Cayley as n^{n-2} for n vertices. Efficiency classes of some algorithm types are listed in Table 5.1 on the following page. The MST is a member of the exponential efficiency class. Some exponential efficiency class problems can be 'reduced' or approximated in 'polynomial' time. A problem, or

Table 5.1:

BASIC ASYMPTOTIC EFFICIENCY CLASSES. From [45]

Class	$n = 10^2$	Name	Comments
$\log n$	6.6	*logarithmic*	Typically, a result of cutting a problem's size by a constant factor on each iteration of the algorithm.
n	10^2	*linear*	Algorithms that scan a list of size n (e.g., sequential search) belong to this class.
$n \log n$	6.6×10^2	n-log-n	Many divide and conquer algorithms including mergesort and quicksort in the average case, fall into this category.
n^3	10^6	*cubic*	Typically, characterizes efficiency of algorithms with three embedded loops. Several non-trivial algorithms from linear algebra fall into this category.
2^n	1.3×10^{30}	*exponential*	Typical for algorithms that generate subsets of an n-element set. Often, the term "exponential" is used in a broader sense to include this and larger orders of growth as well.
$n!$	9.3×10^{157}	*factorial*	Typical for algorithms that generate all permutations of an n-element set.

rather algorithm, is able to be solved in polynomial time when the number of elementary operations can be completed in $O(n^k)$ steps, when n is the **complexity** of the input and k a non-negative integer [2]. For the algorithm to be completed in a reasonable order of growth time, **k needs to be small**. In terms of computational complexity theory, the MST is an NP-complete problem. The algorithm can be solved in non-deterministic polynomial (NP) time by a non-deterministic Turing Machine [1, 4]. Prim and Kruskal's optimized MST algorithms are considered polynomial time algorithms [4, 3].

5.5 Minimum Spanning Trees

A MST is a type of spanning tree, and so let us revisit more formally these MST tree properties.

DEFN: Let $G = \langle V, E \rangle$ be a connected, undirected graph, with a vertex set, V, and an edge set, E, with a cost function, $c(e)$, for mapping edges to real numbers; a spanning tree is a subset of E such that there is a unique path between any two vertices in V and the cost of a spanning tree is

the sum of its edges.

The following two lemmas are the basis for many spanning tree algorithms [33].

lemma 1: Let $G = \langle V, E \rangle$ be a connected, undirected graph and $S = \langle V, T \rangle$ be a spanning tree for G. Then

(a) $\forall v_1$ and v_2 in V, the path between v_1 and v_2 in S is unique and

(b) if any edge $E - T$ is added to S, a unique cycle results.

lemma 2: Let $G = \langle V, E \rangle$ be a connected, undirected graph, and $c(e)$ a cost function on its edges. Let $\{(V_1, T_1), (V_2, T_2), \ldots, (V_k, T_k)\}$ be any spanning forest for G with $k > 1$. Let $T = U^k{}_{i=1} T_i$. Suppose $e = (v, u)$ is an edge of lowest cost in $E - T$, such that $v \in V_1$ and $u \notin V_1$. Then there is a spanning tree for G which includes $T \bigcup \{e\}$ and is of as low a cost as any spanning tree for G that includes T.

The proofs for these two lemmas are in [33] on pages 172-173.

An n vertex graph, G, has $1/2(n^2 - n)$ edges and we know from Cayley that the number of possible spanning trees for n vertices as n^{n-2}. It is **lemma 2** that yields a MST for G. To have a MST of G which is unique, it is a sufficient condition that no two lengths are equal, though Prim [32] proved by introducing a small variation, $\in$, to be made vanishingly small, into one of the lengths of any of the MST's, that a unique MST must still exist. We, however, assert that each cost is unique. We also stipulate that the costs are $\in \mathcal{R}^+$, knowing that MST's of $-ve$ lengths are also legitimate constructs [32]. The space in which the vertices are embedded need not be Euclidean, and the costs need not satisfy the triangle inequality[11]. All increasing $+ve$ symmetric functions[12] of the edge lengths are maximized, and all decreasing $+ve$ symmetric functions[13] are minimized. The product of the edge lengths, and the square root of the sum of the squares of the edge lengths

[11] Let us say a *metric* is a way of measuring a property, distance, d, or convergence, between objects in a set, a *metric space*, X. Given a non-empty set, X, a metric on X is a function d from $X \times X \mapsto R$ such that these 3 axioms are satisfied:

1. $\forall x, y \in X, d(x, y) \geq 0,$ and $d(x, y) = 0$ if and only if $x = y$.
2. $\forall x, y \in X, d(x, y) = d(y, x)$.
3. $\forall x, y, z \in X, d(x, z) \leq d(x, y) + d(y, z)$(triangle inequality).

Euclidean space satisfies these three axioms [55].

[12] don't need to be $+ve$ for $-ve$ lengths

[13] don't need to be $+ve$ for $-ve$ lengths

of the MST are likewise minimized, and the sum of the reciprocals of the edge lengths and the product of the arc cotangents are maximized [32].

Given that there are n^{n-2} spanning trees for graph, G, mapping the MST is a considerable task, especially as n increases, and is considered an impractical last resort. Historically, many different MST algorithms were tried in an effort to sidestep this obstacle.

The most successful of these algorithms tended to convoke 2 simplifying principles. These principles are very close to the above lemmas, but have been recast to better understand the proofs of our two classic algorithms.

Simplifying Principles: Given that all edge costs, $c(e)$, are distinct,

1. **cycle property:** Let C be any cycle and let f be the maximum cost edge belonging to C. Then the MST, T^*, does not contain f.

2. **cut property:** Let S be any subset of vertices and let e be the minimum cost edge with exactly one end-point in S. Then the MST, T^*, contains e.

Figure 5.3 illustrates these two principles.

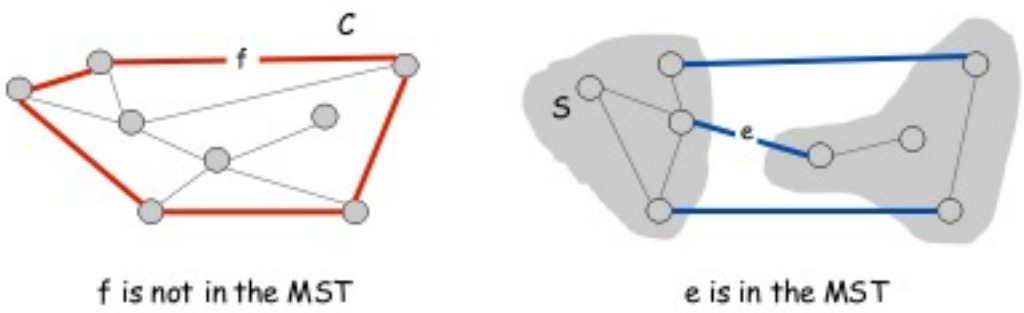

Figure 5.3: **left:** Let C (red) be a cycle,then the edge of **maximum** cost, f, does not belong to a MST. **right:** Let S be a subset of vertices with three edge endpoints (blue) in S. If e has the only **minimum** cost endpoint in S, then e belongs to the MST. From [52].

Figure 5.4 on the following page illustrates a proof by contradiction of the cycle property and Figure 5.5 illustrates a proof by contradiction of the cut property.

The MST algorithms based on the simplifying principles are termed ‘greedy’ and are applied

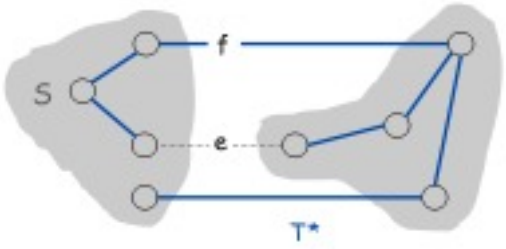

Figure 5.4: **Cycle property Proof:** Suppose the edge e does not belong to T*. Adding e to T* creates a (unique) cycle C in T*. Some other edge in C, say f, has exactly one endpoint in S. T = T* $\cup \{e\} - \{f\}$ is also a spanning tree. Since cost(e)<cost(f), cost(T)<cost(T*). This is a contradiction. From [52].

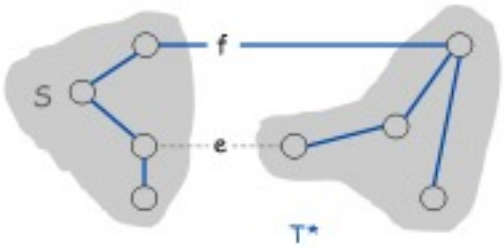

Figure 5.5: **Cut property Proof:** Suppose f belongs to T*. Deleting f from T* disconnects T*. Let S be one side of the cut. Some other edge in C, say e, has exactly one endpoint in S. T = T* $\cup \{e\} - \{f\}$ is also a spanning tree. Since cost(e)< cost(f), cost(T)< cost(T*). This is a contradiction. From [52].

to optimization problems. They are denominated 'greedy' algorithms because with each iteration to choose the next edge, there is a 'greedy' grab for the cheapest alternative, in the reasonable[14] expectation that this sequentially local thrust will eventuate in the global optimal solution to the problem.

We discuss two such classic MST's, Prim's [32] and Kruskal's [34]. Other greedy algorithms include Dijkstra's 'shortest path problem' [38], which can be adapted to become a MST **reverse-delete** algorithms and those of Borůvka [34] and Yao [39]. The greedy technique finds its basis in the abstract combinatorial structure called a 'matroid' [45].

[14]This greedy approach does not work for all optimization problems, but it can be useful as an approximation technique for thus far intractible problems such as the Travelling Salesman Problem.

5.6 Two Classic Algorithms

These algorithms approach the MST problem from reverse aspects. Prim starts with a single vertex to build his MST and Kruskal starts with every vertex a MST. Prim's algorithm is the basis for Bentley and Friedman's Single Fragment Fast MST [31] and Kruskal's concept of single vertex fragments could be claimed as a contribution to their concept of a Multi Fragment Fast MST. As already mentioned, our Yggdrasil MST references Bentley and Friedman's work.

5.6.1 Prim's Algorithm

Prim's [32] algorithm was published in the Bell Telephone Research labs technical Journal, and was for 'A shortest connection network'. It is for complete graphs with edge lengths mapped to the real number system. He also extended his algorithm to connected graphs with some edges not permitted, and to a **longest** spanning sub-tree to arrive at some non-trivial generalizations involving symmetric functions as edge lengths.

Prim initiates an unfolding order of connected sub-trees by selecting any root vertex, v_i, in $G = \langle V, E \rangle$, to be in sub-tree, S. He administers the cut property to the root vertex by greedily choosing the nearest vertex (cheapest edge) in the cutset, $V - v_i = V - S$, to be attached to the sub-tree. The new cutset is then searched for the nearest vertex to the new sub-tree, S, and so on, until all vertices in the dwindling cutset have been used and the MST is assembled.

Prim's original paper is referred to in some detail because the Fast MST [31] is grounded in Prim's construction principles. Some of Prim's original terminology has been here retained and some has been altered to better fit graph thoery.

5.6.1.1 Terminology

- **isolated vertex** – a vertex, at which, at a given stage of construction, no connections have yet been made.
- **fragment** – a vertex subset connected by adjacent edges between members of the subset.

- **distance of a vertex from a fragment** – of which it is not an element, is the minimum of its distances from the original vertices comprising the fragment.
- **isolated fragment** – is a fragment to which, at a given stage of construction, no external connections have been made.
- **A nearest neighbour to a vertex** – is a vertex whose distance from the specified vertex is at least as small as any other.
- **A nearest neighbour of a fragment** – is a vertex whose distance from the specified fragment is at least as small as that of any other.

The two fundamental construction principles (**P1 and P2**) for the MST are now stated as follows. From now on the term nearest neighbour is referred to as **nn**.

5.6.1.2 Construction

P1: Any isolated vertex can be connected to a **nn**.

P2: Any isolated fragment can be connected to a **nn** by the shortest available edge.

The validity of P1 and P2 depend essentially on the establishment of two necessary conditions (**NC1 and NC2**) for a MST.

NC1: Every vertex in a MST is directly connected to at least one **nn**.

NC2: every fragment in a MST is connected to at least one **nn** by a shortest available path.

Prim provides 'inductive proof by contradiction' for NC1 and NC2 similar to the already provided proof of the cut property illustrated in Figure 5.5 on page 134. Figure 5.6 on the following page is a very simple illustration of Prim's algorithm.

The basic idea is: expand the MST by adding the nearest (lightest) edge to it and its endpoint.

To maintain a consistent style of presentation, both algorithm's pseudocodes are taken from [45].

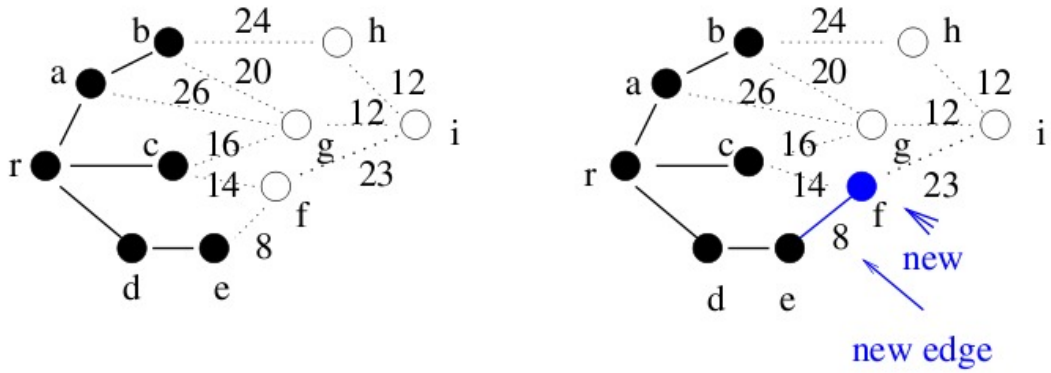

Figure 5.6: Initialize S = any vertex. Repeat n-1 times:
Add to tree the minimum weight edge with one endpoint in S.
Add new node to S. From [51].

5.6.1.3 Prim's Algorithm Pseudocode

// Input: A weighted, connected Graph, $G = \langle V, E \rangle$

// Output: E_T, the set of edges composing a MST of G

$V_T \leftarrow \{v_0\}$ //the set of tree vertices can be initialized with any vertex

$E_T \leftarrow \emptyset$

for $i \leftarrow 1$ **to** $|V| - 1$ **do**

find a minimum weight edge $e^* = (v^*, u^*)$ among all the edges

(v, u) such that v is in V_T and u is in $V - V_T$
$V_T \leftarrow V_T \bigcup \{u^*\}$

$E_T \leftarrow E_T \bigcup \{e^*\}$

return E_T

5.6.1.4 Order of Growth

The time complexity is subject to the type of data structure adopted for the priority queue, Q, of the cutset, $V - S$. The priorities in Q are the shortest edge lengths between the vertices in the cutset and the sub-tree vertices. Q as an unordered array has time complexity $O(n^2)$, since on

each iteration, Q is traversed to find, and delete the minimum edge vertex and possibly update the remaining cutset priorities.

A priority queue which is implemented as a minimum heap is initialized in $O(m)$ time and at each iteration a vertex is extracted in $O(\log n)$ time, thus its time complexity is $O(m \log n)$.

5.6.2 Kruskal's Algorithm

Kruskal [34] derived his algorithm when a 2nd year graduate student. His interest appears to have been piqued by a publication authored by O. Borůvka. Kruskal's MST was built on a finite, connected, complete, acyclic graph with $+ve$ real numbered lengths, although his MST also holds for $-ve$ weights. He also instituted the practice of assigning ∞ to edges 'missing', so that he could, without loss of generality, talk of a complete graph.

Kruskal's strategy is to initiate each vertex in the graph as a MST in itself. Each MST increases with edges ordered in increasing cost unless the insertion of an edge creates a cycle, whereupon the edge is discarded in observance with the cycle property. If this is so, the cutset of the tree is searched in accordance with the cut property, and edge $e = (v, u)$ is assimilated when v is in S, and u is in $S's$ cutset. Illustrations of Kruskal's algorithms are rather long. Two such can be found amongst [54] and [51].

We begin with a forest in G consisting of V trivial MST's, and end with a forest in G consisting of 1 MST.

Kruskal offers up three different methods of construction (A, B and A') and provides a preliminary theorem on graph theory which is in a similar vein to Theorem A in sub-subsection 5.2.1.1 on page 122.

5.6.2.1 Construction

Preliminary Theorem: If G is a connected graph with n vertices, and T is a subgraph of G, then the following conditions are all equivalent:

(a) T is a spanning tree of G.

(b) T is maximal and acyclic[15].

(c) T is minimal[16] connected spanning graph of G.

(d) T is acyclic with $n - 1$ edges.

(e) T is a connected, spanning graph with $n - 1$ edges.

Construction A. Perform the following steps as many times as possible: Among the edges of G not yet chosen, choose the shortest edge which does not form any loops with those edges already chosen.

He proves he has constructed a MST with construction A and an inductive proof by contradiction similar to the cycle proof in Figure 5.4 on page 134 previously given.

5.6.2.2 Kruskal's algorithm pseudocode

// Input: A weighted, connected Graph, $G = \langle V, E \rangle$

// Output: E_T, the set of edges composing a MST of G

sort E in nondecreasing order of the edge weights $w(e_{i_1}) \leq, \ldots, \leq w(e_{i_{|E|}})$

$E_T \leftarrow \emptyset$; $ecounter \leftarrow 0$ //initialize the set of tree edges and its size

$k \leftarrow 0$ //initialize the number of processed edges

while $ecounter < |V| - 1$ **do**

    $k \leftarrow k + 1$

    if $E_T \bigcup \{e_{i_k}\}$ is acyclic

        $E_T \leftarrow E_T \bigcup \{e_{i_k}\}$; $ecounter \leftarrow ecounter + 1$

return E_T

5.6.2.3 Order of Growth

If an efficient union-find [33, 45] algorithm is used for checking whether two vertices belong to the same tree, the time complexity is dominated by the sorting of the edge weights, which is $O(|E| \log |E|)$.

[15] Kruskal uses 'forest' meaning acyclic.

[16] graph does not contain any smaller graph of the same sort

5.7 A Fast Minimum Spanning Tree

MST's have long been deemed impractical for data-sets of any appreciable extent, the computational time involved escalating into many month's duration, even for what might be considered data-sets of relatively modest proportions.

The algorithms of Prim, Kruskal, Dijkstra and others have broad general applications, and have not been engineered to optimize the geometries of the data-set enviroments. For our purposes, our data-set is in 2-D co-ordinate space, although, let it be noted, MST algorithms can be projected by induction into k-D space [31, 30].

It is our desire that an appropriate analysis will reveal large-scale structure in the Pierre Auger Observatory data directions (if it can be detected with the data sets currently available), and so we look for a technique that will accentuate inter-related distances, which can be catagorized as belonging to a common source, or as having passed through regular magnetic fields.

Zahn [30], in his paper on gestalt clusters, maintains the most effective approach to feature or cluster retrieval (pattern recognition) is to design a metric space governed by a distance function, $d(x, y)$, that favours the relationship between feature points, and commits the unrelated points to the background position of noise: 'We shall address ourselves to the problem of detecting inherent separations between subsets (clusters) of a given point set S in a metric space governed by a distance function $\rho(x, y)$'. His reference to the 'geometry' of the metric space, and hence $\rho(x, y)$, is in terms of graph theory and the 2-dimensional linear distances between points. Zahn goes on to define separation between features as depending solely on interpoint distances within the set S. The selection of an appropriate feature space is crucial to pattern recognition algorithms and Zahn [30] specifically mentions a metric space where clusters are identical, or nearly so, as desirable. His language is the language of gaussian statistical means and standard deviations and 'We shall strive for cluster methods that are "determinate" in the sense that detection of a given cluster does not depend on random choices in the detection algorithm and is not sensitive to the order in which points of S are scrutinized'. MSTs, which are determinate[17] by nature, and

[17] it does not matter what order the points are input or which point is chosen initially

invariant under general transformations that preserve the ordering of edge lengths [32, 30], are a frequent, but data-size limited tool, in the machinery of pattern recognition techniques, and Zahn recognized their value. He nowhere addresses the problematic combinatorial nature of MSTs for large data sets. The publications by Kruskal [34] 'On the shortest spanning sub-tree of a graph and the travelling salesman problem' and Prims [32] 'Shortest connection networks and some generalizations', amongst others, are listed in his bibliography.

The Fast MST of Bentley and Friedman [31] builds on the work of Zahn, and incorporates Prim's construction principles to present an algorithm that exploits the 2-dimensional gaussian distribution properties of the points in the metric space in order to **fragment** that space. A MST is generated in each fragment until the number of nearest neighbours, **nn**s, exceeds a given input value, $\mathbf{m_0}$ (this essentially defines how the space is fragmented), whereupon the tree ceases growth. Once all the fragments are populated by their own trees, consideration is given over to the joining of these trees. Some trees will conjoin, and others shall stand alone. There is a resulting economy of computation time, excepting for small data-sets, where any of the MST algorithms mentioned will do.

This landscape of trees and solitary nodes is then scrutinized and may be cut, according to measures based on edge length statistics. The completed composition now enters the final analysis stage, where it is compared with randomly generated landscapes, and its consequent comparison statistical percentage is calculated.

We are co-opting Zahn's principle of exploiting the 'geometry' of our metric space to define desirable feature properties to be retrieved, and have assumed Bentley and Friedman's idea of fragmenting that space into areas where individual preliminary features, or 'sub-trees', are grown and then minimally joined (these features are to be considered preliminary, and will later be 'fine-tuned' to satisfy constraints we hope will accentuate the physics of the regions or space the features occupy).

We have then, instead of dividing our space into fragments with equal numbers of nodes as do Bentley and Friedman [31], assigned a minimum density to define each of our fragments or features. Our MST is constructed in an effort to highlight those areas which are anisotropic,

against a background of randomly generated landscapes where anisotropies will occur, but which, in a statistical sense, set limits on these anisotropic fluctuations.

Our MST introduces previously calculated gaussian point density functions for each node, which smear the node positions across probability densities, and a minimum probability density is set to delimit the fragments boundaries. When we broach the development of our own MST, the language of graph theory (vertices and edges) is changed into language more appropriate for our data-sets; for example, nodes or events and minimum angular distances.

5.7.1 Construction

The Pareto principle described in subsection 5.4.2 on page 129, finds full expression in the **nn** search times of the MST, and so we engage in measures to constrain these searches.

Confining the search for the **nn** to a fragment, involves two alternatives, each one reversing the perspective of the other.

1. To each fragment vertex is stored the distance and position (identity) of its closest isolated vertex.

2. To each isolated vertex is stored the distance and position of its closest fragment vertex.

And so is created a body of potential (virtual) MST edges named **virtual links** which are in a constant state of flux, being continuously updated with each vertex gathered in by the fragment.

Updating in the first strategy consists of:

(a) locating the closest isolated vertex to the new fragment vertex together with -

(b) locating those closest isolated vertices to those previous fragment vertices that had the new fragment node as their **nn**.

The computation time for this alternative is dependent on the number of **nn** calculations, **m**, required per update. If $\mathbf{m} = 1$ then then we have a minimum, if $\mathbf{m} = N$ then we have a maximum. The average value of **m** relies on the vertices' spatial configuration. The computation times range from a maximum = $O(N^2(N-1)/6)$ to a minimum = $O(N(N-1)/2)$ [31].

Updating in the second strategy involves a smallest distance **comparison** between the distance of the isolated vertices to the old fragment versus the distance of the remaining isolated vertices to the new fragment vertex. For M fragment vertices, each update time is $\propto N-M$, for $M \in (1, N-1)$ which delivers a total computational time of $O(N(N-1)/2)$.

It is this second approach that we shall adopt, although Friedman, Bentley and Finkel [47] developed an algorithm, called the k-d tree which, under certain circumstances, conducts **nn** searches in only a portion of the isolated vertices. They also developed their own fast priority queue which restricts the number of **nn** searches. The computational time for these queues whose data structures can be heaps (see sub-subsection 5.2.1.3) or leftist trees, is $O(\log N)$ where N is the number of **fragment** nodes.

Given that computational time is dictated by the algorithm's inner loop, each link insertion, being the outcome of $\mathbf{m} \times \mathbf{nn}$ searches has computational time $O(\mathbf{m} \log \mathbf{n})$. Taking $\bar{\mathbf{m}}$=average of $\mathbf{m}$ over all the links made edges in the MST, the sum total computation time is $O((N-1)\bar{\mathbf{m}} \log N)$. When $\bar{\mathbf{m}} < N/(2 \log N)$, this algorithm has a time advantage over the MSTs of Prim and Dijkstra.

5.7.2 Our Multifragment MST: The Yggdrasil

The Yggdrasil (Ygg), exploits the density properties of data event angular directions which are defined in terms of direction error parameters. Clusters of data, with a minimum of 1 event, that lie within a preset minimum density, present a fragmented metric space of angular distances.

5.7.2.1 Construction

The efficiency of MST algorithms is centred on the number of **nn** searches, $\mathbf{m}$, conducted for each edge merged into the fragment, and its average, $\bar{\mathbf{m}}$, is conditional upon the distribution of the vertices in the metric. The nature of the MST construction (i.e. Prim's construction principles), the **nn** approach, enforces an ordering of edge insertions according to the density gradients of the nodes, unmindful of the original fragment vertex's position. The timeline of edge insertions, also, is a function of the vertices density distribution - there is an accelerated convergence of edge additions towards the fragment nodes' maximum density, and then a deceleration as the fragment

node density decreases.

Bentley and Friedman [31] devised an algorithm that capitalizes on the rapid growth of the MST near regions of high relative density, and curtails its growth when it threatens to become inert by overextending its reach.

The density contours of the points in the metric space and Prim's construction principles lend themselves to a simple multifragment methodology, wherein the metric space is roughly divided into localities, or 'cells', with the same number of points. The number of points in a cell, divided by the cell's volume (area for 2-D space), are used as local density estimates for the points in the cell. Each point is then given a density estimate. Friedman and Bentley include an Appendix on k-d Trees in which, 'each node represents a sub-collection of the points and a *geometrical* partitioning of that sub-collection ... the geometrical partitioning is accomplished by dividing the sub-collection at the median value of one of the co-ordinates. The particular co-ordinate chosen to make the partition is the one which exhibits the greatest range or spread in values'.

In each region, at the point in which the local point density is at its lowest, a MST is seeded. There is an increasing infall of edge additions towards the local density point maximum, until the maximum is reached, and then at a preset juncture, the tree's growth is terminated. These minimum spanning sub-trees will either then be connected, starting with the smallest, or remain as they are. The decision to merge or rest lies in the details of the points' density distributions peculiar to that sub-tree and the adjacent sub-trees.

The multifragment methodology requires

(a) the input of a growth termination factor, $\mathbf{m_0}$.

(b) A maximum density distribution value where any vertices occupying the space with a distribution value $\mathbf{m} \geq \mathbf{m_0}$ greater than or equal to that value, are suppressed in the stage of sub-tree connection.

(c) a quantifiable division of the metric space.

The growth termination factor, $\mathbf{m_0}$, can be given a lower limit as the maximum number of **nn** searches in the regime of rapid tree growth, and an upper limit that prevents the computational time

cost from spiraling. A suitable division of metric space fashioned on the geometry of that space will actually, in itself, provide the growth termination factor. The division of the metric space has a 4-fold significance. It is used in the algorithm's beginning to set the point of lowest local density per partition, the maximum number of **nn** searches in the rapid growth stage hinges on the division, and at the algorithm's end it arbitrates whether the local MSTs will be amalgamated into global clusters or remain separate, and what vertices within a certain density distribution will be suppressed. This suppression further speeds the **nn** searches embarked upon in the joining of the minimum spanning sub-trees. Friedman and Bentleys's Fast MST algorithm is set out in their paper [31]. Our fast MST does not involve a growth termination factor, and the priority queue is the SORT operation of our main programming language IDL. Also no vertices are ultimately left unconnected or not included. We will briefly mention an attempt made to exclude vertices in Chapter 6 on page 151.

For the purposes of analysis of Pierre Auger Observatory CR directions, the fragmentation of the metric space is described by the point-spread gaussian distributions of each event. This development is described in Chapter 6 on page 148. From now on, vertices are referred to as nodes or events. We divide the the 2-D sky into degrees of right ascension, $\mathbf{RA} = [0 : 360]^\circ$, and declination, $\mathbf{dec} = [-90 : 90]^\circ$.

A point-spread gaussian distribution is applied to each event's square degree co-ordinate in the metric space. A set of gaussian densities is then stored for each integer **(RA,dec)** co-ordinate. For instance, there will be events close to one another whose gaussian distributions overlap. These distribution values are summed over their overlapping integer co-ordinate values. There will be groups of co-ordinates where there are no events. Once a global minimum cluster density is assigned (this global minimum is dependent on the memory capacity of IDL, the programming language package used), all co-ordinates with less than that minimum density are considered effectively zero, i.e. the density distributions of the CR events are divided into clusters of densities with a minimum value, separated from other clusters by co-ordinate regions of 'zero' density. A point of lowest local density per cluster is found by locating an integer co-ordinate with the smallest assigned gaussian density value.

5.7.2.2 Order of Growth

Instead of prescribing a boundary value, $\mathbf{m_0}$, for the number of **nn** searches per fragment, individual tree fragments are confined by the global minimum cluster density. Because the CR data-sets we are using are modest in size, the largest data-set being 131 events in the southern hemisphere (the data set before the division of the galaxy into southern and northern hemispheres was 208) in the low-energy range 40 EeV < E ≤ 50 EeV (see Appendix A on page 233), and the fragments generally consist of a minimum of 1 event to a maximum of 13, the predominant growth time of each fragment is $0(N \log N)$.

Finding the **nn** between fragments depends on the separation of the fragments. For our data, separation between fragments varied from quite small angular distances with respect to the angular distances within the fragment, so the fragments growth time for N vertices was $O(N \log N)$ - to large angular distance separations between small event sized fragments, and the growth time for these MSTs was more like the growth time of a Bentley and Friedman MST single fast fragment being $\propto (N-1)\mathbf{\bar{m}_s}$, where $\mathbf{\bar{m}_s}$ is the average of the number of **nn** searches.

It must be noted that growing a MST as one fragment, i.e., disregarding the events gaussian distributions, and using Prim's construction principles, does result in a MST slightly **different** from that of the multi-fragment gaussian density based approach. A Prim MST growth time with an unordered array of N vertices is $\propto O(|N|)^2$ [45] and we deliberately did not proceed with Prim's algorithm for large putative data sets because it became apparent that the process was taking too long and was less precise anyway - our multi-fragment approach based on its co-ordinate event gaussian density distributions, confined within a minimum density, is more accurate for data associated with experimental uncertainties such as the Pierre Auger Observatory CR directions.

We emphasize that we were able to grow Yggs for data sizes up to 1,297 events in 37.08 minutes (see page 150 and also Figure 6.5 on page 154 for an Ygg of 952 events) so this method is suitable for much larger data-sets even though these large data-sets did not eventuate within the time frame of the Pierre Auger Observatory data directions used.

5.7.2.3 Cutting the MST

Once the MSTs are completed, cuts may be applied to isolate preliminary features. We decided (rather than using edge statistical properties, for example, by deleting edges that were a certain size larger than the average size of edges, on the premise that the regions so divided were probably unrelated to one another) to divide the galactic (**l,b**) sky[18], into partitions in the **(l,b)** plane. The galactic disk region latitudes $\in [-15 : 15]^{\circ}$ were singled out, and galactic longitude, **l**, regions where the galactic magnetic fields may be considered chaotic or regular, also were isolated. The remains (branches) of the MST tree structures within those galactic (**l,b**) partitions were subjected to an analysis which is developed in Chapter 6 onwards.

[18]The **(RA,dec)** co-ordinates were transformed into their galactic longitude and galactic latitude co-ordinates, **(l,b)** preserving the order of edge distances.

Chapter 6

Method of Analysis

Our Minimum Spanning Tree (MST) approach to the Pierre Auger Observatory and simulated data analysis is very simple. The only assumptions made are that non-random events will exhibit behaviour contrary to simulated random data distributions. Whatever the appearance or value of the distribution's 'norm', the task is to find ways in which the Pierre Auger Observatory data do not conform within reasonable measures of fluctuation, such as will be found in simulated data-sets. For instance, if the distribution was to be Normal, Pierre Auger Observatory data that either fall outside 3 standard deviations of the simulated data mean ($\lesssim 0.1\%$), or lie between 2 and 3 standard deviations ($\lesssim 5\%$), can be flagged as possibly interesting.

We chose more than one variable to test as non-random. The assembly of each variable's simulated random distribution statistics indicates which Pierre Auger Observatory data should be searched for correlations peculiar to their location in the context of known physics.

6.1 Gaussian Density Contours

The history, theory, and limited use of MSTs in finding the shortest measures between networks of nodes (mathematicians' collective noun for events) is contained in Chapter 5. The minimizing of these event networks could be in terms of distance, or time (often the same), or any other measure desired to be minimized. Here, these events are cosmic ray co-ordinates in the sky minimized over their angular distances. The terms event or node will be used interchangeably. This MST

approach to data analysis can be summarized as follows: density contours of events are created. Feature events are defined according to a set minimum density of their angular distances. Here, we take the 2-Dimensional **(RA,dec)** co-ordinate sky, over 360 Right Ascension (**RA**) degrees by 180 declination (**dec**) degrees and arrange the co-ordinate sky as a (360×180) degree squared grid. To each event is applied a Gaussian density contour, smearing the co-ordinates by an angular resolution, $\sigma = 1.4^\circ$ [20]. Each sky square degree integer co-ordinate is associated with a Gaussian density. Because we are working with Gaussian densities which at their asymptopic limits $\mapsto 0$, and we wish to divide the sky into quantifiable, distinct areas of information, we set a global minimum density, ρ_{min} (dependent on the computer language package's, IDL 4.1, memory capacity), and only work within sky areas with densities such that, $\rho \geq \rho_{min}$. If the integer co-ordinates of events are close enough, the summation of their overlapping density co-ordinate values may well be $\geq \rho_{min}$, and these events merge into more than one event per density contour. We produce a $(360 \times 180)^\circ$ squared sky map of density contours. If events are sparse and isotropic, there can be many single event density contours, none of which overlap. When data-sets are non-isotropic, the more events there are, the more we find contours merging, and event co-ordinates lie in groups defined by the global minimum density boundary. Here, we define such feature groupings of events as clusters. The smallest cluster numbers 1 event.

6.1.1 Clusters and Sub-Trees

Each cluster is considered as an individual MST, or rather, as containing a sub-tree of an entire MST, the Yggdrasil[1] (Ygg pronounced "Igg"). Because MSTs present as combinatorial problems - problems where a small increase in events results in a massive escalation of computation (see Chapter 5), the number of data points able to be manipulated has always been low. By creating sub-tree density clusters, and tackling each cluster as a sub-tree, event numbers can be contained and the computation becomes feasible. One calculates the angular distance between all the sub-tree events, i.e. for each sub-tree event , **n**, there are $\mathbf{n-1}$ angular distances computed. These angular distances are computed for each event in the sub-tree.

[1]The Yggdrasil is the great tree that connects all nine realms of Scandanavian cosmogony. Its trunk forms the Earth's axis.

The minimum distance between all these events is chosen with the caveat that no loops are created in the tree (acyclic), and each path between events is distinct. Thus a tree with **n** events has $\mathbf{n-1}$ edges.

Once each cluster sub-tree is completed, the sub-trees themselves are joined together, according to their minimum angular distances, throughout the Ygg. The no-loop caveat and restriction of n-1 edges per Ygg excludes data points that have already been used twice within their separate sub-tree. i.e. once as the end of a minimum span and once as the start of a minimum span. A node that is already part of the sub-tree structure such as the end of a minimum span, is still considered eligible as a starting point for a minimum distance, together with all nodes completely unconnected to the sub-tree structure. A sub-tree approach can dramatically reduce the number of calculations and time spent on growing Yggs.

I have transcribed 2 entries from when growing trees:

a) for 1,191 events in the $(30-60]$ EeV range in 2009

the sum of calculations for a density MST = $18,270,907$

start time Ygg = 1268287370 s

end time Ygg = 1268288009 s

Total growth time of Ygg =10.65 min.

b) for 1297 events in the $(10-12]$ EeV range in 2009

sum of calculations for density MST = $45,432,014$

start time Ygg = 1268287370 s

end time Ygg = 1268288009 s

Total growth time of Ygg = 37.08 min.

It was decided to allow the maximum number of events in an Ygg to be ~1,300. For example, a MST with 2,000 events was still being computed after 17 hours, after which it was abandoned.

Typical examples of density contours and their subsequently captured sub-trees are shown in Figure 6.1 and Figure 6.2 on page 152. For the sake of continuity and to demonstrate that fairly large event numbers can be manipulated, we have chosen the demonstration MST method to be in the energy range 20 EeV $<$ E $\leq$ 30 EeV, for Pierre Auger Observatory data which were current

until 2/8/2010. Any alternative displays are to highlight some of the problems and rethinking done in the development stage. The fully grown Ygg is displayed in Figure 6.4 on page 153.

Note here that the illustrated sub-trees are captured within the original density contour cuts, $\rho_{\mathbf{min}}$ (see Figure 6.3 on page 153). There were times when the programming language, IDL, reported not enough memory available for energies < 12 EeV and the large Pierre Auger Observatory event numbers, say **n_cut**, to be captured within the original density contour of the event gaussian spreads. Then the density contour minimum, say $\rho_{\mathbf{min}} = \rho_{\mathbf{cut}}$, had to be increased before IDL would capture events within this reset density boundary minimum, i.e. $\rho_{\mathbf{cut}} > \rho_{\mathbf{min}}$. Information lost was in the sense of errors which may be defined in the size of the angular resolution error ($\rho = 1.4^{\circ}$). The core events were retained. However there is already some information lost with the setting of $\rho_{\mathbf{min}}$. Increasing $\rho_{\mathbf{cut}} > \rho_{\mathbf{min}}$ means error margins are not as well defined. No such increase in the original $\rho_{\mathbf{min}}$ density cuts were needed for the Pierre Auger Observatory Yggs for energies > 40 EeV and their simulations. Errors remained the same as those initial errors for $\rho_{\mathbf{min}}$, i.e. $\rho_{\mathbf{min}} \equiv \rho_{\mathbf{cut}}$.

The advantages of making density cuts are potentially 2-fold. Increasing the cluster density cut, whilst still retaining the core event co-ordinates is computationally more efficient for IDL, the language of most of our programs, and, given our goal of obtaining minimum angular distances between events, no relevant information is lost, excepting, as mentioned before, in the sense of errors which may be defined in the terms of the angular resolution.

In those instances when event numbers in sub-trees are too large and the combinatorial calculation is once again a problem, only events occupying the reset densities, $\rho_{\mathbf{cut}} > \rho_{\mathbf{min}}$, are captured for the subsequent MST calculation. The latter is only an advantage when full $(360\times180)^{\circ}$ squared sky exposure is achieved and we may be confident anisotropies isolated don't fall into a limited single region of sky exposure.

The modest numbers in the highest three energy ranges of Pierre Auger Observatory data ensure that all those events are used or 'caught'.

This method builds upon associating events by their densities. Each sub-tree represents an association. Thus events in a sub-tree may all be connected to a single source, such as a supernova,

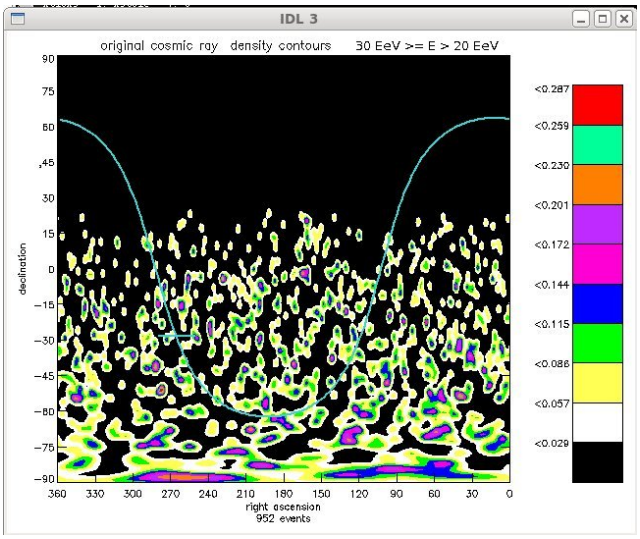

Figure 6.1: Density contours in **RA** and **dec** for all 952 events 20 EeV $< E \leq$ 30 EeV. The events at their co-ordinate positions are not shown (current to 2/8/2010).

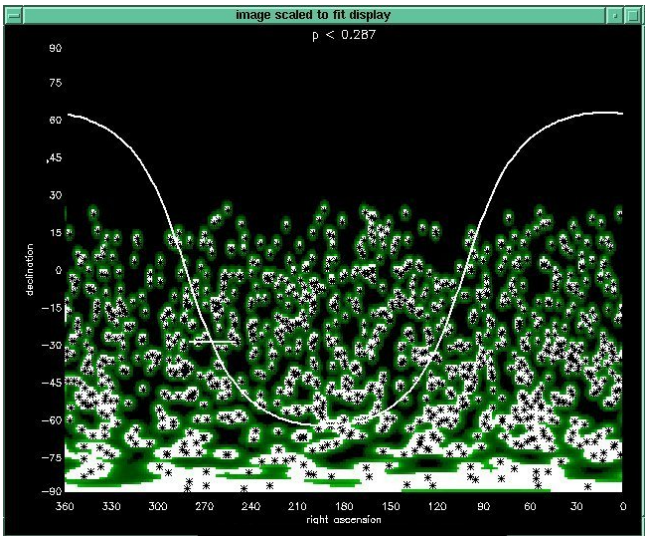

Figure 6.2: Density contours with all 952 captured events in **RA** and **dec** in 20 EeV $< E \leq$ 30 EeV. The events at their co-ordinate positions are denoted by $*$ (current to 2/8/2010).

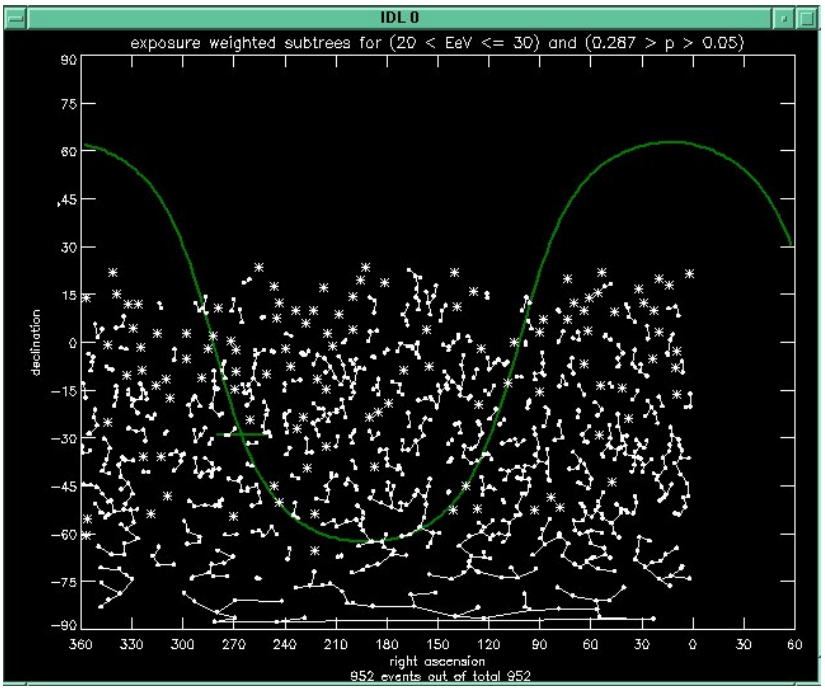

Figure 6.3: Sub-trees of **RA** vs **dec** for all 952 events in 20 EeV < E ≤ 30 EeV. The events at their co-ordinate positions are denoted by ∗ for single event sub-trees and · for sub-trees comprising more than one event (current to 2/8/2010).

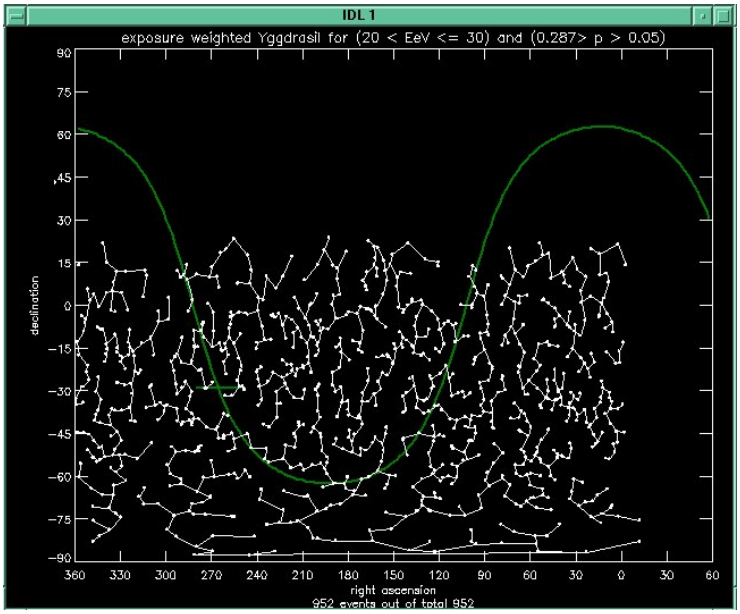

Figure 6.4: Ygg of **RA** vs **dec** for all 952 events in 20 EeV < E ≤ 30 EeV. The events at their co-ordinate positions are denoted by · (current to 2/8/2010).

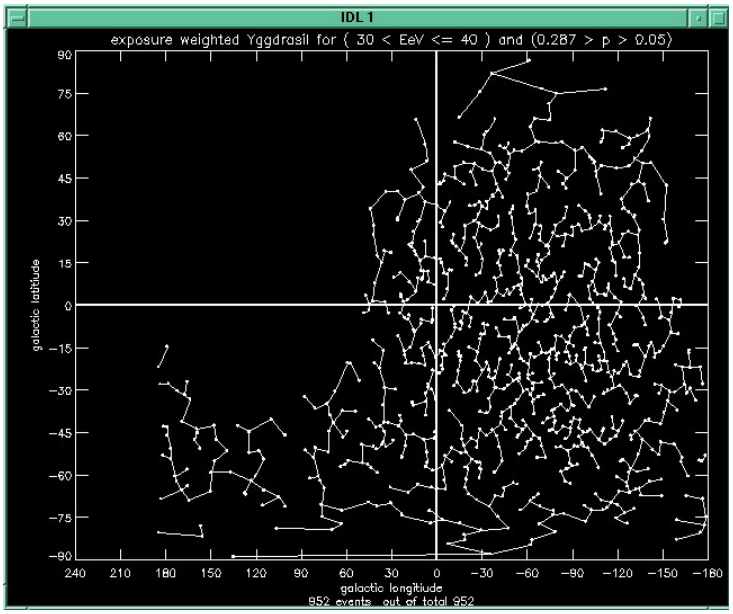

Figure 6.5: Ygg of galactic **l** vs galactic **b** for all 952 events in 20 EeV < E ≤ 30 EeV. The events at their co-ordinate positions are denoted by · (current to 2/08/2010).

or a number of minimally connected sub-trees may represent a collective assocation with a large-scale cause, such as a *Cosmological Compton-Getting* (CC-G) effect [27, 64].

6.2 Divisions in the Galactic Plane and Branches

The Ygg is then divided in the galactic **(l,b)** plane, along galactic latitude **b** ranges, **b**Bands for co-ordinates which may be chosen to be considered significant, for example, as in the 'south' and 'north' latitudes of the galactic disk, the **b**Band galactic latitudes $[-15 : 0]^\circ$ (south) and $(0 : 15]^\circ$ (north). When divisions in the plane are mentioned, these are taken to be divisions in the galactic **(l,b)** plane. The initial **b**Band divisions were $[-90 : -45]^\circ, [-45 : -15]^\circ, [-15 : 15]^\circ, [15 : 45]^\circ$ and $[45 : 90]^\circ$, as seen in the demonstration examples of the MST divisions in the plane. We are still guessing about the extent and nature of the galactic halo, and so, apart from the galactic disk latitude divisions, the remaining divisions are arbitrary. We note that the Pierre Auger Observatory exposure is maximized in the galactic co-ordinate $(\mathbf{l}, \mathbf{b})$ area $(-90, 0)^\circ$ to $(-15, -60)^\circ$. A contour plot taken from [120] picks out Pierre Auger Observatory exposure in shades of blue and is seen

in Figure 6.6.

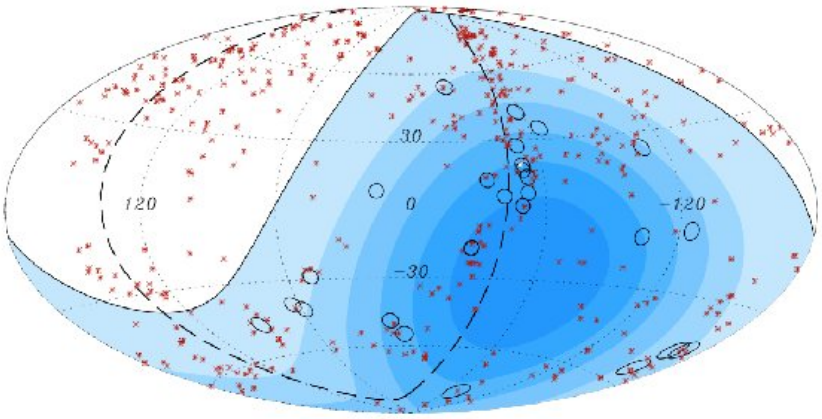

Figure 6.6: Skymap of original Pierre Auger Observatory CR directions for 27 events with E > 57 EeV. The events are circled in black. Solid line - the border of the field of view of the Pierre Auger Observatory, surrounding contours of relative exposure. The darker the contour shade of blue, the larger the relative exposure. From [120].

These divisions in the plane partition the Ygg into 'branches' within the divisions. An exception is made when an Ygg event inside a Band is connected to a single event that lies outside the Band cut. That single event then becomes a part of the branch.

6.2.1 Branch Thetas

When one supposes minimally connected events - events that lie within the same cluster or, for example, a sequence of contiguous clusters, are connected because of a common source or cause, and one considers the Ygg as a minimally connected random walk, the angles, Θ, of the Ygg branches in the plane which enjoy a non-random distribution, may indicate a non-random cause, or source.

The branch-walk Θs are calculated as follows[2].

For a 'vertical' walk - a **b**Walk, in the plane where $\mathbf{l} = 0$ and $\mathbf{b} \in [-90, 90]^\circ$, one lists the starts and ends of each branch span in the **b** ordinate and then one finds the maximum, $\mathbf{b_{max}}$, and

[2]The Ygg sub-trees were first found in their **(RA,dec)** co-ordinates and then these co-ordinates were transformed into their galactic latitude and longitude co-ordinates, **(l,b)**. Angular distances in right angle co-ordinate systems remain invariant.

minimum, $\mathbf{b_{min}}$, of that list. The **b**Walk is then taken as

$$\mathbf{b}\text{Walk} = |\mathbf{b_{max}} - \mathbf{b_{min}}|$$

For a 'horizontal' walk - a **l**Walk, in the plane where $\mathbf{b} = 0$ and $\mathbf{l} \in [-180, 180]^\circ$, one lists the starts and ends of each branch span in the **l** ordinate. Each start/end is multiplied by the cosine of its **b** start/end ordinate.

One then finds the maximum, $\mathbf{l_{max}}$, and minimum, $\mathbf{l_{min}}$, values of these spans.

The **l**Walk is thus

$$\mathbf{l}\text{Walk} = |\mathbf{l_{max}} - \mathbf{l_{min}}|$$

These galactic **l**Walk and **b**Walks are also referred to as galactic **l/b** branch-walk-pairs and their co-ordinates as galactic **l/b** branch-walk-pair/Regions of Interest (ROIs) values.

The Θ of a branch's deviation from 'vertical' or 'horizontal' is taken as

$$\Theta = |\arctan(\mathbf{b}\text{Walk}, \mathbf{l}\text{Walk})|$$

We take angular magnitudes to lie between $[0 - 90]^\circ$. The most obvious angles to be non-random are when events, within a cluster, or belonging to several contiguous clusters, line up horizontally or vertically, such that their angular random walk is at the **extremes** of these scalar angular co-ordinates between $(0 - 90)^\circ$. That is, a horizonal or vertical alignment of events displays a non-random pattern. Such a non-random, artificially inserted, horizontally aligned series of events is seen in Figure 6.7 on the next page

6.2.1.1 The Use of Branch Θs

Coherent magnetic fields present in patterns. The simplest patterns are depicted as magnetic field *lines*. Analysis of data current at the time of writing, indicates that the coherent galactic magnetic fields (GMFs) are 'vertical and/or horizontal', or, rather, azimuthal and/or poloidal, [21, 22, 23, 24, 25]. We thus have a special interest in 'vertical' or 'horizontal' deflections.

We may expect, within the discipline of coherent large-scale GMFs, UHECRs being accel-

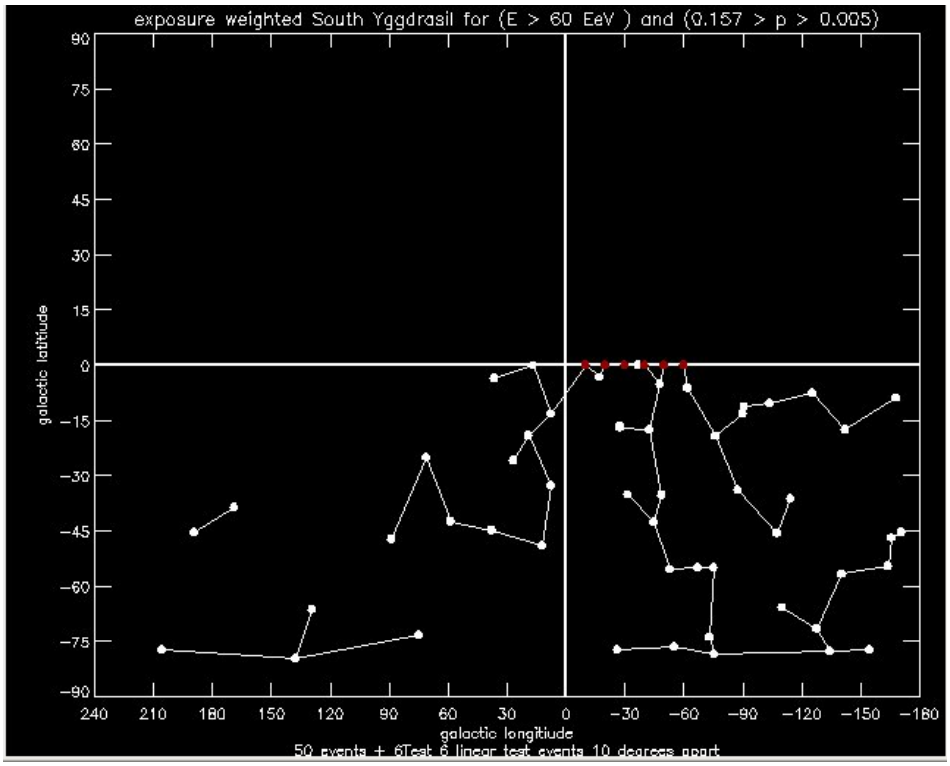

Figure 6.7: One horizontal, non-random linear structure of 6 events placed within 50 Pierre Auger Observatory events for E > 60 EeV is shown in red (current to 8/11/2012).

erated by the same source in roughly the same direction, will, upon encountering magnetic field lines, suffer a deflection associated with the angle between their path, field direction (light nuclei and heavy nuclei will have different deflection angles) and their energy. If the UHECR energies are high enough, the UHECRs suffer very little deflection and continue (towards us), in which case the direction of the UHECR would be close to the direction of the source and various energy-energy ordering techniques (see Chapter 4, subsection 4.3.3 on page 97 [114, 166, 167, 168, 169, 144, 194]) may be applied to UHECR branches if the branches are shown to be statistically significant.

Light UHECRs travelling with an angle δ, between the CR's velocity, $\mathbf{v}$, and the magnetic field, $\tilde{\mathbf{B}}$, produce a gyroradius, r_G, such that

$$r_G \approx \frac{E}{Zec|B|}$$

where E is the energy of the CR, Z is the CR atomic number and $\mathbf{B}$ the magnetic field. e is an elementary charge and c is the speed of light. The subsequent angular deflection of these light UHECRs , $\Delta\Omega$, at a distance L, at right angles to a regular magnetic field, $\mathbf{B}$, is approximately [194]

$$\Delta\Omega \approx 13^\circ \frac{40\text{EeV}}{E/Z} \frac{L}{2\text{kpc}} \frac{|\mathbf{B}|}{5\mu\text{G}} \tag{6.1}$$

If a sequence of single source UHECRs are heavy, the angular deflection of the UHECRs would be larger. Consequently, a sequence of same nuclei, similar energy, single source CRs should trace a linear distribution of event directions.

The angular deflection of UHECRs is of interest with respect to the **Type 1** distributions and the magnitude set for non-random or **extreme** branch Θs (see Chapter 7, section 7.1 on page 182).

The importance of branch Θ magnitudes is predicated on azimuthal or poloidal magnetic fields, and the galactic divisions in the plane being relevant. Whilst the linear nature of CR deflections in the GZK energy regime will hold, the linear patterns may be hidden by divisions in the plane and we also cannot assign the type of CR nuclei as being light or heavy with great confidence.

It has been established that heavy CR nuclei can be delivered without disintegration by Seyfert galaxies within a reconfigured Hillas limit for energies $\sim 10^{20}$ eV, whereas protons could not be

accelerated beyond ~5×10^{19} eV [11]. It has also been established that the angular deflections of light or heavy nuclei in the same magnetic field look much the same when they possess similar rigidities. For example, the deflection of 60 EeV heavy nuclei would be similar to the deflection of 2 EeV protons in our GMF. However, heavy nuclei deflection signatures cannot always be characterized as straight-forward light nuclei signatures magnified by Z and exhibiting energy ordering, $1/E$, features. The appearance of several images in the sky from a single heavy nuclei source is peculiar to the geometry of the magnetic fields the source CRs traverse [169]. At this stage of fallow appreciation of GMFs and extra-galactic magnetic fields we are largely unable to impute the light or heavy nature of UHECRs. Photo-disintegration of medium extra-galactic nuclei would void their appearance, but the two extremes of unknown light or heavy nuclei remain. Given this, light or heavy nuclei with small/large angular deflections are likely candidates in the UHECR regime. This MST method of tracing large-scale structures or isolating point sources may help clarify the orientation, regions, magnitude and nature (regular or chaotic) of magnetic fields, whether galactic or extra-galactic, or, at GZK energies, help to locate sources.

We call the triangular region covered by a branch Θ, a possible region of interest (ROI), and when ROIs from different Θs overlap, we attempt to assign an underlying astrophysical cause and we consider these Θ ROIs interesting. For the purposes of this thesis, we take the area the branch Θs cover to be directed right angled triangles with the angle Θ measured from the 'anchoring bottom' horizontal **l**Walk = $|\mathbf{l_{max}} - \mathbf{l_{min}}|$ versus the vertical **b**Walk = $|\mathbf{b_{max}} - \mathbf{b_{min}}|$, the **b**Walk being set on the right-hand-side of the **l**Walk. Our branch Θ theory has branch Θs undirected, our Ygg is scalar and our interest in the branch Θs was primarily to see how close these Θs values were to 0° (horizontal) or 90° (vertical), but we decided to follow the convention as previously just set out.

To have a more accurate appreciation of the region the branch Θs cover, we should consider the rectangular area covered by all 'directions' of our branch Θs. An example of overlapping triangular branch Θs ROIs is seen in Figure 6.8.

Figure 6.9 on page 162, and Figure 6.10 on page 163, feature MST branches for data in the same energy range, E > 60 EeV, but gathered at different dates. Figure 6.9 is for Pierre Auger

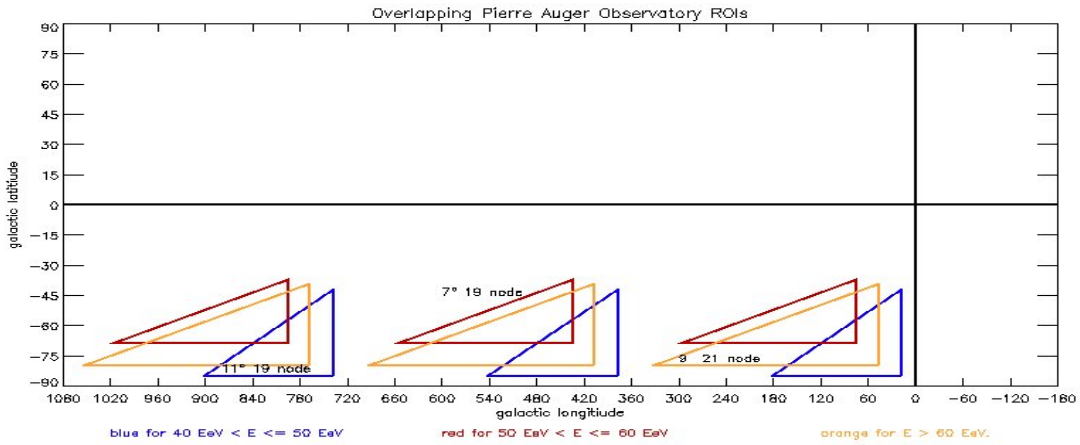

Figure 6.8: Flat skymap of Pierre Auger Observatory branch Θs with overlapping ROIs (over 2.5 cycles in galactic longitude, **l**). For energies 40 EeV < E ≤ 50 EeV, 50 EeV < E ≤ 60 EeV and E > 60 EeV. Blue for the low-energy range. Red for the mid-energy range. Orange for the GZK-energy range. Branches are **b**Band divisions in the plane. See also page 159. (current to 8/11/2012).

Observatory data current until 2/8/2010. Figure 6.10 is for all the Pierre Auger Observatory data current until 8/11/2012, and includes the data in Figure 6.9.

The disposition of these branches and their branch Θs and branch node sums at these two different times, led to a rethink on aspects of the generation of simulated data and overall analysis of Pierre Auger Observatory data (see section 6.4 on page 164).

6.2.2 Branch Node Sums

One aspect of interesting non-random **extreme** behaviour of branch Θs would be a large number of nodes/events in a branch. This is also the case if the Pierre Auger Observatory's branch node sum is considered as an independent property (as in terms of the **Type 2** distributions (see below)) and the Pierre Auger Observatory's branch node sum displays as **extreme** against this distribution, i.e. ≤ 5% of the simulated **Type 2** distributions branch node number is above the Pierre Auger Observatory branch node number. The reasons for a statistical significance may lie in the origins, whether galactic or extra-galactic, in the regions of the sky that these branches occupy.

6.3 Shuffled Distribution Types

So, where is a Pierre Auger Observatory branch Θ, and/or its branch node number seen to be unusual, or rather **extreme**? To this end, it was important to investigate the configuration of simulated random distributions of branches and the nodes of which they are composed. The shuffled distributions (see subsection 2.7.1 on page 55 and section 6.5 on page 166) chosen to work with were –

Distribution Types

Type 1: Distributions of the node number composing a branch Θ. Could we demonstate that **extreme** branch Θs were in any way distinct from all the other branch Θs possible? And what limits do we place on the magnitude of **extreme** branch Θs?

Type 2: Distributions of the frequency of node numbers in branches. Was there a preference in a branch's node number? What would we consider as an **extreme** branch node sum?

Type 3: Distributions of the sum of branch nodes per shuffle in each **b**Band. Were the sum of branch nodes per shuffle approximately constant (uniform) or non-uniform? i.e., is the total node sum of all the branches in each node shuffle within of the one standard deviation of the distribution's mean? When would the totalled sum of Pierre Auger Obsertvatory branch nodes in a **b**Band be considered **extreme**?

Given that the branches have been divided along **b**Bands, each Band features a distribution. i.e., in the **b**Band divisions of $[-90 : -45]^\circ, [-45 : -15]^\circ, [-15 : 15]^\circ, [15 : 45]^\circ$ and $[45 : 90]^\circ$, there are three distribution types for each one of the **b**Bands. A more detailed discussion of each of these distribution type simulations is found in subsection 6.5.1 on page 168. The quantifying of at which value a distribution type will set a Pierre Auger Observatory variable as '**extreme**' is set in relation to the Pierre Auger Observatory's data in Chapter 7 in section 7.1 on page 182.

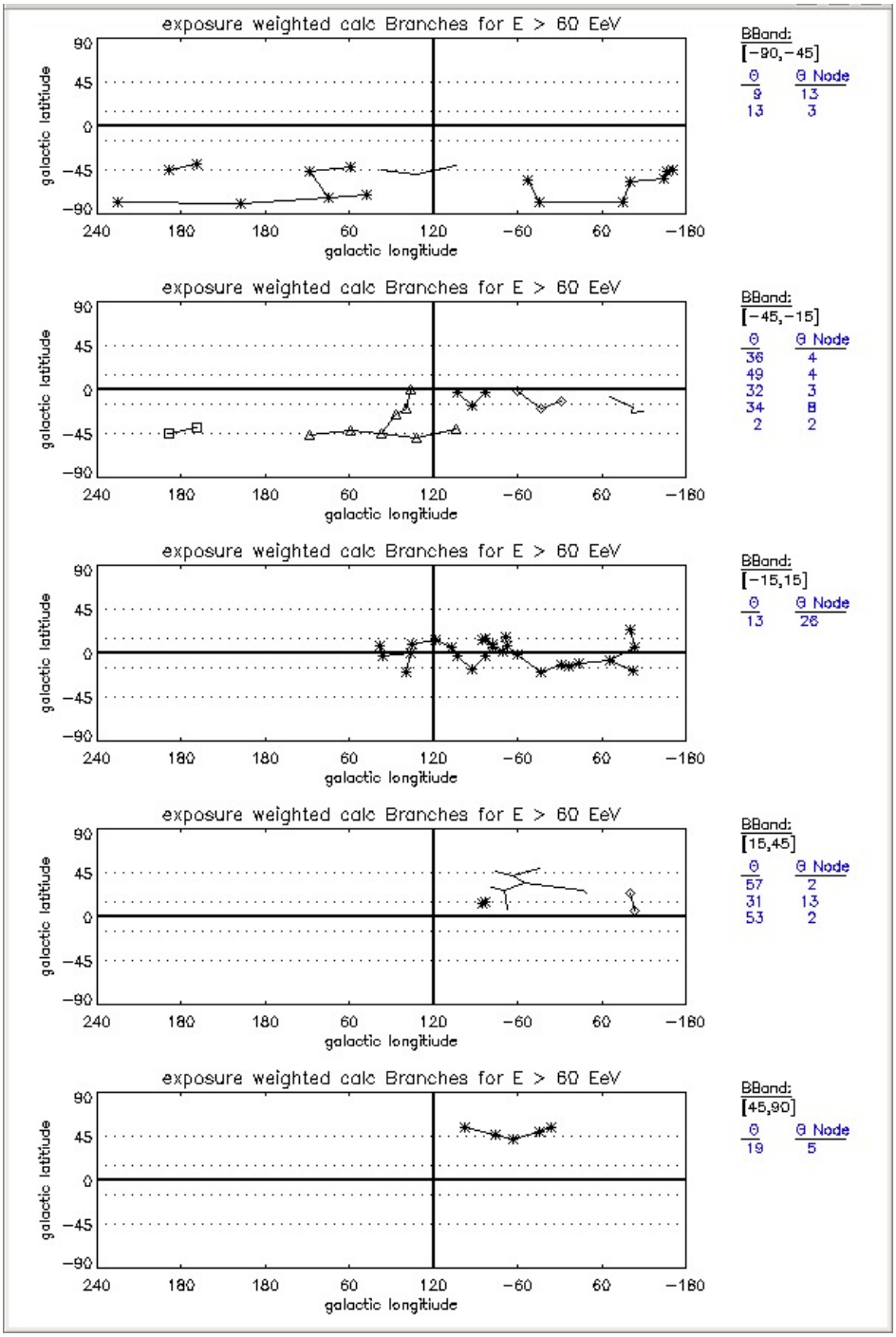

Figure 6.9: Pierre Auger Observatory data branches of the Ygg in galactic **l** vs galactic **b** for events E > 60 EeV. The value of each **b**Band's branch Θ (see 155), with their node numbers, Θ Node , are listed on the RHS of the **b**Band plots. Different symbols represent different branches. **Extreme** Θs are 9° with 13 nodes and Θ = 13° with 3 nodes in the **b**Band [−90 : −45]°. Θ = 2° with 2 nodes in the **b**Band [−45 : −15]°. Θ = 13° with 26 nodes in the **b**Band [−15 : 15]°, and Θ = 19° with 5 nodes in the **b**Band [45 : 90]° (current to 2/8/2010).

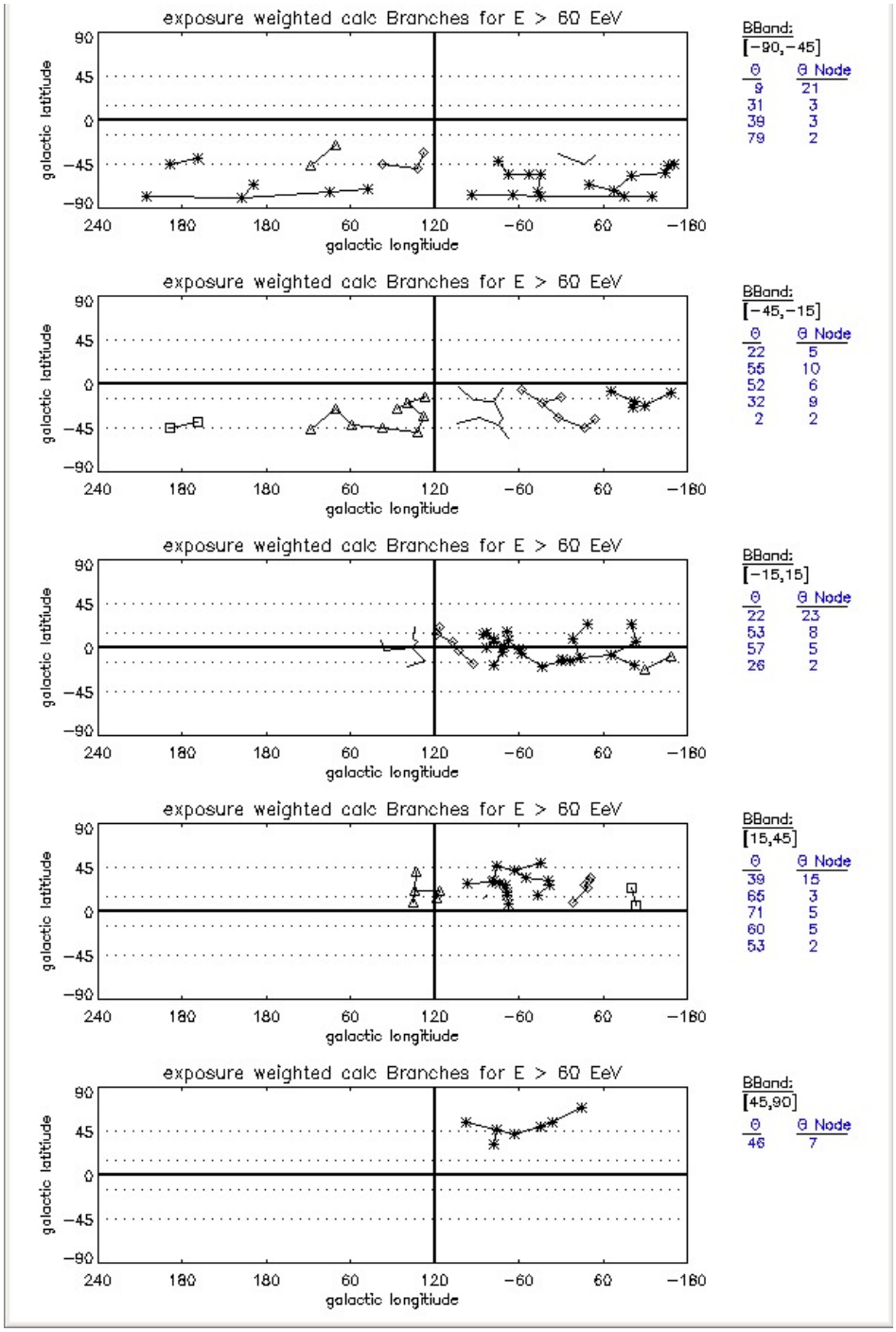

Figure 6.10:
2012 extension to 2010 data. The value of each **b**Band's branch Θ (see 155), with their node numbers, Θ Node, are listed on the RHS of the **b**Band plots. Different symbols represent different branches. **Extreme** Θs are 9° with 21 nodes in the **b**Band $[-90 : -45]°$, and Θ = 2° with 2 nodes in the **b**Band $[-45 : -15]°$. **Note in particular the changes to the configuration of branches in the bBand** $[-15 : 15]°$ **for Pierre Auger Observatory data current until 2/8/2010 (see Figure 6.9 on the previous page). The decomposition from the single 2010 branch's extreme Θ of** 13° **to a 2012 sequence of four non-extreme branch Θ values, together with fewer branch nodes was thought to be too important for a continuation of the MST analysis in its form at the time (see section 6.4 on the following page).** There were no more **extreme** Θs for the higher **b**Bands $[15 : 45]°$ and $[45 : 90]°$ (current to 8/11/2012).

6.4 Revision of the Galactic Disk bBand Division of $[-15 : 15]^\circ$

Figure 6.9 on page 162 is an example of the **b**Band branches for EeV > 60 for data current to 2/8/2010. The interesting results span the galactic disk, with **bBand** $[-15 : 15]^\circ$ and are in keeping with an initial expectation (see section 2.7 on page 53 of Chapter 2) of the galactic disk being a region of interest. The 2010 branch Θ of the sole branch in this **b**Band, $[-15 : 15]^\circ$, is 13°, and its node number is 26. The entire number of events for E > 60 EeV at that time was 59. However, when an Ygg was then grown for the 95 data events current to 8/11/2012, the "small **extreme** Θ valued" branch in this particular **b**Band dissolved. The MST had changed with the addition of more data. The branches in the **b**Bands $[-45 : -15]^\circ, [-15 : 15]^\circ, [15 : 45]^\circ$ had all been reconfigured. In particular, the *single* **extreme** branch across the galactic disk **b**Band with $\Theta = 13^\circ$ and 26 nodes of the 2010 data, was now four branches, the smallest 2012 branch Θ having $\Theta = 22^\circ$ and 23 nodes. While this branch's node sum was still substantial, its branch $\Theta = 22^\circ$ was no longer **extreme** and was considerably larger than the branch $\Theta = 13^\circ$ of the original 2010 data. We had expected the 38 events in the 2012 data of this **b**Band to have shown a smaller **extreme** branch Θ with more nodes than the 2010 branch if this **b**Band division in the plane had relevance.

The **b**Bands of this Pierre Auger Observatory data-set are shown in Figure 6.10 on the previous page.

The dilution of results across the galactic disk with **b**Band $[-15 : 15]^\circ$ (contrast Figure 6.9 with Figure 6.10) led to another look at the **b**Bands themselves. We were aware that the galactic magnetic fields were reversed across the disk [25, 22], and when it was further discovered that two giant magnetized radio outflows, or lobes, emanating from the central 200 pc galactic centre were north and south of the galactic plane ([26] - published in January 2012) we realized that, for the UHECRs we were analysing, the inclusion of these lobes may well be pivotal. Extra-galactic CRs with energies around the GZK limit may have directions connected with these lobes. Perhaps the lobes are the site of a further acceleration of extra-galactic CRs. We also have believed that none of the previously recognized galactic sites (see, amongst many, [19, 16, 91]), of CR

generation and acceleration of UHECR's were capable of contributing to the GZK CR energies we encounter [17, 18]. The galactic lobes recently revealed (and associated features such as Fermi bubbles), are not yet exhaustively surveyed, and may yield physical properties capable of galactic GZK accelerations. Maybe a better approach to the analysis of these events would be to divide the galaxy into north ($\mathbf{b} > 0°$) and south ($\mathbf{b} \leq 0°$) **b**Bands, and grow separate Yggs. The **b**Bands were subsequently changed to $[-90 : -45]°, [-45 : -15]°, [-15 : 0]°$ for the southern galactic directions and $(0 : 15]°, [15 : 45°], [45 : 90]°$ for the northern galactic directions.

The simulated data (see section 6.5 on the following page) are taken from the co-ordinates of Pierre Auger Observatory events within certain energy ranges related to the GZK energy limit. The energy ranges are regarded as generalized limits for light and heavy CR nuclei. Thus we have, for simulated events to which we can ascribe to distances within the GZK limits, an energy range of $E > 60$ EeV. GZK onset energy limits are between $50 \text{ EeV} < E \leq 60$ EeV. Ultra-high energies which we consider pre-GZK limited are $40 \text{ EeV} < E \leq 50$ EeV. We did not attempt to optimize the simulated data (see Chapter 4 section 4.3 on page 89), because the primary thrust of this analysis does not prefigure GMF models or large-scale source distributions. Also, this analysis method is new and we wished to expose any general possibilities that the method may suggest. For instance, small-scale sources, such as Cen A [194], would not necessarily exhibit events in the angular extremes we have mentioned. Events in limited energy ranges associated with small sources, via this method, would probably be connected in the primary density clusters, i.e. they would be connected by proximity and within a sub-tree cluster. In such a case, the number of nodes in a sub-tree itself could be significant. The Θ values of a branch which connects sub-tree clusters could obfuscate significant proximity within clusters which may be evident in the number of events of the sub-tree cluster. Also, we may see branches with non-extreme Θs composed of, say, a *line*[3] of any angle Θ, together with a number of possibly related or unrelated events in the branch which, after all, has its range within preset **b**bands and **l**bands. Branches with large node numbers may be atypical against their simulated **Type 2** and **Type 3** distributions (see sub-subsection 6.5.1.2 and sub-subsection 6.5.1.2) which do not include the branch Θs as a variable. Pattern recognition algorithms such as searching for **extreme** branch Θs are an aid, but

[3] regular magnetic fields may be expressed in the form of all the conic sections [12]

do not replace a visual review (see section 7.3 on page 188). Another possibility is that a source, accelerating CRs across energy ranges, may possibly exhibit relationships spanning these energy ranges. We may find, in the *same* general area of the sky, **extreme**/non-extreme branch Θ's/ROIs in *different* energy ranges. CRs above the GZK limit would presumably be less deflected than CRs below the GZK limit (see subsection 4.2.3, page 92 on). We note that calculations for light and heavy "GZK nuclei" now quantify their individual GZK limits as roughly the same [59].

So, we look for Pierre Auger Observatory event branches which manifest as atypical with respect to branches in simulated distributions. These distributions were chosen to explore the appearance of random data branches with respect to the number of events per branch and/or the branch's walk angle. Whilst large (lots of events) 'horizontal' and 'vertical' Pierre Auger Observatory branches may indicate large-scale sources, we still find interesting branches which display as atypical against distributions, but whose Θs may not be suggestively **extreme**. If these Θ branch-random-walk/ROIs values correlate with known astronomical objects, for example, Cen A $[-50.5°, 19.4°]$, or one of the galaxies in the Virgo cluster such as M87 $[283.7°, 74.49°]$, or the Sombrero galaxy $[298.5°, 51.1°]$[4], or if atypical branches dominate the same area of sky across energy ranges, then such branches warrant further study. A branch might be considered atypical if its node composition number largely outstrips the simulated node composition number in a distribution. **Extreme** Θs with only 2 or 3 nodes in an energy range may be found significant if **extreme**/non-extreme Θs in other energy ranges manifest similar **b**Walk and **l**Walk/ROIs co-ordinates.

6.5 Generation of Simulated Data-Sets

Simulated data-sets, numbering 1200, for each of the low, mid and GZK energy ranges already mentioned, were taken from the reconstructed directions of the Pierre Auger Observatory data. These data sets are termed 'shuffles'. The individual azimuth, local sidereal time, and zenith $< 60°$, were randomly coupled using the IDL RANDOMU function. Our code ensures that each

[4]None of the Pierre Auger Observatorys branches **l**Band random-walk co-ordinates were as far 'left' as the Sombero galaxy's galactic longitude - the Pierre Auger Observatory exposure was negligible here.

direction is unique, and those repeated are rejected. A MST calculation will halt (stop in its tracks) if co-ordinates are repeated. The considerations dictating the assemblage of simulated data are discussed in subsection 2.7.1 on page 55 which covers the Pierre Auger Observatory prescription. Here, no assumptions of isotropy of background data events within zenith ranges are made. Whilst the data numbers for all our specified energy ranges are low, we do not introduce any reconstructed directions from lower energies ∼(1 - 40) EeV to make up a statistically more satisfying larger data-set of 3,000 as suggested in [13]. As already mentioned, we would like to refine the GZK energy limit, and inclusion of simulated data from lower EeV energy ranges may have consequences. For example, above the GZK energy limit, there may be a *Cosmological Compton-Getting* effect [27] from the perspective of the different reference frames of this, our galaxy, and the Super-Galactic Plane.

The separation of Pierre Auger Observatory data into galactic north ($\mathbf{b} > 0°$) and galactic south ($\mathbf{b} \leq 0°$) galactic co-ordinates was carried through in the simulations. To be clear, for each energy range, the original southern and northern Pierre Auger Observatory reconstructed data directions were used to generate 1,200 shuffles purportedly in each of the galactic southern and northern 'hemispheres', and Ygg's grown.

Shuffling reconstructed galactic north and galactic south directions comprising local sidereal time, azimuth and zenith does not result in purely galactic north and purely galactic south shuffle event directions. The azimuth variable of original north directions can result in reconstructed directions that are sometimes in the southern galactic hemisphere and original south directions can result in reconstructed directions that are sometimes in the northern galactic hemisphere. A shuffle example is displayed in Figure 6.11. An anisotropic spread of Pierre Auger Observatory data will still yield anisotropic background sets. Furthermore, the apparent reflection of distributions about the galactic longitudes seen in any of the three distribution types (see subsection 6.5.1 on the following page), supports the separation of Pierre Auger Observation data into northern and southern hemispheres. When the three types of distributions are assembled, they are plotted in the order of the southern galactic hemisphere distribution first, and the northern galactic hemisphere distribution second.

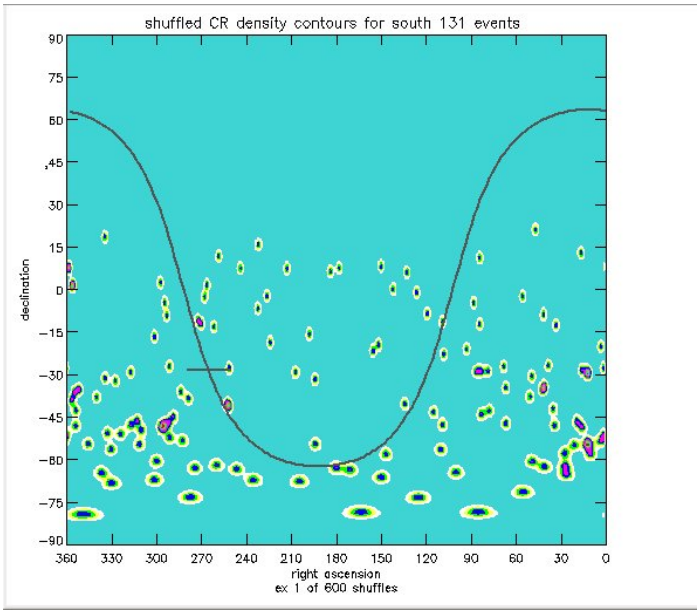

Figure 6.11: Example of southern galactic shuffled density contours of **RA** vs **dec** for 131 events in the energy range 30 EeV < E ≤ 40 EeV. As is evident, not all the reconstructed **RA** vs **dec** values remain in the southern galactic hemisphere (current to 8/11/2012).

No prior incorporation of GMF models or the direct implanting of source directions in signal data to optimize the search parameters made for 'no signal' data-sets (see section 4.1 on page 87). We simply wished to test if the Pierre Auger Observatory data results for the distributions were exceptional, and to present hypotheses in the case of positive outcomes. Non-significant outcomes could indicate that this MST branch analysis is not more powerful than conventional analysis techniques for non-specific scenarios.

The generated background data are assumed to possess the same experimental uncertainties as data read from the Pierre Auger Observatory and the background density of events statistically follows the Pierre Auger Observatory's exposure.

6.5.1 Distributions of Simulated Data and Statistics

Three distinct distributions were decided upon. Each distribution is an aspect of a branch's coupled relationship between Θs and event numbers, or the frequency properties of branch event numbers or the total sum of branch nodes for each shuffle.

We had no *a priori* knowledge of these random distributions. The description of each distribution type given in sub-subsection 6.5.1.1, sub-subsection 6.5.1.2 and sub-subsection 6.5.1.3 below, are for galactic **b**Band divisions in the plane with no divisions of galactic longitude in the plane.

6.5.1.1 Type 1 Distribution - random walk branch Θ

The first distributions are of the random walk angle, Θ, of branches and their node number. This particular distribution type exhibits patterns dependent on the Pierre Auger Observatory exposure (event numbers) and the relationship between a branch's **l**Walk, and the **b**Bands which these **l**Walk galactic longitude shifts occupy. An example of a shuffled branch node vs Θ distribution is seen for the lowest energy range in Figure A.7 on page 239. These **Type 1** distributions are for galactic **b**Band divisions in the plane with no galactic **l**Band divisions in the plane, excepting of course our galactic **(l,b)** divisions of the plane into two separate 'southern' and 'northern' hemispheres. **Type 1** distributions for the three energy ranges are to be found in Appendix A, Figure A.7 on page 239 (low). Appendix B, Figure B.7 on page 257 (mid). Appendix C, Figure C.7 on page 275 (GZK).

The dependence of the random walk angle Θ with event number per **b**Band, is skewed in favour of small angles with small node numbers at the poles, smoothing out to wider, more constant distributions at the low galactic latitudes. The northern hemisphere distributions appear to be reflections about the galactic disk of their southern counterparts but with fewer data. The Pierre Auger Observatory exposure is greatest in the galactic latitude range $\mathbf{l} \in [-45 : -15]^\circ$ and we can also see in the southern galaxy the same pattern distributions emerging, reflected relative to the galactic plane.

Remembering, that a branch's Θ, is defined by

$$\Theta = |\arctan(\mathbf{b}\text{Walk},\mathbf{l}\text{Walk})|$$

and that MSTs are undirected (see section 5.1) we see that the minimum possible value of $\Theta \rightarrow 0^\circ$, and the maximum possible value $\rightarrow 90^\circ$. The further apart the **l**Walk longitude shifts are within their respective **b**Bands, the closer Θ will be to 0°. The maximum angle a galactic longitude shift can subtend is 180°, and the minimum angle a galactic longitude shift can assume is 0°. The

maximum angle a **b**Walk can straddle depends on their latitude cuts which yield $45°, 30°$ and $15°$. As long as the **b**Walk branch distance is much larger than the **l**Walk branch distance, we will have Θ's $\rightarrow 90°$ (the longitudinal, **l** divisions in the galactic (**l,b**) plane provide untrammeled opportunities for branch Θ's $\rightarrow 90°$) . Equally, providing the **l**Walk branch distance is much larger than the **b**Walk distance, we will have branch Θs $\rightarrow 0°$.

At the galactic poles, with their contracted "linear" separation between longitudes, large longitude shifts are more common than, say, at the galactic disk. Coupled with the Pierre Auger Observatory exposure, it is easier at the **b**Band poles to have branches with many nodes stretching over large galactic longitudes, than with **b**Bands closer to, and surrounding the galactic disk. In regions closer to the galactic disk, we see it is still unusual to have **extreme** branch Θ's, but the branch Θ's do tend to be more evenly distributed and their node numbers are lower. This is because of the imposed geometry of the galaxy's shape which is incorporated into the start and finish of the **l**Walk numerator being multiplied by the cosine of its **b**Walk start and finish denominator (see section 6.2.1 on page 155). The **l**Walks close to the galactic disk span larger "linear distances".

6.5.1.2 Type 2 Distribution - event numbers per branch

We have frequency distributions of the number of events per branch. An example is seen in Figure 6.13 on page 177.

As is evident in the Appendices A, B and C , the maximum frequency for the number of nodes per branch is at the node number either two or three. The first node value, 2, is the minimum number of nodes to form a branch and the second node value of 3, is the minimum number of nodes + 1 to form a branch. These randomized distributions demonstrate that branch composition is for predominantly low numbers of nodes. Each distribution's mean is close to its most probable maximum frequency value, the mode, especially for the two most heavily populated **b**Bands, $[-15 : 0]°$ and $(0 : 15]°$. This convergence indicates a tendency towards a central distribution with a pronounced +ve skew provided by the many large node number, but low frequency outliers. In the distributions of the south/north galactic poles there was a wider spread of minimum freqency shuffled nodes numbers.

6.5.1.3 Type 3 Distribution - sum of branch nodes per shuffle

We have distributions of the sum of branch events per shuffle in each **b**Band. An example is seen in Figure 6.14 on page 178.

In terms of the IDL code the bimodel distribution is so marked because each **b/l**Band division is considered, in order, in a loop *outside* the inner loop of 1,200 shuffles. As is evident, when the node count of all branches in a shuffle occupying a **b**Band is totalled, the distribution across the south and north **b**Bands are bi-modal, but only the sense of the relative event numbers available for these 'hemispheres'. The south/north distributions are fairly constant across the **b**Bands with respect to event numbers available to these **b**Bands. For the southern distributions, the fairly constant sum of node numbers per shuffle diminishes as we go further north. For the northern distributions the fairly constant sum of node numbers per shuffle diminishes as we go further south. The means and standard deviations of the shuffles, taken across the entire 1200 shuffles, are inappropriate as can be seen in a peak value and mean that are not closely tied. Despite this, we will just note the Pierre Auger Observatory data results have total branch node sums per shuffle in each **b**Band which are consistently higher than their simulation analogues.

Whilst none of the 3 distribution types is a normal Gaussian, the Gaussian standard distribution, σ, is retained as a useful parameter, as is the mean, μ. The distribution's median is used as an alternative parameter when we take a closer look at a distribution's charactistics by way of applied galactic latitude and longitude filters (see section 6.2 on page 154).

Although we only record the means of all the distribution types for unfiltered galactic latitudes and longitudes and the medians of the filtered galactic latitudes and longitudes, we note that *within* each of the three unfiltered distribution types were means and medians that were generally very close, indicating a trending towards central distributions. The simulated data across energy ranges and within latitude/longitude regions are all similar. The second distribution type shows the most common frequency of branch event numbers to be ~ 2, the minimum number for a branch to exist. At the galactic pole **b**Bands, we see, a wider spread of larger branch node numbers compared with the galactic disk regions where the maximum frequency of the branch node numbers approaches the mean of the branch node numbers. In the third distribution type, there are fairly regular dis-

tributions of total branch event numbers per shuffle depending on whether the **b**Bands are in the 'south' or are in the 'north' galactic hemisphere. It is the **Type 1** distribution, with the coupling of the branch Θs and the branch nodes, and the **Type 2** distribution with the isolation of branch node numbers that we focus on. From the distribution types 2 and 3, we see the random behaviour characteristics of a branch's node number and the sum total of events in a shuffle. It is evident that in many instances the Pierre Auger Observatory's branch node numbers are higher than the simulated branch node numbers, especially for the second distribution. This could indicate a clustering of events denoting a common source.

It is in the first distribution type where we see the connection between branch node numbers and branch Θs. Here the limits of what we term **extreme** branch Θs are set by *twinned* initial conditions (see 7.1 on page 182) and we can nominate, according to which **b**Band we look at, what are typical branch numbers for branch Θs. The second, more independent distribution type is where we set the limit of the sole initial condition of what we term **extreme** branch node sums. All the distribution types are included in their respective energy ranges in Appendix A (low), Appendix B (mid) and Appendix C (GZK).

Each of the 3 types of distributions in the **b**Band divisions in the plane have been subjected to longitude cuts. Two of the longitude cuts are between, in the first case or filter, $\pm[-135 : -45]^\circ$, and then in the second filter ($[-135 : 135]^\circ$ and $[-45 : 45]^\circ$). The reason behind these longitude partitions is that the measured regular galactic magnetic azimuthal field is stronger in the inter-spiral arm spaces and we wish in a general way to confirm the effect of this. The galaxy's azimuthal magnetic field may be at right angles in the cuts $[-135 : 135]^\circ$ **and** $[-45 : 45]^\circ$. Magnetic fields are expected to be more chaotic in the longitude divisions of $\pm[-135 : -45]^\circ$ [29].

6.5.1.4 No type 3 Distribution

In our following discussion of distribution types and and latitude filters we concentrate on the first and second distribution types. As has already been noted, the third distribution type does not yield much useful information, beyond, in some instances, a general confirmation that the total number of branch nodes per shuffle per **b**Band for the Pierre Auger Observatory results are greater than

their simulation counterparts.

6.5.1.5 Galactic bBands with No Filters

- see Appendices, Figure A.7 on page 239, Figure B.7 on page 257 and Figure C.7 on page 275.

The **Type 1** distribution has a more pronounced peak, in increasing energy range, between $34° - 38°$(low), to $31° - 36°$(mid), to $32° - 36°$(GZK) in the two galactic polar **b**Band divisions of $[-90 : -45]°$ and $[45 : 90]°$, and a wider, more heavily populated branch Θ value spread in the intervening **b**Band divisions. The **Type 2** distribution has a pronounced skewed central frequency value of between $2 - 3$ nodes over all the **b**Band divisions in the plane.

6.5.1.6 Galactic bBand Filter $[-135 : -45]°$ and $[45 : 135]°$

- see Appendices, Figure A.10 on page 242, Figure B.10 on page 260 and Figure C.10 on page 278.

For the **Type 1** set of distributions of this galactic longitude filter, both the south and north distributions possess the same relative concentration of **l**Band branches in the **l**Band $[-135 : 135]°$ and also in the **l**Band $[-45 : 45]°$. Distribution branches from **l**Band $[-135 : 135]°$ provide the small valued **extreme** branch Θs, some with large node sums,$\sim$(40-60), in the lower southern **b**Bands. For the two disk **b**Bands, the small branch Θs have all but disappeared. The branch contribution from either **l**Band cut is roughly the same and the branch Θs have spread to moderate branch node numbers, $\sim$(15-20), and non-extreme Θs, although we note a uniform distribution of small branch node numbered **extreme** branch Θs between $(69 - 90)°$.

6.5.1.7 Galactic bBand Filter $[-135 : 135]°$ and $[-45 : 45]°$

- see Appendices, Figure A.13 on page 245, Figure B.13 on page 263 and Figure C.13 on page 281.

In the **Type 1** set of distributions for the galactic longitude filter cuts of $[45 : 135]°$ and $[-135 : -45]°$, there is a minor contribution of branches from the **l**Band $[45 : 135]°$, which decreases as the **b**Bands move further north and disappears at the **b**Band $[45 : 90]°$. The galactic longitude cuts of $[-135 : -45]°$ have non-extreme branch Θs with modest node numbers, $\sim$(15-30).

6.5.1.8 Galactic bBand Filter $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$

- see Appendices, Figure A.16 on page 248, Figure B.16 on page 266 and Figure C.16 on page 284.

Another longitude cut is to identify branches with Θ **l**Walks that subtend at least 90° longitude in the co-ordinate system we use. i.e. galactic longitudes from 180° to 0° left of galactic latitude, **b**= 0°, and them from 0° to −180° to the right of galactic latitude, **b**= 0°. These are the branches with the largest and smallest extent, with branch longitude differences > 90° and ≤ 180°. **Extreme** Θs with node numbers greater than or equal to 0.1% in their simulated distribution comparison percentages, $\%_S$, would be very interesting. An **extreme** branch Θ with a large comparative node number in a small (**l,b**) region could indicate a source - whether primary, as in a point-source, or secondary, as in a coherent GMF acting on CRs. An **extreme** branch Θ with a large comparative node number in a large (**l,b**) area would probably indicate a secondary source if not a large-scale source such as the supergalactic plane. This filter's branch Θ distributions has very few branches across the galactic disk regions, $[-15 : 0]°$ and $(0 : 15]°$. In the **b**Bands contiguous to the galactic polar regions, and the polar regions themselves, there are small **extreme** branch Θs with large node sums,~(40 - 60) nodes. The bulk of small **extreme** branch Θs with large node sums over unfiltered galactic longitudes have been contributed by branches that subtend 90°. There are virtually no large **extreme** branch Θs with nodes of any number. The distributions are reflected about galactic latitide **b**=0. Pierre Auger Observatory branch Θs appear in the two lowest **b**Bands $[-90 : -45]°$, $[-45 : -15]°$ and in the highest **b**Band $[45 : 90]°$.

The **(RA,dec)** position and energy of measured CR's has associated experimental errors (errors in angular resolution). In Pierre Auger Observatory data, for energies > 10 EeV, there is an uncertainty of ~ 12% [194] and 68% of the position co-ordinates lie within 1.4° of the true source direction. All simulated data uncertainties are as the Pierre Auger Observatory's. The Pierre Auger Observatory data formally analysed is current to 8/11/2012.

We test whether the Pierre Auger Observatory values of branch Θs, branch node sums or the number of events in a division of the galactic **(l,b)** plane will be greater than or equal to two standard deviations (**extreme**) of our now two types of simulated distributions (we no longer consider the third, see sub-subsection 6.5.1.4 on page 172). We can then look to the known physics

of the galactic **(l,b)** sky for possible explanations. For example, do the Pierre Auger Observatory branch Θ values in the spiral and inter-spiral arms support the presence of regular or chaotic GMFs?

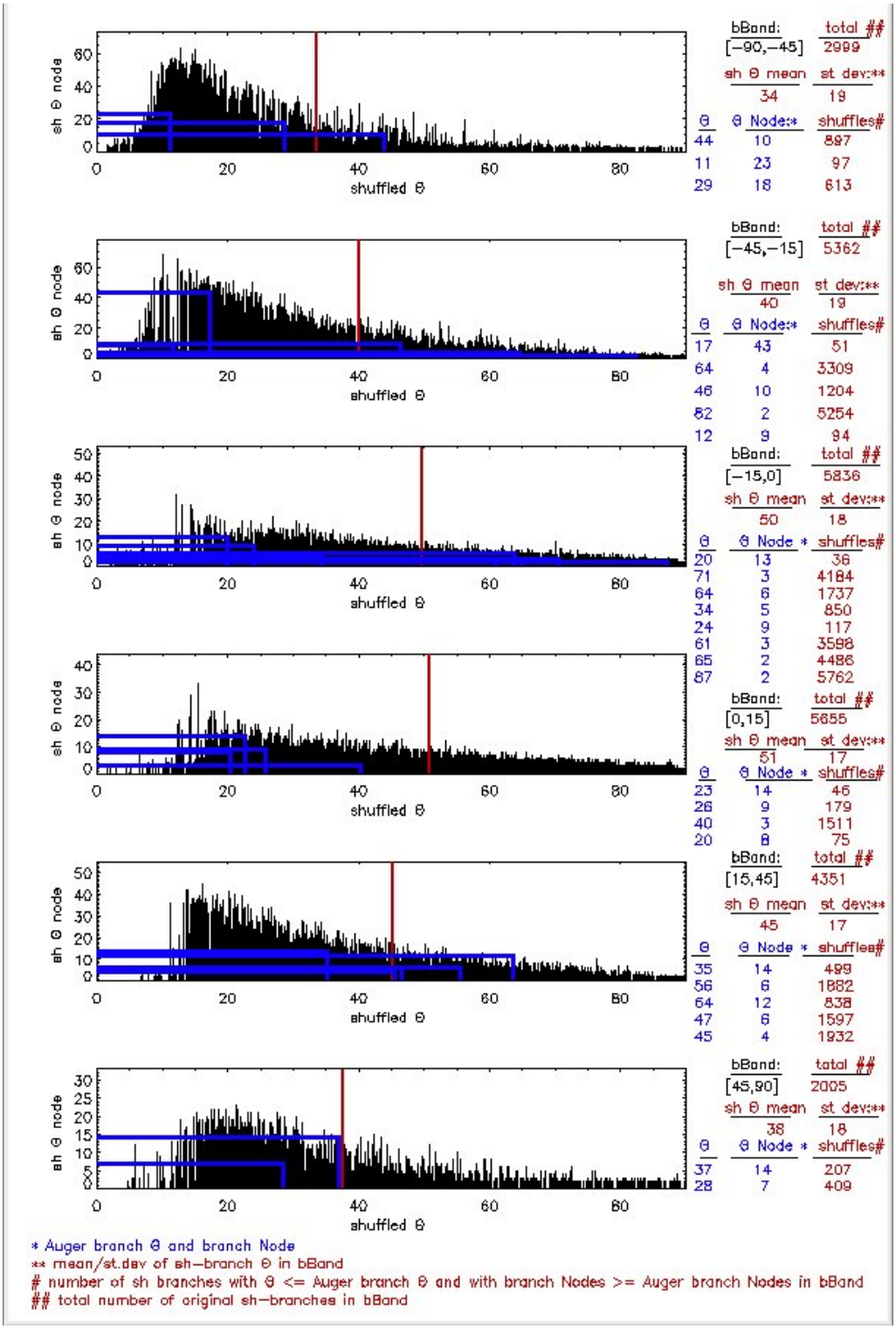

Figure 6.12: **Type 1** distribution of 1200 sets of shuffled branch Θ's (see page 155) numbers vs Θ branch numbers for 40 EeV < E ≤ 50 EeV. The lines picked out in red are the means of the shuffled distributions. The lines picked out in blue are the branch Θ vs branch Θs event number for Pierre Auger Observatory data (current to 8/11/2012).

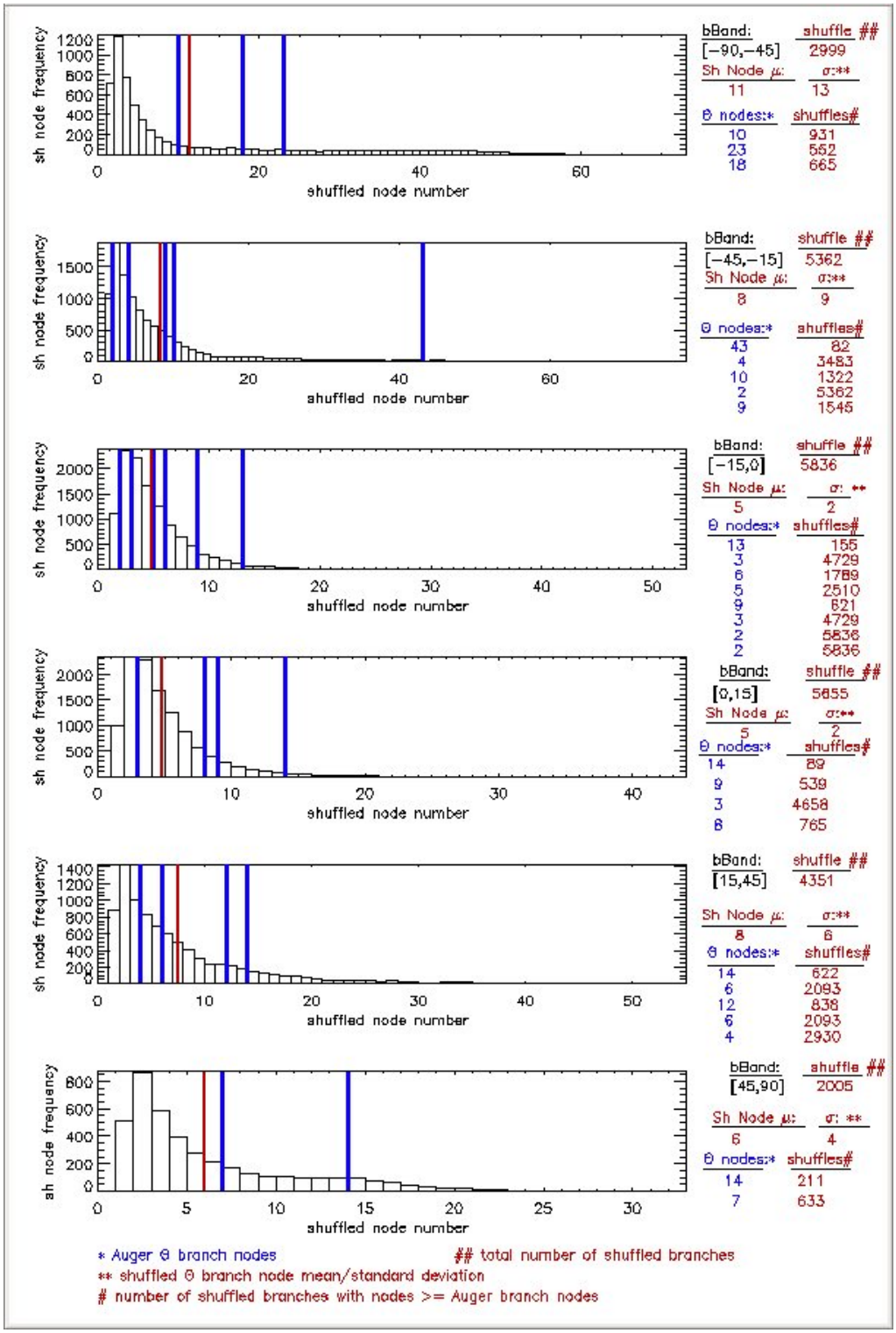

Figure 6.13: **Type 2** distribution of 1200 sets of event numbers vs branch event number frequency for 40 EeV < E ≤ 50 EeV. The lines picked out in red are the means of the shuffled distributions. The lines picked out in blue are the values of Pierre Auger Observatory branch number data (current to 8/11/2012).

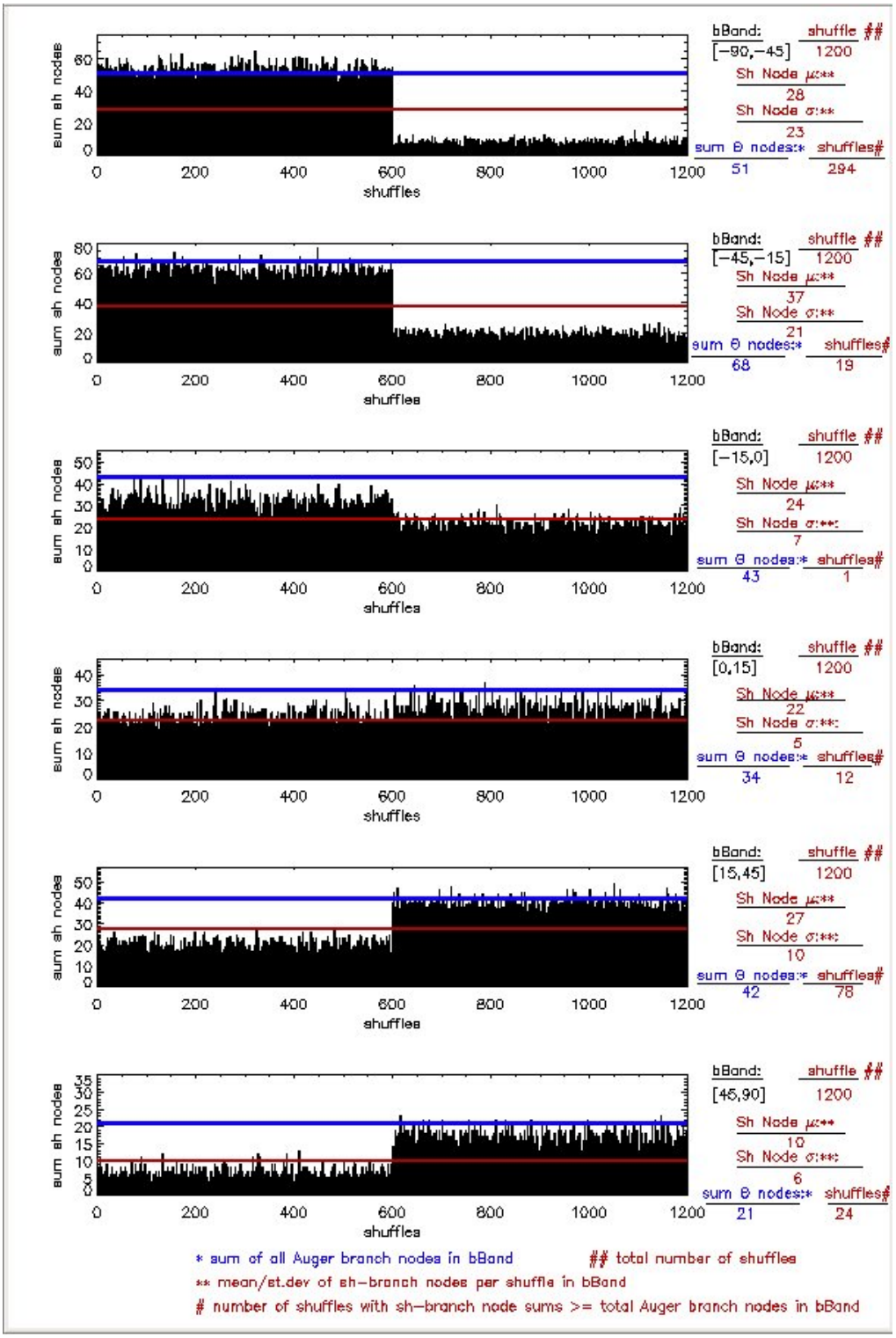

Figure 6.14: **Type 3** shuffled distribution of 1200 sets of shuffles vs shuffled branch sums for 40 EeV $< E \leq$ 50 EeV. The lines picked out in red are the mean of the shuffled distribution. The lines picked out in blue are the total sum of branch events in each **b**band for Pierre Auger Observatory data. The division between contributions of event number sums in the south and north hemispheres **b**bands is reflected about **b** = 0, i.e. event contributions for the first 1200 hundred shuffles are predominantly 'southern' and event contributions for the second 1200 hundred shuffles are predominantly 'northern'. This is because not all of the shuffled co-ordinates recovered from each galactic hemisphere's event directions remain in that hemisphere, there are always shuffled co-ordinate outliers in the opposite galactic hemisphere. That is, when the shuffled co-ordinates are reconstructed from event directions in, for example, the southern hemisphere, there will be a fraction of simulated co-ordinate directions that will lie in the northern hemisphere because of the ± variable azimuth values which are deconstructed from southern event directions yielding both ± values for the simulated declination/galactic latitude co-ordinate values. Such a shuffled example is seen in Figure 6.11 on page 168 (current to 8/11/2012).

Chapter 7

Tabulated Results

The results tabulated in this chapter are from Pierre Auger Observatory UHECR data directions which have been minimally connected to grow 'Yggs' in three energy ranges. 40 EeV < E ≤ 50 EeV (low), 50 EeV < E ≤ 60 EeV (mid) and E > 60 EeV (GZK). A desire to isolate Ygg features which might be considered unique or interesting in the context of 1,200 shuffled simulated randomomized UHECR data directions was followed through with three "distribution types" of the randomized directions, in each of the energy ranges. The physics of the azimuthal and poloidal GMFs led to divisions in the galactic $(\mathbf{l}, \mathbf{b})$ plane where Ygg "branches" might manifest angular values, Θ, consistent with nuclei being deflected in these GMFs or EGMFs. These branch Θ values versus the number of events composing each branch Θ were compared with the randomized "**Type 1** distributions". The number of events composing each branch might also have meaning, independent of that branch's alignment (Θ values). The **Type 2** and **Type 3** distributions were designed to highlight any lack of correlation between the simulated data branch event numbers and the Pierre Auger Observatory branch event numbers. The **Type 2** frequency distributions tackled the question of whether there was a typical maximum frequency of a simulated branch event number, and the **Type 3** distributions were to gauge if the total number of branch events in each shuffle had any consistent characteristics. It soon became obvious that the **Type 3** distributions would yield little useful information because the distributions were initially configured over the entire 1,200 shuffles for each of the **b**Bands of the entire galactic **(l,b)** sky, and the galactic **(l,b)**

sky had subsequently been separated into two hemispheres. The reasons for the **Type 3** distinct bi-modal distributions is discussed in sub-subsection 6.5.1.3 on page 171. Independent of distribution statistics, the **l** and **b** branch-walk-pair values of all Pierre Auger Observatory branches, and the areas in the galactic **(l,b)** sky they encompass, were scrutinized for similarities *across* and *within* energy ranges. An example of the 'area' or 'region' the galactic branch-walk-pair values cover was shown in Figure 6.8 on page 160.

This chapter is organized as follows - there are 16 tables of results listed. The tables are divided into 5 sections. Each section contains three preliminary tables of individual energy ranges considered to be just below, or just above the GZK limit. Any extra-galactic CR sources below the GZK energy limit currently do not have strong limits placed on their distances. The GZK energy limit is so interesting because it does. The energy range, E > 60 EeV (GZK), is so extreme that the GZK limit demands extra-galactic events, but within a region of 200 Mpc.

In the tabulated data, each Pierre Auger Observatory branch Θ and its simulated distribution comparison percentage, $\%_S$, each branch node sum, plus what we term the system of co-ordinates, 'branch-walk-pairs', are recorded. The branch-walk-pairs are a Pierre Auger Observatory branch's **l**Walk maxima and minima in the galactic longitude co-ordinates ($\mathbf{l_{max}}, \mathbf{l_{min}}$), and that branch's **b**Walk maxima and minima in the galactic latitude co-ordinates ($\mathbf{b_{max}}, \mathbf{b_{min}}$). The summarized tabled results of Table 7.6 on page 199 are also presented in the context of galactic filter regions, i.e. galactic **l** and/or **b** divisions in the galactic $(\mathbf{l}, \mathbf{b})$ plane in Tables 7.10, 7.14 and 7.18. The Pierre Auger Observatory branch Θs, and the branch node sums are for those parts of the Ygg MST that lie within these divisions. Any branch nodes included outside the divisions are those single nodes that complete branches of the entire Ygg. Individual energy range tables, such as Table 7.3 on page 193, flag possible **extreme** branch Θs of interest (*), branches with atypical node numbers (!), which we term **extreme** branch node sums[1] and any Pierre Auger Observatory branches with a galactic **l** or **b** branch-walk-pair/ROIs co-ordinate within $\pm 10^\circ$ of either of the galactic 'axes', **l=0** or **b=0** ($\odot$).

A fourth table, such as Table 7.6 on page 199, the final in each section, is an energy comparison table of the three energy range results, and provides a selection of Pierre Auger Observatory

[1] * ! **extreme** defined in section 7.1 on page 182

branches believed to be interesting, not only in the context of individual **extreme** branch Θs, individual **extreme** branch node sums, and a branch-walk-pair co-ordinate within ±10° of either of the galactic 'axes' (a provision anticipating possible effects of poloidal/azimuthal GMFs) but in the galactic **l** and **b** branch-walk-pair values or co-ordinate position 'areas' that the branches span. The introduction of possibilities afforded by considerations of galactic **l** and **b** branch-walk-pair values and their areas is indicated by '?' in the preliminary tables. In the energy comparison tables we note galactic **l** and **b** branch-walk-pair/area relationships *between* branches *across* energy ranges, or branch-walk-pair/area relationships *between* branches *within* the same energy range. Such correlations are marked by a suite of tags[2].

Each energy comparison table is augmented with text citing some reasons for the branch selections. These energy comparison tables are found in Table 7.6 on page 199, Table 7.10 on page 207, Table 7.14 on page 213 and Table 7.18 on page 221. The energy comparison tables have a more complicated system of tags because of several correlation possibilites raised by the galactic **l** and **b** branch-walk-pair/area positions.

Standard divisions chosen in galactic latitude, **b**, are from $[-90 : -45]°$, $[-45 : -15]°$, $[-15 : 0]°$, $(0 : 15]°$, $[15 : 45]°$ and $[45 : 90]°$, and are in place in the tables from Table 7.3 on page 193 to Table 7.18 on page 221. The only galactic latitude, **b**, divisions that have been deliberately selected are the **b**Band divisions across the galactic plane, $[-15 : 0]°$ and $(0 : 15]°$, which is *a priori* to the Pierre Auger Observatory prescription, although, originally, the disk was considered as a single large-scale object between $[-15 : 15]°$ [57]. I divided the galactic plane in this way because the two recently found large radio lobes [26], may be separately responsible for at least some of the accelerations of CRs. In Table 7.3 on page 193 to Table 7.18 on page 221, the longitudinal, **l**, co-ordinate of the galactic plane, was divided into regions or 'filters' where we might see evidence of coherent or non-coherent galactic magnetic fields. The azimuthal GMFs are considered to possess their strongest, most regular, structures in the regions between the galactic spiral arms. The spiral arm GMFs themselves are believed to be more chaotic.

[2]Pierre Auger Observatory branches which are interesting in terms of their **l** and/or **b** branch-walk-pairs, or area positions have several distinguishing tags including ♣, ♠, ♡, ♢.

7.1 The Setting of Extreme Branch Variables

In order to provide useful information, the simulated distribution comparison percentages, $\%_S$, need initial or boundary conditions. What are we comparing, and what limits do we set? We are comparing simulated branch data distributions against Pierre Auger Observatory branch data. Pierre Auger Observatory branch data comprises the end angle, Θ - that a branch's 'random walk' galactic latitude, **b**, and 'random walk' galactic longitude, **l**, make with one another, such that $\Theta = |\arctan(\ \mathbf{b}\text{Walk}, \mathbf{l}\text{Walk}\ )|$ (see subsection 6.2.1 on page 155) - and how many events compose the branch. The area covered by these **l**Walk /**b**Walk co-ordinate arrangements, or branch Θs, are termed 'regions of interest' (ROIs).

The comparison percentage records the percentage, $\%_S$, of the "**Type 1** distributions" simulation branch Theta's, Θ_S, such that $\%_S$ is the percentage of randomized events with $\Theta_S \leq$ the Pierre Auger Observatory branch Θ value. When $\%_S \leq 5\%$ of the Pierre Auger Observatory branch Θ *and* the branch Θ **magnitude** itself is within set boundaries, or when $\%_S \geq 95\%$, *and* the branch Θ **magnitude** itself is within set boundaries, the branch Θ result is termed **extreme** and tagged with *, i.e. **extreme** branch Θs involve twinned initial conditions. An **extreme** branch node number recorded with **!**, indicates a comparison percentage of the "**Type 2** distributions" branch's node sum which is $\leq 5\%$ of the Pierre Auger Observatory branch node sum. The distribution's **Type 2** actual comparison percentage is not recorded, but calculated separately to avoid confusion, and we considered the tag **!** sufficient.

One possible interest is branch Θs (preferably with large node numbers), that tend to the limits of zero degrees or ninety degrees, with respect to galactic $\mathbf{b} = 0^\circ$ and galactic **l** range $\in [0 : 360]^\circ$. This seems to be reasonable because the galactic magnetic fields tentatively 'identified' to the present time, seem to be poloidal [23] or azimuthal.

In the case of small branch $\Theta \rightarrow 0^\circ$, and branch Θs between $(22 - 68)^\circ$, final comparison percentages are the percentage of simulated branches that possess branch Θs less than or equal to the Pierre Auger Observatory branch Θ, while being composed of greater than, or equal numbers of the Pierre Auger Observatory branch node sums. For the case of large branch $\Theta \rightarrow 90^\circ$, final

comparison percentages are the percentage of simulated branches that possess branch Θs greater than, or equal to, the Pierre Auger Observatory branch Θ, while being composed of greater than, or equal numbers of the particular Pierre Auger Observatory branch Θ node sums.

Table 7.1 summarizes the conditions for **extreme** Pierre Auger Observatory branch variables.

Table 7.1: Conditions For Extreme Branch Variables

Branch Variable	condition	$\%_S$
small $\Theta \mapsto 0^\circ$	$\Theta \leq 21^\circ$	$\%_S \leq 5\%$
big $\Theta \mapsto 90^\circ$	$\Theta \geq 69^\circ$	$\%_S \geq 95\%$
branch node sum	distribution branch node sum	$\%_S \leq 5\%$

The caveat of **extreme** branch Θ magnitude has been set at between $(0 - 21)^\circ$ and $(69 - 90)^\circ$, because most of the Pierre Auger Observatory smaller branch Θs with low simulated data comparison percentages were below 21°. The magnitude limits of Pierre Auger Observatory large **extreme** branch Θs provide symmetry over the $(0 - 90)^\circ$ limits of possible branch Θs. For the large **extreme** branch Θs between $(69 - 90)^\circ$, the simulated data comparison percentage statistic, $\%_S$, must be subtracted from 100% to find the final comparison percentages for where the number of shuffled data branch Θs are greater than, or equal to the Pierre Auger Observatory branch Θs, together with a branch node number that is greater than, or equal to the number of Pierre Auger Observatory branch nodes[3].

It is not unusual for Pierre Auger Observatory non-extreme branch Θs, between $(22 - 68)^\circ$, to have branch Θ **Type 1** simulated distributions of $\Theta_S \leq 5\%$, whilst having their branch node number greater than, or equal to the the simulated branch node number. *However*, the twinned **Type 1** distributions initial conditions do require branch Θs to be **extreme**. We chose to regard such cases as possibly interesting only if there were 0.1% (the Pierre Auger Observatory prescription positive result), or less, of the entire **Type 1** distributions simulated branch Θ node sums having greater than, or equal Pierre Auger Observatory branch node sums, i.e. we would then disregard the distributions **Type 1 extreme** Θ twinned initial condition in the individual comparison tables, and tag

[3] All comparison percentages, $\%_S$, are tabled as the number of shuffled data branch Θs that are less than or equal to the Pierre Auger Observatory branch Θs, together with a branch node number that is greater than or equal to the number of Auger Observatory branch nodes, so at the large **extreme** branch $\Theta \sim (69 - 90)^\circ$, the final comparison percentage is $100\% - \%_S$.

the Pierre Auger Observatory branch Θ as also possibly interesting with a ∗. Let it be noted, that although a couple came close, *none* of the **Type 1** distributions of the non-extreme Pierre Auger Observatory branch Θ's proved to have 0.1% or less of their entire simulated branchas node sums greater than, or equal to, the simulations branch node numbers.

The less dependent distributions, **Type 2**, sole initial **extreme** condition, where the simulated data branch node numbers are ≤ 5% of the Pierre Auger Observatory branch node numbers and tagged with '**!**', may indicate clustering effects in the Pierre Auger Observatory branch node number, independent of its branch Θ value.

7.2 Allowed Galactic l and b Branch-Walk-Pair Similarities and Branch Θ Regions of Interest

The tag, **?**, next to the galactic **l**Walk and **b**Walk columns in the individual energy tables, is to signal that Pierre Auger Observatory branches are to be considered candidates for the final energy comparison tables, if they can be seen to demonstrate galactic **l** *and* **b** branch-walk-pairs, or branch Θ ROIs which are similar in the context of known physics and measurement errors. The GMFs are correlated with the shape of the galaxy. Thus we search for similar branch **l**Walk *and* **b**Walk pair values *across* energy ranges, similar branch **l**Walk *and* **b**Walk pair values *within* energy ranges or overlapping (similar) Θ ROIs *across* energy ranges and overlapping Θ ROIs *within* energy ranges. Because the GMFs are poorly understood and unmapped, we allow a certain latitude when we consider the term 'similar'. As a general rule-of-thumb, when we mention ROIs, the area covered by **(l,b)** branch-walk-pairs are also ROIs, but ROIs do not necessarily satisfy the 'similarity' constraints required of **(l,b)** branch-walk-pairs (see Figure 6.8 on page 160).

When comparing one **l** / **b** branch-walk-pair with another **l** / **b** branch-walk-pair *across* energy ranges - if the branch-walk-pair minimum of one branch is ±25° of the branch-walk-pair minimum of the other branch, the branch-walk minima are considered similar. This logic also applies to the branch-walk-pair maxima of the branches being compared. Both the **l** *and* **b** branch-walk-pairs of each branch must satisfy this ±25° condition for their Pierre Auger Observatory branches to be

considered similar and to be included as such in the energy comparison table. If three branches *across* energy ranges are similar, then all three **l** *and* **b** branch-walk-pairs must be within ±25° of one another. These properties may be an indication of a difference between light or heavy nuclei or be connected with the GZK energy limit. A Θ ROI is often easier to visualize, overlapping ROIs are considered as possibly interesting and are flagged by the same tags as similar branch-walk-pair maxima and minima.

This same tag, **?**, indicates a like ±25° search for branches that have similar **l** and **b** walk-pair values *within* the same energy range and reflected *about* the galactic plane **b** = 0, for the galactic **b**Band disk divisions, $([-15 : 0] : (0 : 15])^{\circ}$[4]. We allow both straighforward mirror reflections (M) about galactic **b** = 0 and within the galactic **b**Band divisions of $([-15 : 0], (0 : 15])^{\circ}$ with the galactic longitude range $\mathbf{l} \in ([-180 : 180])^{\circ}$ and diagonal reflections (D) about galactic **b** = 0 and within the galactic **b**Band divisions of $([-15 : 0], (0 : 15])^{\circ}$ with the galactic longitude range $\mathbf{l} \in ([-180 : 180])^{\circ}$. The azimuthal GMF is known to change direction above/below the galactic plane. With regards to branch Θ ROIs, 'similar' branch Θ ROIs can overlap also in a mirror or diagonal reflected manner.

The Pierre Auger Observatory branches of interest are summarized in the energy comparison tables, Table 7.6 on page 199, Table 7.10 on page 207, Table 7.14 on page 213 and Table 7.18 on page 221.

We noticed, occasionally, branches with only small node numbers and **extreme** branch Θs, had a similar branch-walk-pairs/ROIs with the branch-walk-pairs/ROIs of **extreme** branch Θs with large node numbers *across* some other energy ranges. Any similar **extreme** branch Θ branch-walk-pairs/ROIs, regardless of their branch node sums, are included in the comparative energy tables with ♣, ♠.

Any **extreme** or non-extreme branches that have apparently *reflected* **l** and **b** branch-walk-pairs/ROIs about the galactic disk **b**Band regions, $([-15 : 0], (0 : 15])^{\circ}$, *within* an energy range are tagged with $\heartsuit 1, 2..., \diamond 1, 2$ .... These reflection tags have an extra M or D next to them to denote mirror or diagonal reflections.

[4]remember, if the end node of an Ygg branch is outside any galactic **b**Band divisions, that node outside the galactic disk division is included as part of the branch within the said galactic **b**Band division

Any non-extreme branch Θs or branch node sums that have branch-walk-pairs/ROIs similar to the branch-walk-pairs/ROIs of **extreme** branch Θs *across* energy ranges, are to be tagged by † in these tables. Random-seeming branches in one energy range, correlating with **extreme** branch Θs in another, may be a sign of changing CR energy dependent larmor radii for the same type of CR. We do, also, contend that if Pierre Auger Observatory non-extreme branches have similar branch-walk-pairs/ROIs *across* energy ranges, that these branches should be retained as interesting and we tag them as $\surd$ in the energy comparison tables. Since these branches exist, areas having similar galactic branch-walk-pairs/ROIs *across* energy ranges may indicate regional relationships. For example, in certain of the galactic longitude regions, similar non-extreme branches could indicate non-coherent GMFs. These possibilities cannot be supported by our statistics. We are unable to take branch-walk-pair/ROIs statistics, because while the simulated distributions calculated branch Θs and branch node sums were considered branch variables, no output was generated for the simulated **l/b**Walk-pairs used to calculate these variables. The possible meanings of such similarities are at their most obvious when the Pierre Auger Observatory data are separated and sorted within the various divisions or 'filters' of the galactic $(\mathbf{l}, \mathbf{b})$ plane.

7.2.1 Individual Energy Range Tables

To recap, each individual energy range comparison table has four tags. * - to indicate an **extreme** branch Θ, **?** - to allow **l/ b** branch-walk-pair values and Θ ROIs of Pierre Auger branches to be scanned for similar branch-walk-pair values and ROIs. **!** indicates an **extreme** branch node sum and ⊙ flags an **l** or **b** branch-walk-pair co-ordinate within $\pm 10^{\circ}$ of the galactic axes **l=0** or **b=0**.

7.2.2 Final Energy Comparison Tables

The final energy comparison tables have several tags to acknowledge certain relationships between Pierre Auger Observatory branches that may prove interesting and/or to flag positive results. The tags are summarized in Table 7.2 on the next page and explained in more detail below.

All **extreme** branch Θs remain tagged * and have tags, ♣, ♠ - to flag **extreme** branch Θs with similar branch walk-pairs/ROIs *across* energy ranges.

Table 7.2: Tags and Their Conditions

Tag	Condition
**	data current until 8/11/2012.
NA	**Extreme** Θ comparison percentage, $\%_S$, not applicable.
*	Pierre Auger Observatory **extreme** branch Θ of interest.
!	Pierre Auger Observatory **extreme** branch node number of interest.
♣, ♠	Pierre Auger Observatory **extreme** branch Θs with similar branch-walk-pair values/ROIs *across* energy ranges.
♡, ♢	Pierre Auger Observatory similar mirror (M) or diagonal (D) reflected branch-walk-pair values/ROIs in the same energy range, and within the galactic disk **b**Band divisions of $([-15:0],(0:15])^\circ$.
†	**Extreme** branch with similar branch-walk-pairs/ROIs to a non-extreme branch, *across* energy ranges.
√	Non-extreme branches *across* energy ranges with similar branch-walk-pairs/overlapping ROIs.
⊙	Pierre Auger Observatory branch with a **l** or **b** branch-walk-pair co-ordinate within $\pm 10^\circ$ of the galactic axis **l=0** or galactic axis **b=0**.
★	Pierre Auger Observatory **extreme** branch Θ with a positive result.

All **extreme** branch node sums remain tagged **!**

Branches that have an **l** or **b** branch-walk-pair co-ordinate within $\pm 10^\circ$ of the galactic axis **l=0** or galactic axis **b=0** are marked by ⊙. This only happens for one set of two similar branches of the Pierre Auger Observatory results with **b**Band divisions in the plane, for the galactic longitude filters, $[-135:135]^\circ$ and $[-45:45]^\circ$ in Table 7.10 on page 207 (and, naturally, in the first energy comparison table, Table 7.6 on page 199, in the unfiltered galactic **l**Bands). This pair of similar branches are in the low and mid-energy ranges and have branch Θ values that incline more to the vertical of the galactic **(l,b)** plane. These branches have $\Theta = 64^\circ$ with **l/b**Walk co-ordinates of $(-8,-2)^\circ, (-22,-2)^\circ$ (low), and $\Theta = 70^\circ$ with **l/b**Walk co-ordinates of $(0,4)^\circ, (-19,-6)^\circ$ (mid). We note these branches lie in the inter-spiral filter and the disk **b**Band $[-15:0]^\circ$ where we expect azimuthal GMFs.

The tag, †, indicates a similarity between the **l/b** branch-walk-pairs/ROIs of a non-extreme branch Θ and an **extreme** branch Θ *across* energy ranges.

Similar non-extreme and/or **extreme** branches that might be somehow linked - *within* the *same* energy range are marked with ♡, ♢ - if there is a reflected **l** *and* **b** Mirror (M) or diagonal (D) branch-walk-pair/ROI relationship about the galactic disk **b**Bands of $([-15:0],(0:15])^\circ$.

The tag, $\surd$ - indicates a similarity between galactic **l/b** branch-walk-pair/ROIs values of non-extreme branches *across* energy ranges.

Finally, ★ - picks out those Pierre Auger Observatory **extreme** branch Θs with a positive comparison percentage of $\leq 0.1\%$. The presence of "NA" in a comparison percentage box is to show that the **extreme** branch Θs twinned initial conditions are not met, and the comparison percentage is to be ignored as a factor in the retention of that branch in the table.

All individual energy tables are unaccompanied by text. Footnotes explain the meaning of tags. Each final energy table, however, has some text to explain the reasons for some of the interesting branch decisions.

The chapter is concluded by a discussion on any overall interesting or positive results, accompanied by galactic longitude and latitude, (**l**, **b**), plots of the Pierre Auger Observatory branch Θs or ROIs. There is also a galactic longitude and latitude, (**l**, **b**), plot of the Pierre Auger Observatorys CR events current to 8/11/2012, in the energy range 12 EeV $<$ E $\leq$ 15 EeV. We provide this plot to demonstrate the limits of the Pierre Auger Observatory exposure and how it correlates with the Pierre Auger Observatory branch Θs of interest.

A cursory mention of a visual review of Pierre Auger Observatory branches is included in this discussion (see section 7.3). It was clear after implementing branch cuts, that there were a number of branches with some embedded linear patterns that did not make their branch Θs **extreme**.

7.3 How Effective is the Yggdrasil in Capturing Lines?

As has been remarked, a pattern recognition algorithm in this case will only go so far. The branches of an Ygg are predicated on fairly arbitrary divisions in the galactic (**l**, **b**) plane and patterns are easily lost because of these divisions. How does the Ygg perform under visual scrutiny? Are real straight lines able to be discerned? We inserted ten straight lines into the Pierre Auger Observatory data for CR events E > 60 EeV. Half of the events were for straight lines with nodes $\sim 5^\circ$ apart and the other half were for straight lines with nodes $\sim 10^\circ$ apart. Almost all the $\sim 5^\circ$ apart events were captured in sequence by the Ygg algorithm, and all the lines were easily seen. Most of the straight line events $\sim 10^\circ$ apart were captured as a linear phenonomen, but sometimes with three linear

events out of sequence as other nearby Pierre Auger Observatory data were caught because their angular distance from the line events were < 10°. The linear pattern was still discernable visually. Two examples are displayed below. One example is of all six events 10° apart being captured in Figure 7.1. Another has only a couple of events 10° apart being captured (Figure 7.2 on the next page). In the second example, the linear pattern is not picked up by the Ygg MST.

A relationship between the total angular length, L, of an individual Ygg, and a maximum angular distance, Y, between, say, an array of six linear events between 5° – 10° apart, could be calculated and visually checked to gain a more precise appreciation of the maximum possible separation of line events which would be picked up by an Ygg. A statistical optimization process of inserting linear arrays of events (5 – 10)° apart into shuffled MST data would clarify the maximum possible 'linear' event separation sensitivity of Yggs and may be considered for future work.

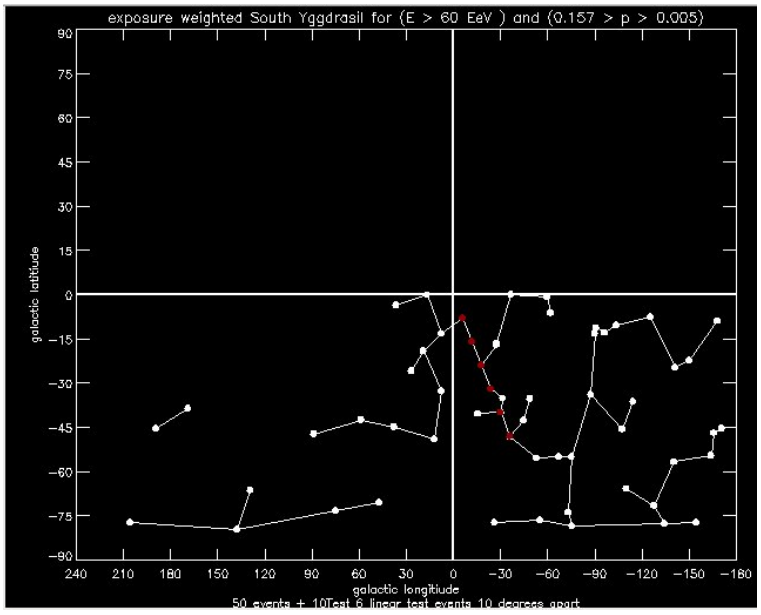

Figure 7.1: Quite good capture of six linear events spaced at 10° intervals apart, indicated in red, inserted into Pierre Auger Observatory data for E > 60 EeV (current to 8/11/2012).

There are many linear structures within the Ygg which are not apparent if one relies only on a branch's **extreme** Θ value. Some instances are displayed in Figure 7.3 and Figure 7.4. All Yggs evince evidence of linear structures. It was more useful, in the first instance, to divide the Ygg into branches, in areas where we have an appreciation of the physics of the region, such as GMFs, and

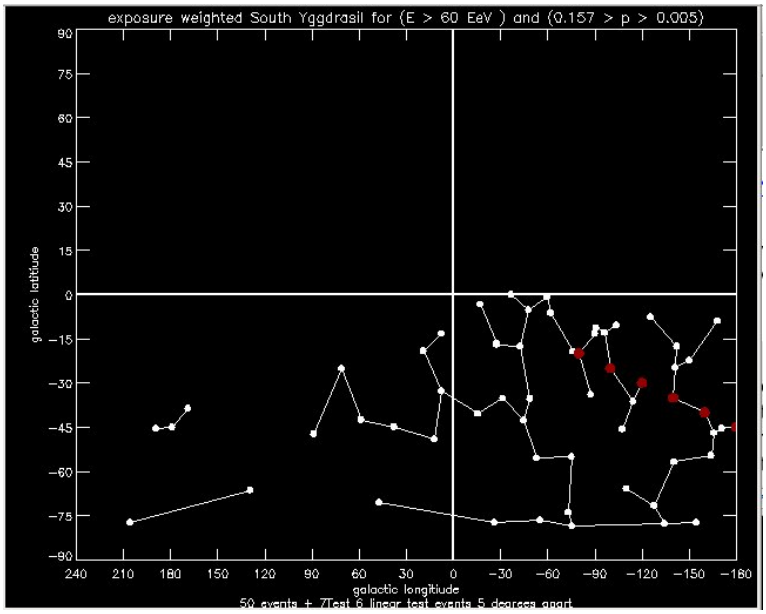

Figure 7.2: Poor capture of six linear events spaced at 10° intervals apart, indicated in red, inserted into Pierre Auger Observatory data for E > 60 EeV (current to 8/11/2012).

then to look for non-random structure over large area's.

We wish to couple Pierre Auger Observatory branches with areas of the galactic **(l,b)** sky where we expect the physics of the region to be somehow singular, e.g. GMF configurations, large-scale objects, even if we don't precisely understand how the physics works. The GC is easily the most prominent feature in our galaxy and we know that 28 neutrino events from the Antarctic IceCube detector with a significance of $\sim 4\sigma$, included an apparent cluster of events near the GC [58]. In a non-statistical manner, as with branch-walk-pairs and branch Θ ROIs, we searched for branches within the ambit of $\pm 30^\circ$ of the GC - $(\mathbf{l_1}, \mathbf{l_2}) \in \pm 30^\circ$ of $\mathbf{l} = \mathbf{0}$ and $(\mathbf{b_1}, \mathbf{b_2}) \in \pm 30^\circ$ of $\mathbf{b} = \mathbf{0}$, and contrived to interpret the branch properties.

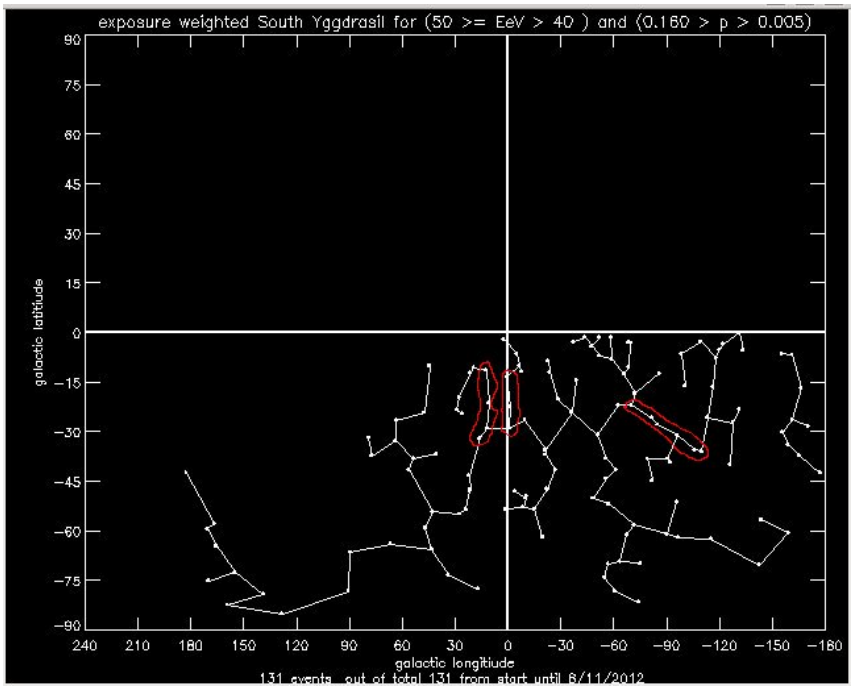

Figure 7.3: Southern Ygg for 40 EeV < E ≤ 50 EeV. Three roughly linear structures among a number of possible linear structures are outlined in red (current to 8/11/2012).

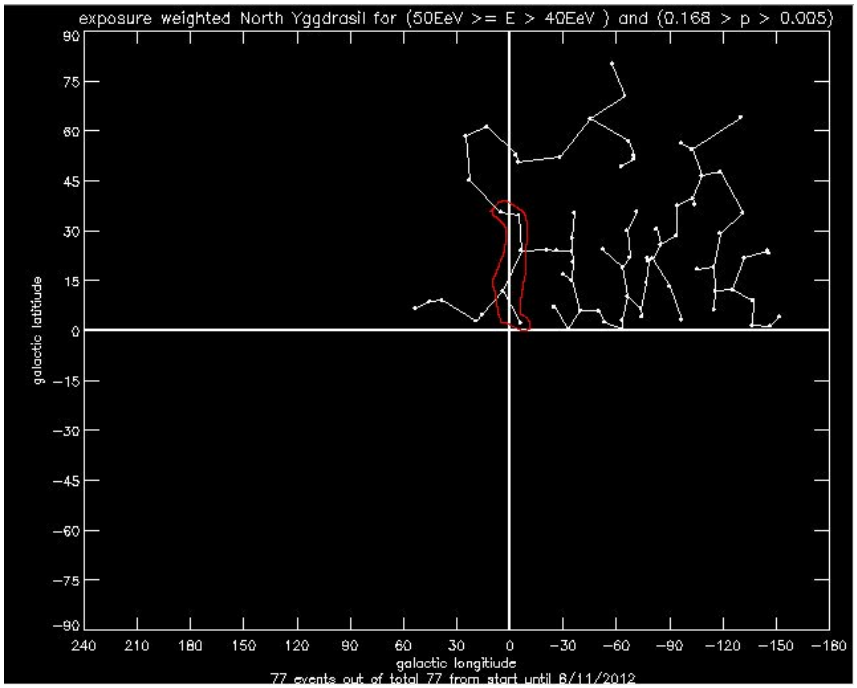

Figure 7.4: Northern Ygg for 40 EeV < E ≤ 50 EeV. One roughly linear structure among a number of possible linear structures is outlined in red (current to 8/11/2012).

7.4 Tables With No Galactic Longitude Divisions

Generalized tables with no galactic **l**Band filters.

7.4.1 40 EeV $<$ E $\leq$ 50 EeV

All distributions types 1, 2, and 3 plots are found in Appendix A on page 233.

distributions **Type 1**: Figure A.7 on page 239.

distributions **Type 2**: Figure A.8 on page 240.

distributions **Type 3**: Figure A.9 on page 241.

The individual energy table is Table 7.3 on the following page.

7.4.2 50 EeV $<$ E $\leq$ 60 EeV

All distributions types 1, 2 and 3 are found in Appendix B on page 251.

distributions **Type 1**: Figure B.7 on page 257.

distributions **Type 2**: Figure B.8 on page 258.

distributions **Type 3**: Figure B.9 on page 259.

The individual energy table is Table 7.4 on page 194.

7.4.3 E $>$ 60 EeV

All distributions types 1, 2 and 3 are found in Appendix C on page 269.

distributions **Type 1**: Figure C.7 on page 275.

distributions **Type 2**: Figure C.8 on page 276.

distributions **Type 3**: Figure C.9 on page 277.

The individual energy table is Table 7.5 on page 195.

Table 7.3: 40 EeV < E ≤ 50 EeV: No Filters.

SOUTH AND NORTH branch Θ and branch node values over					
bBands and entire galactic longitude lBand regions					
bBand°	**40 EeV < E ≤ 50 EeV ****				
	Θ°	**node**	$\%_S$	**(l_1, l_2)?**	**(b_1, b_2)?**
[−90 : −45]					
	44	10	29.91	(-27,1)	(-62,-41)
	11 *	23	3.23	(17,183)	(-85,-42)
	29	18	20.44	(-159,-48)	(-82,-44)
[−45 : −15]					
	17 *	43 !	0.95	(-131,22)	(-50,-8)
	64	4	61.71	(21,29)	(-25,-12)
	46	10	22.45	(41,79)	(-54,-10)
	82 *	2	97.99	(-101,-98)	(-16,-6)
	12 *	9	1.75	(167,204)	(-58,-7)
[−15 : 0]					
	20 *	13 !	0.62	(-86,-37)	(-18,-1)
	71	3	71.69	(-28,-23)	(-20,-9)
	64	6	29.76	(-8,2) ⊙	(-22,-2)
	34	5	14.56	(11,28)	(-21,-11)
	24	9	2.00	(-133,-98)	(-16,0)
	61	3	61.65	(-166,-155)	(-17,-6)
	65	2	76.85	(-39,-36)	(-24,-14)
	87 *	2	98.73	(45,47)	(-24,-10)
[0 : 15]					
	23	14 !	0.83	(-78,-25)	(0,21)
	26	9	3.22	(-152,-115)	(1,22)
	40	3	27.20	(-97,-81)	(3,22)
	20 *	8	1.35	(-6,53)	(2,24)
[15 : 45]					
	35	14	11.47	(-40,23)	(6,45)
	56	6	43.25	(-71,-52)	(10,36)
	64	12	19.26	(-106,-77)	(13,47)
	47	6	36.70	(-131,-105)	(12,48)
	45	4	44.40	(-146,-126)	(12,24)
[45 : 90]					
	37	14	10.32	(-70,25)	(36,80)
	28	7	20.40	(-131,-97)	(35,64)

** Data current to 8/11/2012.

* Pierre Auger Observatory **extreme** branch Θ of possible interest.

! Pierre Auger Observatory **extreme** branch node number of possible interest.

? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges, or *within* energy ranges about the galactic disk **b**Band divisions of $([-15:0],(0:15])^\circ$.

⊙ Pierre Auger Observatory branch of possible interest with a **l** or **b** branch-walk-pair co-ordinate within $\pm 10^\circ$ of the galactic axis **l**/**b** = **0**.

Table 7.4: 50 EeV < E ≤ 60 EeV: No Filters.

SOUTH AND NORTH branch Θ and branch node values over bBands and entire galactic longitude lBand regions					
bBand°	**50 EeV < E ≤ 60 EeV ****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]					
	7 *	19	< 0.062	(75,297)	(-69,-37)
	45	8	44.11	(-51,8)	(-78,-31)
	23	2	45.29	(49,64)	(-47,-40)
[−45 : −15]					
	30	14	9.77	(-104,-38)	(-52,-10)
	59	5	38.82	(-24,9)	(-48,-6)
	59	6	33.11	(-169,-133)	(-50,-12)
	27	2	27.52	(153,155)	(-47,-37)
	33	9	14.27	(31,104)	(-47,-16)
[−15 : 0]					
	53	6 !	10.09	(-112,-98)	(-25,-2)
	62	3	59.61	(-60,-54)	(-23,-4)
	58	2	64.61	(-133,-133)	(-20,-12)
	70	2	83.84	(0,4) ⊙	(-19,-6)
[0 : 15]					
	27	6 !	2.63	(-100,-52)	(8,30)
	55	2	56.35	(-119,-119)	(15,16)
	80	2	92.79	(-148,-143)	(12,17)
	31	4	8.25	(2,30)	(2,18)
[15 : 45]					
	55	8	19.42	(-33,30)	(3,65)
	21 *	11 !	2.21	(-150,74)	(11,38)
	63	6	33.80	(-72,-42)	(14,56)
[45 : 90]					
	37	5	38.19	(-72,-6)	(42,69)

** data current to 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])^\circ$.
⊙ Pierre Auger Observatory branch of possible interest with a **l** or **b**
branch-walk-pair co-ordinate within ±10° of the galactic axis $\mathbf{l}/\mathbf{b} = \mathbf{0}$.

Table 7.5: E > 60 EeV: No Filters.

SOUTH AND NORTH branch Θ and branch node values over bBands and entire galactic longitude lBand regions					
bBand°	**E > 60 EeV ****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]					
	9 *	21	0.058	(47,334)	(-80,-39)
	31	3	53.68	(-114,-88)	(-46,-34)
	39	3	61.62	(8,38)	(-49,-33)
	79 *	2	95.91	(71,89)	(-47,-25)
[−45 : −15]					
	22	5	11.78	(-168,-125)	(-25,-8)
	25	19 !	7.74	(-53,89)	(-56,-3)
	52	6	32.31	(-114,-62)	(-46,-6)
	2 *	2	0.13	(169,189)	(-45,-39)
[−15 : 0]					
	11 *	12 !	< 0.027	(-142,-37)	(-19,0)
	54	2	57.24	(-27,-17)	(-16,-3)
	33	4	11.57	(7,36)	(-19,0)
	26	2	9.37	(-168,-150)	(-22,-9)
[0 : 15]					
	40	7	8.37	(-52,-34)	(2,17)
	33	4	11.39	(12,38	(5,22)
	54	3	46.94	(-13,-2)	(6,22)
	77	2	89.96	(-109,-98)	(10,24)
	53	2	54.75	(-143,-141)	(6,23)
[15 : 45]					
	15 *	27 !	0.11	(-143,15)	(6,50)
	65	3	67.43	(-37,-34)	(13,17)
[45 : 90]					
	46	7	20.13	(-105,-22)	(32,69)

** Data current to 8/11/2012. * Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.• Pierre Auger Observatory
? compare Pierre Auger Observatory branch-walk-pairs/ROIs*across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15:0],(0:15])^\circ$.

7.4.4 Energy Comparison Table 1

In Table 7.6 on page 199, across all three energy ranges, are all the Pierre Auger Observatory branch variables calculated for branches cut according to the **b**Band divisions in the plane, **b**Band $[-90:-45]^\circ, [-45:-15]^\circ, [-15:0]^\circ, (0:15]^\circ, [15:45]^\circ, [45:90]^\circ$.

This table can be considered a preliminary energy comparison table, with no applied galactic longitude filters that may differentiate between very long or short (in their galactic longitude difference) Pierre Auger Observatory branches, or respond to whether branch Θs are consistent with inter-spiral-arm GMFs or the spiral-arm GMFs. Because there are so many results in this initial energy comparison table, along with some multiple similarities and reflections, we don't attach tags (e.g. $\clubsuit, \diamond, \dagger, \surd$) to the text below as it becomes confusing. It must also be noted, the tags attached to branches are not necessarily identical to the tags attached to the same branches that appear in the galactic filter tables, Table 7.5.4 on page 204, Table 7.6.4 on page 212 and Table 7.7.4 on page 219. The similarities between **extreme** branch Θs/ROIs *across* energy ranges are denoted ($\clubsuit, \spadesuit$). The similarities between an **extreme** branch Θ/ROIs *across* energy ranges and a non-extreme branch Θ are denoted ($\dagger$), and the similarities between non-extreme branch Θs/ROIs *across* energy ranges are denoted ($\surd$), but do not necessarily have the identical numbering systems when galactic filters are applied. This also applies to similar (branch Θs/ROIs) branches reflected in a mirror (M) or diagonal (D) fashion *within* energy ranges about galactic latitude, $\mathbf{b} = \mathbf{0}^\circ$, and in the galactic disk regions which are denoted ($\diamond, \heartsuit$).

- In the **b**Band, $[-90:-45]^\circ$, **all** the small **extreme** Θ Pierre Auger Observatory branches of interest, in the three energy ranges, have overlapping ROIs. In order of increasing energy range, the branch Θs are $\Theta = 11^\circ$ with 23 nodes (low), $\Theta = 7^\circ$ with 19 nodes (mid) which is claimed as a positive result, and $\Theta = 9^\circ$ with 21 nodes (GZK), also with a positive result. The GZK branch covers the largest change in galactic longitude.

- In the **b**Band $[-45:-15]^\circ$, there are three small **extreme** branch Θs in the low-energy range. The small **extreme** branch Θ of 17° has an **extreme** node sum of 43. The small **extreme** branch $\Theta = 12^\circ$ with 9 nodes has **l** branch-walk-pair values of $(167, 204)^\circ$ which are similar

to the GZK small **extreme** branch $\Theta = 2^\circ$ with branch-walk-pair values of $(169, 189)^\circ$. In the mid-energy range, there is a non-extreme branch $\Theta = 59^\circ$ with 6 nodes and a big comparison percentage, but its **l/b** branch-walk-pair values are similar to the Pierre Auger Observatory branch $\Theta = 22^\circ$ with 5 nodes in the GZK-energy range.

- In the **b**Band $[-15 : 0]^\circ$, we claim a positive result with the small **extreme** branch $\Theta = 11^\circ$ in the GZK-energy range. This branch also has an **extreme** branch node number.

  In this **b**Band there are two similar branches which also have their **l** branch-walk-pair **co-ordinate** within $\pm 10^\circ$ of the galactic axis **b=0** and one of their **l** branch-walk-pair ordinates within $\pm 10^\circ$ of the galactic axis **l=0**. The branches are, respectively, - in the low-energy range, a branch with $\Theta = 64^\circ$ and 6 nodes, - and in the mid-energy range, a branch with $\Theta = 70^\circ$ and 2 nodes. Both these branches tend more to a vertical inclination with respect to the galactic plane, $\mathbf{b} = 0^\circ$, which is unusual, because most branches have either small **extreme** Θs or non-extreme Θs. These two branches also have their **l/b**Walk co-ordinates within $\pm 30^\circ$ of the GC.

- In the **b**Band $(0 : 15]^\circ$, all of the branches of interest have, like most the branches in the **b**Band $[-15 : 0]^\circ$ south, a galactic **b**Walk ordinate close ($\pm 10^\circ$) to the galactic axis **b=0**. About these two galactic disk **b**Bands, *within* **each** energy range, are reflected **l** and **b** branch-walk-pair values. Some reflections *within* energy ranges have more than one reflection counterpart. We hope to untangle some of these reflection relationships, mirror (M) or diagonal (D), when we consider the galactic **b**Band filters.

- In the **b**Band $(15 : 45]^\circ$, and the GZK-energy range, we have a branch with a small **extreme** branch $\Theta = 15^\circ$ and an **extreme** branch node sum of 27 nodes. For the GZK-energy range, and with the lesser Pierre Auger Observatory exposure in these galactic latitudes, a small **extreme** branch Θ composed of so many nodes is interesting.

This energy comparison table has many branches with interesting properties, including similarities or ROIs of all three types ($\clubsuit, \dagger, \surd$) *across* energy ranges. There are three branches with positive results, and there are similar mirror and diagonal reflected branches *within* energy ranges about

the galactic disk **b**Band divisions. The only **b**Band division with no interesting branches is the **b**Band $[45 : 90]^{\circ}$ where the Pierre Auger Observatory exposure is limited.

In all energy comparison tables, we have included the complete Pierre Auger Observatory branch results whether tagged or untagged. All similar branches (branch-walk-pairs/ROIs) are considered interesting. Non-similar branches with two or more tags are also to be considered interesting.

Some branches are tagged as similar (branch-walk-pairs/ROIs) multiple times. Once the galactic longitude and latitude filters have been applied, some similaries may be disappear, but not always.

Table 7.6: Energy Comparison Table 1. No Filters.

possible correlating SOUTH AND NORTH Pierre Auger Observatory branch Θ and branch node values over bBands and entire galactic longitude lBand regions **															
bBand°	**40 EeV < E ≤ 50 EeV**					**50 EeV < E ≤ 60 EeV**					**E > 60 EeV**				
	Θ°	**node**	**%$_S$**	**(l_1, l_2)**	**(b_1, b_2)**	**Θ°**	**node**	**%$_S$**	**(l_1, l_2)**	**(b_1, b_2)**	**Θ°**	**node**	**%$_S$**	**(l_1, l_2)**	**(b_1, b_2)**
[−90 : −45]															
	11 * ♣1	23	3.23	(17,183)	(-85,-42)	7 * ♣1 ★	19	< 0.062	(75,297)	(-69,-37)	9 * ♣1 ★	21	0.058	(47,334)	(-80,-39)
	44 √1	10	NA	(-27,1) •	(-62,-41)	45 √1	8	NA	(-51,8) •	(-78,-31)	79 * †1	2	95.91	(71,89)	(-47,-25)
	29	18	NA	(-159,-48)	(-82,-44)	23 †1	2	NA	(49,64)	(-47,-40)	31	3	NA	(-114,-88)	(-46,-34)
											39	3	NA	(8,38) ∘	(-49,-33)
[−45 : −15]															
	17 *	43 !	0.95	(-131,22)	(-50,-8)	59 √1	6	NA	(-169,-133)	(-50,-12)	22 √1	5	NA	(-168,-125)	(-25,-8)
	12 *	9	1.75	(167,204)	(-58,-7)	59	5	NA	(-24,9)	(-48,-6)	25	19 !	NA	(-53,89)	(-56,-3)
	82 *	2	97.99	(-101,-98)	(-16,-6)	30 √2	14	NA	(-104,-38)	(-52,-10)	2 *	2	0.13	(169,189)	(-45,-39)
	64	4	NA	(21,29)	(-25,-12)	27	2	NA	(153,155)	(-47,-37)	52 √2	6	NA	(-114,-62)	(-46,-6)
	46 √3	10	NA	(41,79)	(-54,-10)	33 √3	9	NA	(31,104)	(-47,-16)					
[−15 : 0]															
	24 √1 ⋄1M	9	NA	(-133,-98)	(-16,0)	53 √1 ♡2M	6 !	NA	(-112,-98)	(-25,-2)	11 * ★	12 !	< 0.027	(-142,-37)	(-19,0)
	20 * ♡1M	13 !	0.62	(-86,-37)	(-18,-1)	70 √2 √3	2	NA	(0,4) ⊙	(-19,-6)	26 √4 ⋄1M	2	NA	(-168,-150)	(-22,-9)
	87 *	2	98.73	(45,47)	(-24,-10)	62 √5	3	NA	(-60,-54)	(-23,-4)	33 √6 ♡1M ⋄2M	4	NA	(7,36)	(-19,0)
	64 √2 √8	6	NA	(-8,-2) ⊙	(-22,-2)	58 ♡1M ⋄1M	2	NA	(-133,-133)	(-20,-12)	54 √7 √8 √9 ♡2M ♡3D ⋄3D	2	NA	(-27,-17)	(-16,-3)
	34 √3 √6	5	NA	(11,28)	(-21,-11)										
	61 √4	3	NA	(-166,-155)	(-17,-7)										
	71 √7	3	NA	(-28,-23)	(-20,-9)										
	65 √5 √9	2	NA	(-39,-36)	(-24,-14)										
[0 : 15]															
	23 ♡1M	14 !	NA	(-78,-25)	(0,21)	27	6 !	NA	(-100,-52)	(8,30)	77 √2 √4	2	NA	(-109,-98)	(10,24)
	26 ⋄1M	9	NA	(-152,-115)	(1,22)	31 †1 √5	4	NA	(2,30)	(2,18)	40 ♡2M	7	NA	(-52,-34)	(2,17)
	20 * †1 √1	8	1.35	(-6,53)	(2,24)	55 √2 √5 ♡1M ♡2M	2	NA	(-119,-119)	(15,16)	33 †1 √1 √5 ♡1M ⋄3D	4	NA	(12,38)	(5,22)
	40 √4	3	NA	(-97,-81)	(3,22)	80 √3 ⋄1M	2	NA	(-148,-143)	(12,17)					
											54 √5 ♡3D ⋄2M	3	NA	(13,2)	(6,22)
											53 √3 √5 ⋄1M	2	NA	(-143,-141)	(6,23)
[15 : 45]															
	35 √1	14	NA	(-40,23)	(6,45)	21 *	11 !	2.21	(-150,74)	(11,38)	15 *	27 !	0.11	(-143,15)	(6,50)
	56 √2	6	NA	(-71,-52)	(10,36)	55 √1	8	NA	(-33,30)	(3,65)	65	3	NA	(-37,-34)	(13,17)
	64	12	NA	(-106,-77)	(13,47)	63 √2	6	NA	(-72,-42)	(14,56)					
	47	6	NA	(-131,-105)	(12,48)										
	45	4	NA	(-146,-126)	(12,24)										
[45 : 90]															
	37	14	NA	(70,25)	(36,80)	37	5	NA	(-72,-6)	(42,69)	46	7	NA	(-105,-22)	(32,69)
	28	7	NA	(-131,-97)	(35,64)										

** data current until 8/11/2012. NA-**Extreme** Θ comparison percentage %$_S$ not applicable. * Pierre Auger Observatory **extreme** branch Θ of interest. ! Pierre Auger Observatory **extreme** branch node number of interest.

♣1 Pierre Auger Observatory **extreme** branch Θs with similar branch-walk-pair/ROIs *across* energy ranges.

♡1, 2, 3, ⋄1 Pierre Auger Observatory similar mirror (M) or diagonal (D) reflected branch-walk-pair/ROIs values in the same energy range, and within the galactic disk **b**Band divisions of ([−15 : 0], (0 : 15])°.

†1 **Extreme** branch with similar branch-walk-pair/ROI to a non-extreme branch, *across* energy range.

√1, √2, √3, √4, √5, √6, √7, √8, √9 Non-extreme Θ's *across* energy ranges with similar branch-walk-pair/ROIs values. ⊙ - Pierre Auger Observatory branch with a **l** or **b** branch-walk-pair co-ordinate within ±10° of the galactic axis **l**/**b** = **0**.

★ Pierre Auger Observatory **extreme** branch Θ with a positive result.

7.5 Galactic Longitude Filters, both $[-135 : 135]^\circ$ and $[-45 : 45]^\circ$

These galactic longitude filters are associated with the positioning of the galaxy's inter-spiral arms where the azimuthal GMFs are believed to be regular.

7.5.1 40 EeV $<$ E $\leq$ 50 EeV

All distributions types 1, 2 and 3 are found in Appendix A.

distributions **Type 1**: Figure A.10 on page 242.

distributions **Type 2**: Figure A.11 on page 243.

distributions **Type 3**: Figure A.12 on page 244.

The individual energy table is Table 7.7 on the following page.

7.5.2 50 EeV $<$ E $\leq$ 60 EeV

All distributions types 1, 2 and 3 are found in Appendix B.

distribution **Type 1**: Figure B.10 on page 260.

distribution **Type 2**: Figure B.11 on page 261.

distribution **Type 3**: Figure B.12 on page 262.

The individual energy table is Table 7.8 on page 202.

7.5.3 E $>$ 60 EeV

All distributions types 1, 2 and 3 are found inAppendix C.

distributions **Type 1**: Figure C.13 on page 281.

distributions **Type 2**: Figure C.14 on page 282.

distributions **Type 3**: Figure C.15 on page 283.

The individual energy table is Table 7.9 on page 203.

Table 7.7: 40 EeV < E ≤ 50 EeV: Inter-Spiral Arms.

SOUTH AND NORTH branch Θ and branch node values over bBands					
and over galactic filter regions l ∈ ([−135 : 135]° **and** [−45 : 45]°)					
bBand°	**40 EeV < E ≤ 50 EeV****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]					
	11 *	23	3.17	(17,183)	(-85,-42)
	44	10	17.64	(-27,1)	(-62,-41)
[−45 : −15]					
	12 *	9	1.72	(167,204)	(-58,-7)
	64	4	17.68	(21,29)	(-25,-12)
[−15 : 0]					
	61	3	21.83	(-166,-155)	(-17,-6)
	71	3	26.94	(-28,-23)	(-20,-9)
	64	6	8.60	(-8,2) ⊙	(-22,-2)
	34	5	3.60	(11,28)	(-21,-11)
	65	2	29.39	(-39,-36)	(-24,-14)
[0 : 15]					
	0	0	0	0	0
[15 : 45]					
	35	14 !	0.37	(-40,23)	(6,45)
[45 : 90]					
	0	0	0	0	0

** data current until 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])^\circ$.
⊙ Pierre Auger Observatory branch of possible interest,
with a **l** or **b** branch-walk-pair co-ordinate within ±10° of the galactic axix **l/b=0**.

Table 7.8: 50 EeV < E ≤ 60 EeV: Inter-Spiral Arms.

SOUTH AND NORTH branch Θ and branch node values over bBands					
and over galactic filter regions l ∈ ([−135 : 135]° **and** [−45 : 45]°)					
bBand°	**50 EeV < E ≤ 60 EeV****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]	7 *	19	< 0.062	(75,297)	(-69,-37)
[−45 : −15]	27	2	12.11	(153,155)	(-47,-37)
	59	5	8.32	(-24,9)	(-48,-6)
[−15 : 0]	70 *	2	26.55	(0,4) ⊙	(-19,-6)
[0 : 15]	80	2	31.62	(-148,-143)	(12,17)
	31	4	1.52	(2,30)	(2,18)
[15 : 45]	55	8 !	0.55	(-33,30)	(3,65)
[45 : 90]	0	0	0	0	0

** data current until 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])°$.
⊙ Pierre Auger Observatory branch of possible interest,
with a **l** or **b** branch-walk-pair co-ordinate within ±10° of the galactic axix **l/b=0**.

Table 7.9: E > 60 EeV: Inter-Spiral Arms.

SOUTH AND NORTH branch Θ and branch node values over bBands					
and over galactic filter regions l $\in ([-135:135]^\circ$ **and** $[-45:45]^\circ)$					
bBand°	**E > 60 EeV****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
$[-90:-45]$					
	9 *	21	0.058	(47,334)	(-80,-39)
	39	3	31.99	(8,38)	(-49,-33)
$[-45:-15]$					
	2 *	2	0.13	(169,189)	(-45,-39)
$[-15:0]$					
	26	2	3.43	(-168,-150)	(-22,-9)
	54	2	21.49	(-27,-17)	(-16,-3)
	33	4	3.84	(7,36)	(-19,0)
$[0:15]$					
	53	2	17.92	(-143,-141)	(6,23)
	33	4	2.78	(12,38)	(5,22)
	54	3	14.31	(13,2)	(6,22)
$[15:45]$					
	65	3	18.80	(-37,-34)	(13,17)
$[45:90]$					
	0	0	0	0	0

** data current to 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15:0],(0:15])^\circ$.

7.5.4 Energy Comparison Table 2

In this, the second energy comparison table, Table 7.10, we apply the inter-spiral galactic longitude, **l**, filters of $[-135 : 135]^\circ$ and $[-45, 45]^\circ$. We expect the GMF to be azimuthal and regular for at least the galactic disk **b**Bands.

- In the **b**Band $[-90 : -45]^\circ$, we see three Pierre Auger Observatory small **extreme** overlapping ROI branches Θs (♠1). The mid and GZK-energy range branch Θs have positive results (★). These same three branches also qualify for the **b**Band filter $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90^\circ$, (♠), where we try to isolate branches close together possibly with point sources, or branches that are reasonably long across their galactic longitude differences that may be attributed to large-scale sources or causes.

- In the **b**Band $[-45 : -15]^\circ$, there are two small **extreme** branch Θs with overlapping branch Θ ROIs. Namely, in the low-energy range, a branch with $\Theta = 12^\circ$ with a node sum of 9, and in the GZK-energy range a branch with $\Theta = 2^\circ$ and a node sum of 2. There is a similarity in their $(\mathbf{l_1}, \mathbf{l_2})$ branch-walk-pair values. The $(\mathbf{b_1}, \mathbf{b_2})$ branch-walk-pairs do diverge slightly outside the $\pm 25^\circ$ limit, but perhaps we should allow a greater latitude in divergence when we compare GZK events to those at our lower energies, given the uncertainties surrounding the GZK-energy range value for light and heavy nuclei (whether the energy ranges are approximately the same or of different orders). These 2 nuclei are also not necessarily both heavy or both light. We note that the GZK branch node sum of two is the minimum value a branch can be, but its small **extreme** branch $\Theta = 2^\circ$ is smaller than than the low-energy range branch with $\Theta = 12^\circ$, in accordance with our expectations of poloidal magnetic field angular deflections of light CRs in the GZK energy region.

- In the **b**Band $[-15 : 0]^\circ$, there are two similar non-extreme branch Θs , one in the low-energy range and one in the mid-energy range that have **both** galactic **l** branch-walk-pair co-ordinates within $\pm 10^\circ$ of the galactic axis **l=0**, and each a galactic **b** branch-walk-pair ordinate within $\pm 10^\circ$ of the galactic axis **l=0** ($\sqrt{2}, \odot, \bullet$). These branches are, respectively, a branch with $\Theta = 64^\circ$ with a node sum of 6 (low-energy range - see Appendix A : Figure A.5)

and a branch with $\Theta = 70°$ with a node sum of 2 (mid-energy range. See Appendix B: Figure B.5 on page 255). It may be that the GC has an effect on the CR nuclei that compose these two branches. The sole mid-energy range branch, ($\sqrt{4}$), is also similar to another low-energy range branch. Namely, the branch with $\Theta = 34°$ and a node sum of 5 ($\sqrt{4}$). Looking at the unfiltered plots of these branches we prefer the similarity between the two branches with both their **l**Band branch-pair-walks within $\pm 10°$ of the galactic axis **l=0** ($\sqrt{2}$). In the GZK-energy range there is a branch which may be part of a similar set with each of three low-energy range branches. This GZK branch has $\Theta = 54°$ and a node sum of 2 ($\sqrt{2}, \sqrt{3}, \sqrt{5}$) (see Appendix C: Figure C.5 on page 273) . We exclude the similar low-energy range ($\sqrt{2}$) branch, because we have already coupled this branch with the sole mid-energy range branch with $\Theta = 70°$ and a node sum of 2. We exlude the similarity between the GZK branch with $\Theta = 54°$ branch ($\sqrt{5}$) and the low-energy range branch with $\Theta = 65°$ ($\sqrt{5}$) because the GZK branch with $\Theta = 54°$ is also similar to the low-energy range branch with $\Theta = 71°$ ($\sqrt{3}$) and both these branches have their **l/b** walk-pair co-ordinates within $\pm 30°$ of the GC. In all, there are up to five similar non-extreme branches *across* the energy ranges in this **b**Band and three of these branches have **l** and **b** branch-walk-pairs within $\pm 30°$ of the GC.

- In the **b**Band $(0 : 15]°$, there are only GZK Pierre Auger Observatory branches. In the case of similar reflected branches *within* energy ranges about the galactic $\mathbf{b} = 0°$ plane for the galactic disk **b**Band divisions, all the GZK branches have reflected relationships ($\diamond 1M, \heartsuit 1M, \heartsuit 2D$).

- In the **b**Band $[15 : 45]°$, for the low and mid-energy ranges, there are two similar branches with **extreme** node sums ($!, \sqrt{1}$).

- In the **b**Band $[45 : 90]°$, the galactic **b**Band division in the plane, where there is low exposure, there are no Pierre Auger observatory branches.

This galactic **b**Band filter has some interesting results. The south galactic pole's **b**Band division has three similar small **extreme** branch Θs *across* all energy ranges. Small **extreme** branch Θs are

expected for this **b**Band filter if the CR events composing a branch were to encounter a poloidal MF. The existence of the GZK **extreme** branch Θ intimates this poloidal field may by extra-galactic, but the south galactic pole low and mid-energy range branches too are similar, suggesting that there may be a re-accelerating galactic source in the regions of branch similarity.

In the galactic **b**Band disk divisions, six of the total nine branches have **l** and **b** branch-walk-pairs within $\pm 30^\circ$ of the GC. The GMFs around the GC region which houses a black hole should be different from the azimuthal and poloidal GMFs in the disk regions. The limit of the GC influence of $\pm 30^\circ$, was chosen because many of the Pierre Auger Observatory branches have **l** and **b** branch-walk-pairs within this boundary. Our understanding of the poloidal and azimuthal GMFs in the disk, and how they interact with one another, let alone the GC, is rudimentary.

This galactic **b**Band filter selects directions associated with coherent azimuthal GMFs for the galactic disk regions. Possibly those non-extreme reflected mirror (M) or diagonal (D) GZK branch events about the galactic disk regions, $\mathbf{b}([-15:0],(0:15])^\circ$, are connected to the south and north galactic lobes (re-accelerated CR nuclei) and/or connected with the reversal of directions of the GMF above and below the galactic plane.

Table 7.10: Energy Comparison Table 2. Inter-Spiral Arms.

possible correlating SOUTH AND NORTH branch Θ and branch node values over bBands															
and over galactic longitude filters $[-135:135]^\circ$ **and** $[-45:45]^\circ$ **															
bBand°	**40 EeV < E ≤ 50 EeV**					**50 EeV< E ≤ 60 EeV**					**E > 60 EeV**				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$	$(\mathbf{b_1}, \mathbf{b_2})$	**Θ°**	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$	$(\mathbf{b_1}, \mathbf{b_2})$	**Θ°**	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$	$(\mathbf{b_1}, \mathbf{b_2})$
$[-90:-45]$	11 * ♠1 44	23 10	3.17 NA	(17,183) (-27,1) ○	(-85,-42) (-62,-41)	7 * ♠1 ★	19	< 0.062	(75,297)	(-69,-37)	9 * ♠1 ★ 39	21 3	0.058 NA	(47,334) (8,38) ○	(-80,-39) (-49,-33)
$[-45:-15]$	12 * 64	9 4	1.72 NA	(167,204) (21,29)	(-58,-7) (-25,-12)	27 59	2 5	NA NA	(153,155) (-24,9)	(-47,-37) (-48,-6)	2 *	2	0.13	(169,189)	(-45,-39)
$[-15:0]$	61 √1 71 √3 64 √2 34 √4 65 √5	3 3 6 5 2	NA NA NA NA NA	(-166,-155) (-28,-23) (-8,2) ⊙ (11,28) (-39,-36)	(-17,-6) (-20,-9) (-22,-2) (-21,-11) (-24,-14)	70 √2 √4	2	NA	(0,4) ⊙	(-19,6)	33 ♡1*M* √4 26 √1 ◇1*M* 54 √2 √3 √5 ♡2*D*	4 2 2	NA NA NA	(7,36) (-168,-150) (-27,-17)	(-19,0) (-22,-9) (-16,-3)
$[0:15]$											33 ♡1*M* 53 ◇1*M* 54 ♡2*D*	4 2 3	NA NA NA	(12,38) (-143,-141) (13,2)	(5,22) (6,23) (6,22)
$[15:45]$	35 √1	14 !	NA	(-40,23)	(6,45)	55 √1	8 !	NA	(-33,30)	(3,65)	65	3	NA	(-37,-34)	(13,17)
$[45:90]$															

** data current until 8/11/2012. NA-**Extreme** Θ comparison percentage $\%_S$ not applicable. * Pierre Auger Observatory **extreme** branch Θ of interest.! Pierre Auger Observatory **extreme** branch node number of interest.
♣1, ♠1 Pierre Auger Observatory **extreme** branch Θs with similar branch-walk-pair/ROIs *across* energy ranges
♡1, 2, ◇1 Pierre Auger Observatory similar mirror (M) or diagonal (D) reflected branch-walk-pairs/ROI values in the same energy range, and within the galactic disk **b**Band divisions of $([-15:0], (0:15])^\circ$.
√1, √2, √3, √4, √5 Non-extreme branches *across* energy ranges with similar branch-walk-pair values/ROIs.
⊙ Pierre Auger Observatory branch with a **l** or **b** branch-walk-pair co-ordinate within $\pm 10^\circ$ of the galactic axis **b/l=0**. ★ Pierre Auger Observatory **extreme** branch Θ with a positive result.

7.6 Galactic Longitude Filters, both $[-135 : -45]^\circ$ and $[45 : 135]^\circ$

These galactic longitude filters are positioned in the galaxy's spiral arms where the GMFs are believed to be chaotic.

7.6.1 40 EeV $<$ E $\leq$ 50 EeV

all distributions types 1, 2 and 3 are found in Appendix A.

distributions **Type 1**: Figure A.13 on page 245.

distributions **Type 2**: Figure A.14 on page 246.

distributions **Type 3**: Figure A.15 on page 247.

The individual energy table is Table 7.11 on the following page.

7.6.2 50 EeV $<$ E $\leq$ 60 EeV

All distributions types 1, 2 and 3 are found in Appendix B.

distribution **Type 1**: Figure B.13 on page 263.

distribution **Type 2**: Figure B.14 on page 264.

distribution **Type 3**: Figure B.15 on page 265.

The individual energy table is Table 7.12 on page 210.

7.6.3 E $>$ 60 EeV

All distributions types 1, 2 and 3 are found in Appendix C.

distribution **Type 1**: Figure C.7 on page 275.

distribution **Type 2**: Figure C.8 on page 276.

distribution **Type 3**: Figure C.9 on page 277.

The individual energy table is Table 7.13 on page 211.

Table 7.11: 40 EeV < E ≤ 50 EeV: Spiral Arms.

SOUTH AND NORTH branch Θ and branch node values over bBands and over galactic longitude filters [-135:-45]° and [45:135]° **					
bBand°	**40 EeV < E ≤ 50 EeV**				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]	0	0	0	0	0
[−45 : −15]	82	2	23.46	(-101,-98)	(-16,-6)
[−15 : 0]	24	9 !	0.21	(-133,-98)	(-16,0)
[0 : 15]	40	3	6.61	(-97,-81)	(3,22)
[15 : 45]	56 64 47	6 12 ! 6	8.50 1.86 6.62	(-71,-52) (-106,-77) (-131,-105)	(10,36) item (13,47) (12,48)
[45 : 90]	28	7 !	1.45	(-131,-97)	(35,64)

** data current to 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])^\circ$.

Table 7.12: 50 EeV < E ≤ 60 EeV: Spiral Arms.

SOUTH AND NORTH branch Θ and branch node values over bBands					
and over galactic longitude filters $[-135 : -45]^\circ$ **and** $[45 : 135]^\circ$					
bBand°	**50 EeV < E ≤ 60 EeV****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
$[-90 : -45]$	23	2	1.24	(49,64)	(-47,-40)
$[-45 : -15]$	0	0	0	0	0
$[-15 : 0]$	53	6 !	1.71	(-112,-98)	(-25,-2)
	62	3	13.56	(-60,-54)	(-23,-4)
	58	2	15.52	(-133,-133)	(-20,-12)
[0 : 15]	27	6 !	0.25	(-100,-52)	(8,30)
	55	2	13.05	(-119,-119)	(15,16)
[15 : 45]	0	0	0	0	0
[45 : 90]	0	0	0	0	0

** data current to 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])^\circ$.

Table 7.13: E > 60 EeV: Spiral Arms.

SOUTH AND NORTH branch Θ and branch node values over bBands					
and over galactic longitude regions $[-135 : -45]^\circ$ **and** $[45 : 135]^\circ$					
bBand°	**E > 60 EeV ****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
$[-90 : -45]$	31 79	3 2	3.51 16.88	(-114,-88) (71,89)	(-46,-34) (-47,-25)
$[-45 : -15]$	52	6	2.13	(-114,-62)	(-46,-6)
$[-15 : 0]$	0	0	0	0	0
$[0 : 15]$	77	2	31.32	(-109,-98)	(10,24)
$[15 : 45]$	0	0	0	0	0
$[45 : 90]$	0	0	0	0	0

** data current to 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])^\circ$.

7.6.4 Energy Comparison Table 3

In the energy range comparison table, Table 7.14, there are three interesting branches (i.e. branches with two or more tags).

- In the **b**Band $[-15 : 0]^\circ$, the low-energy range has a branch with an **extreme** branch node sum of 9 and the almost small $\Theta = 24^\circ$. Its **l** and **b** branch-walk-pair values are similar to a branch in the mid-energy range, which has a branch $\Theta = 53^\circ$ with an **extreme** branch node sum of 6.(!, $\sqrt{1}$).

- In the **b**Band $(0 : 15]^\circ$, there is one branch in each energy range and they are all interesting. The branch in the low-energy range is similar to the branch in the GZK range ($\sqrt{1}$). The branch in the mid-energy range has an **extreme** branch node sum (!).

All the five branches are interesting in the sense of their **Type 2** simulated distributions, and there are two sets of similar branches *across* two energy ranges. There are no Pierre Auger Observatory branches with both their branch-walk-pairs within $\pm 30^\circ$. of the GC. There are no reflected branches around the galactic disk **b**Band divisions of $[-15 : 0]^\circ$ and $(0 : 15]^\circ$. This makes sense. In chaotic azimuthal GMFs we would not expect symmetries about the galactic disk. All the other branches have either one tag, or none, and so are not interesting.

Table 7.14: Energy Comparison Table 3. Spiral Arms.

possible correlating SOUTH AND NORTH															
branch Θ and branch node values over															
over galactic longitude regions [−135 : −45]° **and** [45 : 135]° **															
bBand°	**40 EeV < E ≤ 50 EeV**					**50 EeV< E ≤ 60 EeV**					**E > 60 EeV**				
	Θ°	**node**	**%**	**(l_1, l_2)**	**(b_1, b_2)**	**Θ°**	**node**	**$\%_S$**	**(l_1, l_2)**	**(b_1, b_2)**	**Θ°**	**node**	**%**	**(l_1, l_2)**	**(b_1, b_2)**
[−90 : −45]						23 √1	2	NA	(49,64)	(-47,-40)	31	3	NA	(-114,-88)	(-46,-34)
											79 √1	2	NA	(71,89)	(-47,-25)
[−45 : −15]	82	2	NA	(-101,-98)	(-16,-6)						52	6	NA	(-114,-62)	(-46,-6)
[−15 : 0]	24 √1	9 !	NA	(-133,-98)	(-16,0)	53 √1	6 !	NA	(-112,-98)	(-25,-2)					
						62	3	NA	(-60,-54)	(-23,-4)					
						58	2	NA	(-133,-133)	(-20,-12)					
[0 : 15]	40 √1	3	NA	(-97,-81)	(3,22)	27	6 !	NA	(-100,-52)	(8,30)	77 √1	2	NA	(-109,-98)	(10,24)
[15 : 45]	56	6	NA	(-71,-52)	(10,36)										
	64	12 !	NA	(-106,-77)	(13,47)										
	47	6	NA	(-131,-105)	(12,48)										
[45 : 90]	28	7 !	NA	(-131,-97)	(35,64)										

** data current until 8/11/2012. NA-**Extreme** Θ comparison percentage $\%_S$ not applicable.
! Pierre Auger Observatory **extreme** branch node number.
√1 Non-extreme branches *across* energy ranges with similar branch-walk-pairs/ROIs.

7.7 Tables with the Galactic Longitude Divisions of $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90^\circ$

The purpose of this longitude filter is to select those Pierre Auger Observatory branches, **extreme** or otherwise, that may be very close together, i.e. at the ends of the galactic longitudes, **l**= $\pm 180^\circ$, or to find those Pierre Auger Observatory branches that are as wide apart as possible, i.e. those long branches whose galactic longitude walk-pair difference, $\Delta|\mathbf{l_1} - \mathbf{l_2}| \rightarrow 90^\circ$. Very long branches composed of many nodes may indicate large-scale sources and those short branches with many nodes may signal clustering from single point-sources. As always, we search for branches that possess similar **l** and **b** branch-walk-pair values *across* energy ranges or similar reflected (M/D) **l** and **b** branch-walk-pair values in the two galactic disk **b**Band divisions.

7.7.1 40 EeV $<$ E $\leq$ 50 EeV

All distributions types 1, 2 and 3 are found in Appendix A.

distributions **Type 1**: Figure A.16 on page 248.

distributions **Type 2**: Figure A.17 on page 249.

distributions **Type 3**: Figure A.18 on page 250.

The individual energy table is Table 7.15 on page 216.

7.7.2 50 EeV $<$ E $\leq$ 60 EeV

All distributions types 1, 2 and 3 are found in Appendix B.

distributions **Type 1**: Figure B.16 on page 266.

distributions **Type 2**: Figure B.17 on page 267.

distributions **Type 3**: Figure B.18 on page 268.

The individual energy table is Table 7.16 on page 217.

7.7.3 E > 60 EeV

All distributions types 1, 2 and 3 are found in Appendix C.

distributions **Type 1**: Figure C.16 on page 284.

distributions **Type 2**: Figure C.17 on page 285.

distributions **Type 3**: Figure C.18 on page 286.

The individual energy table is Table 7.17 on page 218.

Table 7.15: 40 EeV < E ≤ 50 EeV: $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90^\circ$

SOUTH AND NORTH branch Θ and branch node values over bBands					
over galactic longitude regions with $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90^\circ$					
bBand°	**40 EeV < E ≤ 50 EeV ****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]					
	11 *	23	5.05	(17,183)	(-85,-42)
	29	18	31.70	(-159,-48)	(-82,-44)
[−45 : −15]					
	17 *	43 !	0.95	(-131,22)	(-50,-8)
	12 *	9	1.73	(167,204)	(-58,-7)
[−15 : 0]					
[0 : 15]					
[15 : 45]					
[45 : 90]					
	37	14	NA	(-70,25)	(36,80)

** data current to 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])^\circ$.

Table 7.16: 50 EeV < E ≤ 60 EeV: $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$.

SOUTH AND NORTH branch Θ and branch node values over bBands					
over galactic longitude regions with $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ **					
bBand°	**60 EeV ≥ E > 50 EeV**				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]	7 *	19	<0.062	(75,297)	(-69,-37)
[−45 : −15]	0	0	0	0	0
[−15 : 0]	0	0	0	0	0
[0 : 15]	0	0	0	0	0
[15 : 45]	0	0	0	0	0
[45 : 90]	0	0	0	0	0

** data current to 8/11/2012.

* Pierre Auger Observatory **extreme** branch Θ of possible interest.

? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges, or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])°$.

Table 7.17: E > 60 EeV: $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$.

SOUTH AND NORTH branch Θ and branch node values over bBands					
over galactic longitude regions with $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$					
bBand°	**E > 60 EeV ****				
	Θ°	**node**	$\%_S$	$(\mathbf{l_1}, \mathbf{l_2})$?	$(\mathbf{b_1}, \mathbf{b_2})$?
[−90 : −45]	9 *	21	0.058	(47,334)	(-80,-39)
[−45 : −15]	25 2 *	19 2	7.74 0.13	(-53,89) (169,189)	(-56,-3) (-45,-39)
[−15 : 0]	11 *	12 !	<0.027	(-142,-37)	(-19,0)
[0 : 15]	0	0	0	0	0
[15 : 45]	15 *	27 !	0.11	(-143,15)	(6,50)
[45 : 90]	0	0	0	0	0

** data current to 8/11/2012.
* Pierre Auger Observatory **extreme** branch Θ of possible interest.
! Pierre Auger Observatory **extreme** branch node number of possible interest.
? compare Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges,
or *within* energy ranges about the galactic disk **b**Band divisions of $([-15 : 0], (0 : 15])°$.

7.7.4 Energy Comparison Table 4

This energy comparison table has some interesting results, in terms of similarity, small **extreme** branch Θs, and branches with positive results. There are ten Pierre Auger Observatory branches in all, and eight of those branches have small **extreme** Θs. Eight of the ten Pierre Auger Observatory branch results have branch-walk-pair values that are *almost* similar *across* energy ranges and roughly in the galactic, **(l,b)**, quadrant of $\mathbf{l} \in [-180 : 0]^\circ$, $\mathbf{b} \in [-90 : 0]^\circ$. While this quadrant is where the Pierre Auger Observatory's exposure is the greatest, the fact that similarities in galactic **l/b** Walks occur *across* energy ranges, suggests areas where we might have a common cause or source. Three of the four small **extreme** GZK-energy range branches lie in this quadrant and so we might invoke a large-scale moving astronomical object with a poloidal MF or an extra-galactic poloidal MF.

- In the **b**Band $[-90 : -45]^\circ$, there are three overlapping Θ ROIs. Pierre Auger Observatory branches *across* energy ranges also have small **extreme** branch Θs ($*, \spadesuit$) and two of the branches have positive results ($\star$).

- In the **b**Band $[-45 : -15]^\circ$, all the three branches have small **extreme** branch Θs ($*$). Two of the branches are in the low-energy range and the other is in the GZK-energy range. In particular the low-energy range branch with small **extreme** $\Theta = 17^\circ$ ($*$) also has an **extreme** branch node sum of 43 (!). The other two branches, one in the low-energy range with small **extreme** branch $\Theta = 12^\circ$ ($*$), and the branch in the GZK-energy range with small **extreme** branch $\Theta = 2^\circ$ ($*$ but with only 2 nodes) both have **l** branch-walk-pair values within $\pm 25^\circ$ but their **b** branch-walk-pair values are outside the $\pm 25^\circ$ similarity requirement. Their branch Θ ROIs overlap.

- In the **b**Band $[-15 : 0]^\circ$ and the GZK-energy range, there is a branch with a small **extreme** $\Theta = 11^\circ$ ($*$), and a positive result ($\star$).

- In the **b**Band $[15 : 45]^\circ$ and the GZK-energy range, is a branch with a small **extreme** $\Theta = 15^\circ$ ($*$) and an **extreme** node sum of 27 (!).

The uniqueness of the two rather long GZK branches in southern galactic disk **b**Band $[-15:0]^{\circ}$ and the northern **b**Band $[15:45]^{\circ}$, may each be connected to a large-scale galactic source, re-accelerating CR nuclei in that region, or some feature of the supergalactic plane.

Of the ten Pierre Auger Observatory results, four of the branches have galactic longitude differences moderately far apart. There are no reflected Pierre Auger Observatory branches *within* energy ranges about the galactic disk regions. There are also no Pierre Auger Observatory branches with both **l** and **b** branch-walk-pairs/ROIs within $\pm 30^{\circ}$ of the galactic axes **l/b=0**.

Table 7.18: Energy Comparison Table 4. $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$

possible correlating SOUTH AND NORTH branch Θ and branch node values over bBands															
and over galactic longitude regions with $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ **															
bBand°	**40 EeV < E ≤ 50 EeV**					**50 EeV< E ≤ 60 EeV**					**E > 60 EeV**				
	Θ°	**node**	**$\%_S$**	**$(\mathbf{l_1}, \mathbf{l_2})$**	**$(\mathbf{b_1}, \mathbf{b_2})$**	**Θ°**	**node**	**$\%_S$**	**$(\mathbf{l_1}, \mathbf{l_2})$**	**$(\mathbf{b_1}, \mathbf{b_2})$**	**Θ°**	**node**	**$\%_S$**	**$(\mathbf{l_1}, \mathbf{l_2})$**	**$(\mathbf{b_1}, \mathbf{b_2})$**
[−90 : −45]	11 * ♠ 29	23 18	5.04 NA	(17,183) (-159,-48)	(-85,-42) (-82,-44)	7 * ♠ ★	19	< 0.062	(75,296)	(-69,-37)	9 * ♠ ★	21	0.058	(47,334)	(-80,-39)
[−45 : −15]	17 * 12 *	43 ! 9	1.56 2.26	(-131,22) (167,204)	(-50,-8) (-58,-7)						2 *	2	0.13	(169,189)	(-45,-39)
[−15 : 0]											11 * ★	12	< 0.027	(-142,-37)	(-19,0)
[0 : 15]															
[15 : 45]											15 *	27 !	0.11	(-143,15)	(6,50)
[45 : 90]	37	14	NA	(-70,25)	(36,80)										

** data current until 8/11/2012. NA-**Extreme** Θ comparison percentage $\%_S$ not applicable.
* Pierre Auger Observatory **extreme** branch Θ.
! Pierre Auger Observatory **extreme** branch node number.
♠ Pierre Auger Observatory **extreme** branch Θs with similar branch-walk-pairs/ROIs *across* energy ranges.
★ Pierre Auger Observatory **extreme** branch Θ with a positive result.

7.8 Exposure, Positive Θs and Θs from Similar Branches

The flat skymap of Pierre Auger Observatory exposure (over 2.5 cycles in galactic longitude, **l**), in the energy range 12 EeV < E ≤ 15 EeV (see Figure 7.5), serves as a comparison companion to the Pierre Auger Observatory branches of interest which are illustrated in Figure 7.6 on the following page, Figure 7.7 on the next page and Figure 7.8 on page 224. The branches of interest roughly fit the Pierre Auger Observatory's exposure. Inclusion of all the interesting branches of interest was not practical.

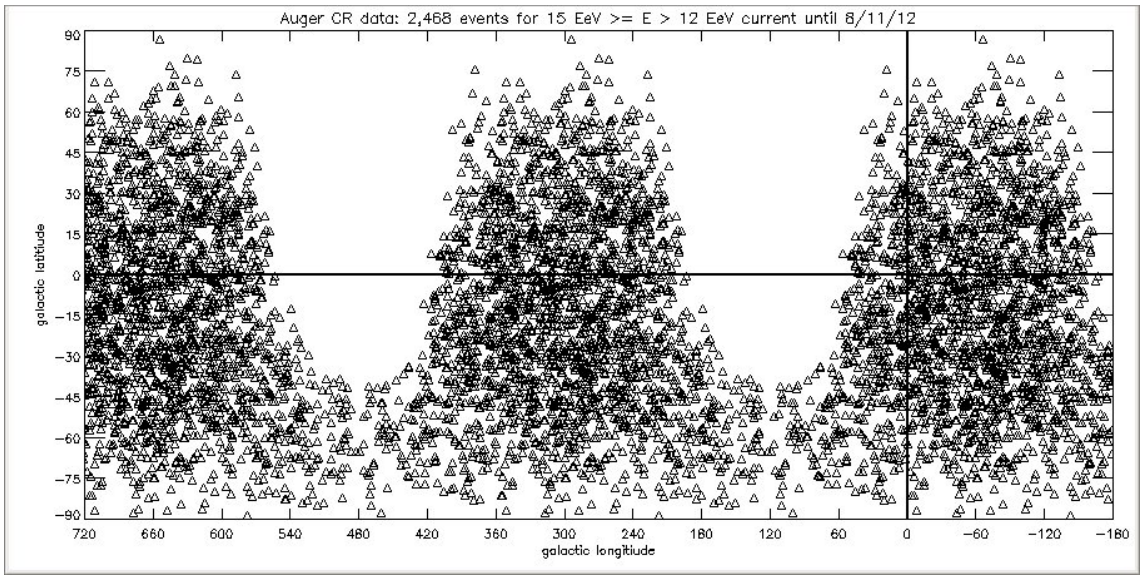

Figure 7.5: Flat skymap of Pierre Auger Observatory data (over 2.5 cycles in galactic longitude, **l**), for CR events 12 EeV < E ≤ 15 EeV (current to 8/11/2012).

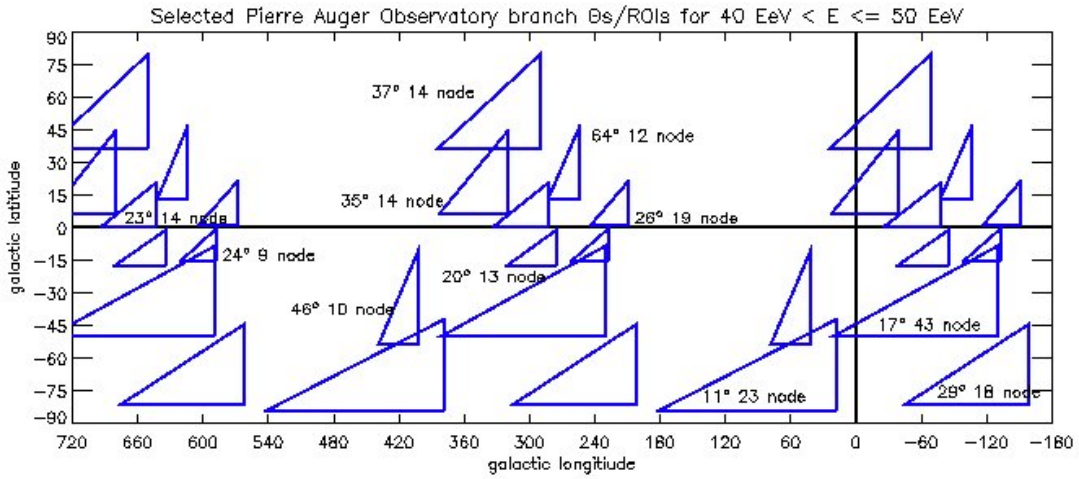

Figure 7.6: Flat skymap of Pierre Auger Observatory selected BBand branches of interest (over 2.5 cycles in galactic longitude, **l**). For 40 EeV < E ≤ 50 EeV. See also page 159.(current to 8/11/2012).

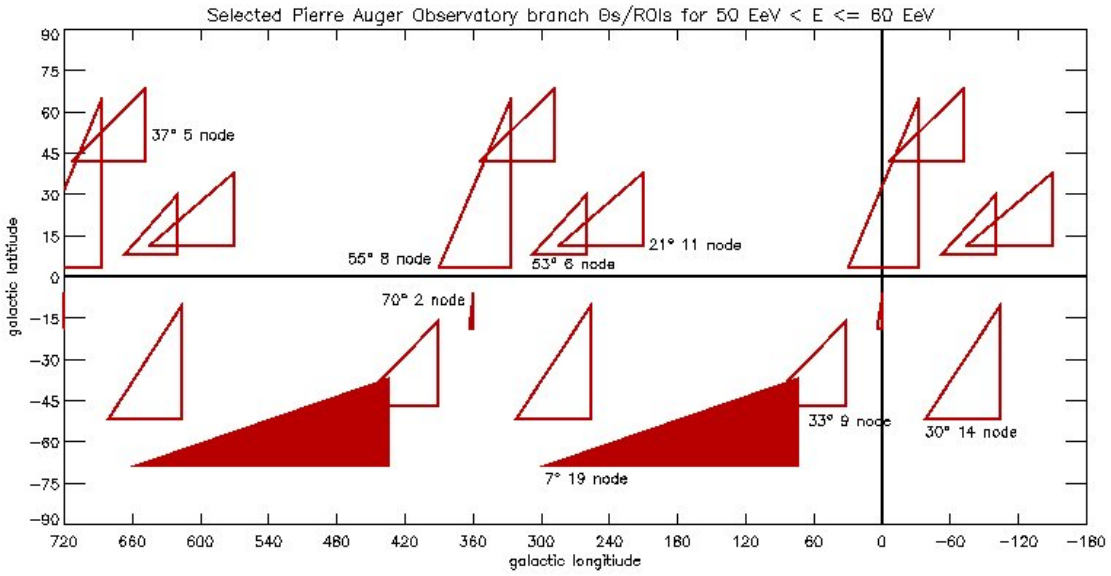

Figure 7.7: Flat skymap of Pierre Auger Observatory selected branches of interest (over 2.5 cycles in galactic longitude, **l**) . For 50 EeV < E ≤ 60 EeV. Filled in symbols are positive results. See also page 159.(current to 8/11/2012).

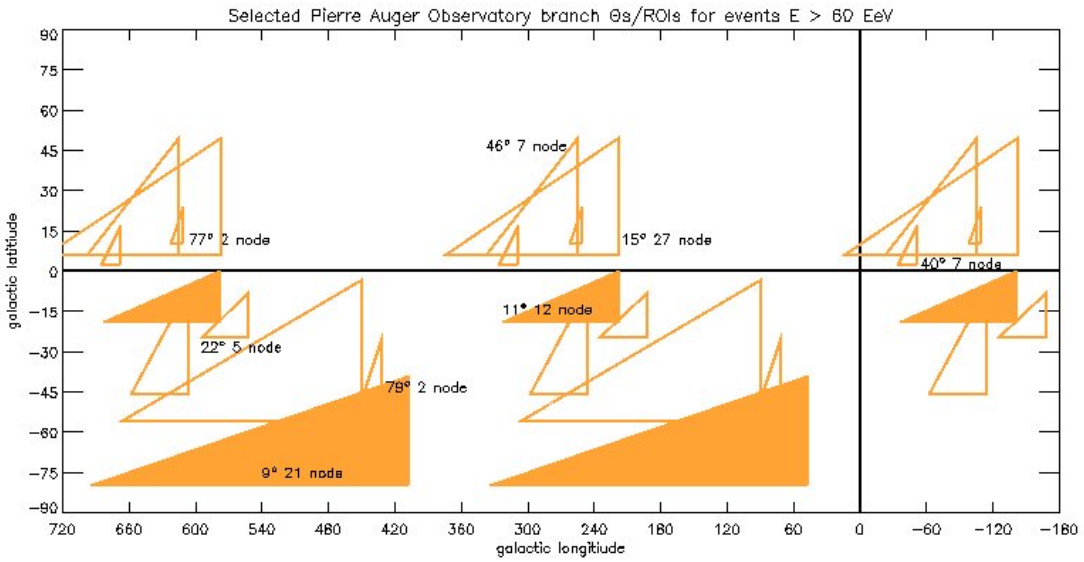

Figure 7.8: Flat skymap of Pierre Auger Observatory selected branches of interest (over 2.5 cycles in galactic longitude, **l**). For E > 60 EeV. Filled in symbols are positive results. See also page 159.
(current to 8/11/2012).

Chapter 8

Conclusion

We have developed a method of analysis of Pierre Auger Observatory CR directions. We associate events by virtue of the angular proximity of their arrival directions, and within three energy ranges around the GZK energy limit. We exploit the gaussian density contours of CR directions and fragment the density contour 'sky' by setting a minimum density boundary value of the accumulated contours to work within. Each fragment captures all of the CR events that lie within that density boundary in that region of sky. A minimum spanning tree is created by growing 'feature' sub-trees within each density boundary fragment and then joining those sub-trees under the same conditions that governed the feature sub-tree's growth. The joining of the sub-trees results in a single minimally connected tree which we term the Yggdrasil. This type of minimum spanning tree algorithm is much faster than more general, previous, algorithms and the algorithm tailors experimental directional errors of CR events to underpin the growth of the entire tree. We can potentially analyse data in sets of ~1,300 events, although the largest Pierre Auger Observatory data set we analysed was in the energy range, 40 EeV < E ≤ 50 EeV, for 131 CR directions in the 'southern' galaxy.

Each Ygg is 'cut' into branches, along prescribed galactic latitude and longitude filters that we felt may yield information about the physical interactions of UHECRs within the filter regions and any astrophysical phenomena peculiar to the filter regions. Each branch has an ensemble of variables, the branch Θ, branch node sum, and the co-ordinates of each galactic **l** and **b** branch-

walk-pair values. We sought patterns in data directions of CRs which may be the consequence of CR interactions with large-scale objects, point-sources or magnetic fields of a galactic/extra-galactic nature. Making inferences about the reasons for values of some branch variables, and seeing patterns linking other branch variables enabled us to do this.

The GMFs are believed to be both azimuthal and poloidal with respect to the galactic disk. So we sought roughly linear structures that are at right angles or parallel to the galactic disk (**extreme** branch Θs). We also looked for evidence of the weak extra-galactic magnetic fields. Depending on its orientation, a large-scale extra-galactic regular magnetic field, albeit weak, would possibly deliver different linear deflections of UHECR light and heavy nuclei. Considering magnetic field deflections of UHECRs, an UHECR travelling a distance, L, at right angles to a regular magnetic field, **B**, has, for light nuclei, a deflection [194]

$$\Delta\theta \approx 13^{\circ}\frac{40\ \text{EeV}}{E/Z}\,\frac{L}{2\ \text{kpc}}\,\frac{|B|}{5\mu\text{G}} \tag{8.1}$$

The deflection of heavier nuclei would be similar, but larger due to the larger value of Z. Poloidal extra-galactic MFs could still deliver small branch Θs and azimuthal extra-galactic MFs could still deliver large branch Θs. Small extra-galactic MFs can deliver heavy CR nuclei or a larger sized extra-galactic MF could deliver light CR nuclei with a wider angle of deflection. Given the variety of physical circumstances which could deliver heavy nuclei with small deflection angles and light nuclei with large deflection angles, we do not speculate whether CR events are from light or heavy nuclei [169].

If we consider large-scale extra-galactic structures and a diffuse spread of sources within such structures, the GZK distance limit of ∼200 Mpc makes for a feasible delivery of UHECRs, even for a very weak extra-galactic magnetic field. Such branch Θ galactic/extra-galactic scenarios are always made with the realization that the linear structure of branches is easily swallowed up, because the branches are determined by galactic **l**Band and **b**Band divisions in the galactic $(\mathbf{l}, \mathbf{b})$ plane. A visual survey of Yggs was also required. The **l** and **b** branch-walk-pair co-ordinate variables are useful in identifying possible regional relationships between different branches *across* and *within* energy ranges.

When we had a positive **extreme** branch Θ or an **extreme** branch node sum, or a combination of two or more branch variables that are considered noteworthy, i.e. tagged, we considered these branches interesting and looked for physical explanations of their interesting status.

Quite a few interesting results were obtained by growing Yggs and comparing Ygg branch data to the branch data of simulated distributions (comparing distributions **Type 1** and distributions **Type 2**). Further interesting results were obtained from comparing galactic **l/b** branch-walk-pair/ROIs values *across* and *within* energy ranges.

Overall, there were three Pierre Auger Observatory branches with positive branch Θ results ($< 0.1\%$), out of a total 67 branches, which we view as possible evidence of regular magnetic fields or sources, whether galactic or extra-galactic. These three positive Pierre Auger Observatory branch Θs are shown in Figure 8.1.

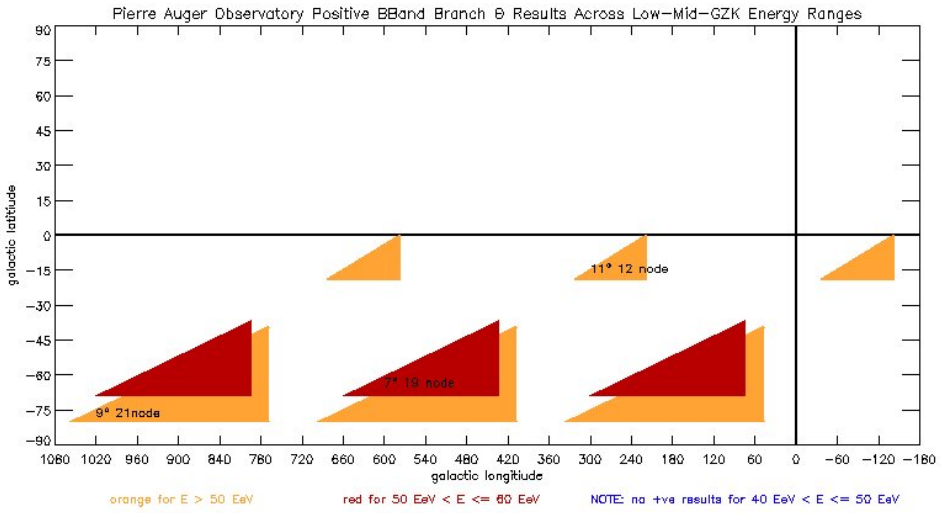

Figure 8.1: Flat skymap of Pierre Auger Observatory branch Θs with a positive result ($< 0.01\%$). For energies 40 EeV $< E \leq$ 50 EeV, 50 EeV $< E \leq$ 60 EeV and $E >$ 60 EeV. Blue for the low-energy range. Red for the mid-energy range. Orange for the GZK-energy range. See also page 159. (current to 8/11/2012).

All of the positive branch Θs were small **extreme** branch Θs, and one may speculate be tentative evidence of a poloidal GMF or a poloidal extra-galactic MF with respect to our galactic disk, involved in the deflection of light or heavy nuclei dependent on the magnitude of the extra-galactic MF.

The nearest large-scale object within the GZK-energy range is part of the super-galactic plane and an extra-galactic magnetic field may well equally explain these small **extreme** branch Θs, particularly GZK CR nuclei. There is also a possibility of UHECR nuclei being somehow galactically re-accelerated up into the GZK regime. If the γ-ray Fermi bubbles (there has been detection of a partially ordered magnetic field in the region of the Fermi bubbles [217]) so closely associated with the south and north galactic lobes are moving, and the bubbles traverse a region of space in an acceptable astrophysically modelled span of time, they may be such re-acceleration sites, and the branch Θs would not necessarily be **extreme**. The collimated magnetized ridges with their magnetic fields of up to $\sim$15 μG winding about the northern and southern galactic radio lobes may be GMF re-acceleration sites. There is speculation these winding ridges are the result of helical fields produced by the dynamo activity of the galactic centre [26].

Two of these three positive, small **extreme** branch Θ results have a similar **l/b** branch-walk-pair values/ROIs with another branch *across* energy ranges.

Two of the positive, small **extreme** Θ results are in the inter-spiral arms, similar to one another (ROIs) and another **extreme** branch Θ/ROIs. These three similar small **extreme** Θ branches also occupy the **l**Band filter of $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ (see Table 7.18 on page 221) and have large **l**Walk differences. The galactic inter-spiral arms are supposed to have uniform GMFs in the galactic disk **b**Band regions. From low-mid-GZK-energy ranges, we have branches with small **extreme** $\Theta = 11°, 7°$ and $9°$. We can suppose that these three branches have been composed of CR nuclei deflected by galactic/extra-galactic poloidal MFs as they occupy the **b**Band $[-90 : -45]°$.

A third positive, small **extreme** branch Θ, occupies the galactic **l**Band filter of $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$, in the **b**Band $[-15 : 0]°$, and is in the GZK-energy range. It has no similar Pierre Auger Observatory branches. This branch would fit the galactic/extra-galactic poloidal field scenario of the deflection of CR nuclei.

In the **spiral** arms (chaotic MFs), there are only three sets of similar Pierre Auger Observatory branches, but there are four branches with **extreme** branch node **sums** and one of the four does have a similar branch in another energy range. The Pierre Auger Observatory branches occupying the spiral arms are in total sixteen, and nine of those branches are for non-extreme branches with

$\Theta > 30°$. The large non-extreme branch Θs, lend support to the belief the galactic spiral arm MFs are chaotic.

There are three sets of Pierre Auger Observatory branches within the GZK-energy range in the **inter-spiral** arms and reflected about the galactic **b**Band disk regions of $[-15 : 0]°$ and $(0 : 15]°$. In the case of reflections, a set is composed of two branches, although sometimes there can be an ambiguity when the $\pm 25°$ caveat for **l/ b**Walk branch-pair-values is valid for more than two branches being part of a reflected set, and we need more information to decide which branch is a part of the reflected branch set of two. Two sets of these **inter-spiral** branches are mirror reflections and the other set is a diagonal reflection. All the branch Θ values of the three sets are non-extreme. One of the two sets of mirror reflections have identical (integer rounded off) branch Θ values, namely one set of two branches has both branches with $\Theta = 33°$. In addition, the GZK **extreme** $\Theta = 33°$ branch is similar to a branch in the low-energy range which has a branch $\Theta = 34°$. This set is mirror reflected with positive **l** branch-walk-pair values. The other set of mirror reflections has negative **l** branch-walk-pair values. A conjecture about the first set of GZK mirror reflections **l/b** in a region of space where the GMF is assumed regular, is that any regular GMF is non-existent or weak or chaotic when both the **l/b** branch-walk-pair values are relatively close to the GC. It may be that the GC has powerful magnetic fields of its own and this has more to do with CR nuclei deflections. If the GZK branches are galactic, they must also be re-accelerated by a moving GMF. The Fermi-bubbles in the radio outflows must move over time. The radio outflows are too energetic not to influence the Fermi-bubbles. Also, it has been assumed the Fermi-bubbles detach from the central region of the galaxy through differential rotation [23]. An extra-galactic location of the GZK nuclei could still be feasible no matter what the orientation the extra-galactic field. The GC could still reasonably play a part in the deflections of the extra-galactic GZK CRs. The second set of mirror reflections have the **b** branch-walk-pair co-ordinate value close to the GC, but their **l** branch-walk-pair co-ordinate values place these branches much further off the GC. The respective branch node sums are the minimum sum required to make a branch. Still, the branch in the galactic **b**Band $[-15 : 0]°$ has a similar branch *across* the low-energy range, albeit with a branch node sum of 3. Regarding the diagonal reflections of branches

about the galactic plane, both have non-extreme branches with $\Theta = 54^\circ$ values and both their **l** and **b** branch-walk-pair values are within $\pm 27^\circ$ of the galactic axes **l/b=0**. The diagonal nature of these galactic disk reflections could possibly suggest a connection to the disk azimuthal MFs, which change rotation across the galactic plane.

An extra-galactic source scenario comprising regular MFs in parts of the supergalactic plane, or the diffuse spread of Seyfert galaxies within the supergalactic plane, both providing heavy UHECRs [61, 11], may explain some of the GZK results obtained by this method when applied to Pierre Auger Observatory events. A connection beween the acceleration/re-acceleration and supply of UHECRs situated in the 'southern' and 'northern' galactic lobes cannot as yet be ruled out, nor can the role of the GC, because we have much to understand about the properties of these large-scale objects and their effects on UHECRs.

These are some of the most conservative co-incident branches found by the Ygg. A cross-correlation with large-scale-object catalogues may provide some clarification either way on the usefulness of this technique for some of the non-positive CR event branches which we have nominated as interesting. The increasing energies of the engineered collisions beween both light and heavy nuclei at CERN will give us a better idea of whether the UHECR nuclei are light or heavy.

Cen A, with galactic $(\mathbf{l}, \mathbf{b})$ co-ordinates$\sim$[-50.5,19.4]$^\circ$, has been identified in a number of studies [109, 194, 168] as a region of anisotropy, or as a translated region of anisotropy of the LSS of the Virgo cluster [168]. Table 8.1 lists four Pierre Auger Observatory branch-walk-pairs/ROIs subtending the galactic Cen A co-ordinates.

Table 8.1: Cen A$\sim [-50.5, 19.4]^\circ$ Branches

EeV	Θ°	nodes	$\%_S$	Tags	$(\mathbf{l_1}, \mathbf{l_2})^\circ$	$(\mathbf{b_1}, \mathbf{b_2})^\circ$	**Band**$^\circ$
$40 < E \leq 50$	23	14	NA	$!, \heartsuit M$	(-78,-25)	(0,21)	**b**Band [0 : 15]
$50 < E \leq 60$	21	11	2.21	$*, !$	(-150,74)	(11,38)	**b**Band [15 : 45]
	63	6	NA	$\surd$	(-72,-42)	(14,56)	**b**Band 15 : 45]
$E > 60$	15	27	0.11	$*, !$	(-143,15)	(6,50)	**b**Band [15 : 45]

* Pierre Auger Observatory **extreme** branch Θ.
! Pierre Auger Observatory **extreme** branch node sum.
$\surd$ similar non-extreme branches *across* energy ranges.
$\heartsuit M$ Pierre Auger Observatory similar mirror reflected branch-walk-pair values
within the same energy range, and about galactic **b**= 0° *within* galactic disk **b**Band divisions of ([-15:0],(0:15])$^\circ$.

Table 8.2: Positive Result Branches and their Interesting Similar Branches

EeV	Θ°	nodes	$\%_S$	Tags	$(\mathbf{l_1}, \mathbf{l_2})^\circ$	$(\mathbf{b_1}, \mathbf{b_2})^\circ$	**Band**°
$40 < E \leq 50$	11	23	3.17	$*, \spadesuit$	(17,183)	(-85,-42)	**b**Band[-90:-45]
$50 < E \leq 60$	7	19	< 0.062	$*, \spadesuit, \bigstar$	(75,297)	(-69,-37)	**b**Band[-90:-45]
$E > 60$	9	21	0.058	$*, \spadesuit, \bigstar$	(47,334)	(-80,-39)	**b**Band[-90:-45]
$E > 60$	11	12	< 0.027	$*, !, \bigstar$	(-142,-37)	(-19,0)	**b**Band[-15:0]

* Pierre Auger Observatory **extreme** branch Θ.
! Pierre Auger Observatory **extreme** branch node number.
♣, ♠ Pierre Auger Observatory **extreme** branch Θs with similar branch-walk-pair/ROIs co-ordinates *across* energy ranges.
within galactic disk **b**Band divisions of $([-15:0], (0:15])^\circ$.
★ Pierre Auger Observatory **extreme** branch Θ with a positive result.

Table 8.2 lists the Pierre Auger Observatory branches with positive results, and those branches whose similarities to the positive result branch co-ordinates/ROIs *across* energy ranges we think make them significant. At this early stage of understanding the role of galactic lobes and Fermi bubbles we don't include as significant those Pierre Auger Observatory branches which are reflected (M) about galactic $\mathbf{b} = \mathbf{0}^\circ$ to positive branch Θ co-ordinate results.

In the **b**Band divisions in the plane, for the **b**Band $[-90 : -45]^\circ$ we have two positive, overlapping ROIs results. One positive result is $< 0.062\%$ for a small **extreme** branch $\Theta = 7^\circ$ with 19 nodes in the mid-energy range ($*, \spadesuit, \bigstar$). The other positive result is $< 0.058\%$ for a small **extreme** branch $\Theta = 9^\circ$ with 21 nodes in the GZK-energy range ($*, \spadesuit, \bigstar$). These two overlapping ROIs, positive results, lend support to the claim that another overlapping ROI, small **extreme** branch $\Theta = 11^\circ$ with 23 nodes in the low-energy range be considered significant ($*, \spadesuit$). See Figure 6.8 on page 160.

Another branch in the **b**Band divisions in the plane, for the **b**Band $[-15 : 0]^\circ$ has a positive result. This branch is unique and the positive result is $< 0.027\%$ for a small **extreme** branch $\Theta = 11^\circ$ with 12 nodes in the GZK-energy range ($*, !, \bigstar$).

We have three small **extreme** branch Θ positive results for the BBand divisions in the plane, with supporting evidence for one other branch with a small Θs. Two of the positive result branches have a branch in the GZK-energy range and another branch in a non GZK-energy range, which might imply the involvement of a poloidal moving GMF. The final singular, small **extreme** branch

Θ positive result lies in the GZK-energy range and is for a fairly long branch across its galactic longitudes $\in [-142, -37]$. This result may be from a poloidal extra-galactic source or cause.

Regarding the simulated distribution statistics quoted, we exercise caution for a variety of reasons. There were no quantative statistics provided as to the statistical significances of similar Pierre Auger Observatory branch-walk-pairs/ROIs *across* energy ranges and these 'similarities' and tags should be regarded as qualitative. There may be a small proportion of simulated distribution branch Θs that are in error when a branch crosses over from $\mathbf{l} \in -180°$ to $\mathbf{l} \in 180°$. We are not sure. Certainly the Pierre Auger Observatory branch Θs are correct (we were able to check them with the aid of their branch plots), but our algorithm may not have covered all the different cases for branch-walk-pair co-ordinates when we have such a galactic longitude $-180°$ crossover to $180°$. Also, there should have been a statistical penalty enacted when we changed from our single 'sky' Yggdrasils to our separate 'north' and 'south' Yggdrasils using the same set of simulated distributions, and **b**Band divisions in the plane, excepting of course the recast **b**Band division of $[-15 : 15]°$ which were separated in **b**Band divisions of $[-15 : 0]°$ and $[0 : 15]°$. This involved tuning our data by acknowleding that the recently discovered radio emission lobes (see section 6.2 on page 154 and section 6.4 on page 164) might have a bearing on how we divide our galactic **(l/b)** sky. This MST method is better viewed as a preliminary branch-walk-pair/ROI tuning study wherein we choose certain selection cuts(filters) of the Yggdrasil to test galactic/extra-galactic magnetic field and CR source models with a view to future *a priori* searches.

A useful extension of this work would be to take purely **l**Band divisions in the galactic 'north' and 'south' **(l/b)** plane. These **l**Band divisions could likely be $\mathbf{l} \in ([-135 : -45], [-45 : 45], [45 : 135], [135 : 225])°$. The central **l**Band division of $[-45 : 45]°$ covers the angular extent of the north and south galactic radio lobes and we may learn more of their behaviours if symmetries involving branch-walk-pairs/ROIs are found.

Appendix A

South and North 40 EeV < E ≤ 50 EeV:

A.1 Yggdrasil Equatorial (RA,dec) Co-ordinates

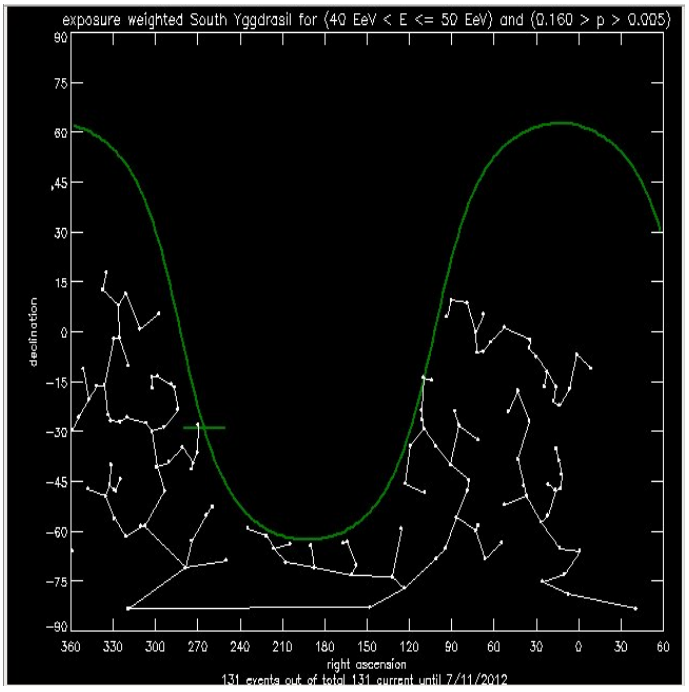

Figure A.1: South Yggdrasil of all events in **RA** vs **dec** for 40 EeV< E ≤ 50 EeV (current to 8/11/2012).

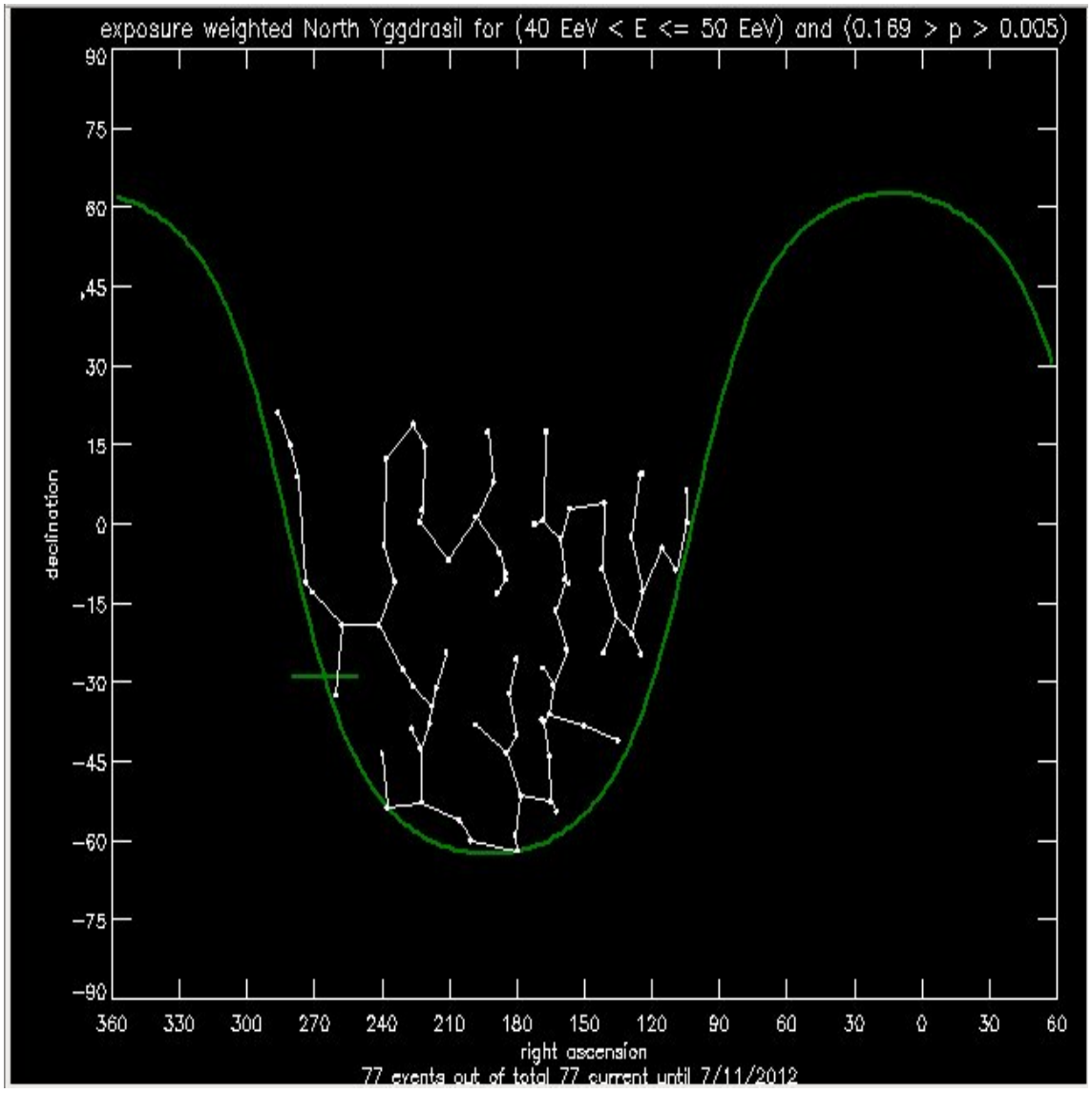

Figure A.2: North Yggdrasil of all events in **RA** vs **dec** for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

A.2 Yggdrasil Galactic (l,b) Co-ordinates.

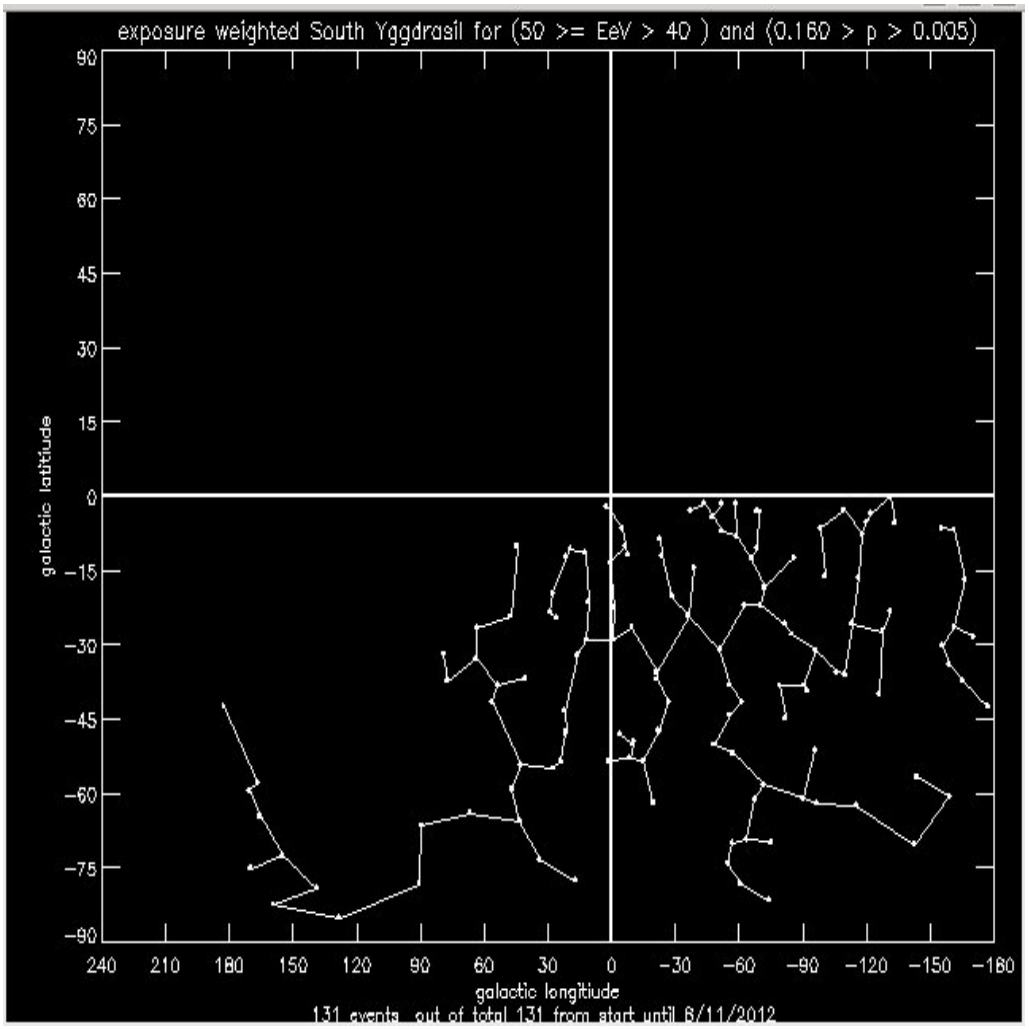

Figure A.3: South Yggdrasil of all events in **l** vs **b** for 40 EeV< E ≤ 50 EeV (current to 8/11/2012).

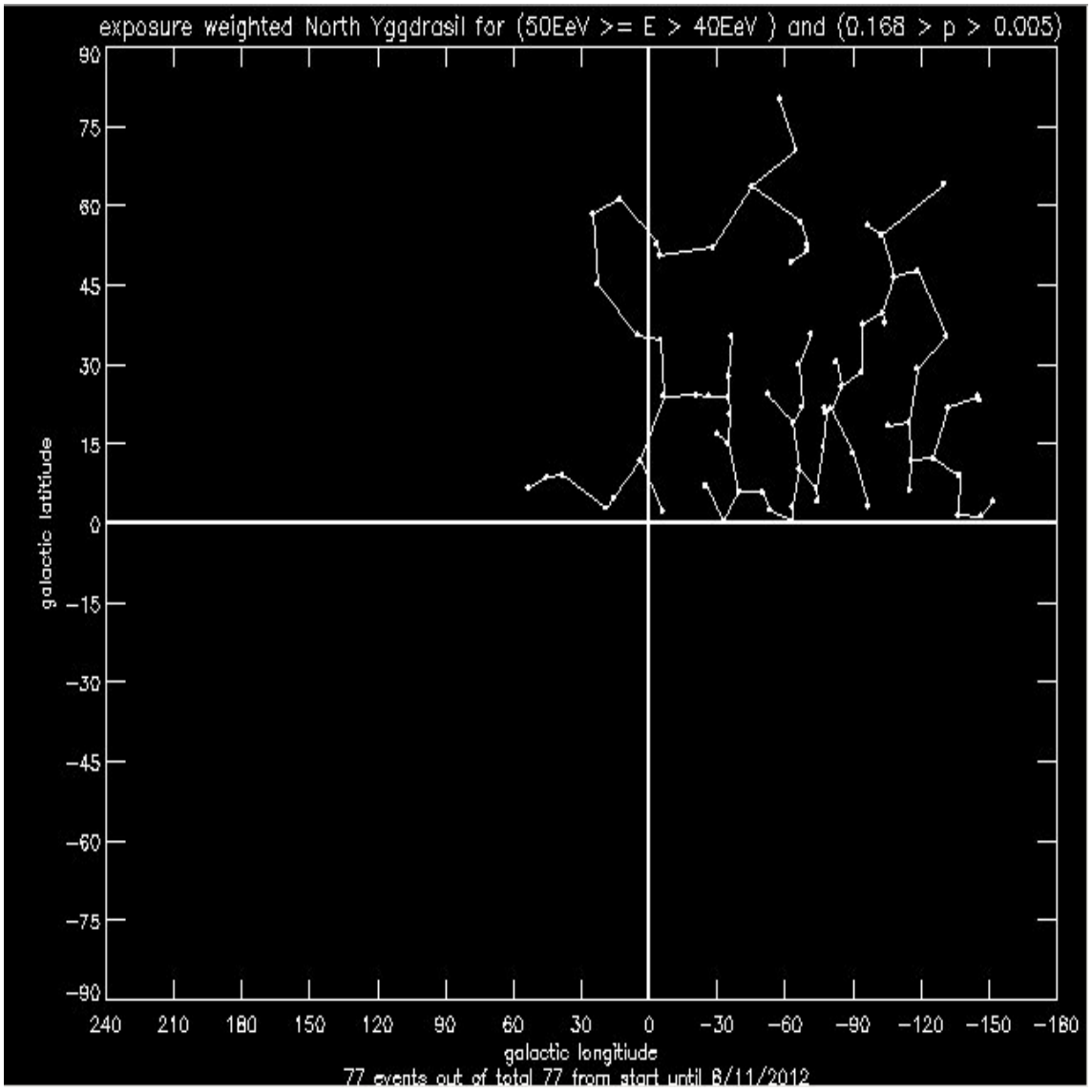

Figure A.4: North Yggdrasil of all events in **l** vs **b** for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

A.3 bBands

The Yggs of Pierre Auger data have been divided into branches along latitude cuts. These branches are seen in Figure A.5 on the following page and Figure A.6 on page 238 , and are followed by 3 lots of distributions flagged with Pierre Auger branch data values.

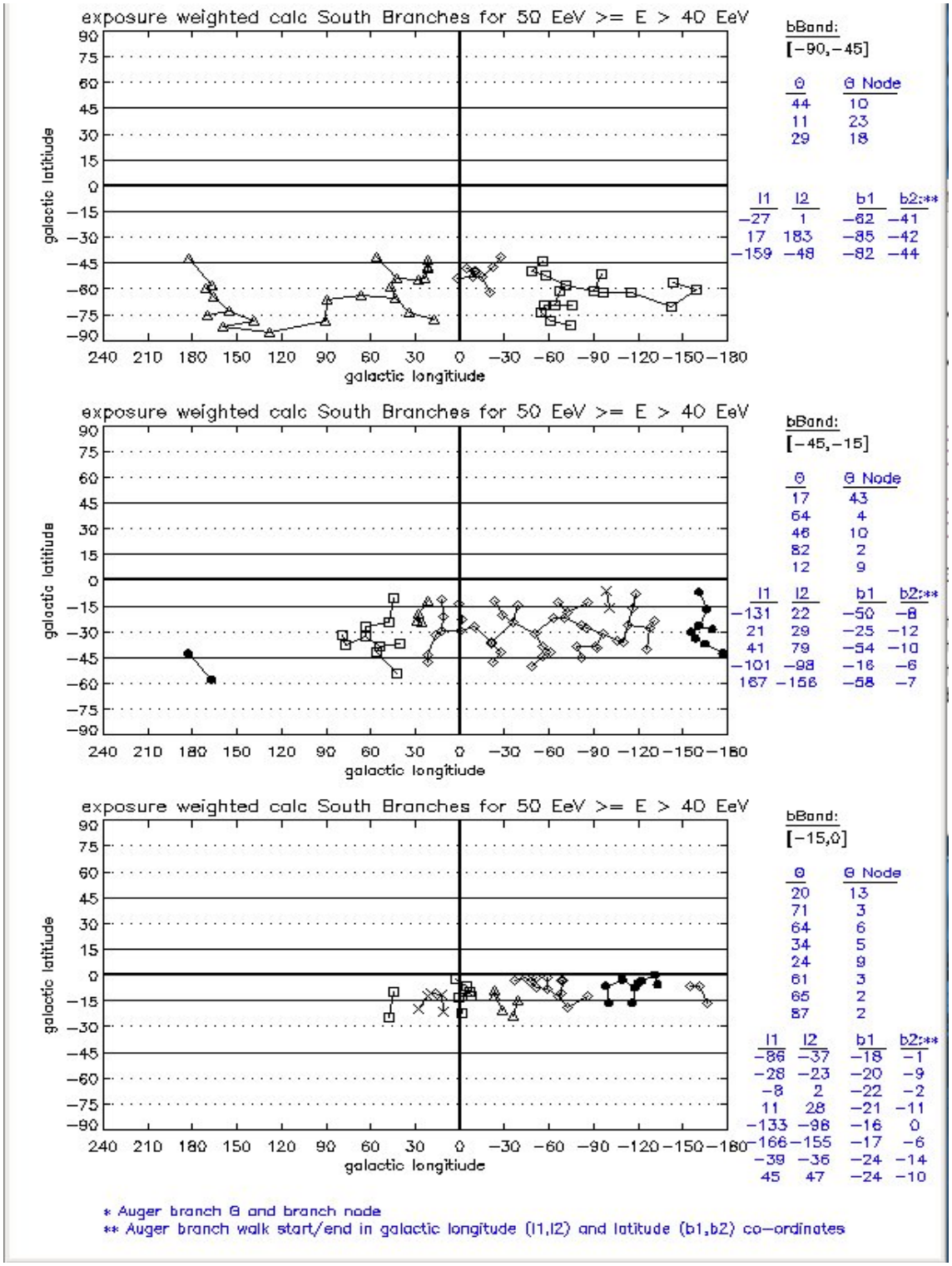

Figure A.5: .
South **b**Band branches for 40 EeV < E ≤ 50 EeV (Each branch denoted by a different symbol). (current to 8/11/2012)

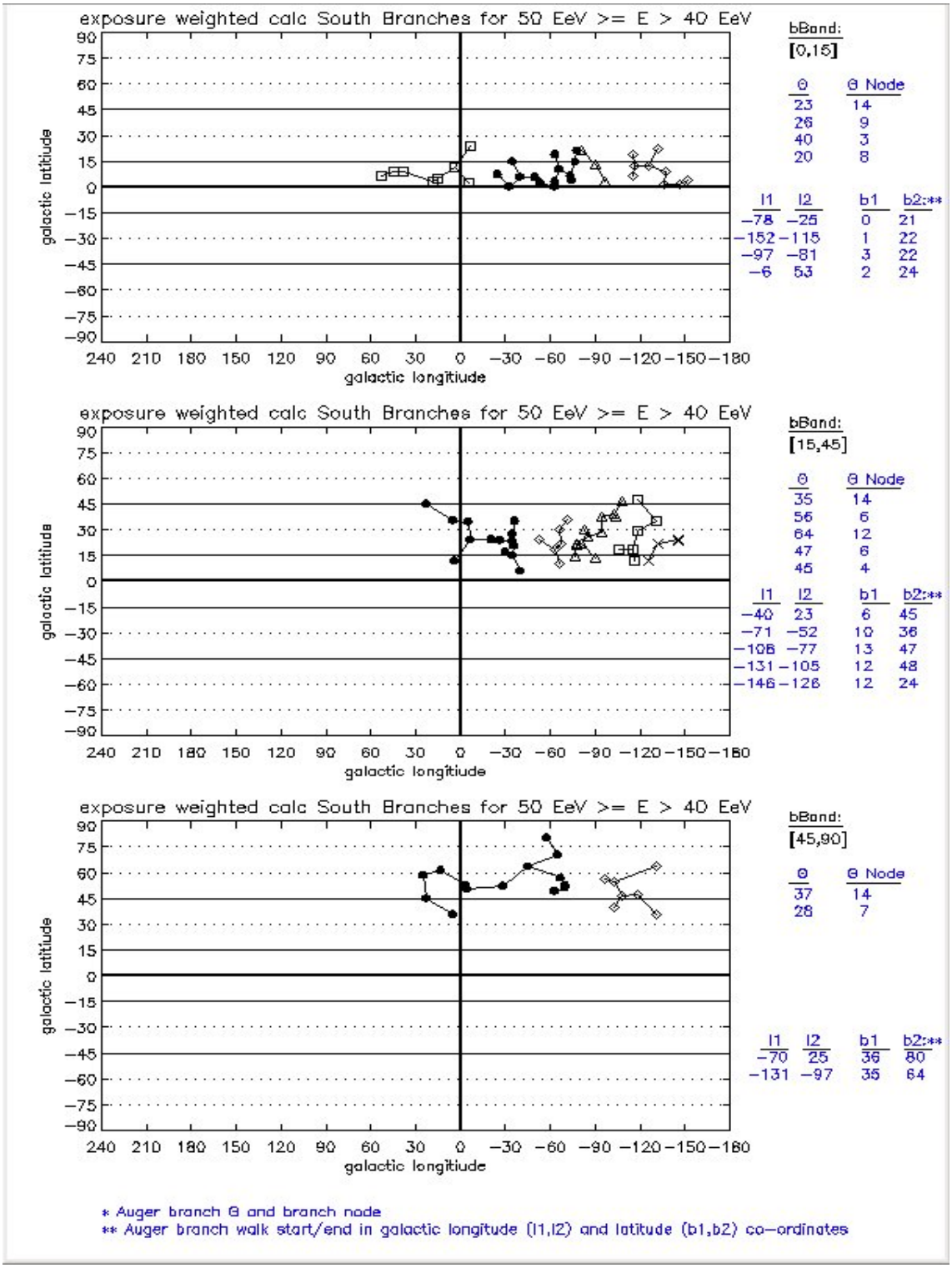

Figure A.6: North **b**Band branches for 40 EeV < E ≤ 50 EeV Each branch denoted by a different symbol. (current to 8/11/2012).

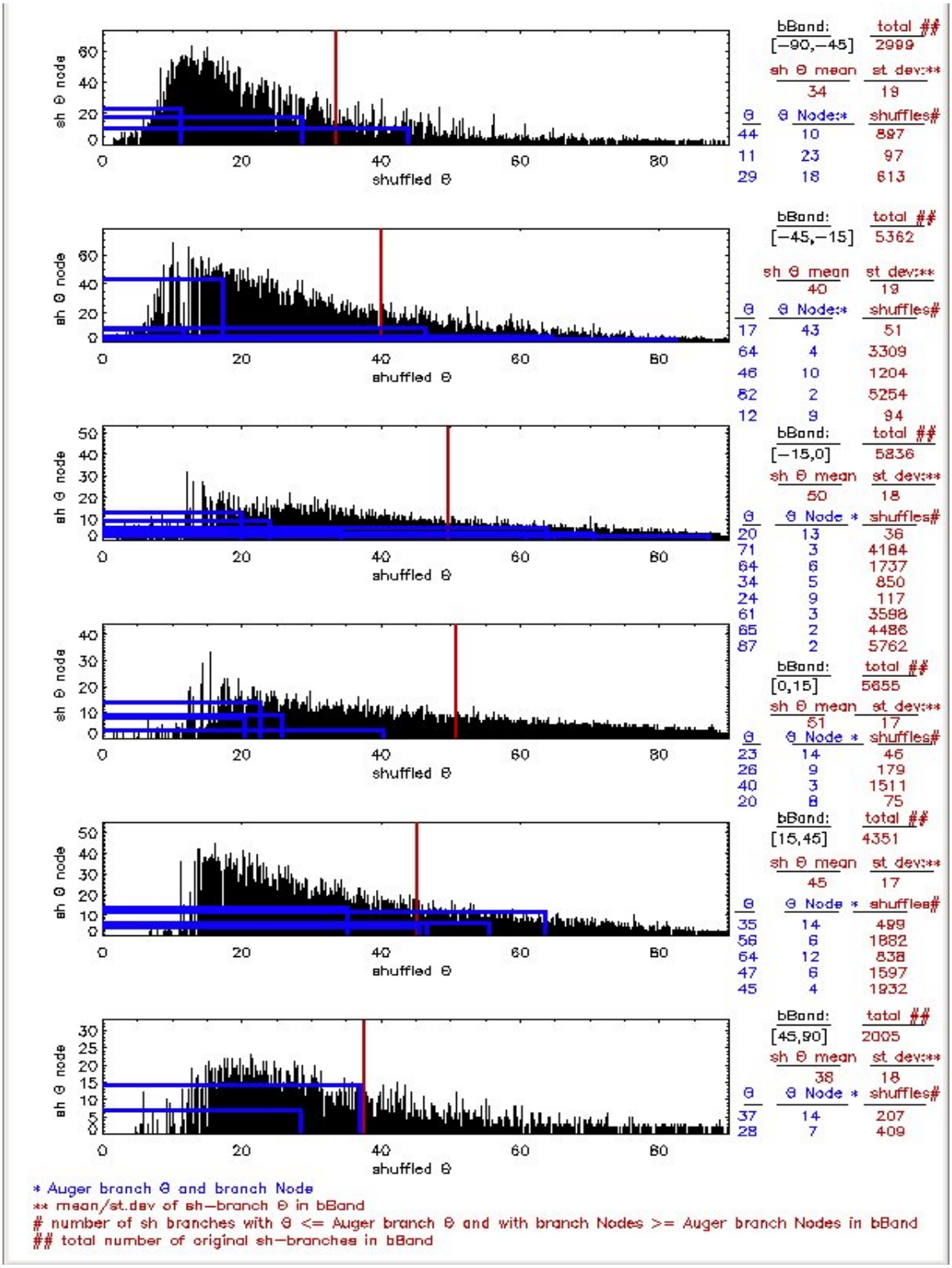

Figure A.7: **b**Band shuffle Θ node vs Θ for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

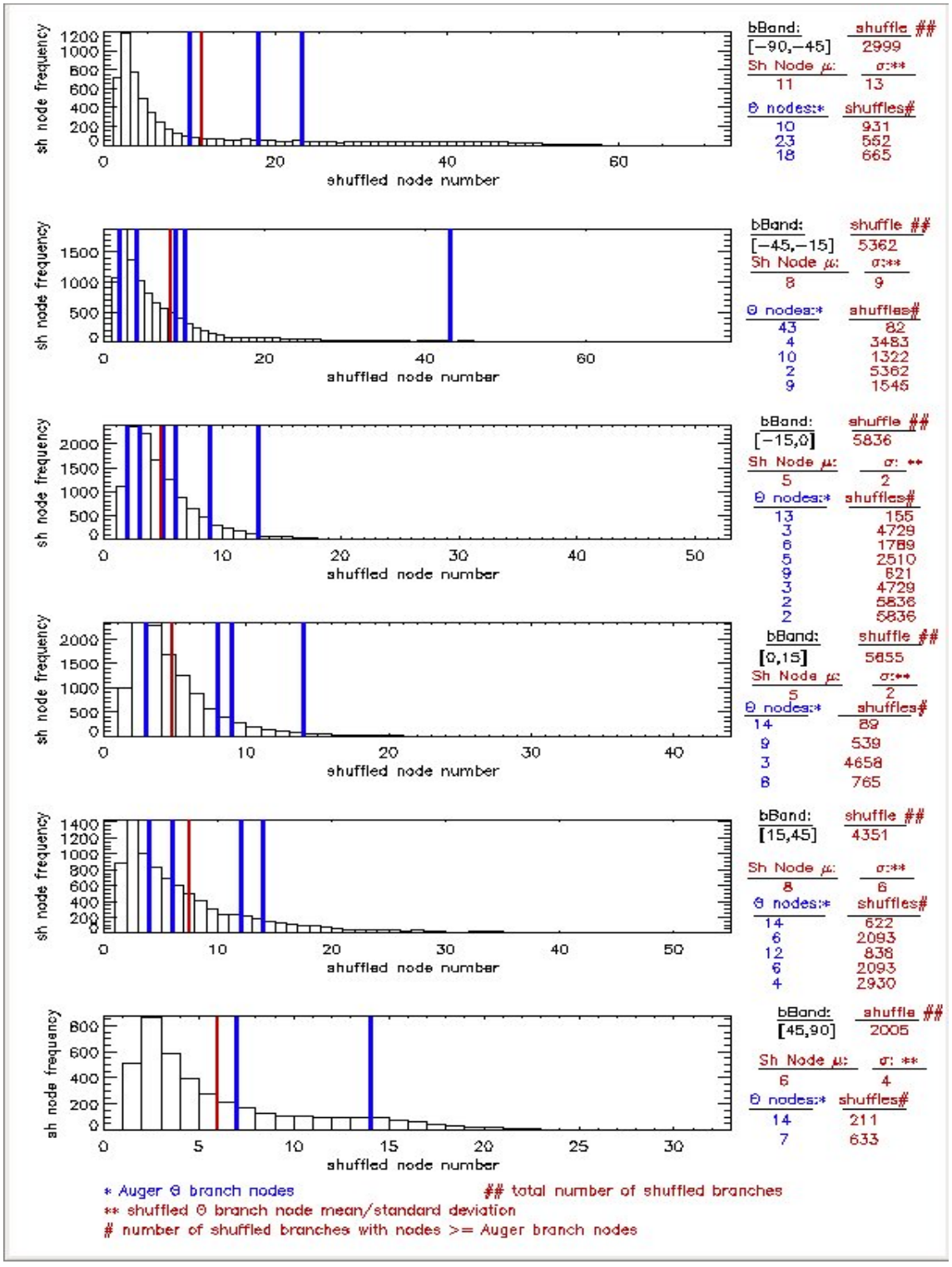

Figure A.8: **b**Band shuffle node frequency vs nodes for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

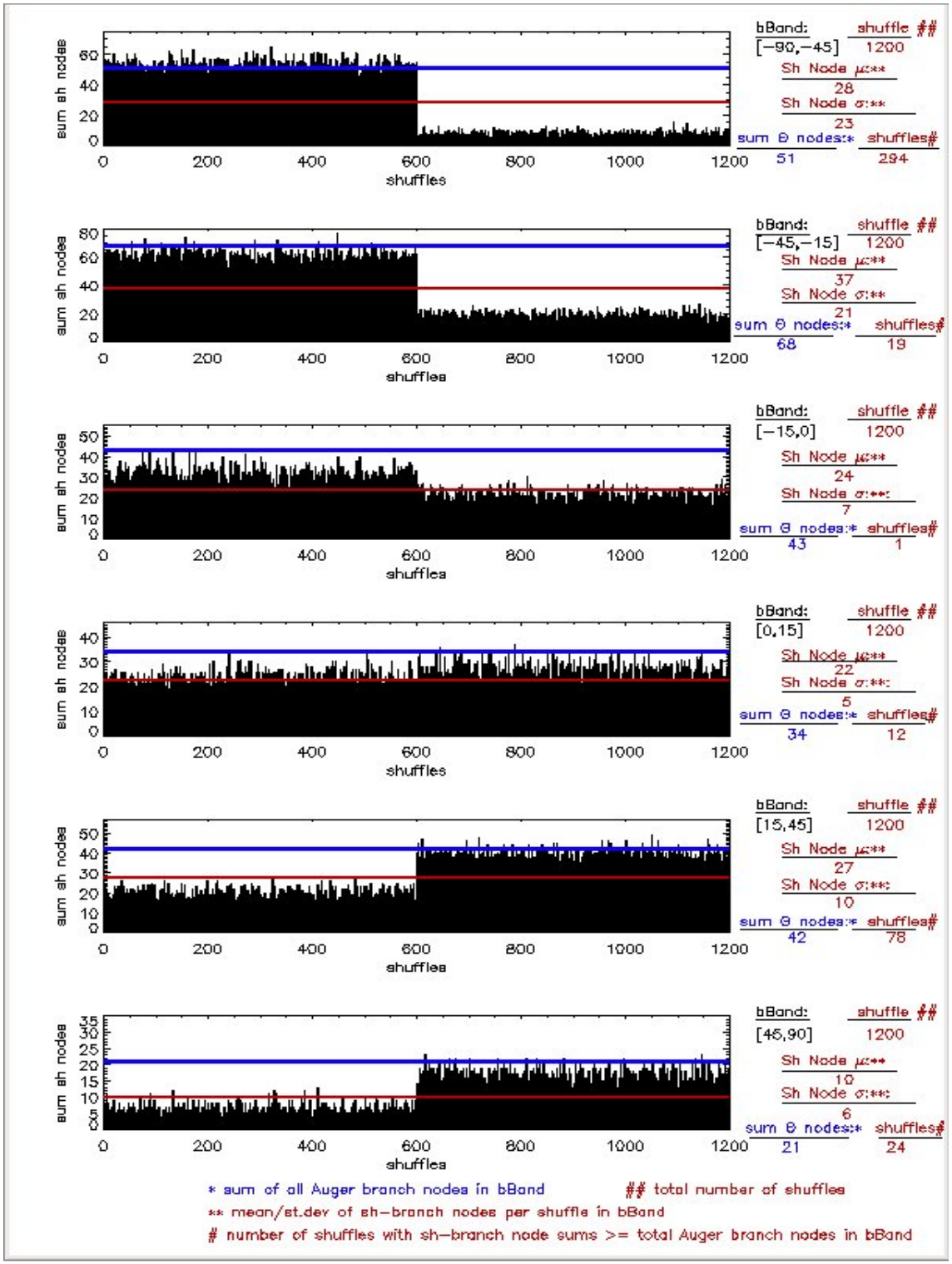

Figure A.9: **b**Band shuffle nodes vs shuffles for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

A.4 Galactic Longitude Quadrants and bBands

bBands ∈ [−135 : 135] **and** ∈ [−45 : 45]

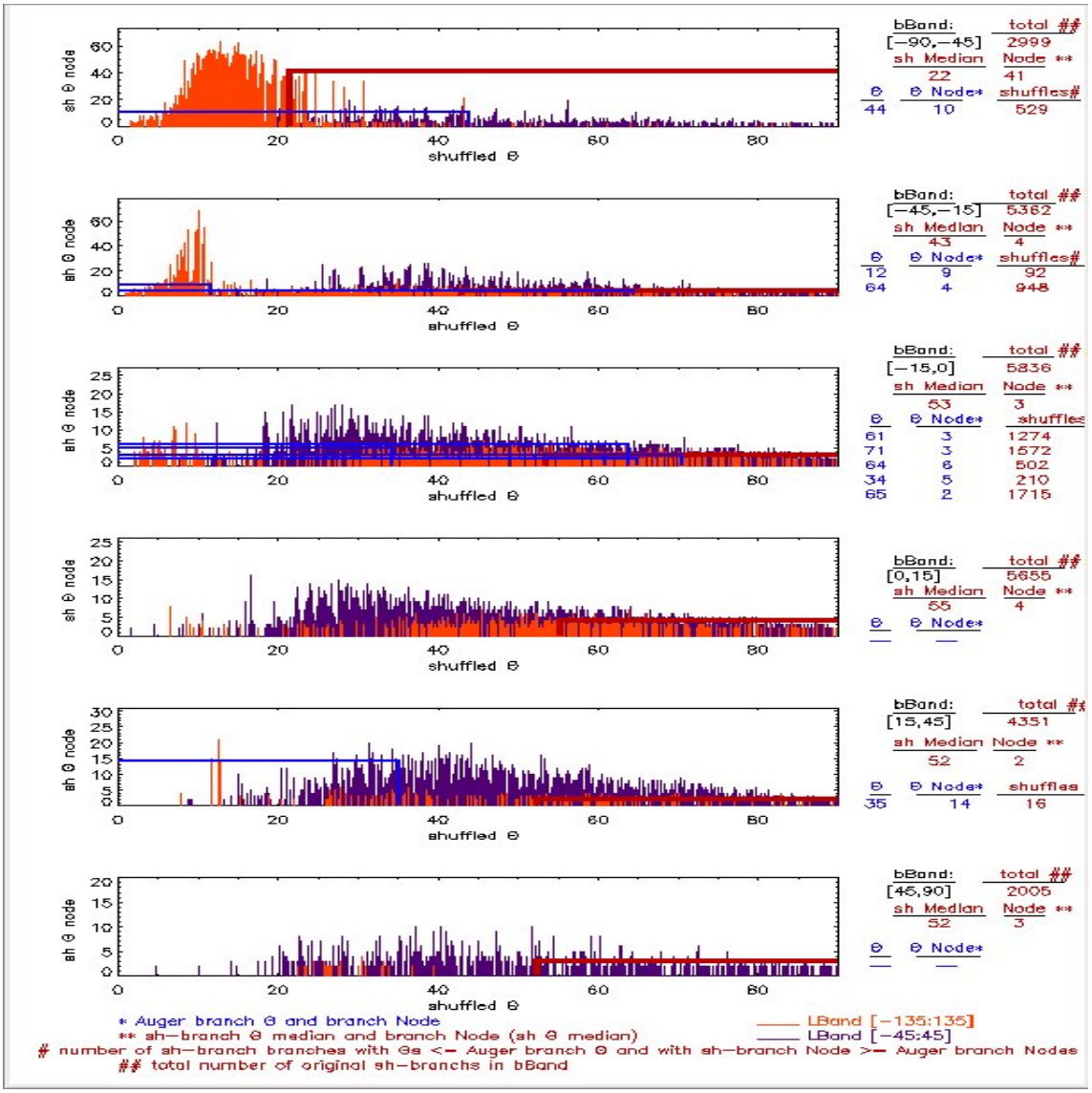

Figure A.10: **b**Band shuffle Θ nodes vs Θ where Gal l ∈ ([−45 : 45], [−135 : 135])° for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

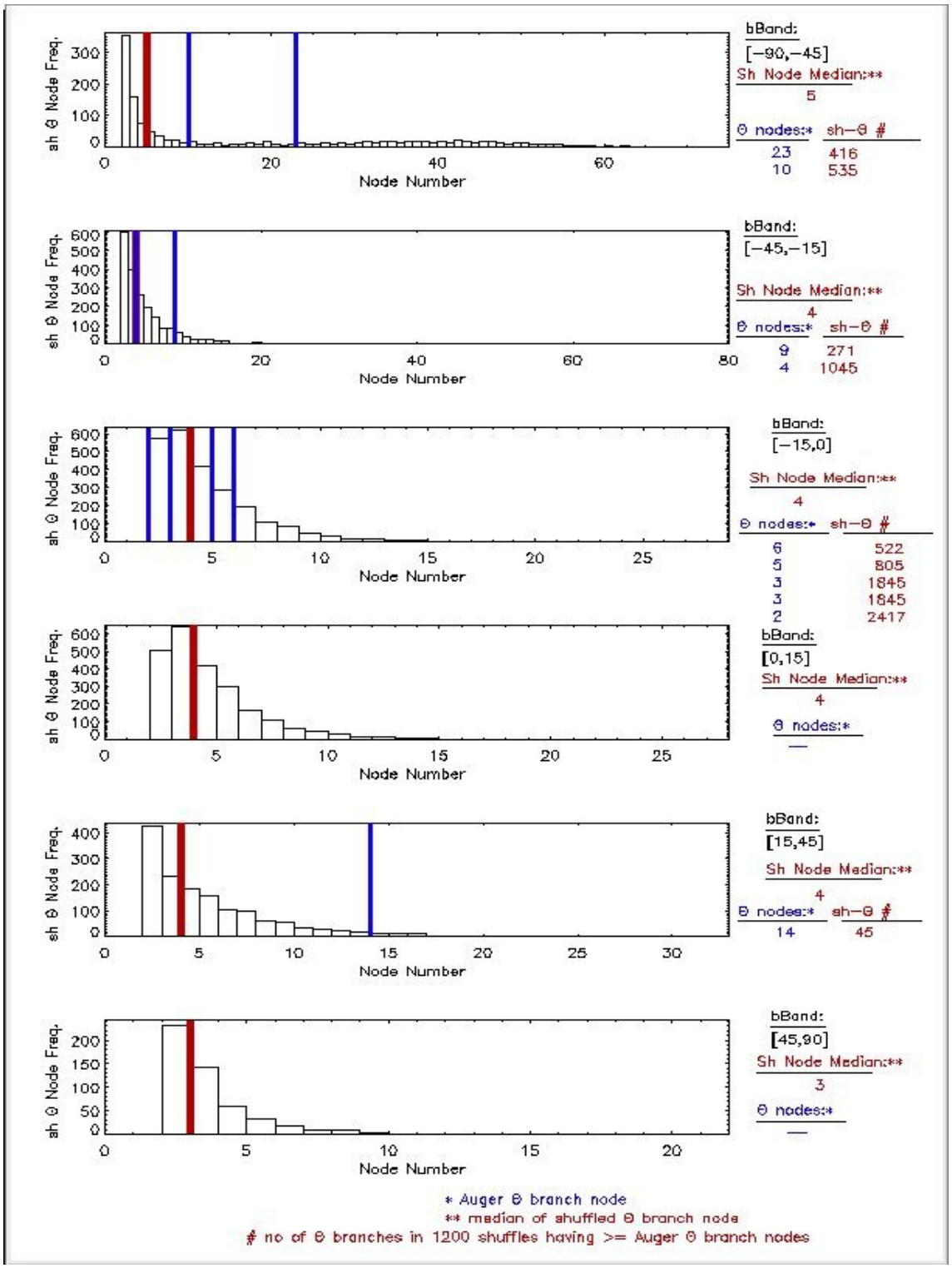

Figure A.11: **b**Band shuffle node frequency vs shuffled node number where Gal l ∈ ([−45 : 45], [−135 : 135])° for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

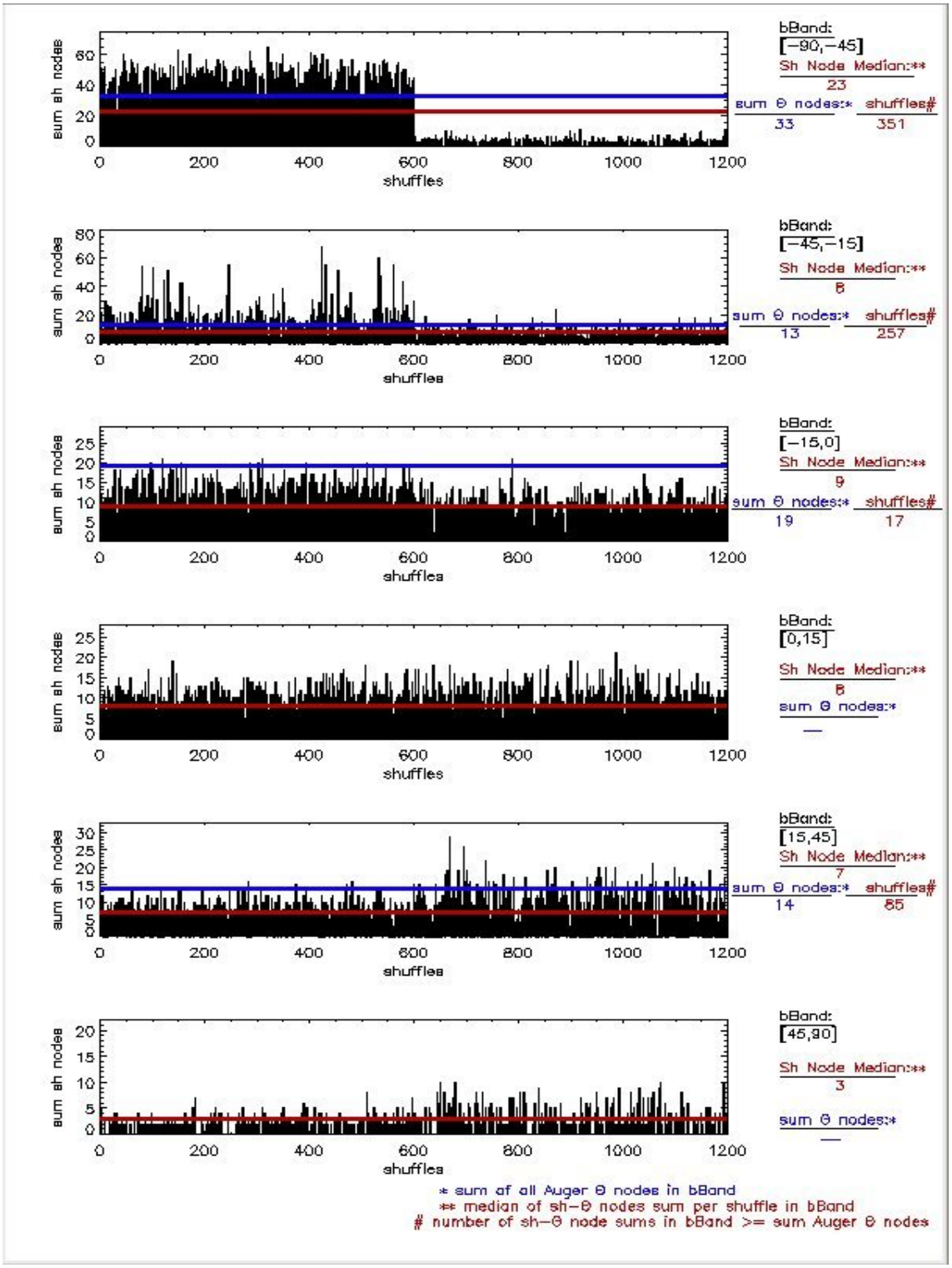

Figure A.12: **b**Band shuffle nodes vs shuffles where Gal $l \in ([-45 : 45], [-135 : 135])^\circ$ for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

bBands ∈ [−135 : −45]° **and** ∈ [45 : 135]°

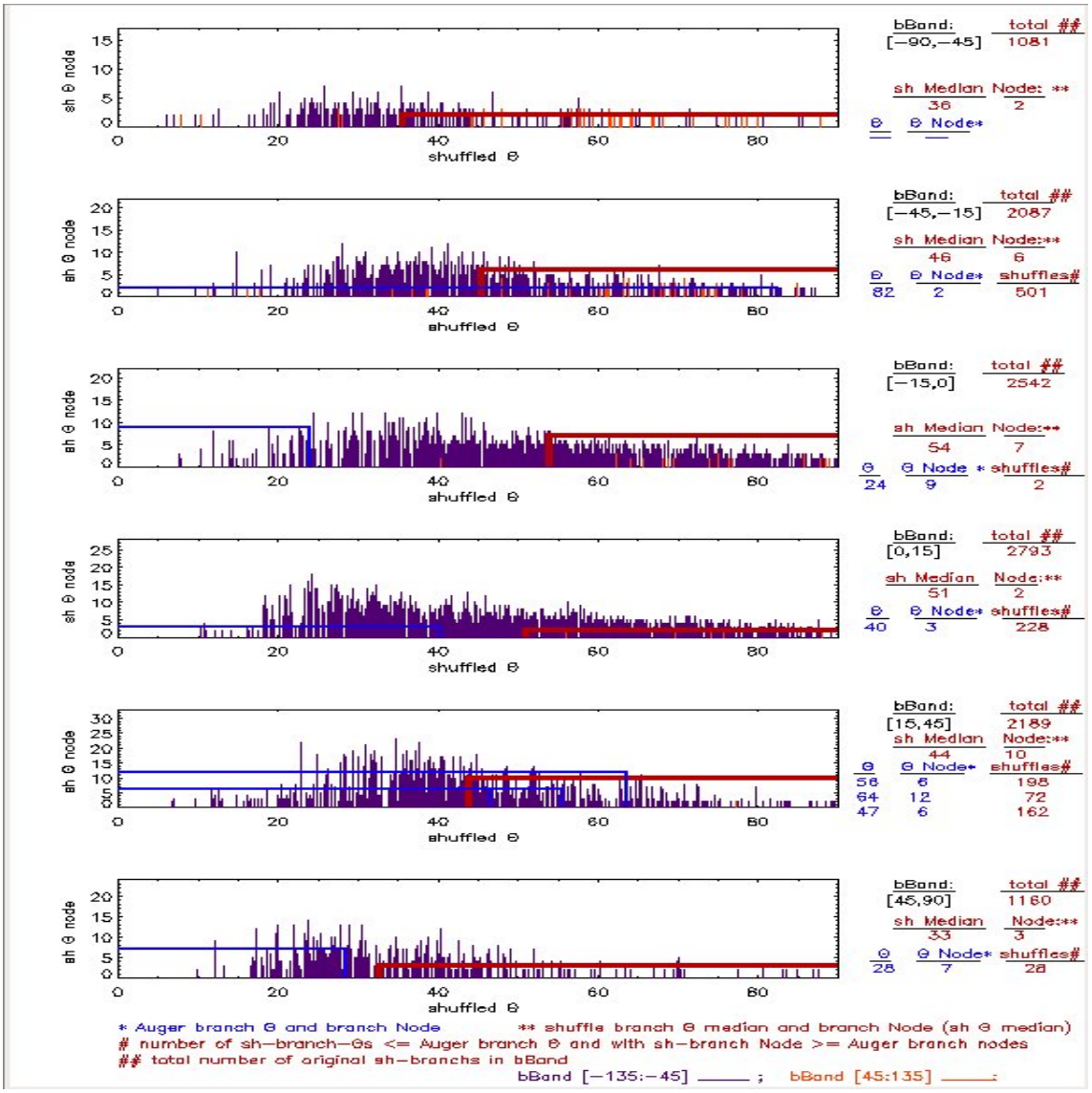

Figure A.13: **b**Band shuffle Θ nodes vs shuffled Θ where Gal l ∈ ([−45 : −135], [45 : 135])° for 40 EeV < E ≤ 50 EeV (current to 8/11/2013).

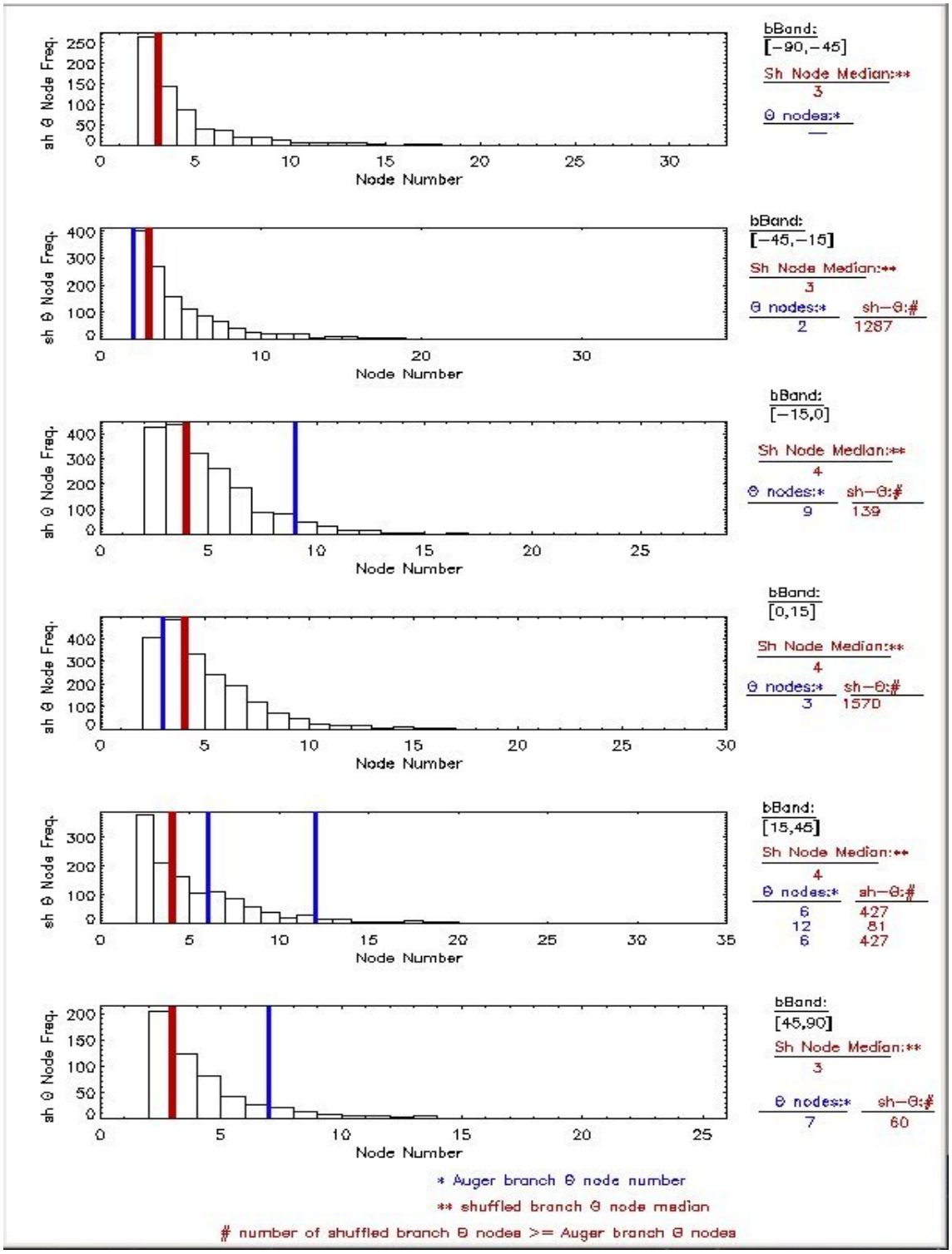

Figure A.14: **b**Band shuffle node frequency vs shuffled node number where Gal l ∈ ([−45 : −135], [45 : 135])° for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

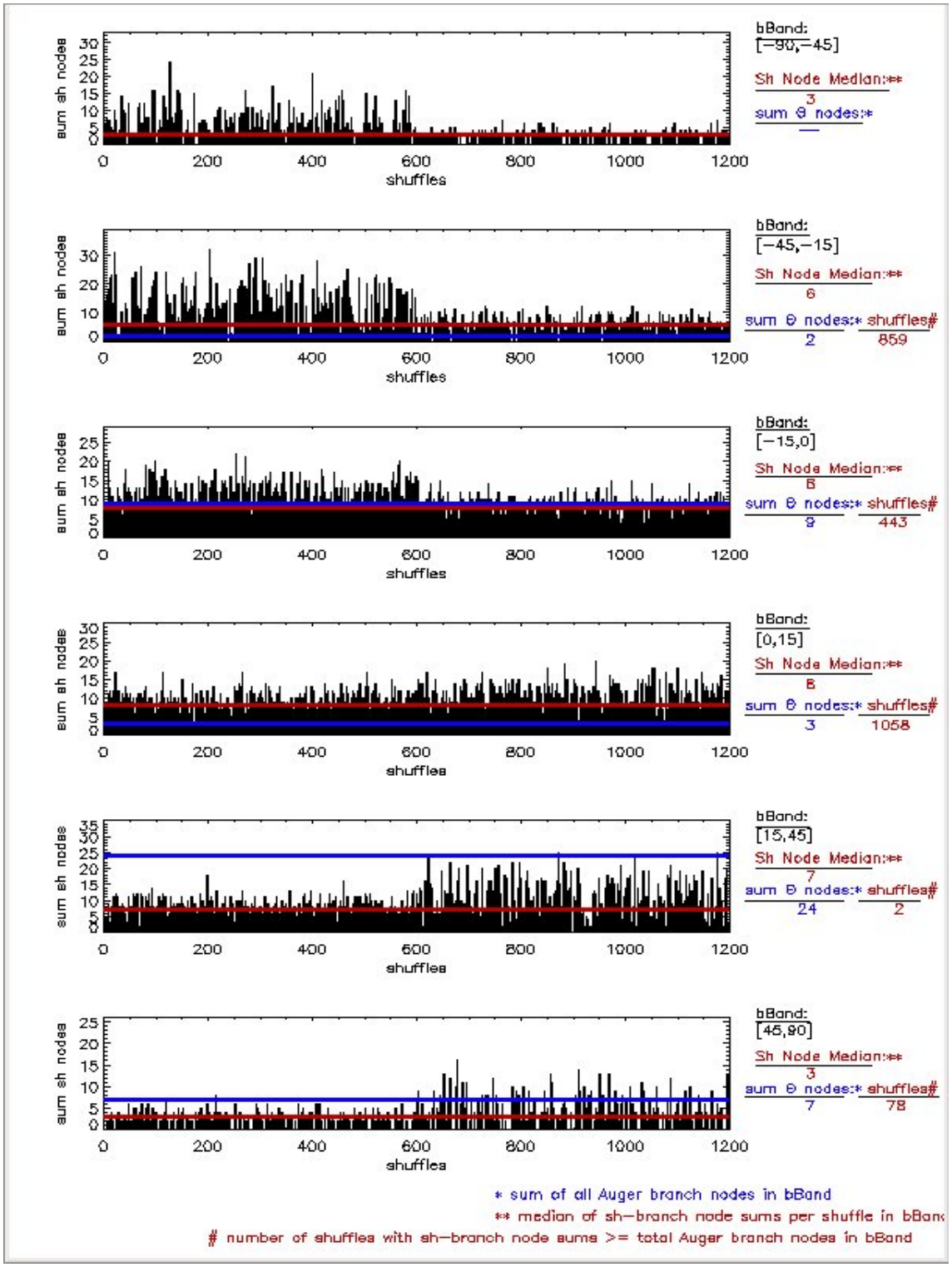

Figure A.15: **bB**and shuffle nodes vs shuffles where Gal $l \in ([-45 : -135], [45 : 135])^\circ$ for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

A.5 bBands where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$

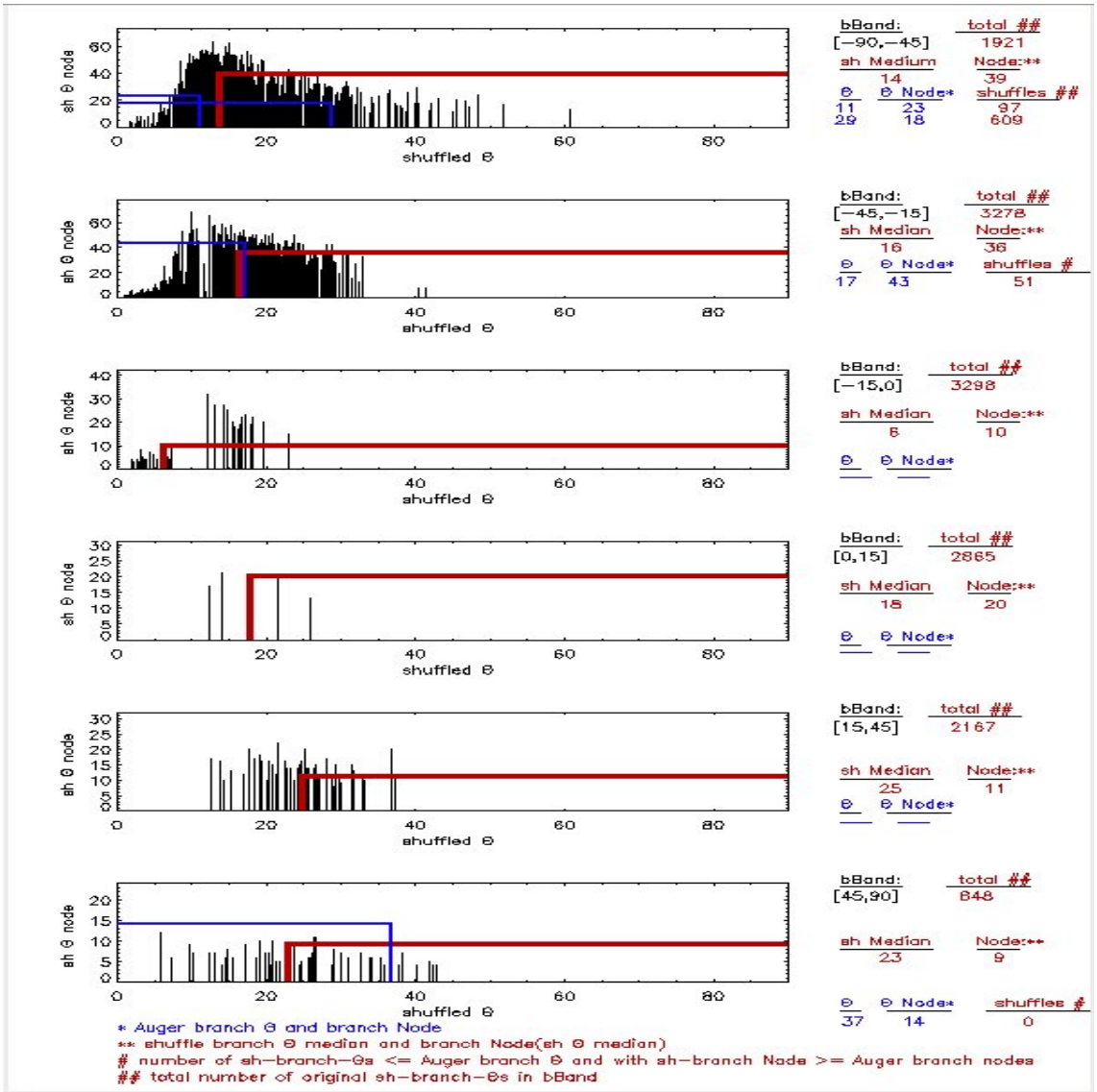

Figure A.16: **b**Band shuffle Θ nodes vs Θ where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ for 40 EeV < E ≤ 50 EeV: (current to 8/11/2012).

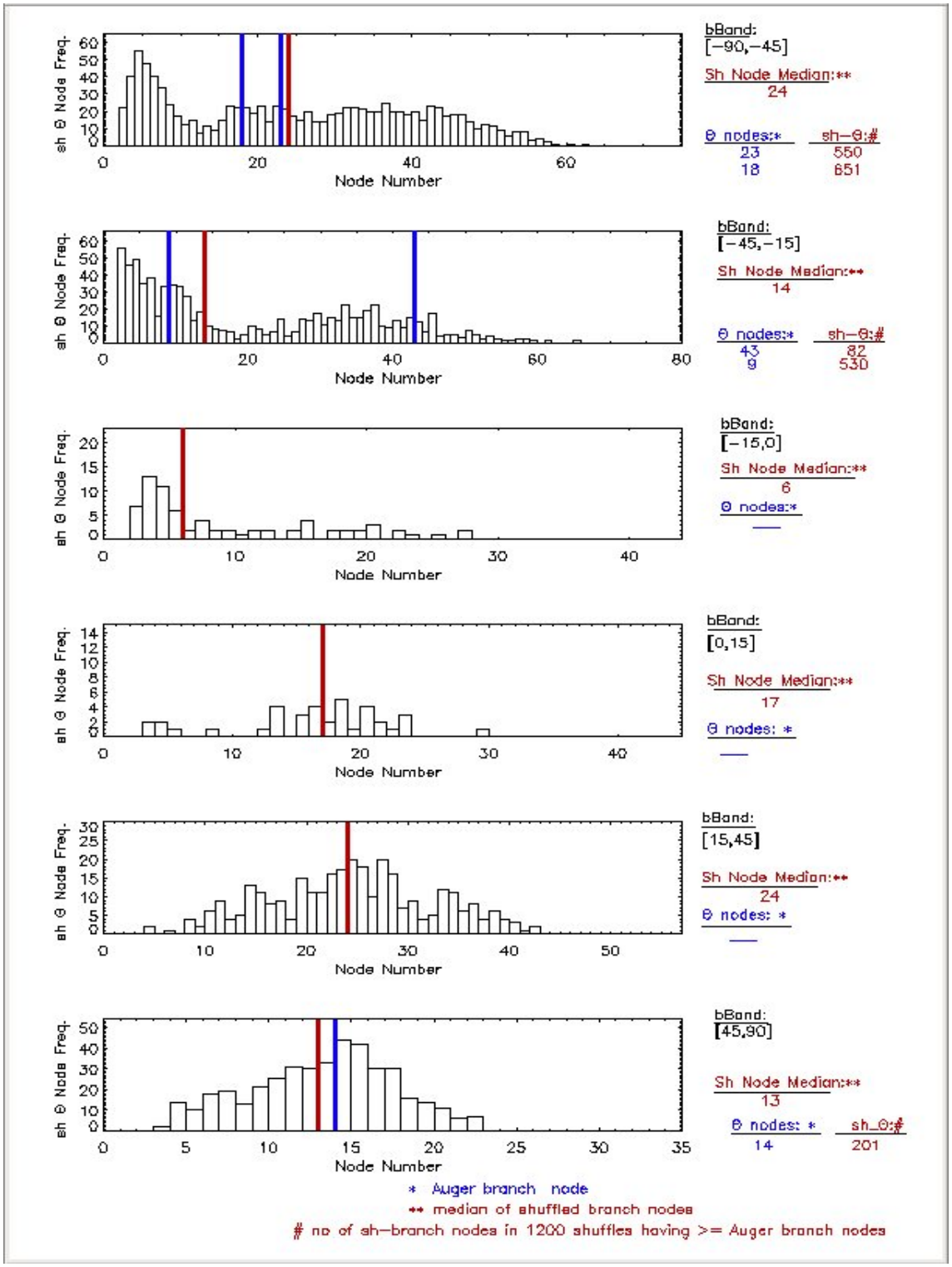

Figure A.17: **b**Band freq shuffle nodes vs shuffle node number where $\Delta|\mathbf{l}_1 - \mathbf{l}_2| > 90°$ for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

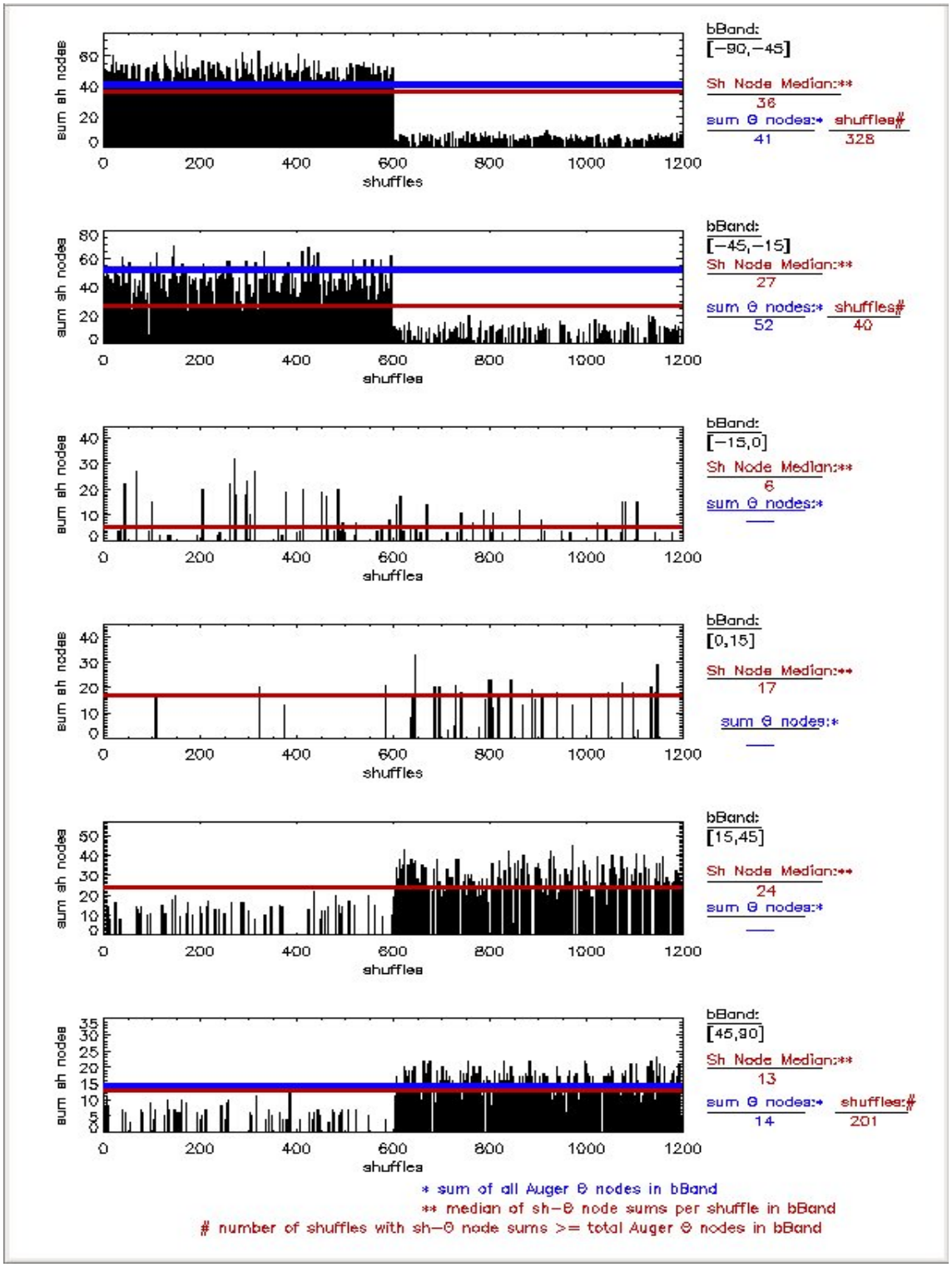

Figure A.18: **b**Band shuffle nodes vs shuffles where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90^\circ$ for 40 EeV < E ≤ 50 EeV (current to 8/11/2012).

Appendix B

South and North 50 EeV < E ≤ 60 EeV:

B.1 Yggdrasil Equatorial (RA,dec) Co-ordinates

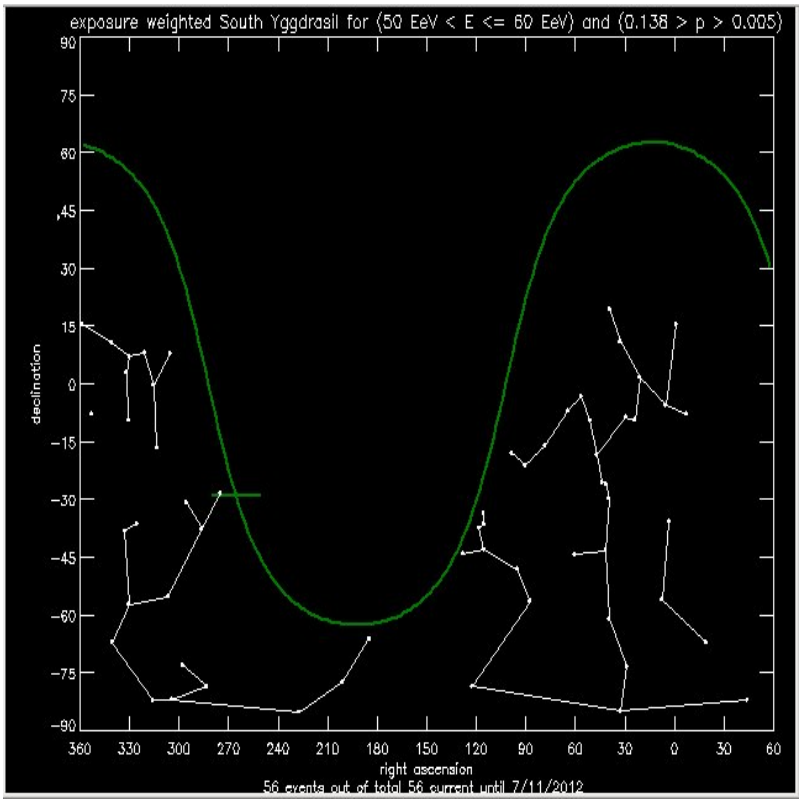

Figure B.1: South Yggdrasil of all events in **RA** vs **dec** for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

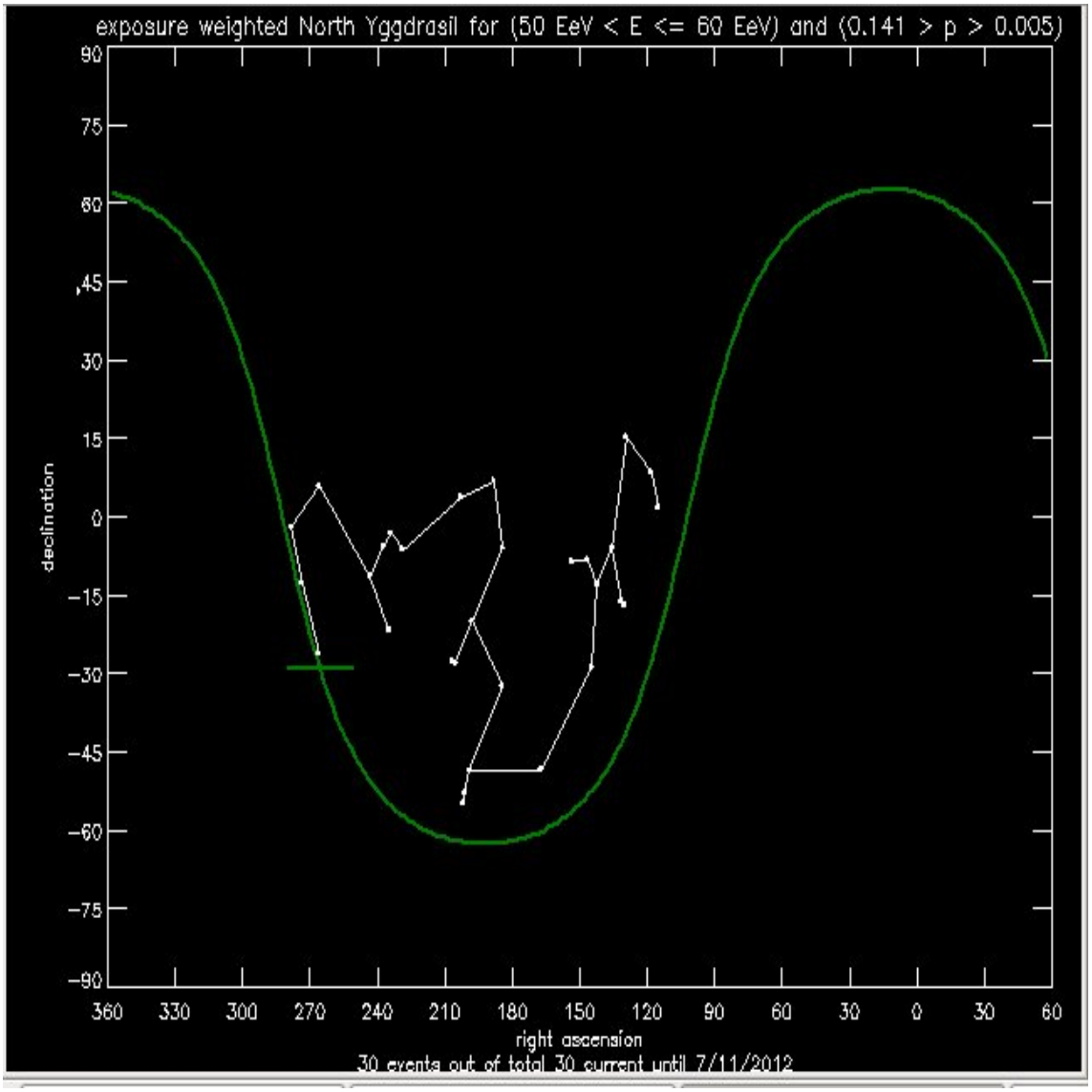

Figure B.2: North Yggdrasil of all events in **RA** vs **dec** for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

B.2 Yggdrasil Galactic (l,b) Co-ordinates

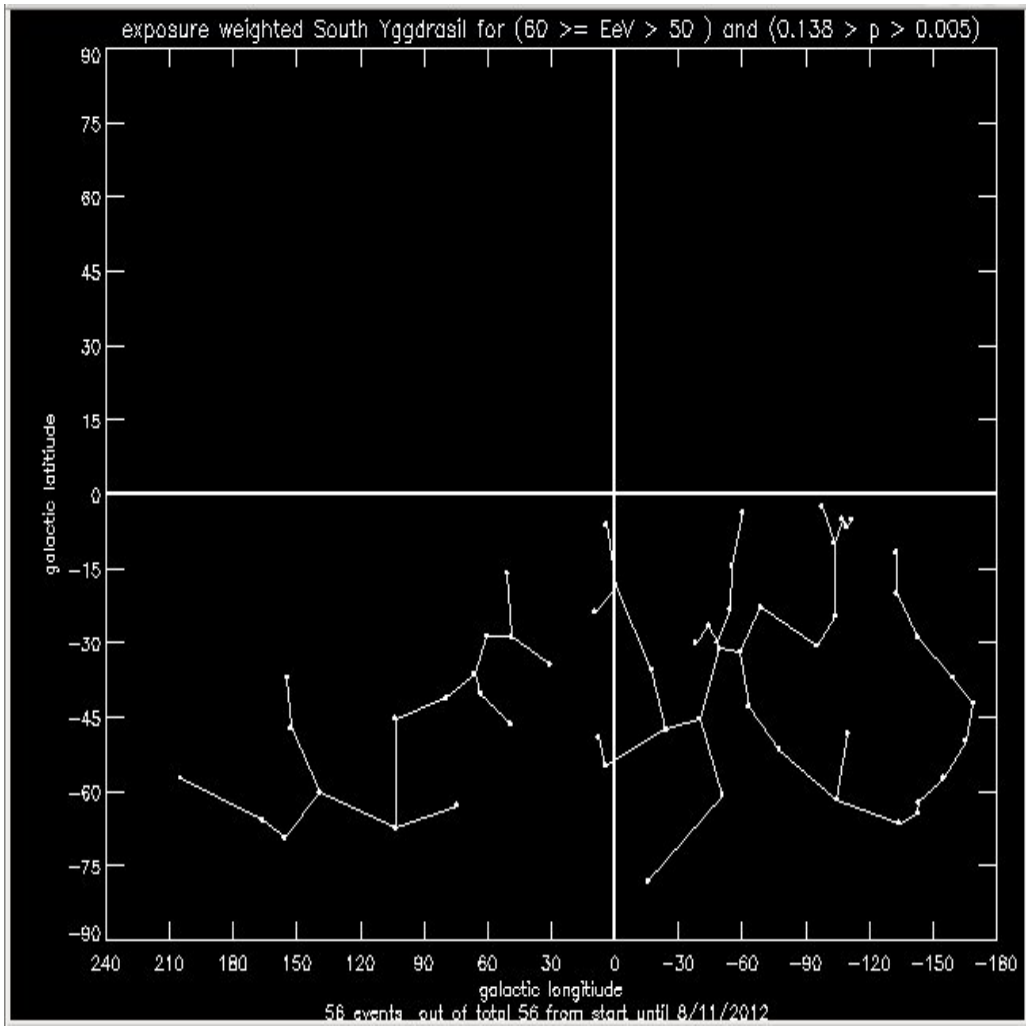

Figure B.3: South Yggdrasil of all events in **l** vs **b** for 50 EeV < $E \leq$ 60 EeV (current to 8/11/2013).

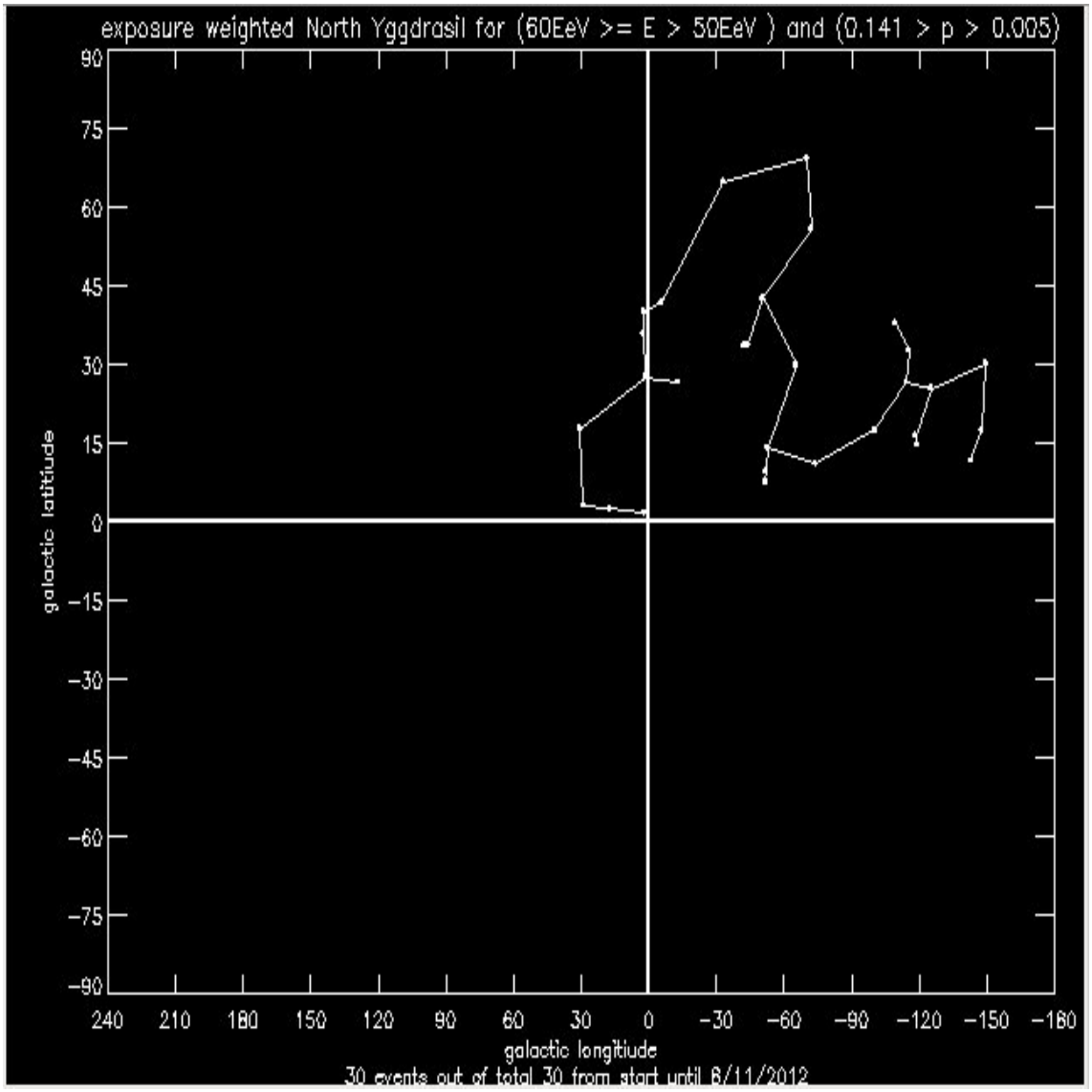

Figure B.4: North Yggdrasil of all events in **l** vs **b** for 50 EeV < E ≤60 EeV (current to 8/11/2012).

B.3 bBands

The Yggs of Pierre Auger data have been divided into branches along latitude cuts. These branches are seen in Figure B.5 and Figure B.6 on the next page, and are followed by 3 lots of distributions flagged with Pierre Auger data values.

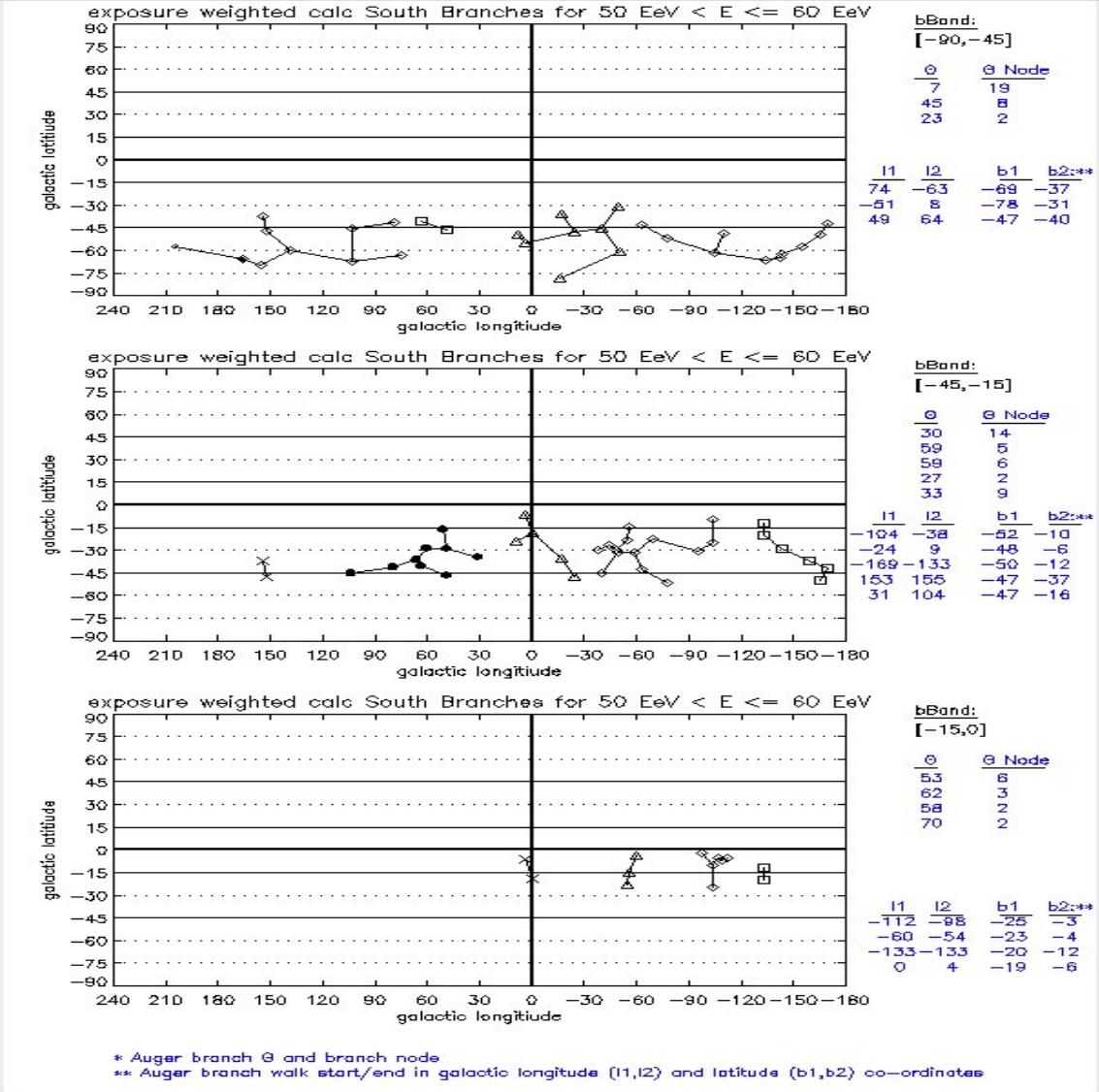

Figure B.5: **South bBand branches for 60 EeV ≥ E > 50 EeV. Each branch denoted by a different symbol. (current to 8/11/2012).**

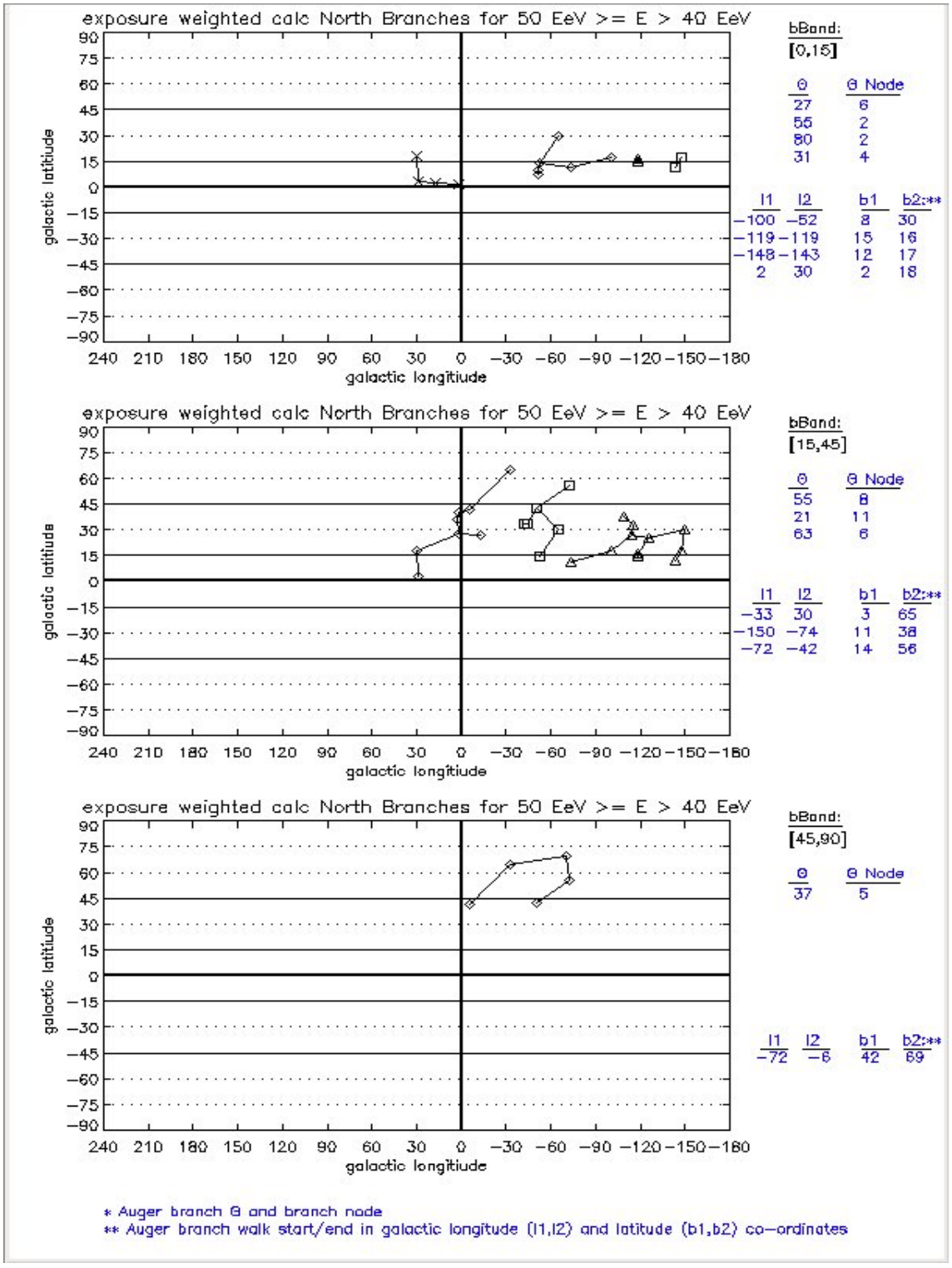

Figure B.6: **North bBand branches for 60 EeV ≥ E > 50 EeV. Each branch denoted by a different symbol. (current to 8/11/2012).**

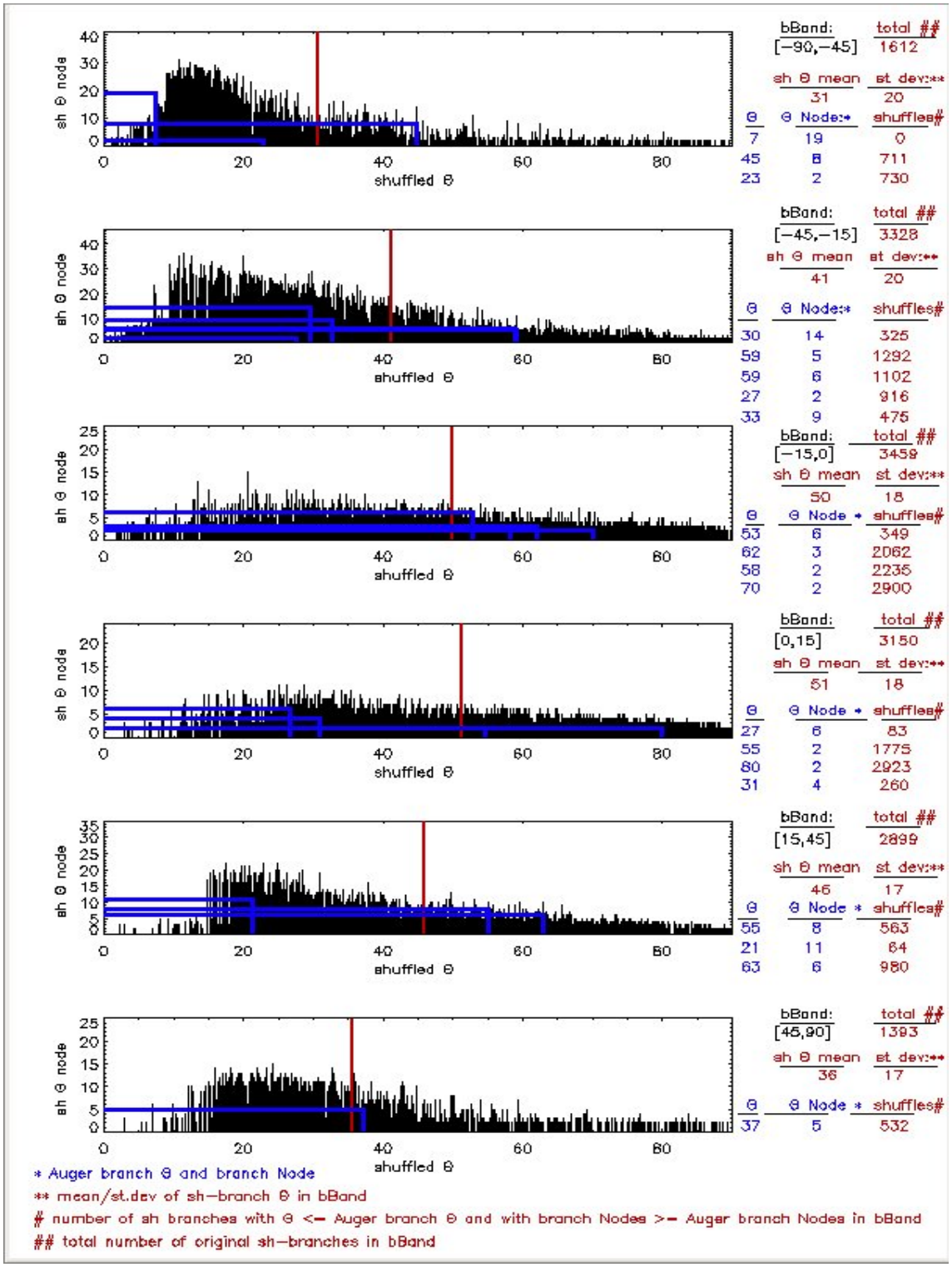

Figure B.7: **b**Band shuffle Θ node vs Θ for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

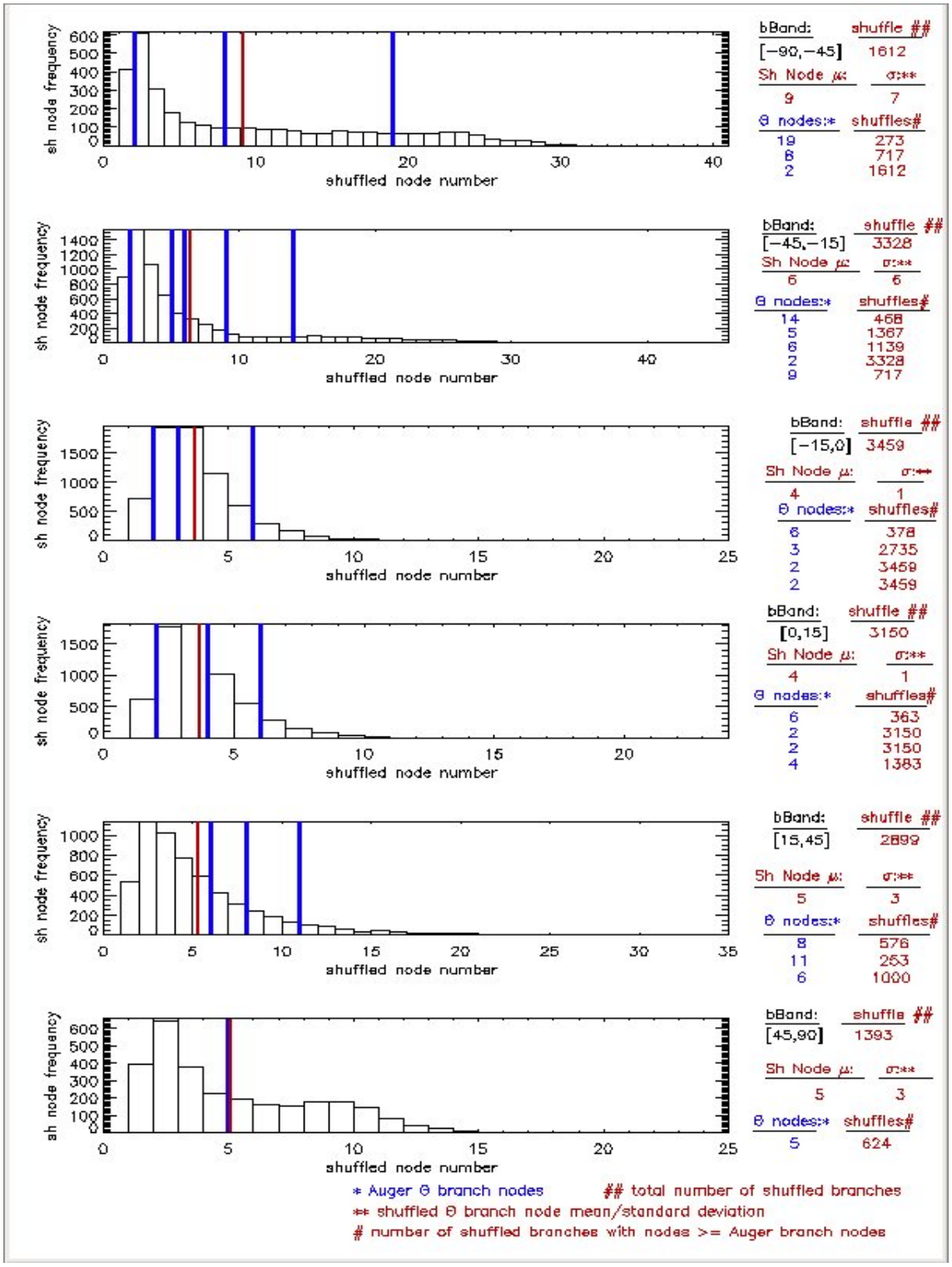

Figure B.8: **b**Band shuffle node frequency vs nodes for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

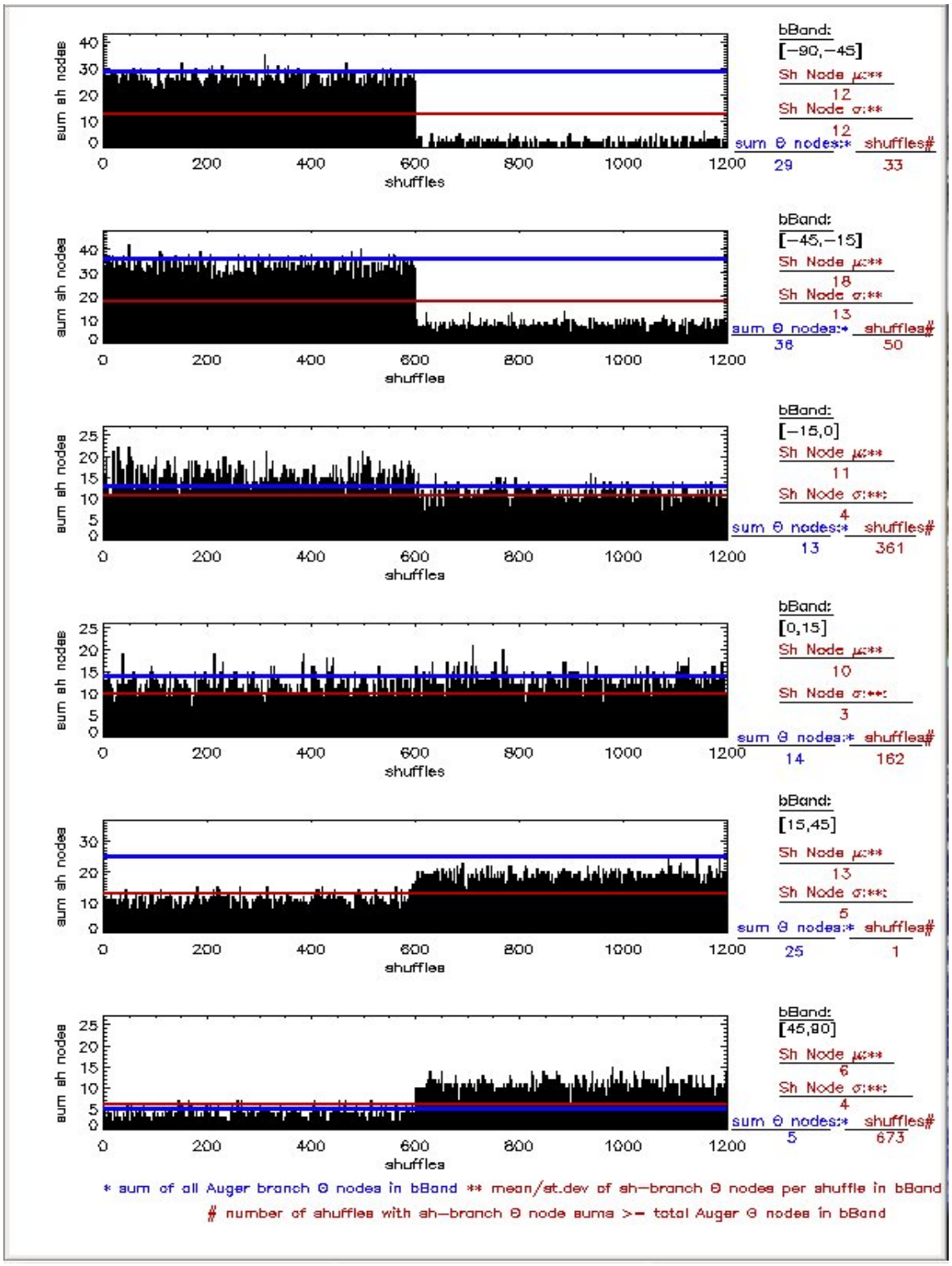

Figure B.9: **b**Band shuffle nodes vs shuffles for 50 EeV > E ≤ 60 EeV (current to 8/11/2012).

B.4 Galactic Longitude Quadrants and bBands

bBands ∈ ([−135 : 135])° **and** ∈ ([−45 : 45])°

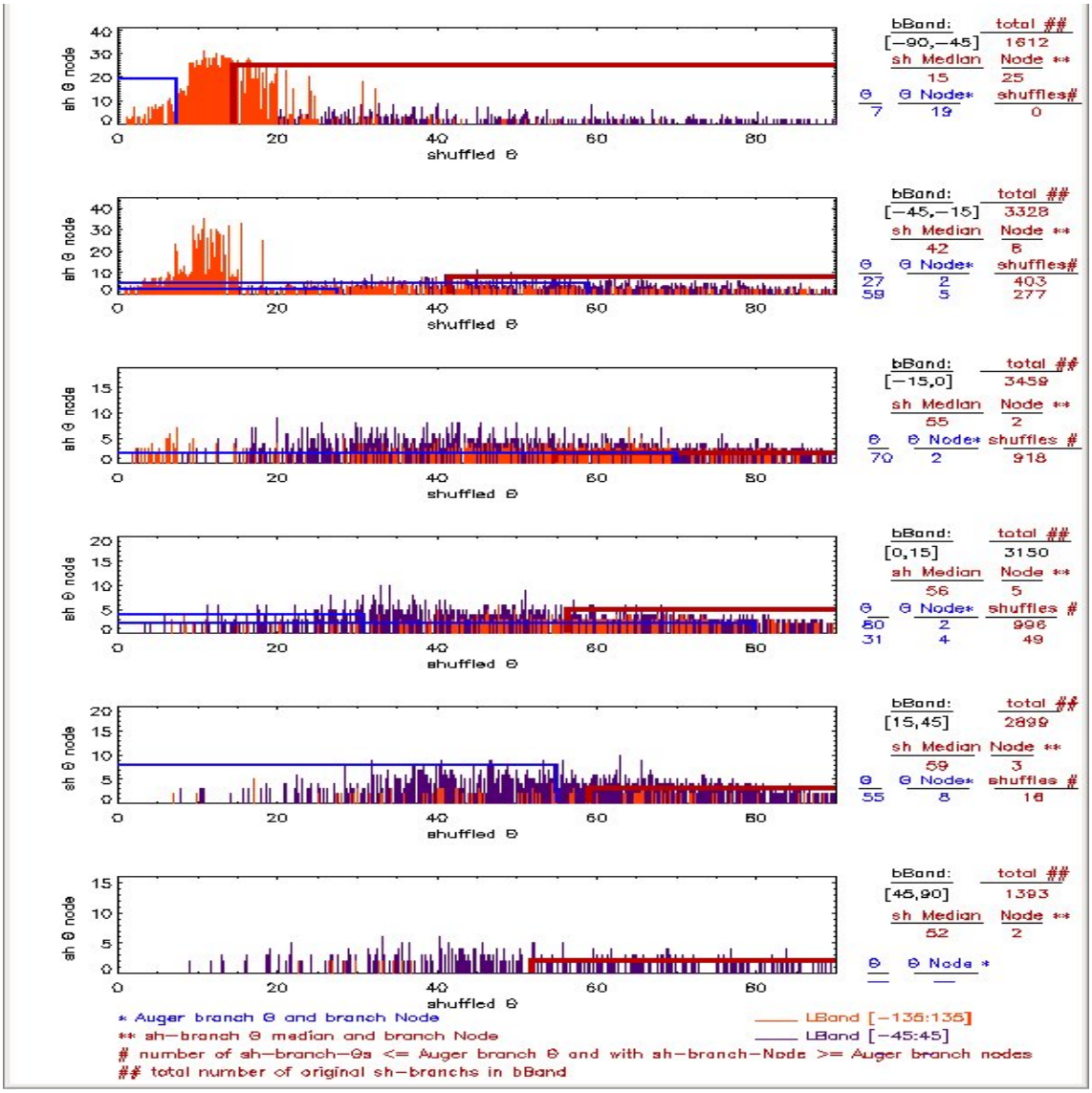

Figure B.10: **b**Band shuffle Θ nodes vs Θ where Gal **l** ∈ ([−45 : 45], [−135 : 135])° for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

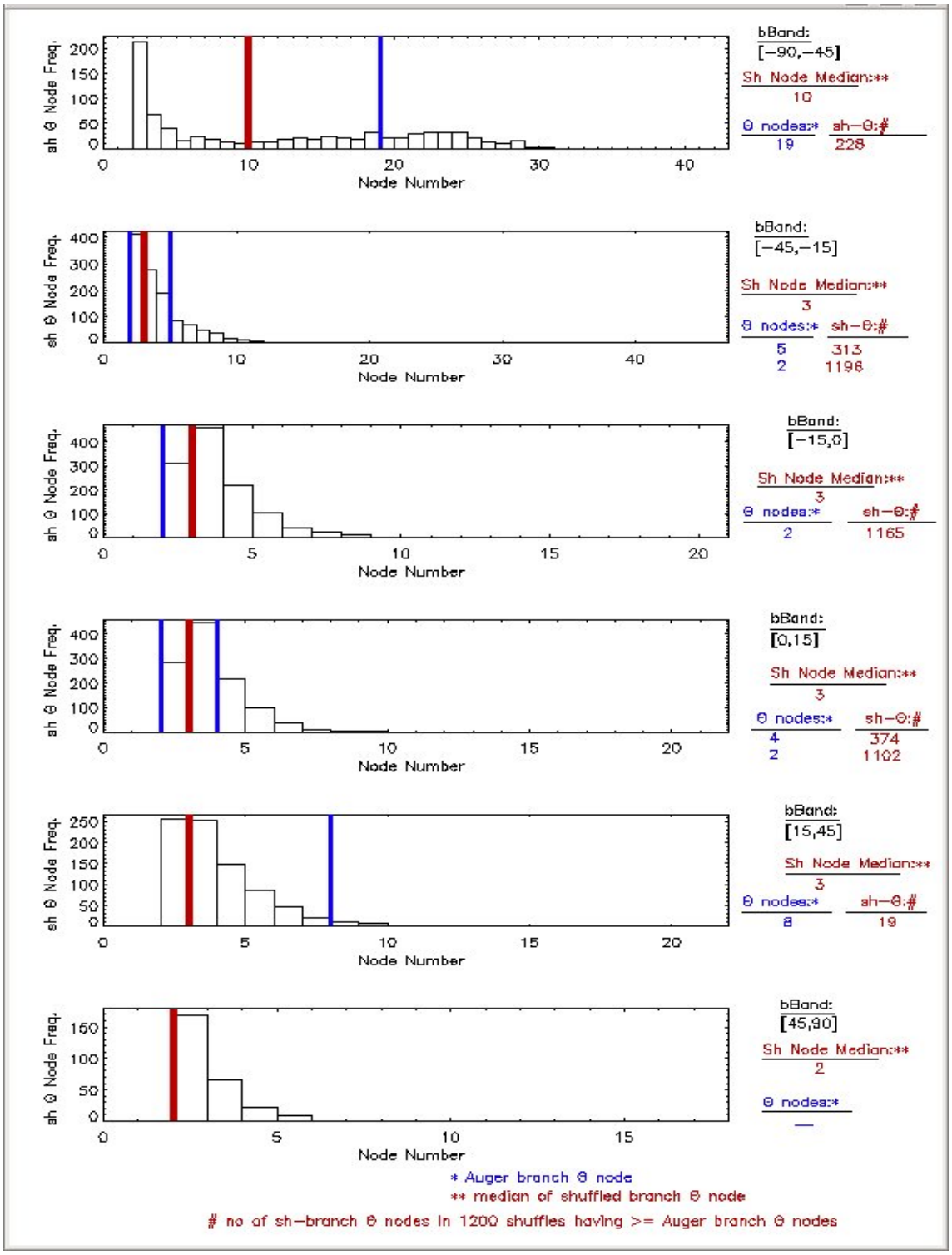

Figure B.11: **b**Band shuffle node frequency vs shuffled node number where Gal l ∈ ([−45 : 45], [−135 : 135])° for 60 EeV ≥ E > 50 EeV (current to 8/11/2012).

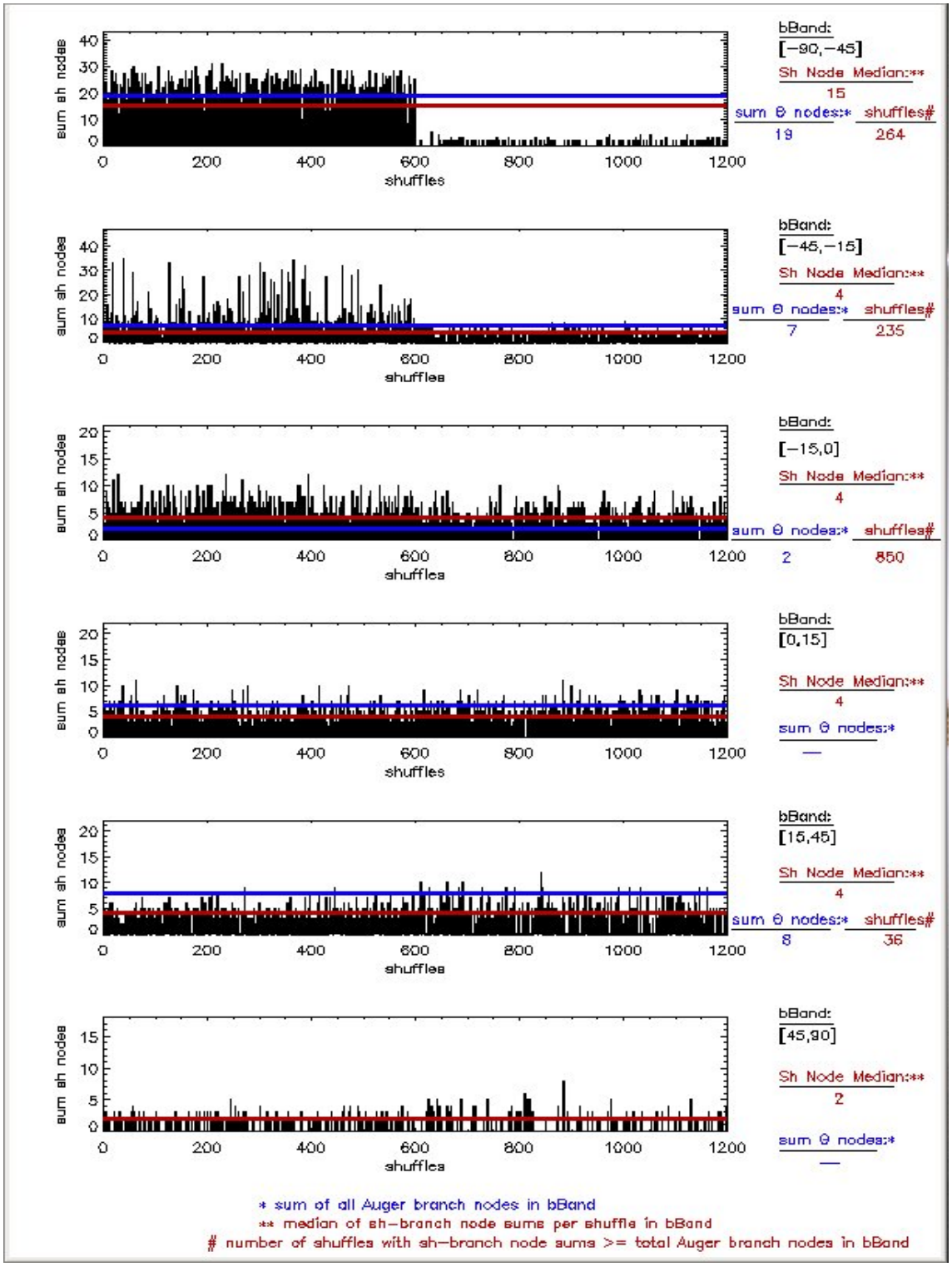

Figure B.12: **b**Band shuffle nodes vs shuffles where Gal $l \in ([-45:45],[-135:135])^\circ$ for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

bBands ∈ ([−135 : −45])° **and** ∈ ([45 : 135])°

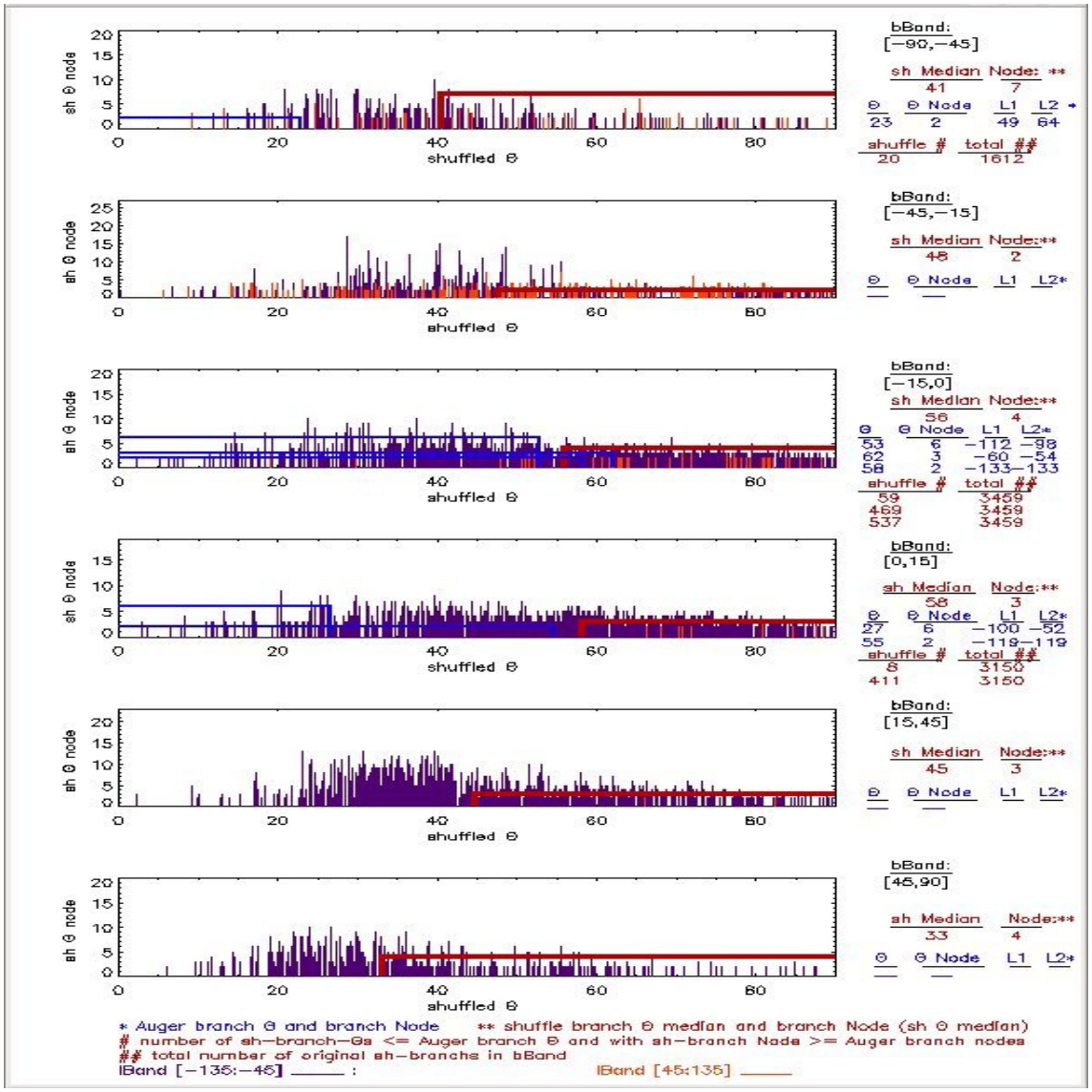

Figure B.13: **b**Band shuffle Θ nodes vs shuffled Θ where Gal l ∈ ([−45 : −135], [45 : 135])° for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

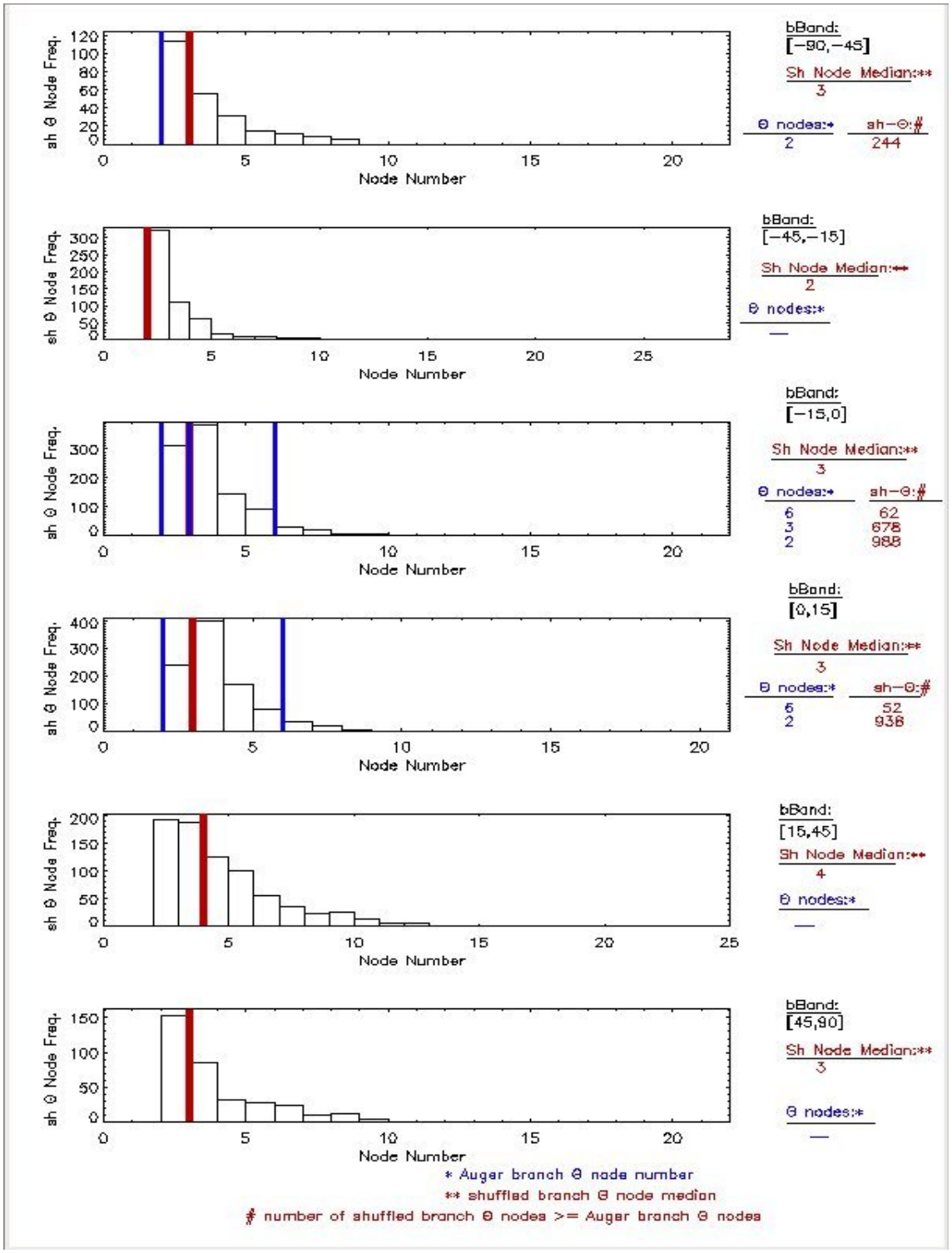

Figure B.14: **b**Band shuffle node frequency vs shuffled node number where Gal l ∈ ([−45 : 45], [−135 : 135])° for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

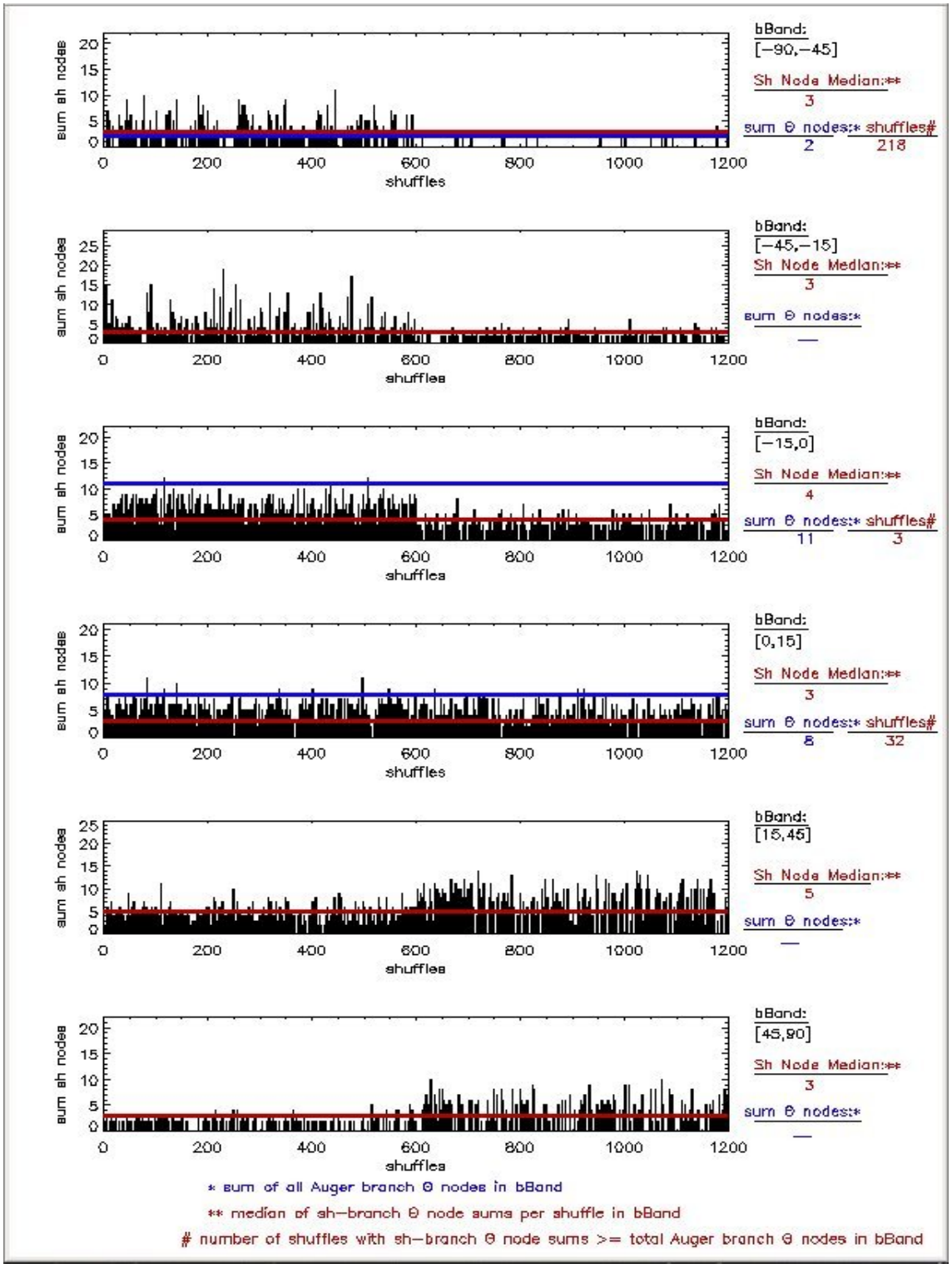

Figure B.15: **b**Band shuffle nodes vs shuffles where Gal l ∈ ([−45 : −135], [45 : 135])° for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

B.5 bBands where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$

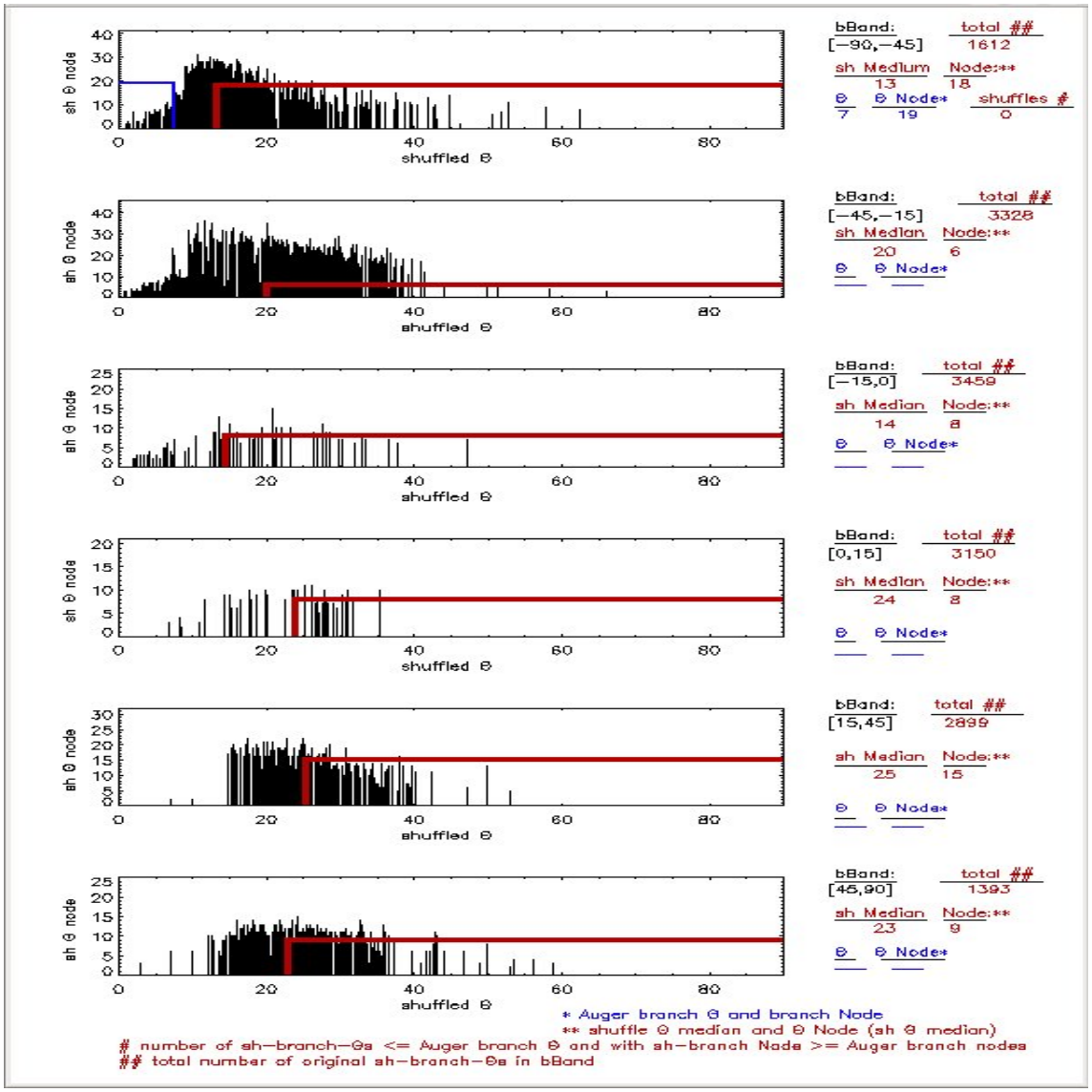

Figure B.16: **b**Band shuffle Θ nodes vs Θ where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

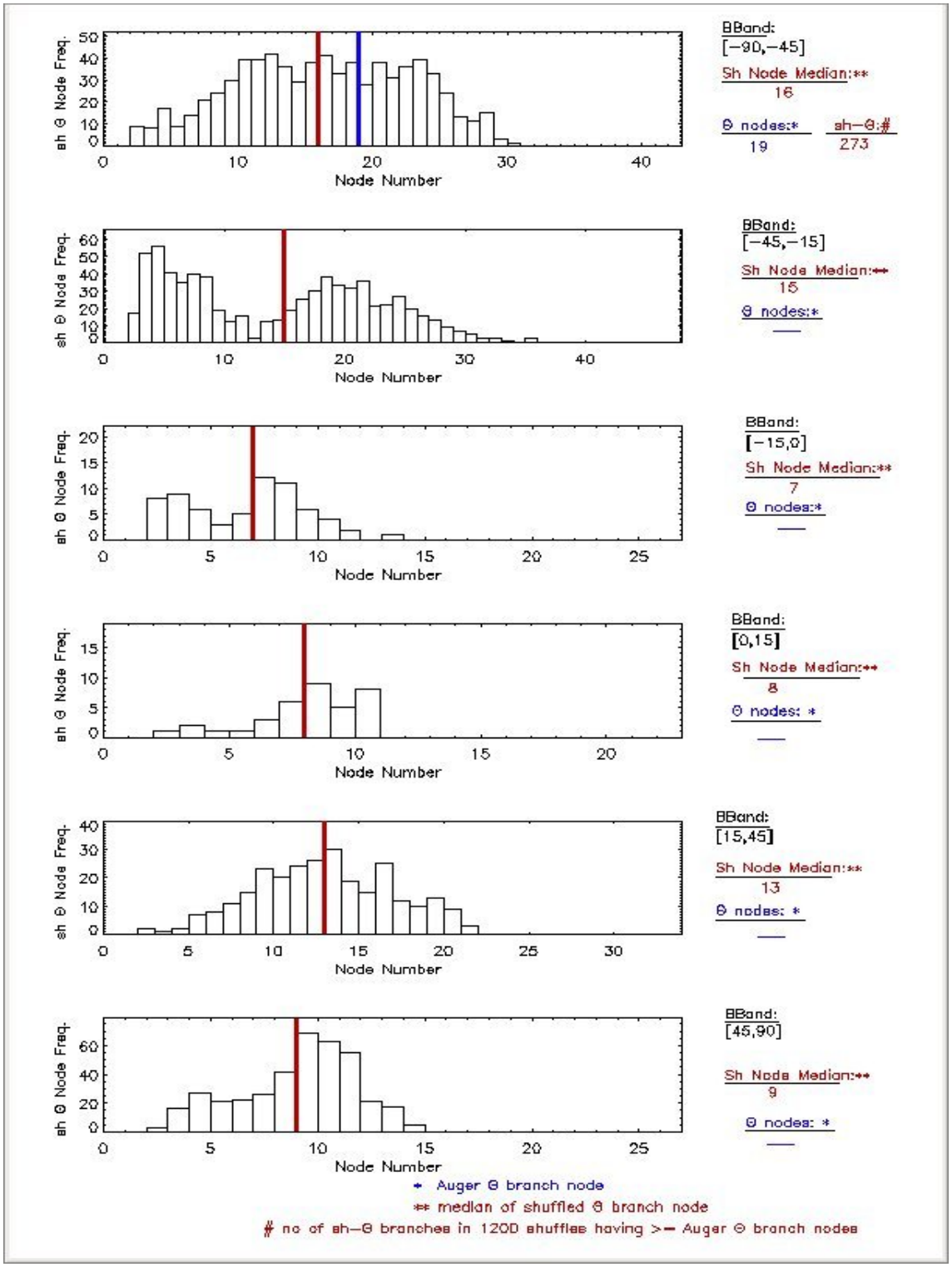

Figure B.17: **b**Band freq shuffle nodes vs shuffle node number where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

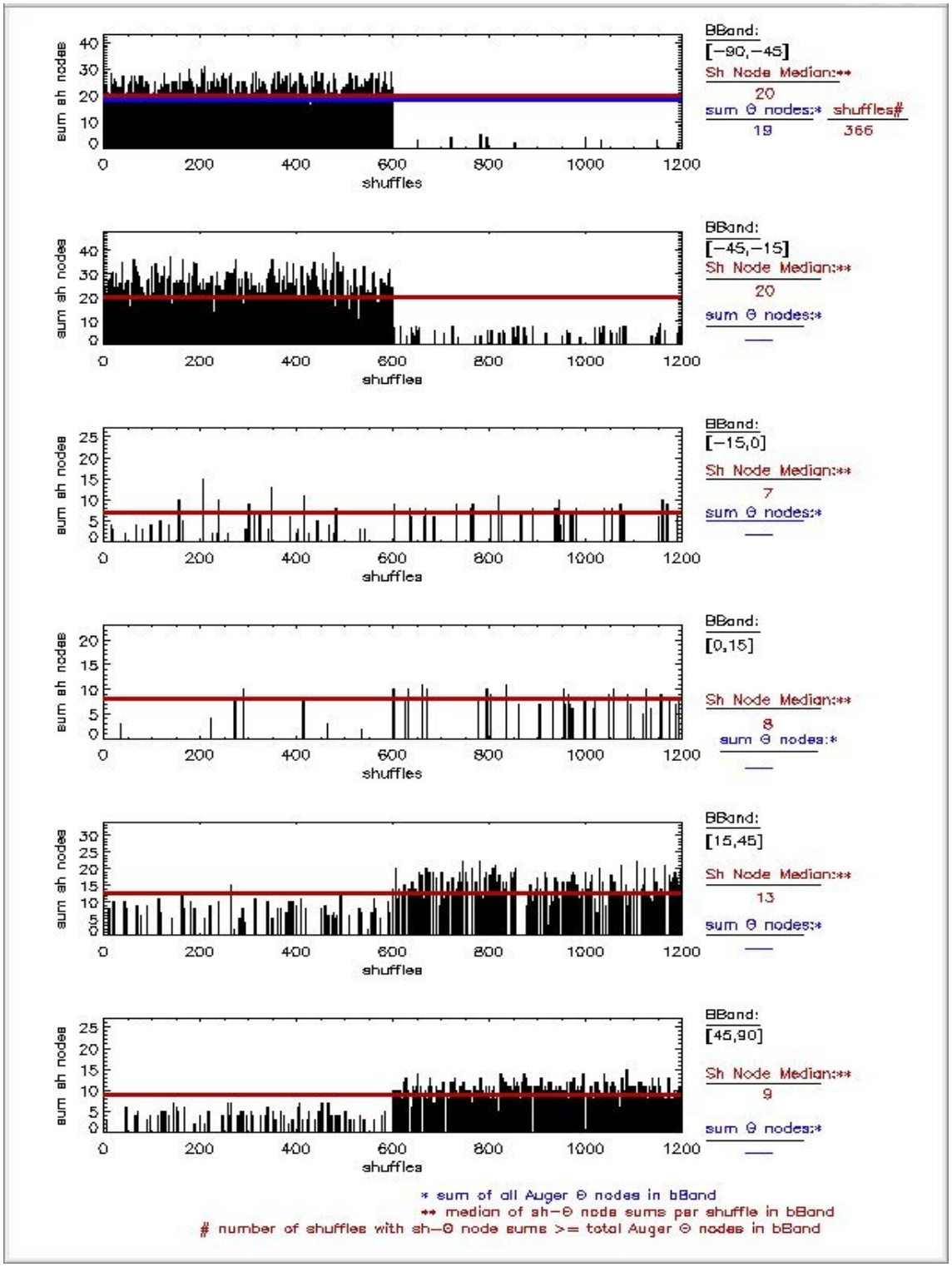

Figure B.18: **b**Band shuffle nodes vs shuffles where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90^\circ$ for 50 EeV < E ≤ 60 EeV (current to 8/11/2012).

Appendix C

South and North E > 60 EeV:

C.1 Yggdrasil Equatorial (RA,dec) Co-ordinates

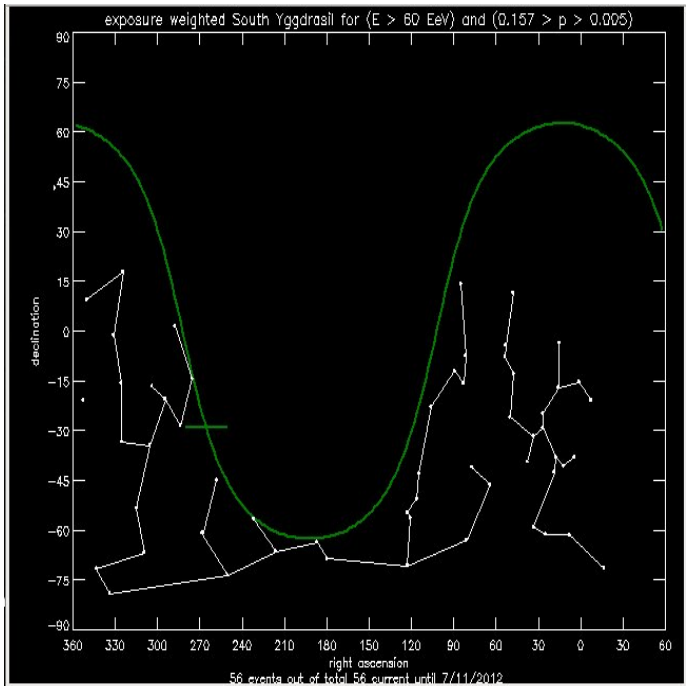

Figure C.1: **South Yggdrasil of all events in RA vs dec for E > 60 EeV (current to 8/11/2012).**

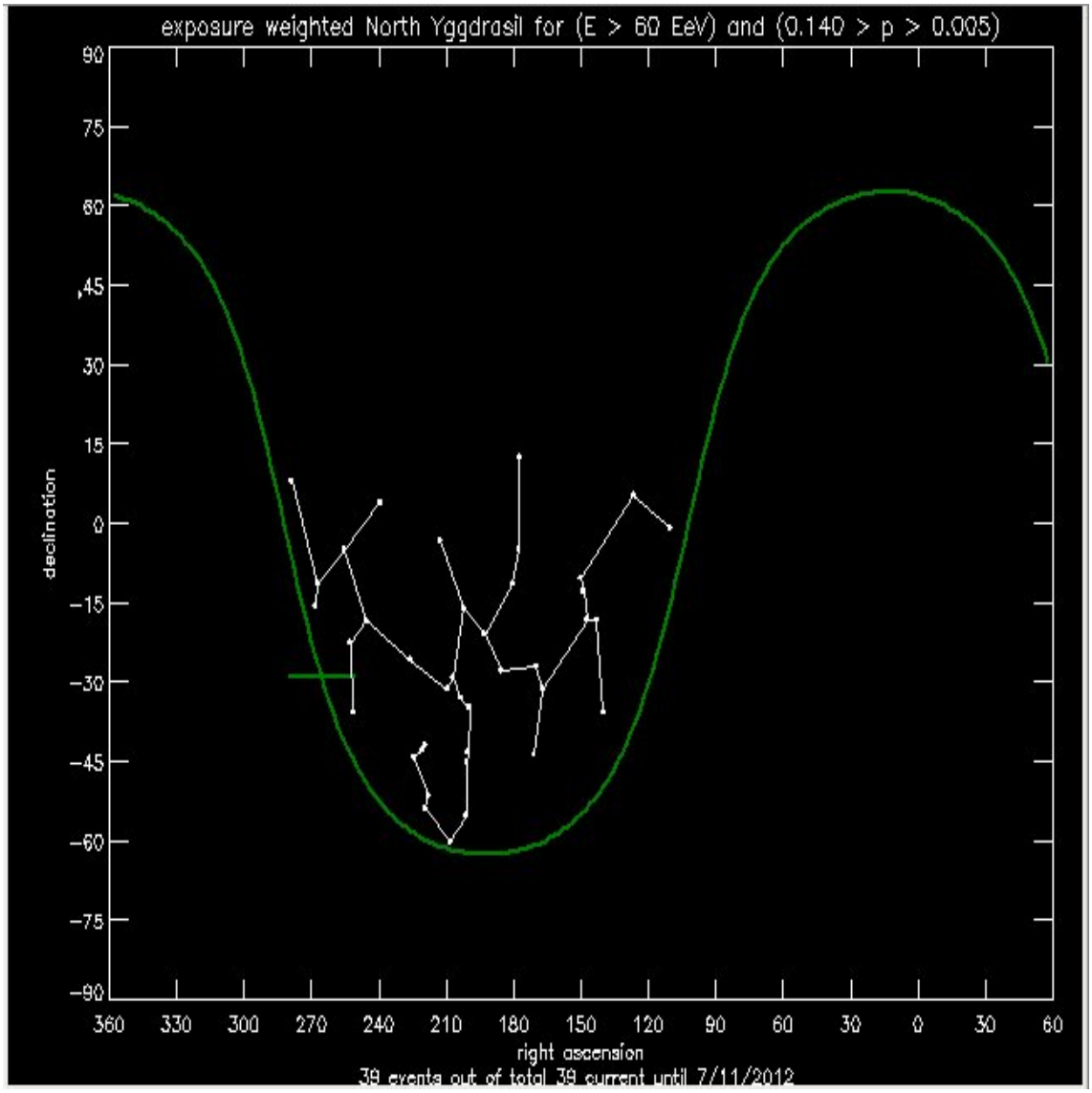

Figure C.2: **North Yggdrasil of all events in RA vs dec for E > 60 EeV (current to 8/11/2012).**

C.2 Yggdrasil Galactic (l,b) Co-ordinates

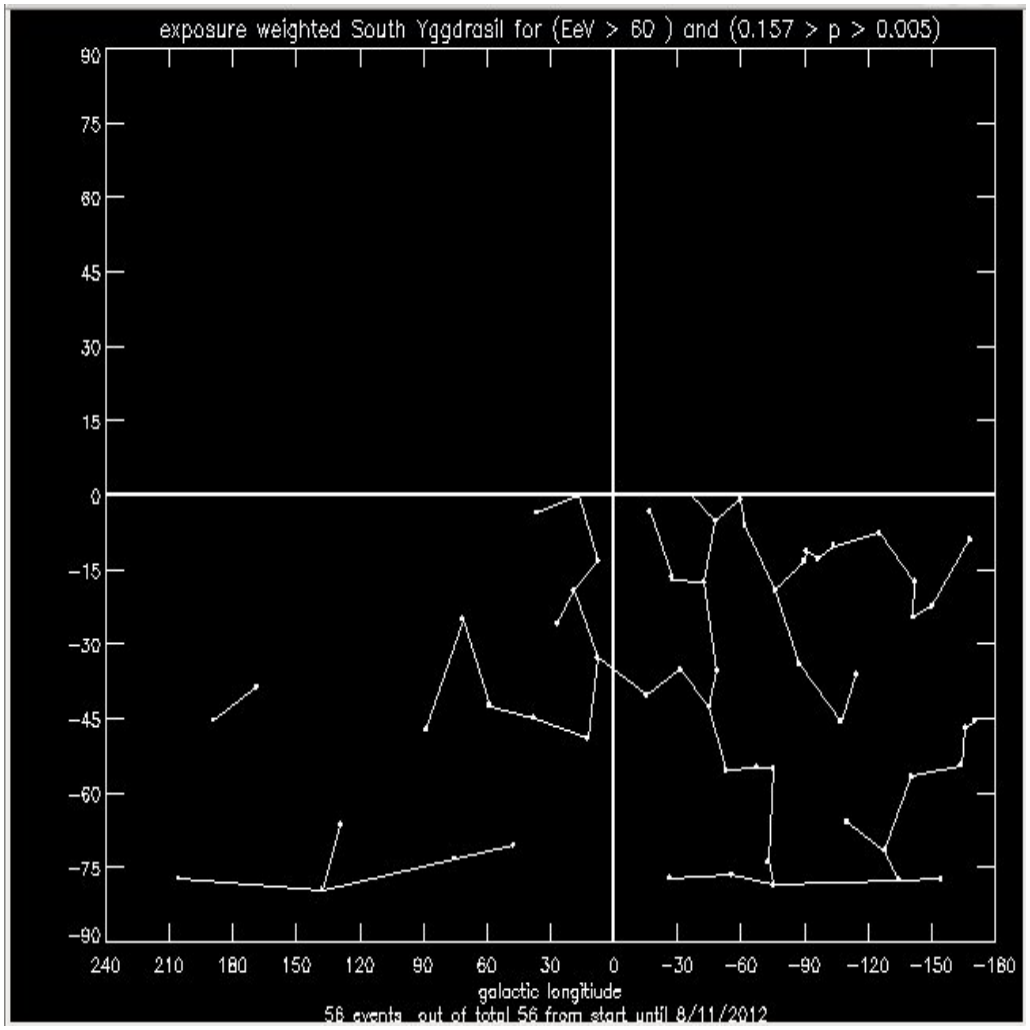

Figure C.3: **South Yggdrasil of all events in l vs b for E > 60 EeV (current to 8/11/2012).**

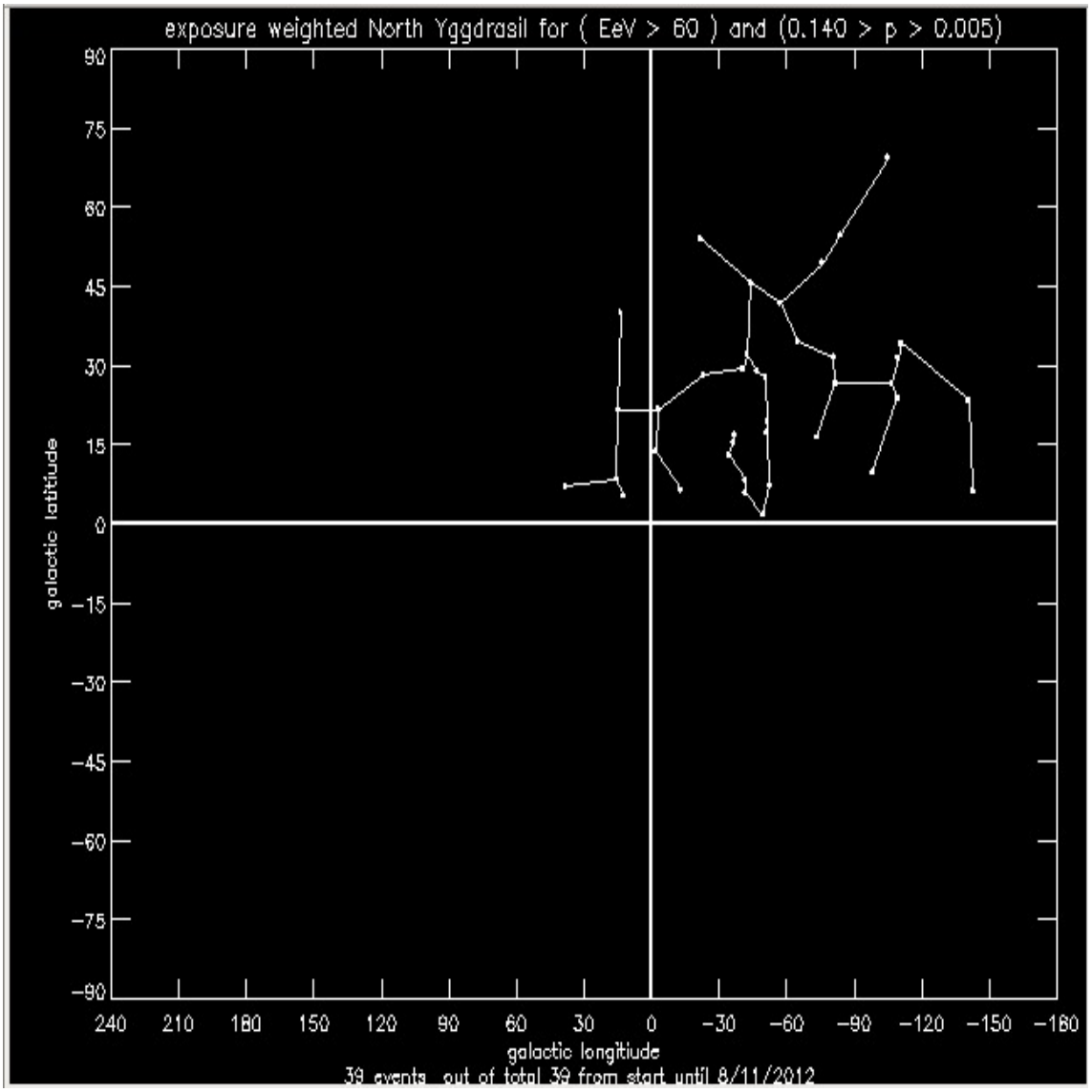

Figure C.4: **North Yggdrasil of all events in l vs b for E** > 60 **EeV (current to 8/11/2012).**

C.3 bBands

The Yggs of Pierre Auger data have been divided into branches along latitude cuts. These branches are seen in Figure C.5 and Figure C.6 on the next page, and are followed by 3 lots of distributions flagged with Pierre Auger data values.

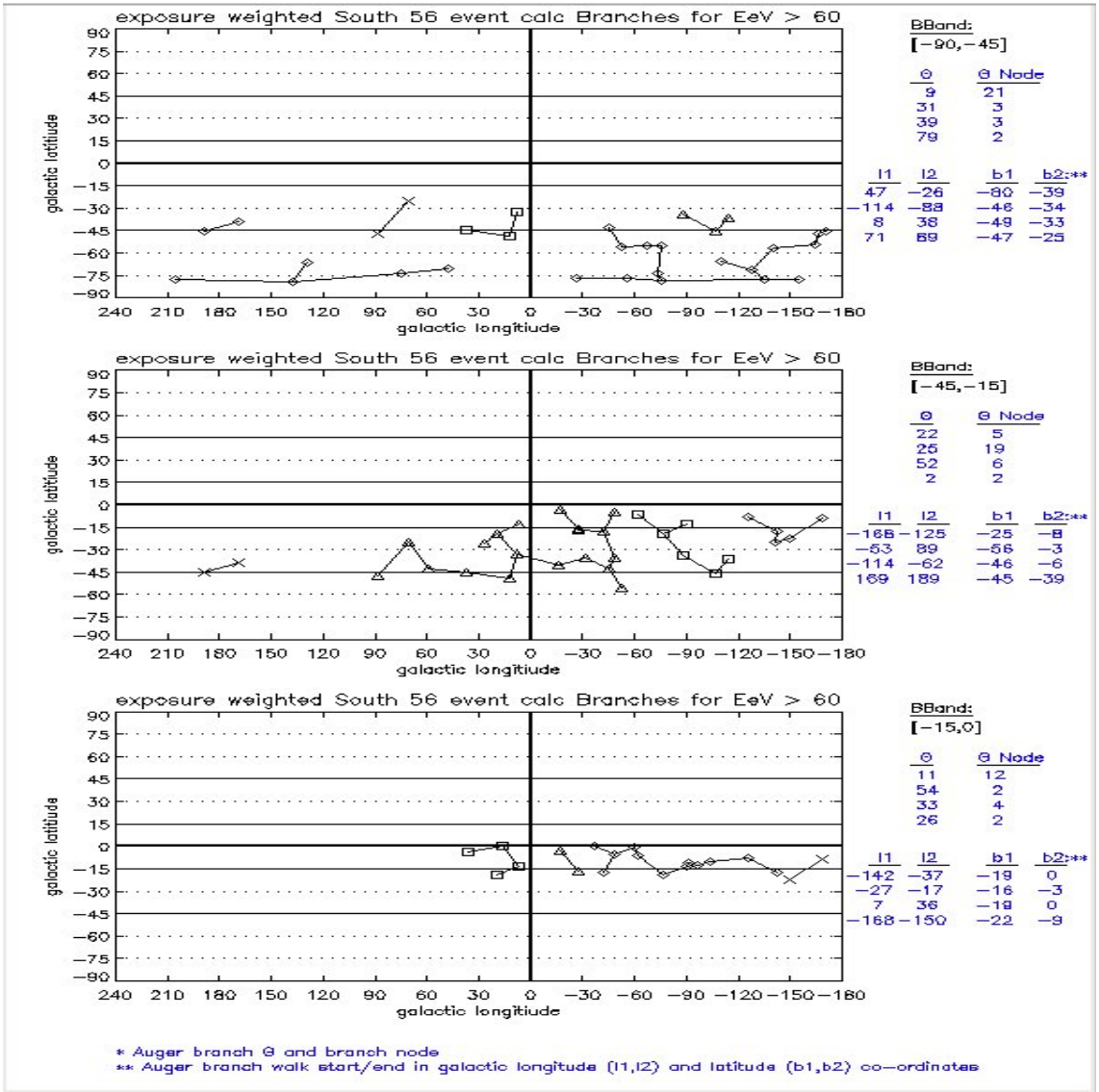

Figure C.5: **South bBand branches for E > 60 EeV. Each branch denoted by a different symbol. (current to 8/11/2012).**

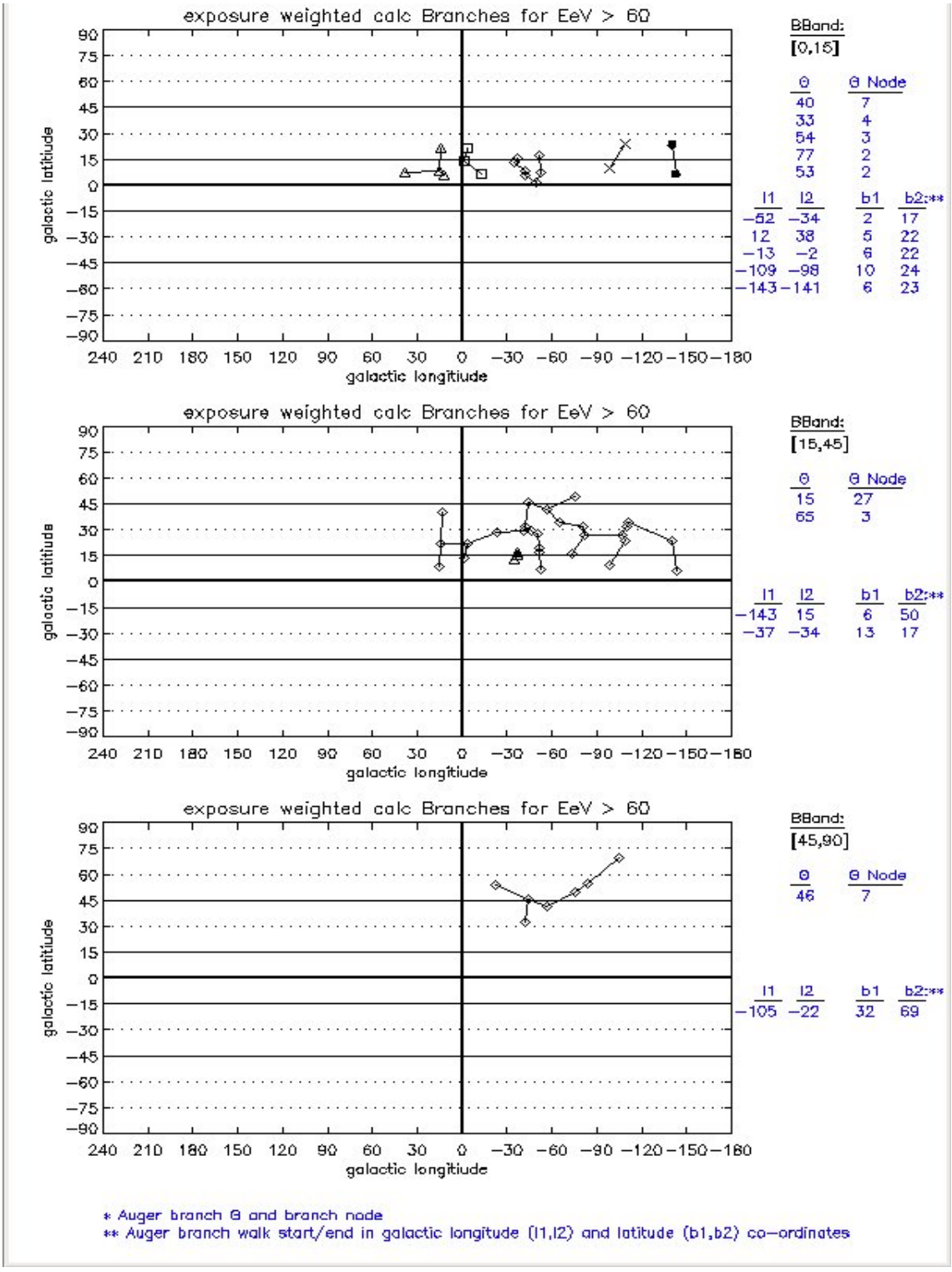

Figure C.6: **North bBand branches for E > 60 EeV. Each branch denoted by a different symbol. (current to 8/11/2012).**

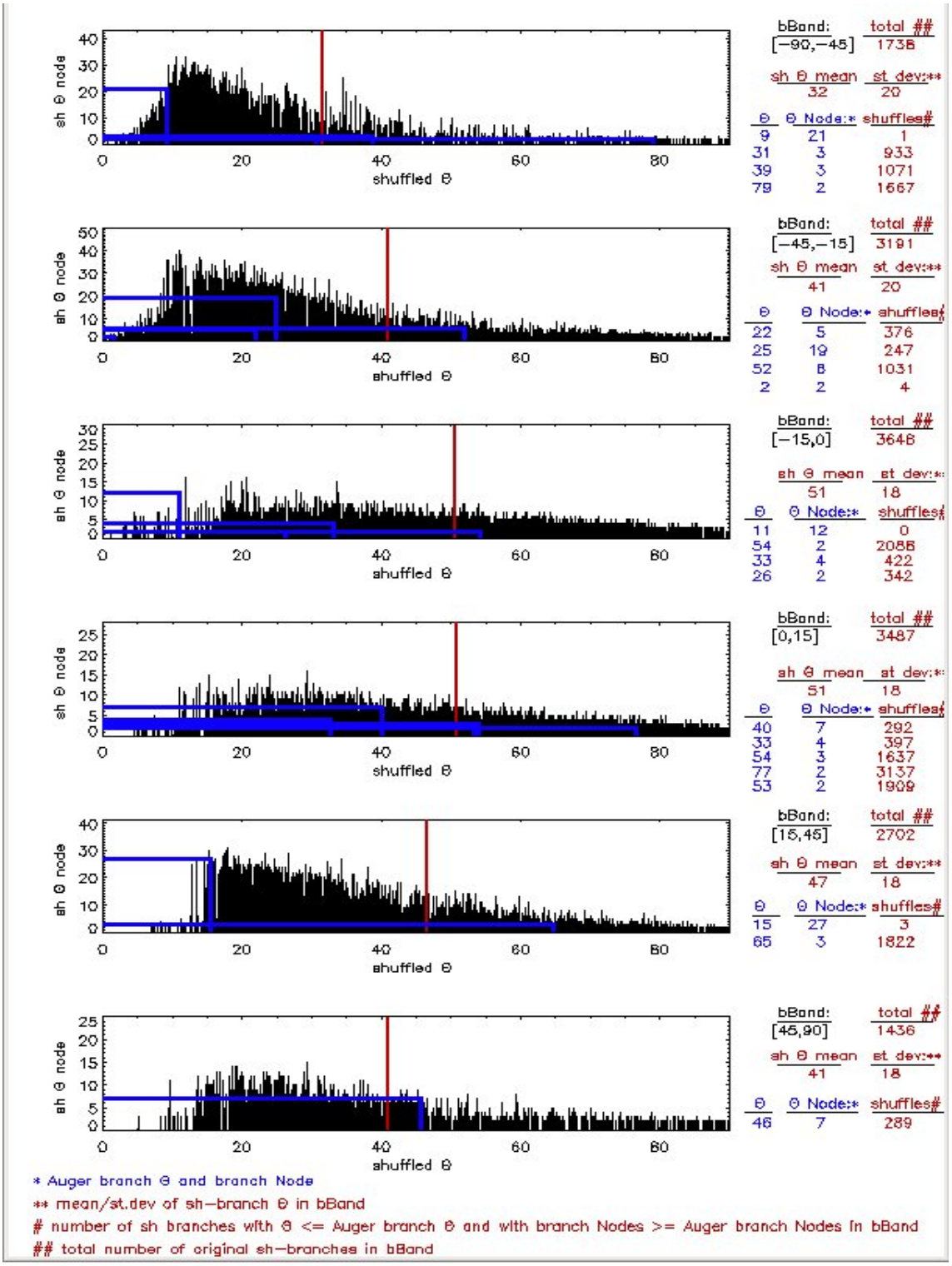

Figure C.7: **bBand shuffle Θ node vs Θ for E > 60 EeV** **(current to 8/11/2012).**

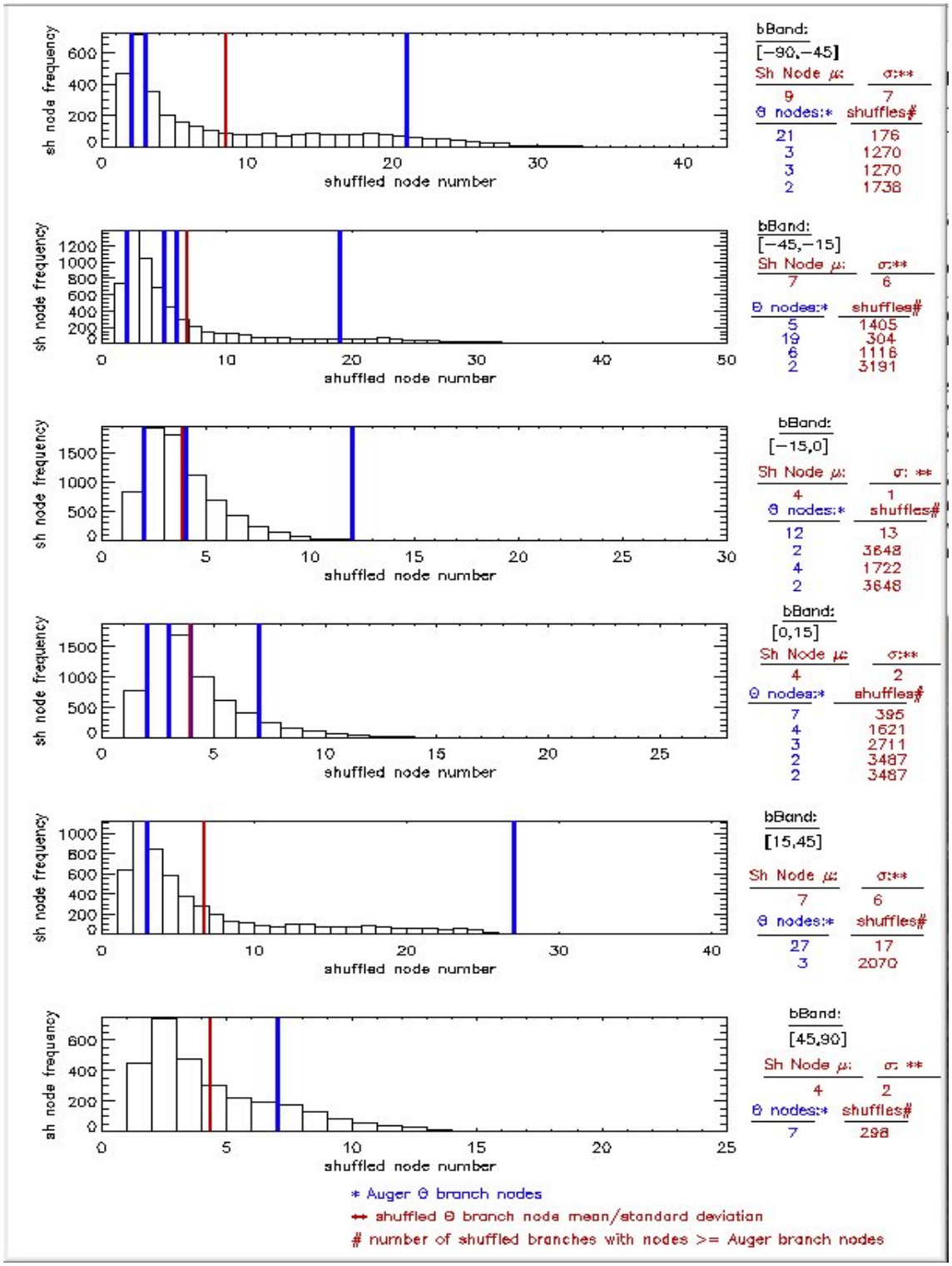

Figure C.8:
bBand shuffle node frequency vs nodes for E > 60 EeV: (current to 8/11/2012).

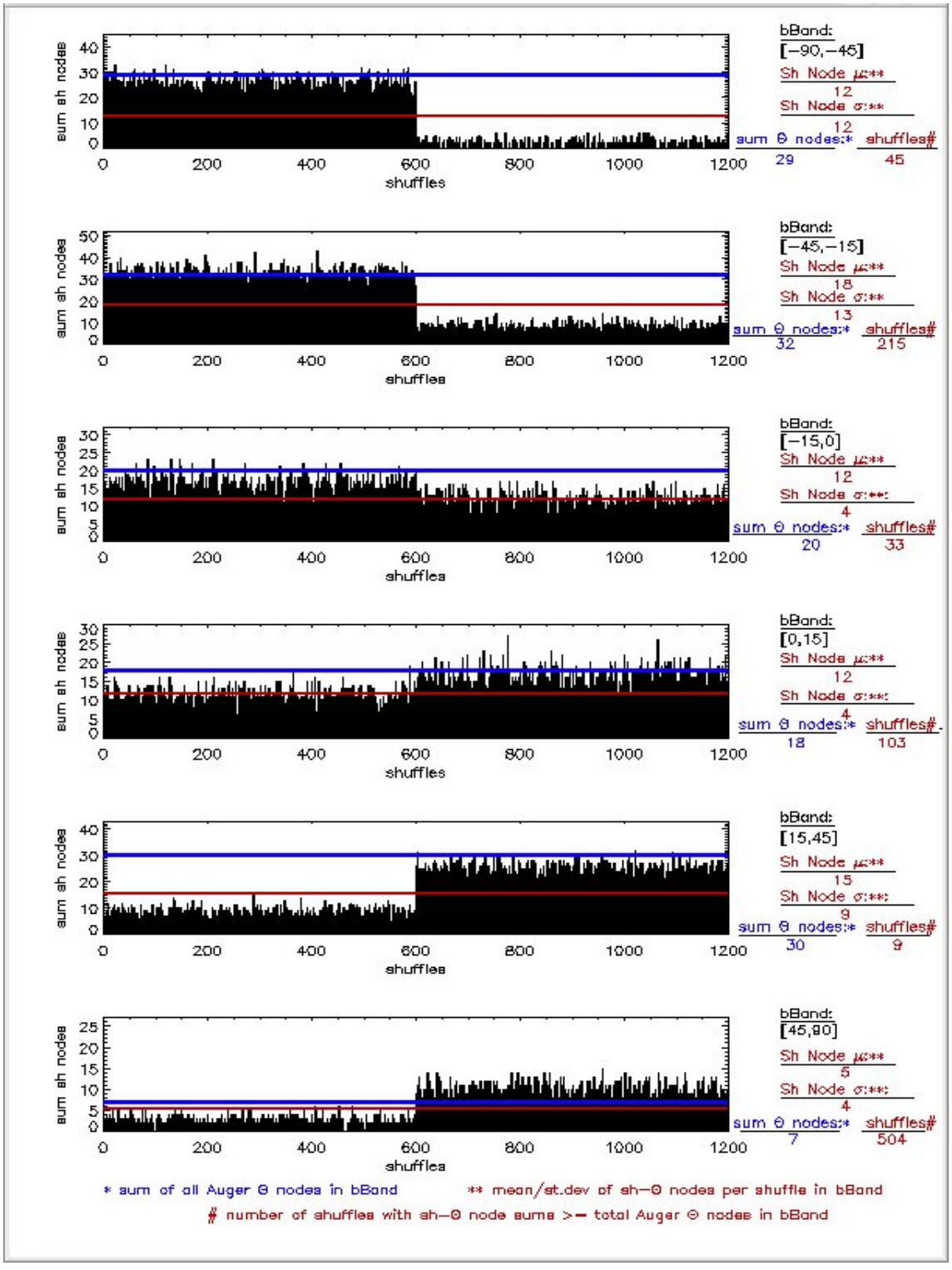

Figure C.9: **b**Band shuffle nodes vs shuffles for E > 60 EeV(current to 8/11/2012).

C.4 Galactic Longitude Quadrants and bBands

bBands $\in ([-135:135])^\circ$ **and** $\in ([-45:45])^\circ$

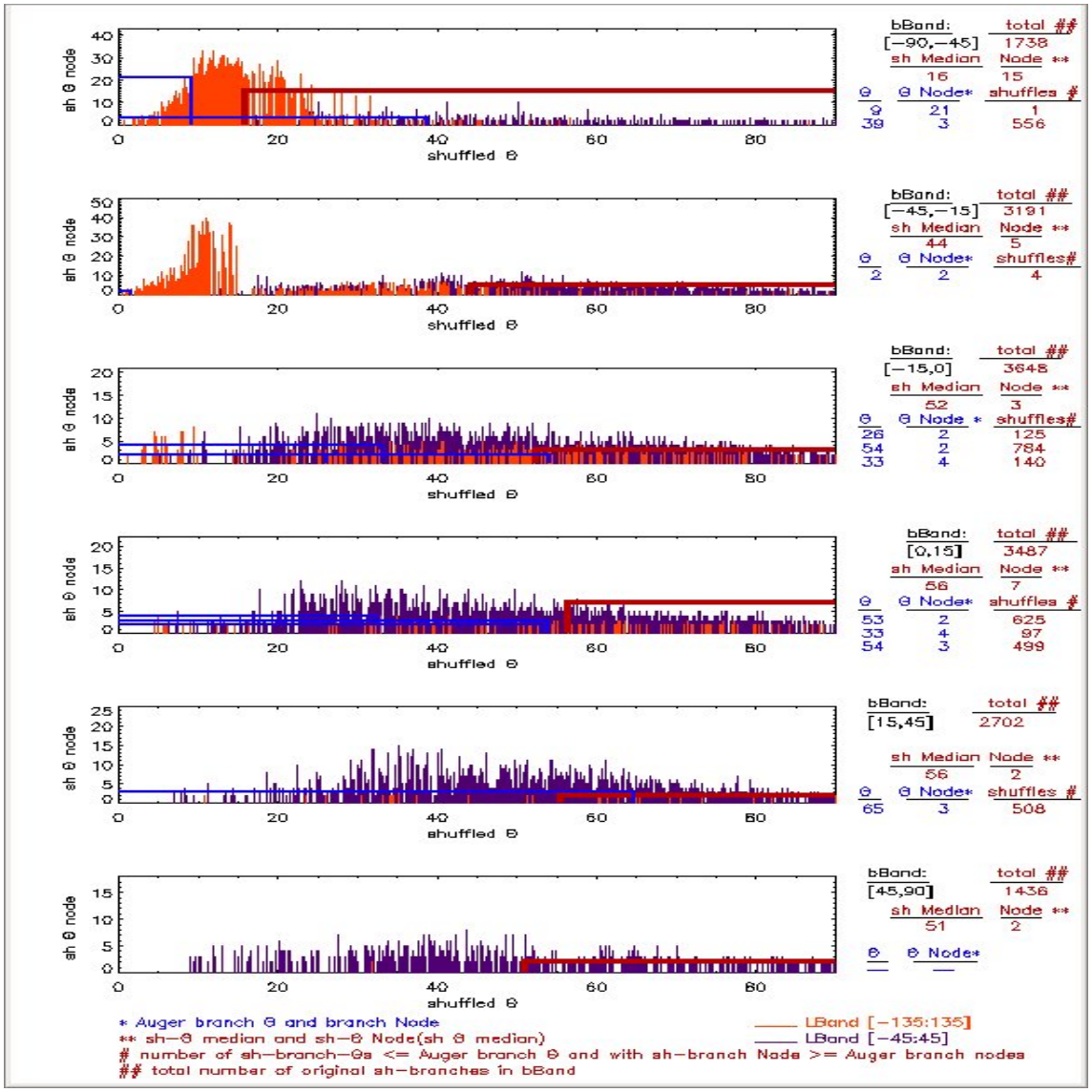

Figure C.10: **b**Band shuffle Θ nodes vs Θ where Gal l $\in ([-135:135]), ([-45:45])^\circ$ for E > 60 EeV
(current to 8/11/2012).

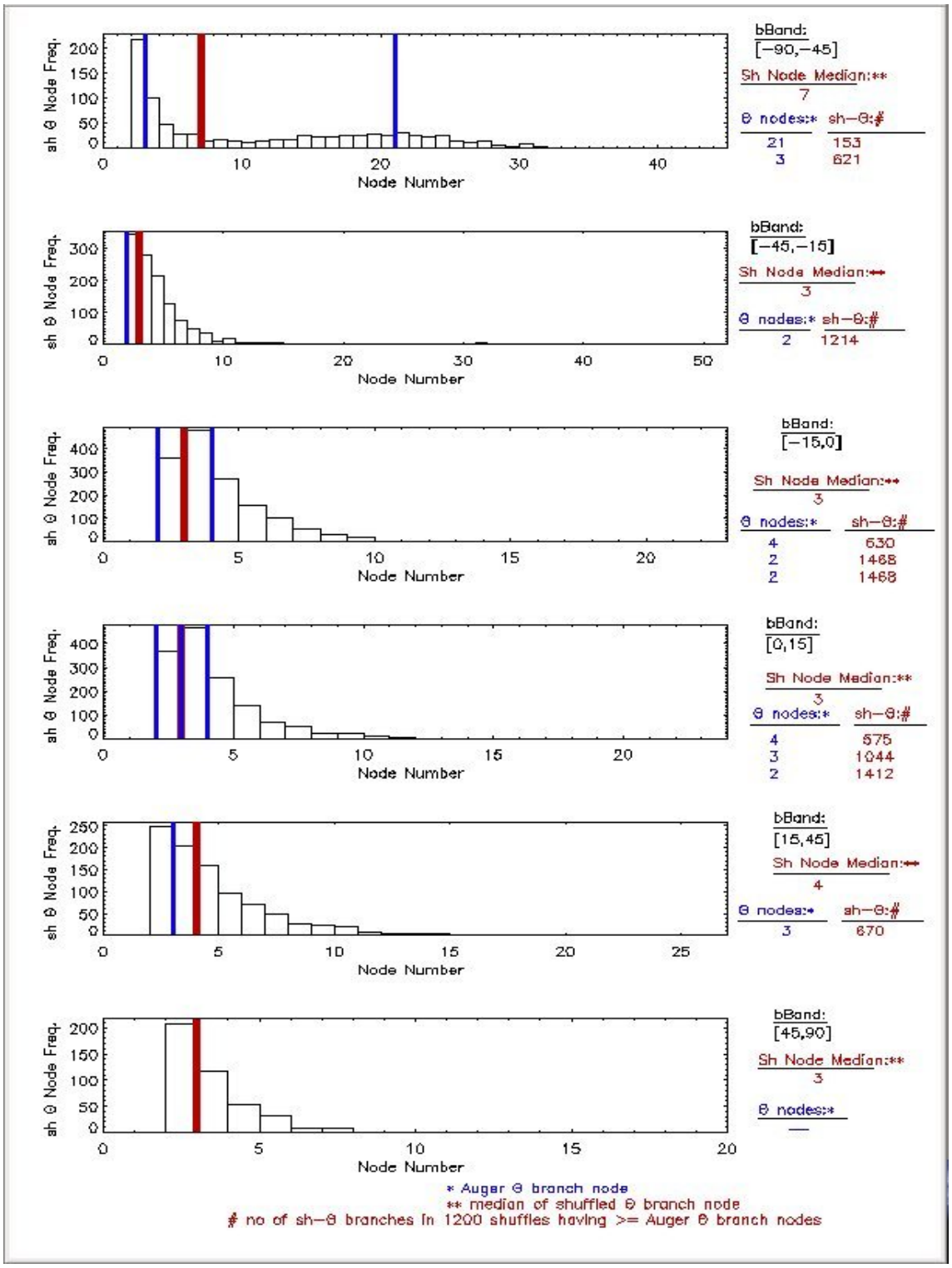

Figure C.11: **b**Band shuffle node frequency vs shuffled node number where Gal l ∈ ([−45 : 45], [−135 : 135])° for E > 60 EeV (current to 8/11/2012).

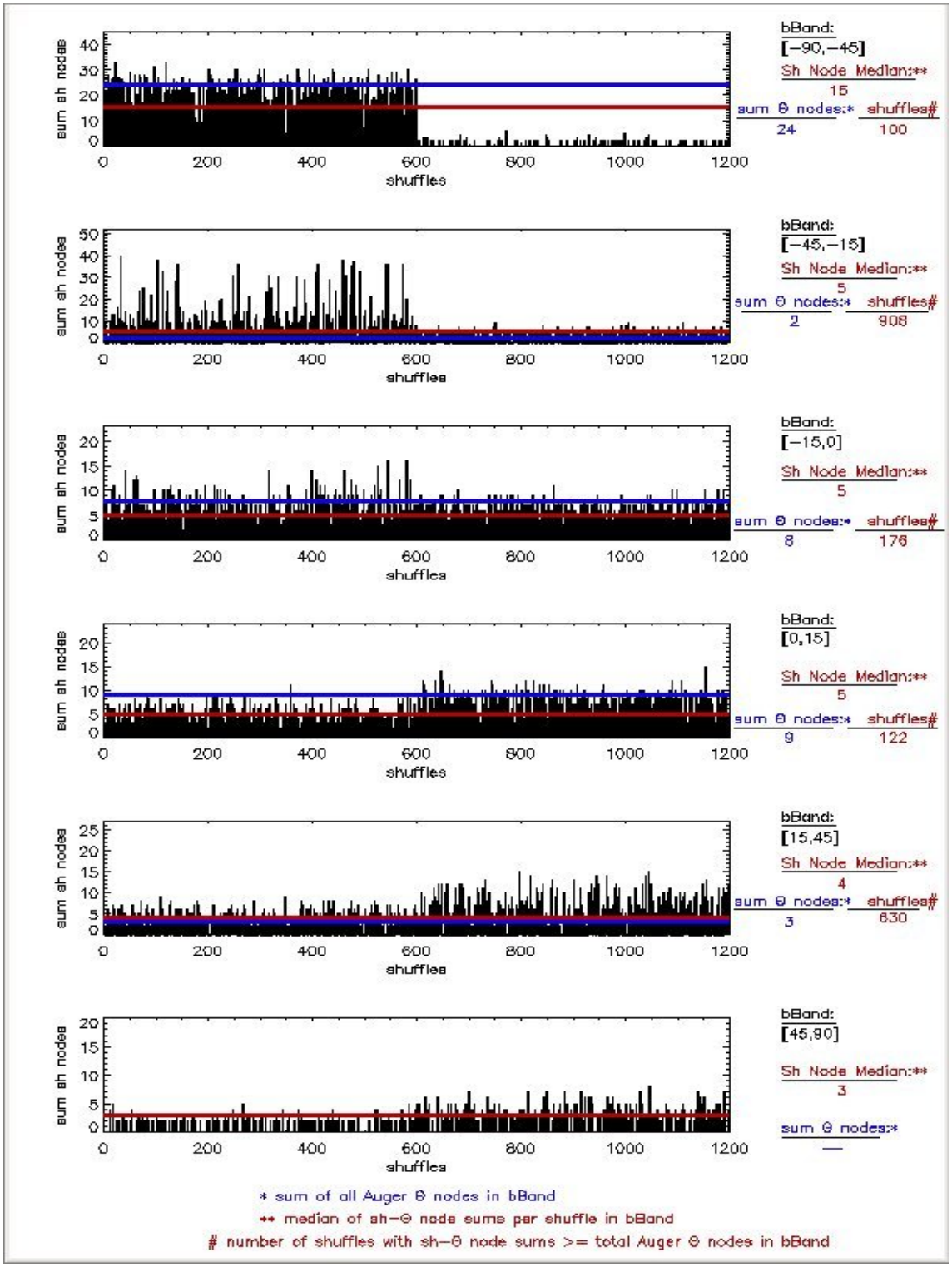

Figure C.12: **b**Band shuffle nodes vs shuffles where Gal l ∈ ([−45 : 45], [−135 : 135])° for E > 60 EeV (current to 8/11/2012).

bBands $\in ([-135 : -45])^\circ$ **and** $\in ([45 : 135])^\circ$

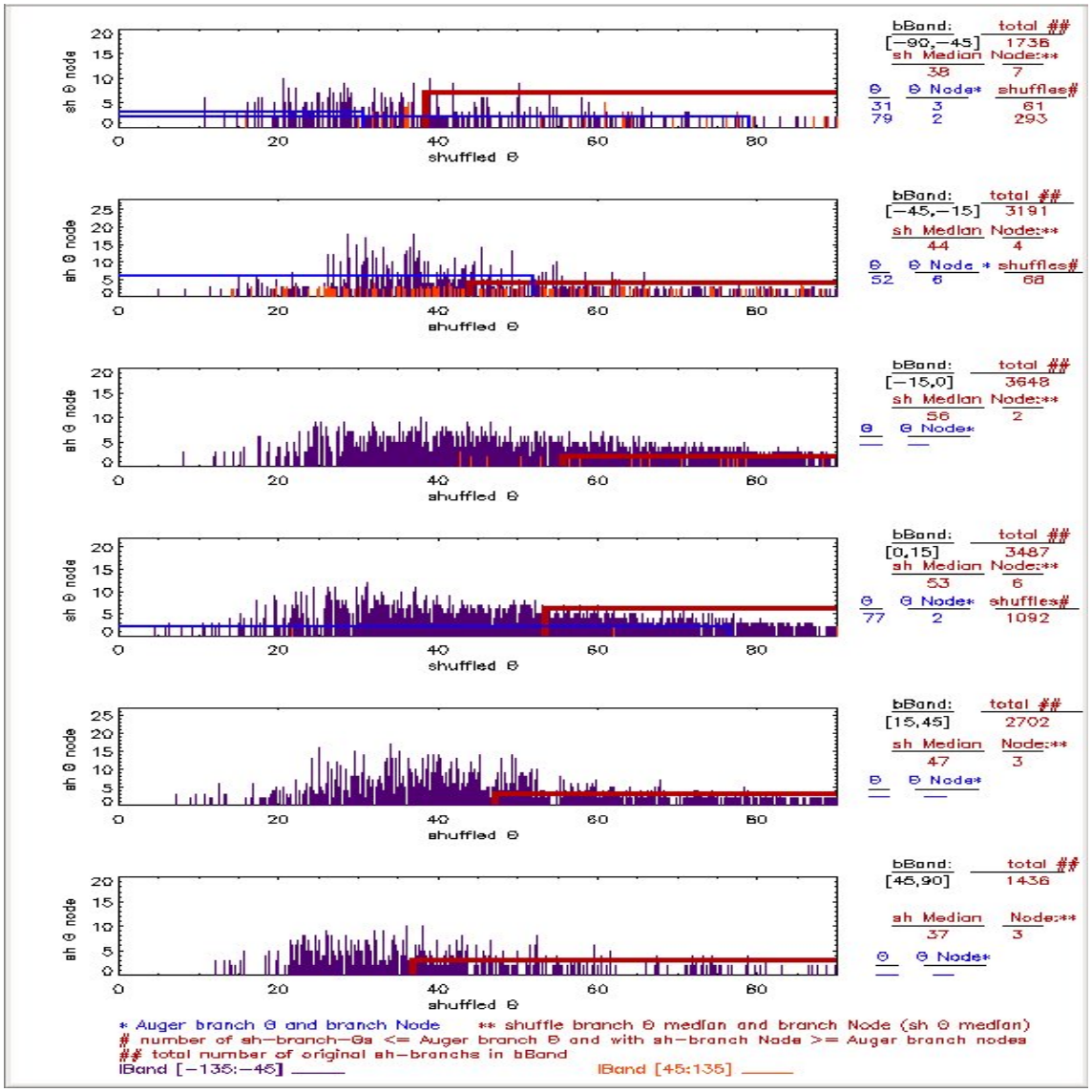

Figure C.13: **b**Band shuffle Θ nodes vs shuffled Θ where Gal l $\in ([-45 : -135], [45 : 135])^\circ$ for E > 60 EeV (current to 8/11/2012).

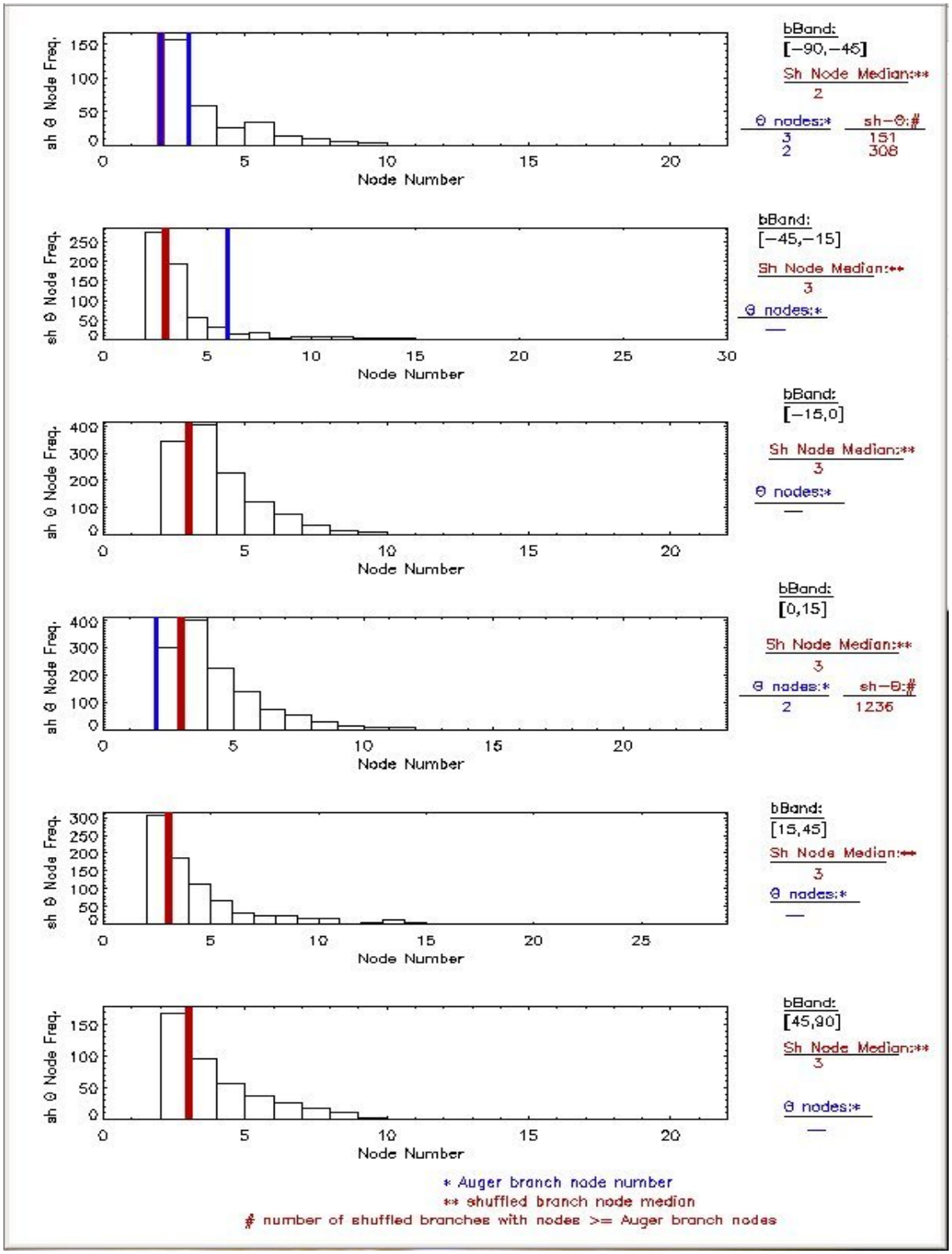

Figure C.14: **b**Band shuffle node frequency vs shuffled node number where Gal $l \in ([-45 : 45], [-135 : 135])^\circ$ for E > 60 EeV (current to 8/11/2012).

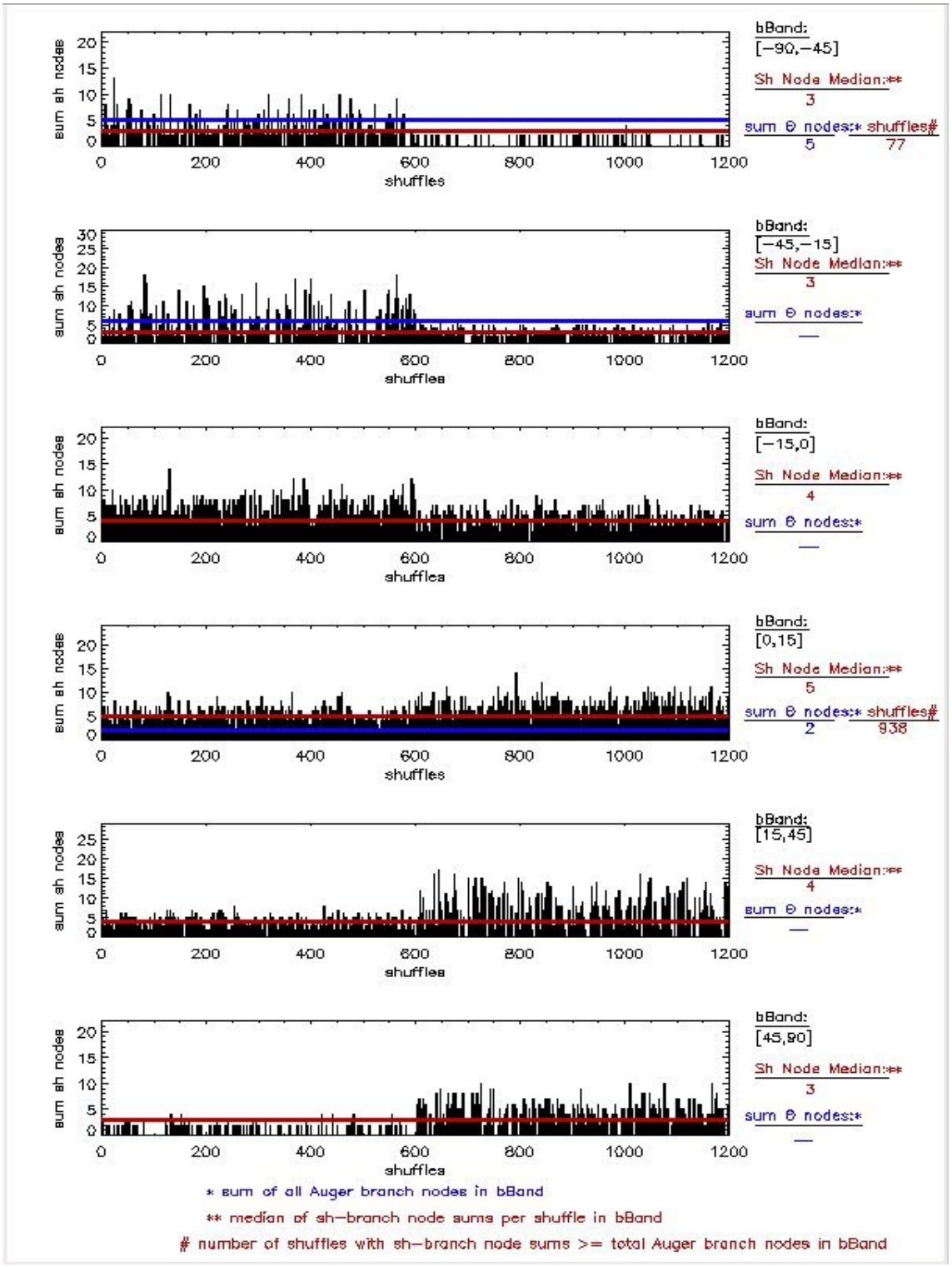

Figure C.15: **b**Band shuffle nodes vs shuffles where Gal l ∈ ([−45 : −135], [45 : 135])° for E > 60 EeV (current to 8/11/2012).

C.5 bBands where $\Delta|(\mathbf{l_1} - \mathbf{l_2})| > 90°$

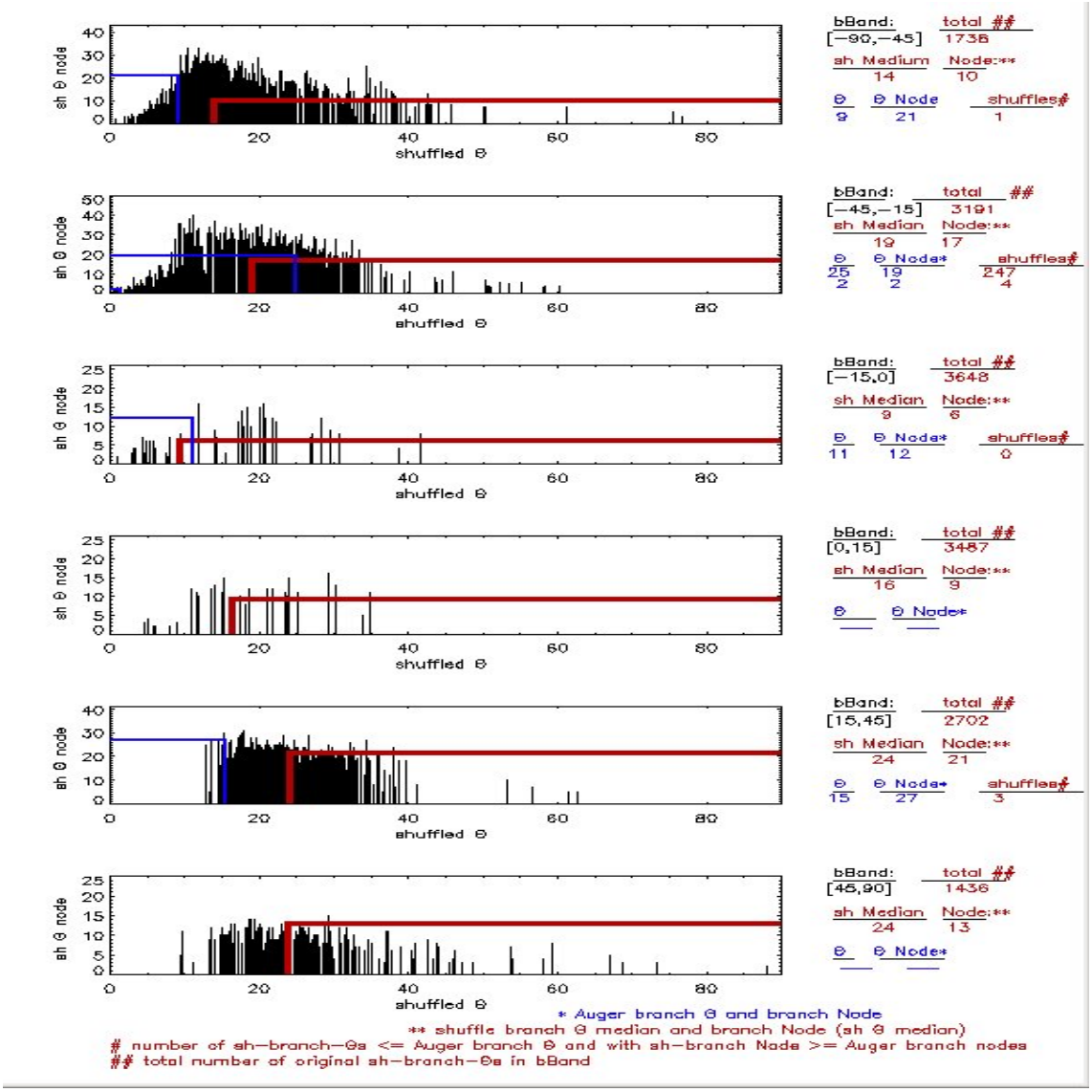

Figure C.16: **b**Band shuffle Θ nodes vs Θ where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ for E > 60 EeV (current to 8/11/2012).

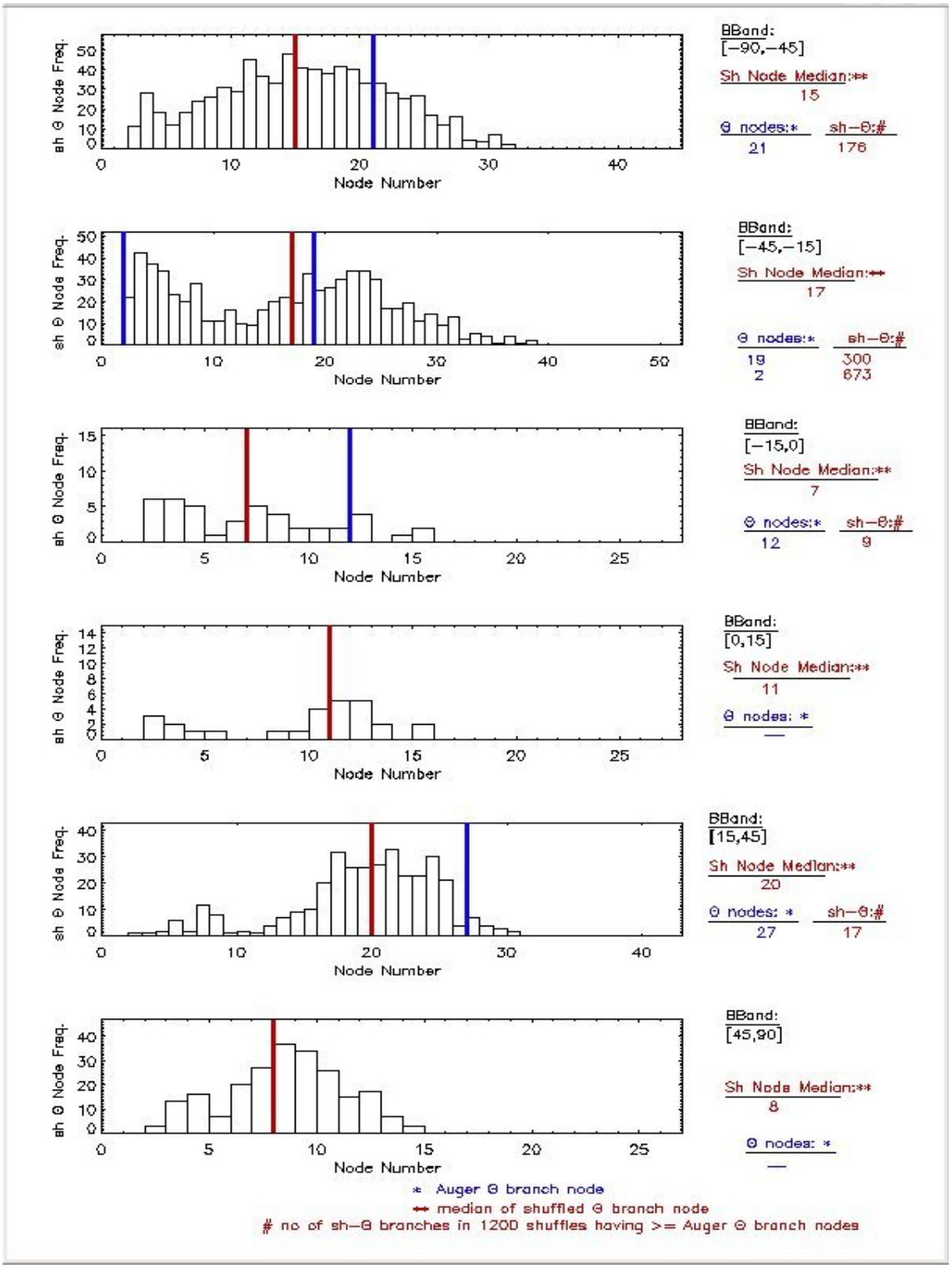

Figure C.17: **b**Band freq shuffle nodes vs shuffle node number where $\Delta|\mathbf{l}_1 - \mathbf{l}_2| > 90°$ for E > 60 EeV (current to 8/11/2012).

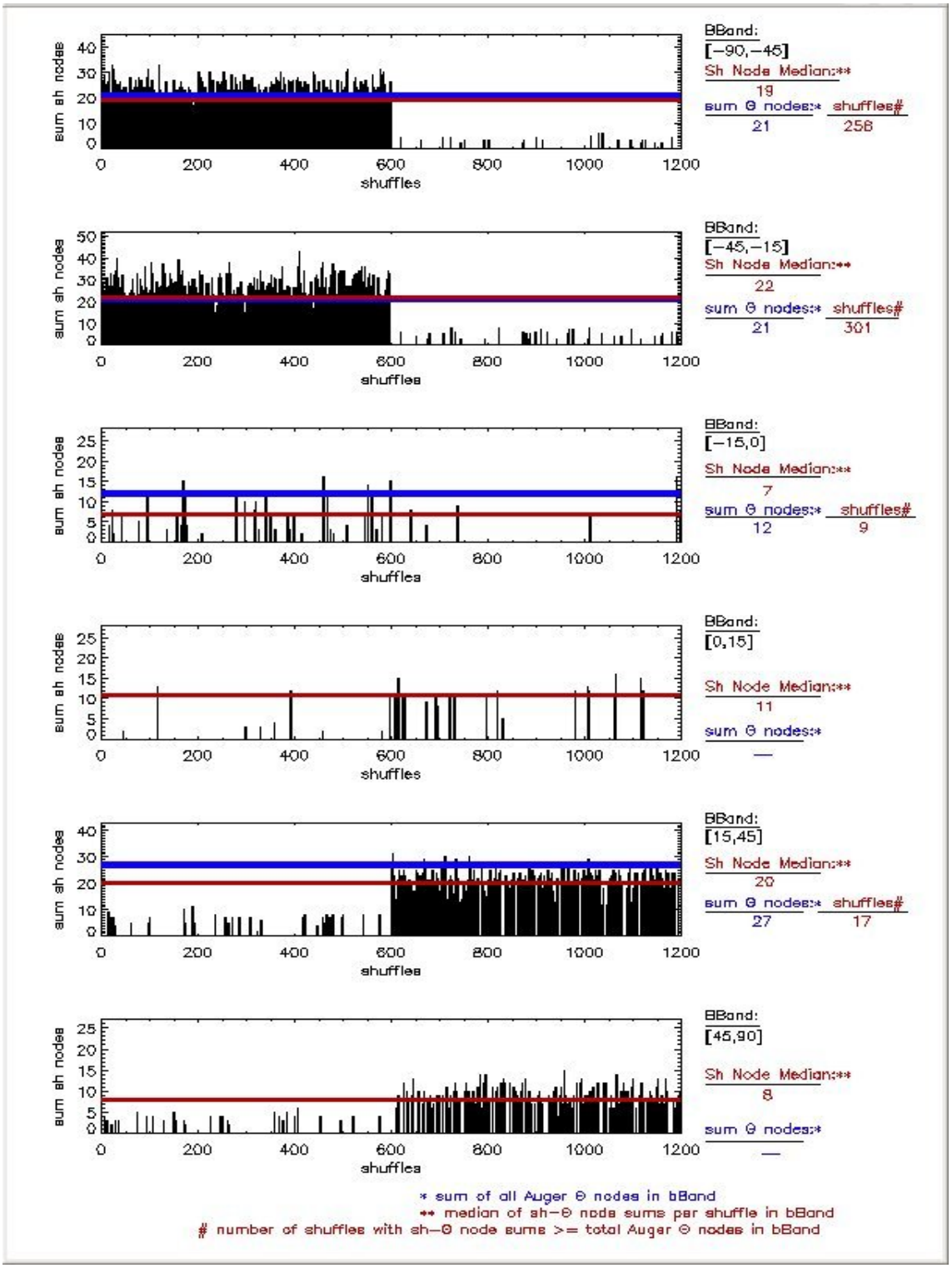

Figure C.18: **b**Band shuffle nodes vs shuffles where $\Delta|\mathbf{l_1} - \mathbf{l_2}| > 90°$ for E > 60 EeV (current to 8/11/2012).

Bibliography

[1] *'NP-Problem'* http://mathworld.wolfram.com/NP-Problem.html.

[2] *'Polynomial Time'* http://mathworld.wolfram.com/PolynomialTime.html.

[3] *'ICS 161: Design and Analysis of Algorithms Lecture notes for March 12, 1996'* https://www.ics.uci.edu/ eppstein/161/960312.html.

[4] cseweb.ucsd.edu/~kube/cls/100/Lectures/lec13.pdf.

[5] M. Fukishima for the TA Collaboration, *'Recent Results from the Telescope Array'* arXiv:1503.06961v1 [astro-ph.HE] 24 Mar 2015.

[6] R. Aloisio, *et al.*, *'Cosmogenic Neutrinos and Ultra-High-Energy Cosmic Ray Models'* arXiv:1505.04020 [astro-ph.HE] 15 May 2015.

[7] C. J. Todero, V. de Souza, P. L. Biermann, *'Cosmic Rays: the Spectrum and Chemical Composition from* 10^{10} *to* 10^{20} *eV'* arXiv:1502.00305v1 [astro-ph.HE] 1 Feb 2015.

[8] K-H. Kampert for the Pierre Auger Collaboration, *'The Pierre Auger Observatory: Selected Results and Future Plans'* arXiv:1404.6515v1 [astro-ph.HE] 19 Apr 2014.

[9] S. P. Bhavsar, *'Bootstrap, Data Permuting and Extreme Value Distributions: Getting the most out of Small Samples'* https://ned.ipac.caltech.edu/level5/Sept01/Bhavsar/paper.pdf.

[10] M. Alpaslan, *'The Cosmic Web Unravelled. A Study of Filamentry Structure in the Galaxy and Mass Assembly Survey'* Ph.D Thesis, http://research-repository.st-andrews.ac.uk/research-repository.st-andrews.ac.uk/.../6/MehmetAlpaslanPhDThesis.pdf.

[11] K. Ptitsyna, S. Troitsky, *'Physical Conditions in Potential Sources of Ultra-High-Energy Cosmic Rays: Updated Hillas plot and radiation-loss constraints'* arXiv:0808.0367v2 [astro-ph] 26 Mar 2010.

[12] T. Dray, C. A. Manogue, *'Electromagnetic Conic Sections'* Oregon State University, Corvallis, 22 May 2001.

[13] *'The Anisotropy Search Program for the Pierre Auger Observatory'* personal communication.

[14] R. Clay, *et al.*, *'Proposal for an Auger Anisotropy Prescription'* personal comminication 2005.

[15] D. Ellison, G. Cassam-Chenaï, *'Radio and X-ray Profiles in Supernova Remnants Undergoing Efficient Cosmic Ray Production'* The Astrophysical Journal, 632, p 920-931, Oct 20 2005.

[16] D. C. Ellison, L. O'C. Drury, J-P. Meyer, *'Galactic Cosmic Rays from Supernova Remnants. II. Shock Acceleration of Gas and Dust'* The Astrophysical Journal, 487, p 197-217, Sep 20 1997.

[17] K. Griesen. *'End to the Cosmic Ray Spectrum? '* Physical Review Letters, Volume 16, Number 17, 25 April 1966.

[18] G. T. Zatsepin, V. A. Kuz'min, *'Upper Limit of the Spectrum of Cosmic Rays'* Journal of Experimental and Theoretical Physics Letters 4: p 78-80, 1966.

[19] P. Blasi, *'On the Origin of Very High Energy Cosmic Rays '* arXiv:astro-ph/0511235v1 8 Nov 2005.

[20] C. Bonifazi, *'Angular resolution of the Pierre Auger Obsevatory'* arXiv:0901.3138v1 [astro-ph.HE] 20 Jan 2009.

[21] H. Takami, *'Propagation of Ultra-high-energy Cosmic Rays in Galactic Magnetic Field'* arXiv:1104.0278v1 [astro-ph.HE] 2 Apr 2011.

[22] J. C. Brown, *'The Magnetic field of the Milky Way Galaxy'* arXiv:1012.2932v1 [astro-ph.GA] 14 Dec 2010.

[23] R. Jansson, G. R. Farrar, *'A New Model of the Galactic Magnetic Field'* arXiv:1204.3662v1 [astro-ph:GA] 16 Apr 2012.

[24] M. S. Pshirkov, P. G. Tinyakov1, P. P. Kronberg, K. J. Newton-McGee, *'Deriving the Global Structure of the Galactic Magnetic Field from Faraday Rotation Measures of Extragalactic Sources'* arXiv:1103.0814v3 [astro-ph:GA] 1 Dec 2011.

[25] J. Han, *'New Knowledge of the Galactic Magnetic Fields'* arXiv:0901.0040v1 [astro-ph:GA] 31 Dec 2008.

[26] Carretti, *et al.*, *'Giant Magnetized Outflows from the Centre of the Milky Way'* arXiv:1301.0512v1 [astro-ph:GA] 3 Jan 2013.

[27] M. Kachelrie, P. D. Serpico, *'The Compton-Getting Effect on Ultra-High Energy Cosmic Rays of Cosmological Origin'* arXiv:astro-ph/0605462v2 30 Aug 2006.

[28] K. Munakata, *et al.*, *'Galactic Anisotropy of* 10 *TeV Cosmic-Ray Intensity Observed by the Tibet Air Shower Array'* Proceedings of the 28th ICRC 2003.

[29] R. Clay. Personal communication.

[30] C. T. Zahn, *'Graph-Theoretical Methods for Detecting and Describing Gestalt Clusters'* SLAC-PUB-672, (Revised), July 1970.

[31] J. L. Bentley, J. H. Friedman, *'Fast Algorithms for Constructing Minimum Spanning Tree in Coordinate Spaces'* SLAC PUB-1665 STAN-CS-75-529 Revised Mar 1977.

[32] R. C. Prim, *'Shortest Connection Networks And Some Generalizations'* Bell System Technical Journal, 1957 Volume 36 p.1389-1401 Nov 1957

[33] A. V. Aho, J. E. Hopcroft, J. D. Ullman, *'The Design and Analysis of Computer Algorithms'*(Addison-Wesley Series in Computer Science and information Processing, 1974).

[34] J. B. Kruskal Jr., *'On the Shortest Spanning Subtree of a Graph and the Travelling Salesman Problem'* Proceedings of the American Mathematical Society p 48-50, 1956.

[35] D. E. Knuth, *'The Art of Computer Programming'* Addison-Wesley Series in Computer Science and information Processing, (Addison-Weslty Publishing Company, Volume 1, 1969).

[36] D. E. Knuth, *'Selected Papers on Analysis of Algorithms'* Centre for the Study of Language and Information, 2000.

[37] D. E. Knuth, *'Big Omicron and Big Omega and Big Theta'* www.phil.uu.nl/datastructuren/10-11/knuth_big_omicron.pdf.

[38] E. W. Dijkstra, *'A Note on Two Problems in Connection with Graphs'*, Numerische Mathematic 1, p 269-271 1959.

[39] J. Kleinberg, E. tárdos, *'Algorithm Design'* (Pearson-Addison-Wesley , WWW, 2005).

[40] E. W.Dijkstra, *'A Short Introduction to the Art of Programming'* WWW, Aug 1971.

[41] N. Wirth, *'Algorithms + Data Structures = Programs'* (Prentice-Hall, Inc, 1976).

[42] J. Cohen,*'Computer-Aided Micro-Analysis of Programs'* Proceedings of 4th Software Engineering, Sep 1979.

[43] Wikipedia, *'optimization'* en.wikipedia.org/wiki/Mathematical_optimization.

[44] M. L. Minsky, *'Computation:Finite and Infinite Machines'* (Prentice-Hall, Inc., 1967).

[45] A. Levitin, *'Introduction to The Design & Analysis of Algorithms'* (Pearson-Addison-Wesley, 2007).

[46] D. R. Hofstadter, *'Gödel, Esher, Bach: An Eternal Golden Braid'* (Penguin Books, 1980).

[47] J. H. Friedman, J. L. Bentley, R. A. Finkel, *'An Algorithm for Finding Best Matches in Logarithmic Expected Time'* ACM Transactions on Mathematical Software, Volume 3, No 3, p 209-226 Sep 1977.

[48] E. Spada, *et al.*, *'Use of the Minimum Spanning Tree Model for Molecular Epidemiologial Investigation of a Nosocomial Outbreak of Hepatitis C Virus Infection'* American Society for Microbiology, May 2004.

[49] A. Han, D. Zhu, *'DNA Computing Model for the Minimum Spanning Tree Problem'* Proceedings of the Eighth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 2006.

[50] http://stackoverflow.com/questions/15866666/what-is-a-acyclic-connected-undirected-graph.acyclic.eps.,

[51] https://www.cs.princeton.edu 04GreedyAlgorithmsII.pdf.,

[52] www.cs.princeton.edu/courses/archive/spr07/cos226/.../mst.pdf.

[53] www.cse.ust.hk/~dekai/271/notes/L07/L07.pdf.

[54] www.comp.dit.ie/rlawlor.

[55] https://greenlees.staff.shef.ac.uk/mas331/MAS331notes2.pdf.

[56] http://pages.cs.wisc.edu/~vernon/cs367/notes/11.PRIORITY-Q.html.

[57] R. Clay for the Pierre Auger Collaboration, *'The Anisotropy Search Program for the Pierre Auger Observatory'* Personal communication.

[58] IceCube Collaboration, *'Evidence for High-Energy Extraterrstrial Neutrinos at the IceCube Detector'* Science Volume 342 22 Nov 2013.

[59] G. Matthaie, V. Versi, *'The Cosmic Ray Energy Specrtum Measured with the Auger Observatory'* Personal communication.

[60] D. Hooper, S. Sarkar, A. M. Taylor, *'The Intergalactic Propagation of Ulta-High Energy Cosmic Ray Nuclei'* Astroparticle Physics 27 p 199-212 2007.

[61] G. Giacinti, M. Kachelrieb, D. V. Semikoz, G. Sigl, *'Cosmic Ray Anisotropy as Signature for the Trasnsition from Galactic to Extragalactic Cosmic Rays'* arXiv:1112.5599 [astro-ph:HE] 16 Jul 2012.

[62] T. J. L. McComb, K. E. Turver, *'The Elongation Rate Theorem and its Application'*, Proceedings of the 17th ICRC, Paris, France, p 234-237 July 13-25, 1981.

[63] R. W. Clay, B. R. Dawson, *et al.*, *'Cosmic Rays from the Galactic Centre'* www.physics.adelaide.edu.au/astrophysics/astro_papers/galcenter.ps.

[64] L. J. Gleeson, W. I. Axford, *'Astrophysics and Space Science'* Volume 2, Issue 4, p 431-437 1968.

[65] T. Wibig, A. W. Wolfendale, *'Cosmic Rays of the Highest Energies'*, Romanian Reports in Physics, Volume 59, No. 2, p 707 2007.

[66] T. K. Gaisser, *'Cosmic-Ray Energy Spectrum from* 10^{15} *to* 10^{18} *eV'*, Letters to Nature, Nature 248, p 122 - 124 8 Mar 1974.

[67] J. Abraham, *et al.*, *'The Fluorescence Detector of the Pierre Auger Observatory'* arXiv:0907.4282v1 [astro-ph.IM] 24 Jul 2009.

[68] R. Ramaty, *et al.*,*'Rain of Fire'* The Sciences November 1999.

[69] T. Stanev, *'High Energy Cosmic Rays'* (Springer-Praxis, 2004).

[70] P. Sommers, *et al.*, *'Mapping the U.H.E. sky in Search of Point-Sources'* Nuclear Physics B.14A p 291-298 1990.

[71] A. A. Watson, *'The mass composition of cosmic rays above* $10^{17} eV'$ based on a talk given at the CRIS 2004 meeting: 'GZK and Surroundings': Catania 31 May - 4 June 2004.

[72] B. Voigt, *'Cosmic Ray Composition'* WWW 7 Jun. 2005.

[73] J. A. Simpson, *'Elemental and Isotropic Composition of the Galactic Cosmic Rays'* Annual Review of Nuclear and Particle Science, p 323-381 1983.

[74] A. N. Witt, *'The Chemical Composition of the Interstellar Medium'* Philosophical Transactions:Mathematical, Physical and Engineering Sciences, Volume 359, no 1787, Origin and early Evolution of Solid Matter in the Solar system p 1949-1959 Published by: The Royal Society Oct 15, 2001.

[75] L. H. Aller, *'Solar Abundances and the Role of Nucleogenesis in the Low-to-Medium Mass Stars in the Galaxy'* Presented at the Suess Sympoium University of California, San Diego Dec 14, 1984.

[76] J. R. Hörandel, *et al.*, *'Galactic Cosmic Rays and the Knee - Results from the KASCADE experiment'* astro-ph/0509253v1 9 Sep 2005.

[77] P.Bhattacharjee, G. Sigl, *'Origin and Propagation of Extremely High Energy Cosmic Rays'* astro-ph/9811011v2 23 Aug 1999.

[78] science.nasa.gov/headlines/y2005/07oct_afraid.htm.

[79] A. M. Hillas, *'Cosmic Rays:Recent Progress and some Current Questions'* arXiv:astro-ph/0607109v2 5 Sep 2006.

[80] A. M. Hillas, *'The Origin of Ultra-High-Energy Cosmic Rays'*, Annual Review of Astronony and Astrophysics 22: p 425-44 1984.

[81] A. M. Hillas, *'Can Diffusuve Shock Acceleration in Supernova Remnants Account for High-Energy Galactic Cosmic Rays?'* Journal of Physics G: Nuclear Particle Physics 31 R95-R131 2005.

[82] H. Rebel, *'What Do We Expect to Learn from Very High Energy Cosmic Ray Observations ?'* arXiv:stro-ph/0109520v1 27 Sep 2001.

[83] T. Antoni, *et al.*, *'KASCADE Measurements of Energy Spectra of Elemental Groups of Cosmic Rays:Results and Open Problems'* arXiv:astro-ph/0505413v1 19 May 2005.

[84] J. R. Hörandel, *et al.*, *'Results from the KASCADE, KASCADE-Grande, and LOPES experiments'*, Journal of Physics: Conference Series **39** p 463-470 2006.

[85] D.Allard, E.Parizot, A.V.Olinto, *'On the Transition from Galactic to Extragalactic Cosmic-Rays:Spectral and Compositional Features from Two Opposite Scenarios'* arXiv:astro-ph/0512345 v1 14 Dec 2005.

[86] D. R. Bergman, J. W. Belz, *'Cosmic Rays: The Second Knee and Beyond'* arXiv:0704.3721v1 [astro-ph] 27 Apr 2007.

[87] J. R. Hörandel , N. N. Kalmykov, A. V. Timokhin, *'The End of the Galactic Cosmic-Ray Energy Spectrum – a Phenomenological View'*, Journal of Physics: Conference Series 47 p 132-141 2006.

[88] J. R. Hörandel, *'On the Knee in the Energy Spectrum of Cosmic Rays'* arXiv:astro-ph/0210453v1 20 Oct 2002.

[89] J. R. Hörandel, *'Overview on Direct and Indirect Measurements of Cosmic Rays'* arXiv:astro-ph/0501251v1 13 Jan 2005.

[90] J. R. Hörandel, *'Models of the Knee in the Energy Spectrum of Cosmic Rays'*, Astroparticle Physics 21 241-265 2004.

[91] J. R. Hörandel, N. N. Kalmykov, A. V. Timokhin, *'Propagation of Super-High-Energy Cosmic Rays in the Galaxy'*, Astroparticle Physics 27 p 119-126 2007.

[92] J. R. Hörandel, *'The Origin of Galactic Cosmic Rays'* http://particle.astro.kun.nl.

[93] bf 1034 J. Blümer, R. Engel, and J. R. Hörandel *'Cosmic Rays from the Knee to the Highest Energies'* Progress in Particle and Nuclear Physics 63(2), p 293-338 2009.

[94] J. Bluümer, R. Engel, J. R. Hörandel, *'Cosmic Rays from the Knee to the Highest Energies'*, arXiv:0904.0725v1 [astro-ph].HE] 4 Apr 2009.

[95] J. R. Hörandel, *'Cosmic Rays at the Highest Energies'* Progress in Particle and Nuclear Physics 64 p 351-359 2010.

[96] J. R. Hörandel, *'Cosmic Rays from the Knee to the Second Knee:* 10^{14} *to* 10^{18} *eV'* arXiv:astro-ph/0611387v1 13 Nov 2006.

[97] A. V. Olinto, *'Cosmic Rays at the Highest Energies'* arXiv:1201.4519v2 [astro-ph.HE] 15 Mar 2012.

[98] V. Berezinsky, A. Gaziziv, S. Grigorieva, *'Signatures of AGN model for UHECR'* arXiv:astro-ph/025v2 4 Oct 2002.

[99] T. K. Gaisser, *'Cosmic Rays and Particle Physics'* (Cambridge University Press 1990).

[100] M. S. Longair, *'High Energy Astrophysics'* (Cambridge University Press Volume 2, second edition 1994).

[101] E. N. Parker. *'Cosmic Rays and Their Formation of a Galactic halo'* Astrophysical Journal, Volume 142 p 584 2 Apr 1965.

[102] V. Ptuskin, *et al.*, *'Spectrum of Galactic Cosmic Rays Accelerated in Supernova Remnants'* Astrophysical Journal p31 718 2010 .

[103] A. W. Strong, I. V. Moshalenko, V.S. Ptuskin, *'Cosmic-Ray Propagation and Interactions in the Galaxy'* Annual Review of Nuclear and Particle Science 57 p 285-327 2007.

[104] R. Aloisio, V. Berezinsky, P. Blasi, A. Gaziziv, S. Grigorieva, B. Hnatyk, *'A Dip in the UHECR Spectrum and the Transition from Galactic to Extragalactic Cosmic Rays'* arXiv:astro-ph/0608219v2 18 Oct 2006.

[105] M. Ahlers, L. A. Anchordoqui, S. Sarkar, *'Neutrino Diagnostics of Ultra-High Energy Protons'* arXiv:0902.3993v1 24 Feb 2009.

[106] The Pierre Auger Collaboration, *'Observation of Anisotropy of the Highest Energy Cosmic Rays from their Correlation with the positions of Nearby Active Galactic Nuclei'*, preprint submitted to Elsevier 22 Oct 2007.

[107] The Pierre Auger Collaboration, *'Observation of the Suppression of the Flux of Cosmic Rays above* 4×10^{19} *eV'* arXiv:0806.4302v1 [astro-ph] 26 Jun 2008. C

[108] The Pierre Auger Collaboration, *'The Fluorescence Detector of the Pierre Auger Observatory'* arXiv:0907.4282v1 [astro-ph.IM] 24 Jul 2009.

[109] The Pierre Auger collaboration, *'Update on the Correlation of the Highest Energy Cosmic rays with Nearby Extragalactic Matter'* arXiv:1009.1855v2 [astro-ph.HE] 29 Sep 2010.

[110] The Pierre Auger collaboration, *'Measurement of the Energy Spectrum of Cosmic Rays above* 10^{18} *eV Using the Pierre Auger Observatory'* arXiv:1002.1975v1 [astro-ph.HE] 9 Feb 2010.

[111] The Pierre Auger Collaboration, *'Properties and Performance of the Prototype Instrument for the Pierre Auger Observatory'* Nuclear Instruments and Methods in Physics Research A Volume 523 p 50-95 May 2004.

[112] The Pierre Auger Collaboration, *'Anistropy Studies around the Galactic Centre at EeV Energies with the Auger Observatory'* Astroparticle Physics Volume 27, issue 4, p 244-253 Apr 2007.

[113] The Pierre Auger Collaboration, *'A search for Anisotropy in the Arrival Directions of Ultra High Energy Cosmic Rays Recorded at the Pierre Auger Observatory'* arXiv:1210.3602 [astro-ph.HE] 12 Oct 2012.

[114] The Pierre Auger Collaboration, *'Search for Signatures of Magnetically-Induced Alignment in the Arrival Directions Measured by the Pierre Auger Observatory'* arXiv:1111.2472v1 [astro-ph.HE] 10 Nov 2011.

[115] The Pierre Auger Collaboration, *'Search for First Harmonic Modulation in the Right Ascension Distribution of Cosmic Rays Detected at the Pierre Auger Observatory'*, arXiv:1103.2721v1 [astro-ph.HE] 14 Mar 2011.

[116] P. Sommers for the Pierre Auger Collaboration *"* Proceedings of the 29th ICRC, Pune, vol.7, p 387 2005.

[117] The Pierre Auger Collaboration, *'Measurement of the Depth of Maximum of Extensive Air Showers above* 10^{18} *eV'* arXiv:1002.0699v1 [astro-ph:HE] 3 Feb 2010.

[118] The Pierre Auger Collaboration, *'The Pierre Auger Observatory II: Studies of Cosmic Ray Composition and Hadronic Interaction models'* arXiv:1107.4804 [astro-ph.HE] 32nd ICRC, Beijing 2011.

[119] The Pierre Auger Collaboration, *'The Fluorescence Detector of the Pierre Auger Observatory'* arXiv:0907.4282v1 [astro-ph.IM] 24 Jul 2009.

[120] The Pierre Auger Collaboration, *'Correlation of the Highest-Energy Cosmic Rays with the Positions of Nearby Active Galactic Nuclei'* arXiv:0712.2843 [astro-ph] 23 Jun 2008.

[121] M. Unger, B.Dawson, *et al.*, *'Reconstruction of Longitudunal Profiles of Ultra-High Energy Cosmic Ray Showers From Fluorescsnce and Cherenkov Light Measurements'* arXiv:0801.4309v1 [astro-ph] 28 Jan 2008.

[122] C. Dobrigkeit for The Pierre Auger Collaboration, *'Recent Results from the Pierre Auger Observatory'* Astronomical Notes Volume 335, issue 6-7, p 573-580 Aug 2014.

[123] R. Šmída for The Pierre Auger Collaboration, *'Results from The Pierre Auger Observatory'*, arXiv:1109.4743v2 [astro-ph.HE] 18 Oct 2011.

[124] K-H. Kampert, M. Unger, *'Measurements of the Cosmic Ray Composition with Air Shower experiments'* arXiv:1201.0018v2 [astro-ph.HE] 19 Feb 2012.

[125] KASCADE-Grande and LOPES Collaboration, *'Recent Results from KASCADE-Grande and LOPES'* astro.uni-wuppertal.de/~kampert/Publications.../Kampert-CRIS08.pdf.

[126] A. Letessier-Selvon, *'Ultrahigh Energy Cosmic Rays'* arXiv:1103.0031v1 [astro-ph.HE] 28 Feb 2011.

[127] A. Letessier-Selvon *'Establishing The GZK Cutoff With Ultra High Energy Tau Neutrinos'* arXiv:astro-ph/0009444v1 27 Sep 2000. .

[128] J. Matthews *'A Heitler model of extensive air showers'* Astroparticle Physics 22 p 387-397 2005.

[129] J. van Buren, *et al.*, *'Muon Size Spectrum measured by KASCADE-Grande'* Proceedings of the 29th ICRC Pune 6, p 301-304 2005.

[130] M. Bertaina, *et al.*, *'Results from KASCADE-grande'* Nuclear Instruments and Methods In Physics Research A 2012.

[131] M. Amenomori, *et al.*, *'The Energy Spectrum of All-Particle Cosmic Rays Around the Knee Region Observed with the Tibet-III Air-Shower Array'* arXiv:0803.1005v1 [astro-ph] 6 Mar 2008.

[132] V. Ptuskin, O. Strelnikova, L. Sveshnikova, *'On Leaky-Box Approximation to GALPROP'* ŁÓDŹ Proceedings of the 31st ICRC 2009.

[133] T. K. Gaisser, T. Stanev, *'24 Cosmic Rays'* Revised August 2011.

[134] R. A. and D. Boncioli, *'Ultra High Energy Cosmic Rays: Anisotropies and Spectrum'* arXiv:1002.4134v2 [astro-ph.HE] 4 Jun 2011.

[135] S. P. Knurenko, A. Sabourov, *'The Depth of Maximum Shower Development and its Fluctuations:Cosmic Ray Mass Composition at* $E_0 \geq 10^{17}$ *eV'* Astrophysical Space Science Transactions 7, p 251-255 2011.

[136] P. F. San-Luis for the Pierre Auger Collaboration, *'The Distribution for Shower Maxima of UHECR Air Showers'* Proceedings of the 32nd ICRC, Beijing 2011.

[137] D. G-Pinto for the Pierre Auger Collaboration, *'Measurements of the Longitudunal Development of Air Showers with the Pierre Auger Observatory'* Proceedings of the 32nd ICRC, Beijing 2011.

[138] F. Salamida, *'Updata on the measurement of the CR energy spectrum above* 10^{18} *eV made using the Pierre Auger Observatory'* Proceedings of the 32nd ICRC, Beijing 2011.

[139] H. P. Dembinski for the Pierre Auger Collaboration, *'The Cosmic Ray Spectrum above* $4 \times 10^{18} eV$ *as Measured with Inclined Showers Recorded at the Pierre Auger Observatory'* Proceedings of the 32nd ICRC, Beijing 2011.

[140] G. Rodriguez for the Pierre Auger Collaboration, *'Reconstruction of Inclined Showers at the Pierre Auger Observatory: Implications for the Muon Content'* Proceedings of the 32nd ICRC, Beijing 2011.

[141] D. Veberic for the Pierre Auger Collaboration, *'The End of the Cosmic Ray Spectrum'* arXiv:1110.0615v2 [astro-ph.HE] 10 Oct 2011.

[142] R. U. Abbasi, *et al.*, *'First Observation of the Greisen-Zatsepin-Kuzmin Suppression'* The High Resolution Fly's Eye Collaboration. arXiv:astro-ph/0703099v2 15 Feb 2009.

[143] R. U.Abbasi, *et al.*, *'Indications of Proton-Dominated Cosmic Ray Composition above* 1.6 *EeV'*, The High Resolution Flys Eye Collaboration. arXiv:0910.4184v2 [astro-ph.HE] 3 Apr 2010.

[144] R. U. Abbasi, *et al.*, *'Alternative Methods to Finding Patterns in HiRes Stereo Data'* arXiv:astro-ph/0702361v1 13 Feb 2007.

[145] R. U. Abbasi, *et al.*, *'A Search for Arrival Direction Clustering in the HiRes-I Monocular Data above* $10^{19.5}$ *eV'* arXiv:astro-ph/0404366v2 27 May 2004.

[146] R. U. Abbasi, *et al.*, *'Search for Point-Sources of Ultra-High-Energy Cosmic Rays Above* 4.0×10^{19} *eV Using a Maximum Liklihood Ratio Test'* Astrophysical Journal p 164-170 2005.

[147] R. U. Abbasi, *et al.*, *'Search for Correlations between HiRes Stereo Events and Active Galactic Nuclei'* arXiv:astro-ph/0804.0302v2 15 Aug 2008.

[148] J. Belz for the HiRes Collaboration, *'Overview of Recent HiRes Results'* Nuclear Physics B (Proc. Suppl.) 190 2009.

[149] P. Sommers, Astropart. Phys. 14 (2001) 271 [arXiv:astro-ph/0004016].

[150] E. G. Berezhko, *'Astroparticle Studies in Yakutsk'* Yu. G. Shafer Institute of Cosmophysical Research and Aeronomy Yakutsk, Russia.

[151] Y. T. Sunesada for The Telescope Array Collaboration, *'Highlights from Telescope Array'* arXiv:1111.2507v1 [astro-ph.HE] 10 Nov 2011.

[152] P. Sokolsky, G. B. Thomson, *'Highest Energy Cosmic Rays and Results from the HiRes Experiment'* arXiv:0706.1248v1 [astro-ph] 8 Jun 2009.

[153] M. Takeda, *et al.*, *'Energy Determination in the Akeno Giant Air Shower Array Experiment'* Astroparticle Physics 19, 447, arXiv:astro-ph/029422 13 Nov 2002.

[154] M. Takeda, *et al.*, *'Small-Scale Anisotropy of Cosmic Rays above* 10^{19} *eV observed with the Akeno Giant Air Shower Array'* arXiv:9902239v2 [astro-ph:] 21 June 2000.

[155] X.-J. Bi, Z. Cao, Ye. Li, Q. Yuan *'Testing Lorentz Invariance with Ultra High Energy Cosmic Ray Spectrum'* arXiv:0812.0121v1[astro-ph] 30 Nov 2008.

[156] F. Aharonian, *et al.*, *'The Energy Spectrum of Cosmic-Ray electrons at TeV Energies'*, arXiv:0811.3894v2 [astro-ph] 6 Jan 2009.

[157] S. Hummera, M. Maltonib, W. Winterc, C.Yagunad *'Energy Dependent Neutrino Flavor Ratios from Cosmic Accelerators on the Hillas plot'* arXiv:1007.0006v3 [astro-ph].HE] 21 Jun 2011.

[158] R. Aloisio, V. Berezinsky, A. Gazizov *'Disappointing model for Ultrahigh-Energy Cosmic Rays'*, arXiv:1105.1111v1 [astro-ph.HE] 5 May 2011.

[159] R. Aloisio R, V. Berezinsky V, A Gazizov, *'Ultra High Energy Cosmic Rays: The Disappointing Model'* arXiv:0907.5194v2 [astro-ph:HE] 16 Feb 2011.

[160] M. Kachelrie *'Anisotropies and Clustering of Extragalactic Cosmic rays'* arXiv:astro-ph/0610862v1 29 Oct 2006.

[161] R. Beck *'Galactic and Extragalactic Magnetic Fields'*, arXiv:0810.2923v4 [astro-ph] 12 Jan 2009.

[162] R. Beck, *et al.*, *'Observation of Galactic Magnetic Fields and Confrontation with Dynamo Models'*, mhd.ens.fr/IHP09/Paris09.pdf.

[163] R. M. Crocker *et al.*, *'The AGASA/SUGAR Anisotropies and TeV Gamma Rays from the Galactic Center: A Possible Signature of Extremely High-energy Neutrons'* arXiv:astro-ph/0408183v1 10 Aug 2004.

[164] R. M. Crocker, *et al.*, *'Steady-state Hadronic Gamma-Ray Emission from 100-Myr-Old Fermi-Bubbles'* arXiv:1312.0692v3 [astro-ph.HE] 17 Aug 2014.

[165] K. C. Anand, R. R. Daniel, S. A. Stevens, *'On the Existence of the Galactic Radio Halo'* Proceedings of the Indian Academy of Sciences, Section A, 68 (5). p 219-227. ISSN 0370-0089 1968.

[166] M. Erdmann, P. Schiffer, *'A Method of Measuring Cosmic Magnetic Fields with Ultra High Energy Cosmic Ray Data'* arXiv:0904.4888v2 [astro-ph:HE] 2 Feb 2010.

[167] P. Schiffer for the Pierre Auger Collaboration, *'Measurement of Energy-Energy-Correlations with the Pierre Auger Observatory'* Proceedings of the 32nd ICRC Beijing 2009.

[168] G. Giacinti, X. Derkx, D. V. Semikoz, *'Search for Single Sources of Ultra High Energy Cosmic Rays on the Sky'* arXiv:0907.1035v2 [astro-ph:HE] 28 Jun 2010.

[169] G. Giacinti, D. V. Semikoz, *'Method to Look for Imprints of Ultrahigh high Energy Nuclei Sources'* arXiv:1011.6333v2 [astro-ph:HE] 8 Apr 2011.

[170] D. V. Semikoz, *'Ultra-High Energy Nuclei Source in the Direction to Virgo Cluster.'* arXiv:1109.3879 [astro-ph:HE] 16 Jul 2012.

[171] W. D. Apel for the KASCADE-Grande Collaboration, *et al.*, *'The Spectrum of High Energy Cosmic Rays Measured with KASCADE-Grande'* arXiv:1206.3834v1 [astro-ph:HE] 18 Jun 2012.

[172] M. Teshima, *et al.*, *'Anisotropy of Cosmic-Ray Arrival Directions at* 10^{18} *eV Observed by AGASA'* Proceedings of ICRC 2001.

[173] D. S. Gorbunov, P. G. Tinyakov, I. I. Tkachev, S. V. Troitsky, *et al.*, *'Evidence for a Connection between the γ-Ray and the Highest Energy Cosmic-Ray Emissions by BL Lacertae Objects'* Astrophysical Journal 577 l p 93-96 2002.

[174] D. Harari, for The Pierre Auger Collaboration, *'Search for Correlation of UHECRs and BL Lacs in Pierre Auger Observatory Data'* Proceedings of the 30th ICRC 4 p 283-286 2008.

[175] N. Hayashida, M. Nagano *et al 'The Anisotropy of Cosmic Ray Arrival Directions around* 10^{18} *eV'* arXiv:9807045v2 [astro-ph:] 15 May 1999.

[176] P. Sokolsky, *'Final Results from the High Resolution Fly's Eye (HiRes) Experiment'* arXiv:1010.2690v2 [astro-ph:HE] 15 Oct 2010.

[177] Y. Tsunesada, *'Highlights from Telescope Array'* arXiv:1111.2507v1 [astro-ph:HE] Proceedings of the 32nd ICRC Beijing Nov 2011.

[178] A. V. Glushkov, *et al.*, *'Muon Content of Ultra-High-Energy Air Showers: Yakutsk Data versus Simulations'* arXiv:0710.5508v2 [astro-ph] 18 Feb 2008.

[179] V. Hess, Physik. Zeitschr, **13** p 1084-1091 1912

[180] L. Nava, S. Gabici, *'Anisotropic CR Diffusion and γ-ray Production Close to Supernova Remnants, with an Application to W28'* arXiv:1211.1668v1 [astro-ph:HE] 7 Nov 2012.

[181] KASCADE-Grande home page http://www-ik.fzk.de/KASCADE_home.html.

[182] H-B. Hu, Q. Yuan, B. Wang, C. Fan1, J-L. Zhan and X-J. Bi, *'On the e^+e^- Excesses and the Knee of the Cosmic Ray Spectra – Hints of Cosmic Ray Acceleration in Young Supernova Remnants'* www.astrophys-space-sci-trans.net /7/issue2.html.

[183] S. Schoo, *'The Energy-Spectrum and Mass Composition of Cosmic Rays'* Karlsruher Institut für Technologie - Institut für Kernphysik 8 Nov 2012.

[184] D. Harari , S. Mollerach, E. Roulet, *'The Toes of the Ultra High Energy Cosmic Ray Spectrum'* arXiv:astro-ph/9906309v2 7 Dec 1999.

[185] D. Harari , S. Mollerach, E. Roulet, *'Signatures of Galactic Magnetic Lensing upon Ultra High Energy Cosmic Rays'* arXiv:astro-ph/0001084 6 Jan 2000.

[186] D. Harari , S. Mollerach, E. Roulet, *'Anisotropies of Ultra-High Energy Cosmic Rays Diffusing from Extragalactic Sources'* arXiv:1312.1366v1 [astro-ph.HE] 4 Dec 2013.

[187] E. N. Parker, *'Cosmical Magnetic Fields* (Clarendon Press Oxford 1979).

[188] M. Prouza, R. Šmída, *'The Galactic Magnetic Field and Propagation of Ultra-High Energy Cosmic Rays'* arXiv:astro-ph/0301765v1 9 Jul 2003.

[189] X. H. Sun, W. Reich, *'The Galactic Halo magnetic Field Revisited'* arXiv:1010.4394v1 [astro-ph:GA] 21 Oct 2010.

[190] K. Ferrier, *'Interstellar Magnetic Fields in the Galactic Center Region'* arXiv:0908.2037v1 [astro-ph.GA] 14 Aug 2009.

[191] M. Haverkom, *'Magnetic Fields in the Milky Way'* arXiv:1406.0283v1 [astro-ph.GA] 2 Jun 2014.

[192] M. S. Sutherland, B. M.Baughman, J.J. Beatty, *'A Method for Constraining Cosmic Magnetic Field Models Using Ultra-High Energy Cosmic Rays: The Field Scan Method'* arXiv:1207.1654v1 [astro-ph.IM] 6 Jul 2012.

[193] T. R. Jaffe, *et al.*, *'Modelling the Galactic Magnetic Field on the Plane in 2D'* arXiv:0907.3994v2 [astro-ph.GA] 29 Jan 2010.

[194] B. J. Whelan, *'A Magnetic Spectrometer Analysis Method for Ultra High Energy Cosmic Ray Data'* Ph.D. thesis Adelaide University Nov 2011.

[195] J. A. Bellido, *et al.*, *'Southern Hemisphere Observations of a* 10^{18} *ev Cosmic Ray Source near the Direction of the Galactic Centre'* Astroparticle Physics 15, 167, 2001 Astroparticle Physics 15, 167, 2001.

[196] J. R. T. de Mello Neto for the Pierre Auger Collaboration, *'Studies of Cosmic Rays at the Highest Energies with the Pierre Auger Observatory'* arXiv:1308.0820v1 [astro-ph HE] 2014.

[197] http://en.wikipedia.org/wiki/Large_Hadron_Collide.

[198] http://en.wikipedia.org/wiki/Centaurus_A.

[199] M. Su, T. R. Slatyer, D. P. Finkbeiner, *'Giant Gamma-Ray Bubbles from* Fermi-*Lat: Active Galactic Nucleus Activity or Bipolar Galactic Wind?'* Astrophysical Journal 724: p 1044-1082 Dec 1210.

[200] D. E. A: E. Nebot del Busto Supervisor: Dr. Jose Del Peso Malagón, *'Electron Energy Reconstruction in the ATLAS Electromagnetic End-Cap Calorimeter using Calibration Hits'* hepexp.ft.uam.es/atlas/DEA/DEA-Nebot.pdf.

[201] J. M. C. Montanus, *'An extended Heitler-Matthews model for the Full Hadronic Cascade in Cosmic Air Showers'* arXiv:1311.0642v3[astro-ph.HE] 5 Mar 2013.

[202] pdg.lbl.gov/2011/reviews/rpp2011-rev-passage-particles-matter.pdf 25 Jun 2014.

[203] Y. Guo, Z.. Feng, Q. Yuan, C. Liu, H. Hu, *'On the Galactic Center Being the Main Source of Galactic Cosmic Rays as Evidenced by Recent Cosmic Ray and Gamma Ray Observations'* arXiv:1101.5192v4 [astro-ph.HE] 23 Jan 2013.

[204] *'A Review of the Ginzburg-Syrovatskii's Galactic Cosmic-Ray Propagation Model and Its Leaky-Box Limit'* ntrs.nasa.gov/archive/nasa/casi.ntrs.../20120013305.pdf.

[205] E. Fermi, *'On the Origin of the Cosmic Radiation'* Physics Review 75 p 1169-1174 1949.

[206] A. A. Watson, *'High-Energy Cosmic Rays and the Greisen-Zatsepin-Kuzmin effect'* arXiv:1310.0325v1 1 Oct 2013.

[207] A.V. Olinto, *et al.*, *'White Paper on Ultra-High Energy Cosmic Rays'* arXiv:0903.0205v1 [astro-ph.HE] 2 Mar 2009.

[208] A. Schulz, *'The Measurement of the Energy Spectrum of Cosmic Rays Above* 3×10^{17} *eV with the Pierre Auger Observatory'* 33rd ICRC Rio de Janerio 2013.

[209] T. Stanev, *'Ultra-high-energy Cosmic Rays and the Large-Scale Structure of the Galactic Magnetic Field'*, Astrophysical Journal, 479, p 290 1997.

[210] *'Hough Transform'* www.cs.jhu.edu/~misha/Fall04/GHT1.pdf.

[211] *'Image Transforms'* homepages.inf.ed.ac.uk/rbf/HIPR2/hough.htm.

[212] Unpublished work by author.

[213] V. Berezinsky, *'Extragalactic Cosmic Rays and Their Signatures'* arXiv:1301.0914 [astro-ph.HE] 20 Apr 1013.

[214] P. Abreu, *et al.*, *'Constraints on the Origin of Cosmic Rays Above* 10^{18} *eV from Large Scale Anisotropy Searches in Data of the Pierre Auger Observatory'* arXiv:1212.3083v1 [astro-ph.HE] 13 Dec 2013.

[215] T. Stanev, *'Ultrahigh Energy Cosmic Rays: Review of the Current Situation'* arXiv:1210.0927v1 [astro-ph.HE] 2 Oct 2012.

[216] G. Sigl, A. Van Vliet, *'PeV neutrinos From Ultra-High-Energy Cosmic Rays'* arXiv:1407.6577v1 [astro-ph.HE] 24 Jul 2014.

[217] M. V. Barkov, V. Bosch-Ramon, *'Formation of Large-Scale Magnetic structures Associated with the Fermi Bubbles'* arXiv:1311.6722v1 [astro-ph.HE] 26 Nov 2013.

[218] R. Abassi *et al.* for the HiRes collaberation *'First Observation of the Greisen-Zatsepin-Kuzmin Suppression'* arXiv:astro-ph/0703099.

[219] S. Boldyrev *'Spectrum of MDH turbulance'* (University of Chicago) astro-ph/0503053; Astrophysical Journal 626, L37, 2005.

[220] A. Haungs for the JEM-EUSO Collaboration *'Physics Goals and Status of JEM-EUSO and its Test Experiments'* arXiv:1504.0259v1 [astro-ph:IM] 10 Apr 2015.