CHARACTERIZING ARRIVAL DIRECTION PROBABILITIES OF ULTRA HIGH ENERGY COSMIC RAYS WITH THE PIERRE AUGER OBSERVATORY AND PROGRESS TOWARD AN IN-SITU CROSS-CALIBRATION OF AUGER AND TELESCOPE ARRAY SURFACE DETECTOR STATIONS

by

SEAN PATRICK QUINN

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Physics

CASE WESTERN RESERVE UNIVERSITY

January, 2018

CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis/dissertation of

Sean Patrick Quinn

candidate for the degree of **Doctor of Philosophy**

Advisory Committee:
Professor Corbin Covault, Chair/Advisor

Professor Frederic Sarazin, Committee Member

Professor Chris Mihos, Committee Member

Professor Wojbor Woyczinski, Committee Member

Professor John Ruhl, Committee Member

Date of Defense
November 28 2017

Table of Contents

List of Tables

List of Figures

Acknowledgements

It is inconceivable that this work would have been completed or even possible without the immense and humbling support provided by many dozens of people. I am deeply grateful and appreciative for my advisor's continued encouragement, scholarly advice, feedback and for providing me with unbelievable opportunities in cosmic ray physics, notably the ability to collaborate in the Pierre Auger Observatory. To be part of such a massive effort studying some of the most compelling unsolved problems in physics and astrophysics is an unparalleled privilege. I have no doubt years spent as a member of this collaboration will be among my fondest professional memories. I have tremendous respect for the hundreds of dedicated collaborators and dozens of staff members, truly amazing folks, who work tirelessly to keep it a world-class observatory. I would also like to thank my adviser for his transparency and unrelenting advocacy, especially during arduous times. My involvement in the Auger@TA project would not have been possible without the warm and generous collaborative support from Fred Sarazin. His advice, availability and close collaboration significantly improved many aspects of this study. Without his coordination of receiving equipment from Argentina and deploying it in Utah as well as writing the white paper to motivate the project, it would not have moved forward.

All of this can also be said for many of our TA colleagues as well. I wish to thank Nonaka Toshiyuki for his extraordinary patience, selflessness, and diligence in helping to set up much of TA share of infrastructure for the project. His willingness and dedication on and off site have been instrumental for the success and continued

operation of this project. I also thank Ryuji Takeishi and Toshihiro Fujii for their assistance in the field and patiently answering my pestering questions about TA procedures and methodologies. I thank Fujii-san for his hospitality in Delta and for his time spent supporting the project with data transfers, MOU discussions and logistics matters. I wish to also express my humble gratitude to Hiroyuki Sagawa for his patient assistance with the project, and sponsorship within the TA collaboration. A warm thank you to all members of the TA collaboration and accept my best wishes/luck for the future.

I would like to profusely thank Ricardo Sato for his instrumental CDAS emulation code and for setting aside time for Skype calls. Without these tools there would be no data from our Cosmo station. Ricardo's documentation about the encoding of the Auger hex dump was also critical. I am deeply grateful for Robert Sobin's technical guidance and assistance for all things electronic. Without Bob's circuit wisdom and breadth of expertise the local trigger system could not have been built. I've learned a lot from Bob and thank him for his patience and excellent collaboration. I would also like to thank Jeffrey Johnson and Ryan Lorek for their excellent support in the field and in the lab. Jeff was instrumental in getting me acquainted with the Auger north systems and provided significant guidance for using the suite of programs to handle data extraction. Jeff is also extremely talented with construction and fabrication and our setup has greatly benefited from his expertise. Speaking of fabrication, I wish to thank Richard Bihary for his wonderful suggestions and patience for enduring my relentless pestering over the years and answering any question I had about building anything and everything. Rick and the physics shop are an incredible

resource that I've been spoiled with. I need to express my gratitude to Laurent Guglielmi for his invaluable software support, and timely response to my majority wall of text, minority short and concise, e-mails and his keen assistance in debugging certain programs along with his lucid explanations and guides. Without Laurent's expertise we would have been starting at ground zero, and there is no way we would have progressed at such a fast pace. I thank Robert Halliday for our stimulating physics and hardware discussions and for his critical role in porting his time tagging firmware from the Zed development board onto the MicroZed, a significant effort. Without Rob's expertise we would have no local trigger timestamps and thus no external trigger for the Auger doublet. I thank Andrew Ferguson for stimulating physics discussions especially relevant to Auger and cosmic rays. Andy's patient assistance with electronics, computing and detectors over the years have undoubtedly made me a better experimentalist. Andy is the only real life Arch Linux user I've met (I bow in deference) and is without a doubt a superior programmer and he was a pleasure to work with. I thank Danielle Lahurd for interesting and insightful physics discussions, computing support, and for her kindness and encouragement over the years. My gratitude to Orlen Wolf for his expert support in the field, especially in terms of critical infrastructure.

I also must thank John Ruhl, Chris Mihos and Wojbor W. for agreeing to serve on my committee and for their valuable feedback.

I also need to recognize Linus Torvalds and the hundreds of developers who contribute to the Linux kernel. Further, I need to thank the hundreds of volunteer developers who work on the Debian operating system, especially the ARM version

for the Raspberry Pi. Additionally, the hundreds of developers working on Python and it's libraries allowed for the detailed analyses performed. With these tools available (for free no less) this project, and much of my graduate work, would not be possible.

Arrival Direction Probabilities of Ultra High Energy Cosmic Rays with the Pierre Auger Observatory and Progress Toward an in-situ Cross-calibration of Auger and Telescope Array Surface Detector Stations

Abstract

by

SEAN PATRICK QUINN

The origin and nature of ultra-high energy cosmic rays is an open question in astrophysics since their discovery in the 1960s. Observed energies of these primary particles can approach greater than $10\times$ the center of mass energies achieved at modern collider facilities, and there is presently no consensus for an astrophysical mechanism capable of reaching these energies. The Pierre Auger Observatory is the world's largest cosmic ray observatory, spanning $\sim$3,000 km^2, dedicated to this problem. After initial acceleration, the primary trajectory from source to solar system is not rectilinear due to magnetic deflection. This dissertation investigates how uncertainties of a modern galactic magnetic field model translate into arrival direction uncertainties for $E > 50$ EeV events using a sensitivity analysis approach. In most cases it's found that uncertainties from **B**-field model dominate compared to observation systematic errors. Furthermore, the angular extent of 1σ arrival direction contours is found to enclose many potential astrophysical objects, making it difficult to isolate individual sources. Implications for anisotropy studies are also briefly discussed.

To enable full-sky coverage, Auger data can be combined with the Telescope Array Project, a similar array operating in the northern hemisphere. In both experiments most data are generated from the surface-detector (SD) array. The TA and Auger experiments use different SD station designs, giving them different sensitivities to extensive air-shower components. We seek to understand and cross-validate these complementary detectors on a hardware level using an *in-situ* approach to observe the same air showers. We describe the technical details associated with installing the detectors, data acquisition, and analysis of signals for this first stage of the Auger@TA project. Integrated signals are compared to a collection of models, and in general are found to agree with expectations. For a subsample of events air-shower parameters are available, and a sophisticated simulation is run to predict the expected response of the two detectors, which is compared to observed data. These simulations appear to show some discrepancy. Characterizing and understanding the discrepancies will be important for taking full advantage of future planned upgrades to both experiments, especially in the context of atomic composition measurements.

Chapter 1: The history of cosmic ray physics

1.1 Organization

In Chapter 1 the historical evolution of the field is presented with a focus on the progression of experimental campaigns through the present day. In Chapter 2 the fundamental physics of cosmic rays is presented along with the paradigms for their study. In Chapter 3 I provide a detailed description of the two largest active cosmic ray observatories and also perform a comparison of their important features. In Chapter 4 I present the methodology and analysis techniques used to investigate the Jansson-Farrar 2012 galactic magnetic field model. In Chapter 5 the Auger@TA project is motivated and introduced followed by a detailed description of the hardware developed and deployed in the field. In Chapter 6 the data analysis scheme is discussed along with a variety of results, such as an empirical MIP-VEM model, a comparison of data to a simplified geometrical expectation, as well looking at simulated vs. observed signals. Chapter 7 describes plans for improving our setup in addition to the next phase of the project. Chapter 8 presents a discussion and interpretation of the Auger@TA project results. Chapter 9 includes miscellaneous material which did not conveniently mesh with the other chapters, such as lab work done to validate hardware and calculations for orphaned projects.

1.2 Preamble

Cosmic ray physics is a multidisciplinary field that aims to understand the origin and nature of highly energetic particles arriving at Earth from the cosmic environment. How particles are accelerated to these energies remains an open, unsolved problem in astrophysics. To understand this phenomenon the current paradigm involves building and operating massive air shower detector arrays, usually with large international collaborations. The highest energy particles, those exceeding 10^{21} eV, are the rarest form of stable matter in the universe and result in marginal particle flux at Earth's surface. Therefore the efficiency of the detector scales with its surface area which motivates mega arrays built by large collaborations.

Although we can't pinpoint sources, there's strong evidence suggesting these particles originate within the galaxy in the low energy regime, and outside the galaxy in the high energy regime. Additionally, sophisticated reconstruction techniques are employed enabling us to identify the composition (or type) of primary particle incident on the Earth's atmosphere. This is strongly contrasted with the birth of the field, where researchers were concerned with basic questions, such as whether these particles were coming from inside the Earth, or the heavens. In this chapter I will present the historical events where the phenomenon was first identified and explored, how it was determined that these particles are coming from space and finally a survey of recent and current experiments.

Figure 1.1: Artist's depiction of a late 18th century electroscope. This model uses gold foils secluded in a glass jar to prevent air currents, moisture and other externalities from disturbing observations. To measure charge the experimenter holds the object close to the top disk which inductively charges the apparatus. The gold foil leaves will be displaced according to the charge magnitude.

1.3 Atmospheric ionization

The earliest known indirect observation of cosmic rays was made by Charles-Augustin de Coulomb circa 1780. During this period the state of the art equipment for electrostatics were electroscopes; see Figure 1.1. These devices operated on the simple principle of repulsion of like charge. The magnitude of charge could be determined by the separation distance between two metal leaves. During his studies he encountered situations where electroscopes would suddenly discharge, it seemed, through the air medium instead of other controlled contacts [1]. At the time it was known that air acted as an insulator, but the primitive nature of the apparatus could not rule out systematic errors. Roughly a century passed before this curiosity was explored further.

Although articles in the literature [2, 3] refer to Michael Faraday's work on the subject, specifically his observation of spontaneous discharge in electroscopes of superior construction, after an examination of the primary source [4] I'm unable to find any explicit description of this work. However, given Faraday's exhaustive work with electromagnetic phenomena, it's certainly plausible the effect was recorded.

A definitive demonstration that air ionization[1] causes spontaneous discharge of the electroscope was shown by William Crookes, whose primary interest was building cathode ray and vacuum tubes. In experiments carried out in 1879 he showed that an electroscope placed in a sealed container *with* air eventually lost its charge but the same apparatus didn't lose its charge when the container was *evacuated* [5]. This observation explains why Coulomb observed spontaneous discharge: it was due to ionization of the air. However, Crookes' experiments offer no explanation for why the air became ionized with no external influence.

The microscopic description of this ionization was provided by J.J. Thomson's work on the electrical conductivity of gases between 1880–1900 where he introduced the terminology and definitions of "polarized molecule, positive atom, negative atom" and "ions" [6]. The work ultimately led to the discovery of the electron, garnering Thomson a Nobel prize, but the origin of the ions remained elusive.

[1] at this time the term 'ionization' had not yet been invented, but we use this terminology here since most readers should be familiar with it

1.4 Discovery of radioactivity

Radioactive materials proved to be a helpful resource for elucidating the studies of ionized gases, but also led researchers down a "garden path." Radioactivity was first observed by Henri Becquerel in experiments with uranium salts. Using a shielded apparatus, he was able to demonstrate that radiation detected on his photographic plate was different than the recently discovered X-rays [7]. This led to a flurry of research into the topic which generated profound new insights, and would help establish the field of nuclear physics. Of particular importance was the discovery that polonium and radium would transmute into new material while simultaneously producing radioactivity [8].

These discoveries significantly influenced the problem of spontaneous air ionization: researchers conjectured that random electroscope discharge was the result of charged particle emission from radioactive particles in the air, or from species in the Earth's surface. This was a classic case of attributing correlation to causation: although the connection between the two areas was compelling, some physicists required further quantitative evidence to establish radioactivity as the source of spontaneous ionization [2].

In the years following Curie *et al.'s* landmark paper [8] multiple European groups built improved electroscopes for the careful study of radioactivity. Combined with better techniques, these instruments provided enough sensitivity to allow for a quantitative measurement of the ionization rate of a sample. It was known that radioactive species emitted charged particles, such as the decay $^{226}\mathrm{Rn} \rightarrow \alpha + ^{224}\mathrm{Rn}$, and the

advanced electroscopes were used in measuring decay rates. A notable experiment in 1899 by Elster & Geitel demonstrated that thick metal shielding lead to a decrease in ionization [3], suggesting the source was outside the container. However, this result wasn't useful for determining if the external source was at, above, or below the surface of the Earth. A short time later in 1901 C.T.R Wilson verified their result and offered the hypothesis of extraterrestrial sources (now known to be correct). Throughout the 1910s numerous efforts were made to determine the directionality of the ionization, but experimental uncertainties prevented robust conclusions. However, these experiments demonstrated the presence, even in shielded environments, of a ubiquitous background radiation. It was quickly proposed that if surface radioactivity was causing the background, then the ionization rate should be a radial function of distance from the Earth's crust.

1.5 Ionization vs. height and depth

During the years of 1909 and 1910 several key advances were made that might be considered the catalysts for the field of cosmic ray physics. The first was the construction of a more precise and portable electrometer by Theodor Wulf. His apparatus used metalized glass wires that could achieve a resolution of 1 Volt. A schematic, taken from [3], appears in Figure 1.2. Wulf put the electrometer through its paces, taking measurements of air ionization throughout Europe, finding consistent results. During his travels it occurred to him that he could use his device to study how ionization varied with height and test the hypothesis of surface radioactivity. If

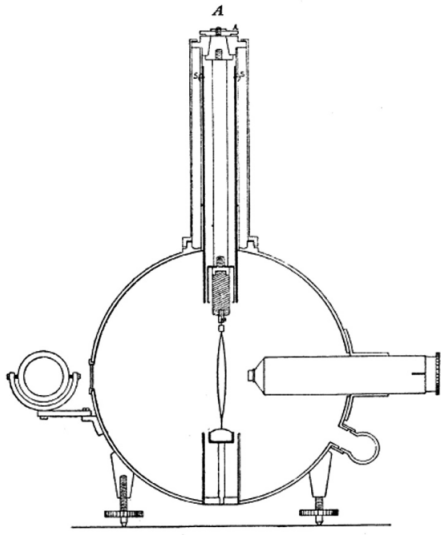


Figure 1.2: Wulf's electrometer drawing. The right cylinder is a microscope used to measure the separation of glass wires. The mirror on the left is used for lighting. Figure from [3].

the background's source was the Earth's crust, as one moves further away (higher up) the radiation must decrease. In papers published in 1909 and 1910 Wulf ascended the Eiffel Tower ($\sim$ 300 m) and showed there was no decrease in ionization relative to ground level [9, 10] that one would expect from the geometrical dilution factor ($\propto r^{-2}$). Still, vis a vis these striking results, Wulf concluded the radiation was emanating from the surface, but was more intense than previously assumed. Wulf's work was held as the most reliable data regarding altitude dependence of background radiation and his discovery was a critical signpost for other researchers.

Another key figure was Domenico Pacini who studied the background at various locations on maritime vessels. His first experiment was to travel 300 m from a beach and measure the ionization rate, which was found to be 2/3 of the rate on the surface [11]. This measurement is interesting because it transparently shows that while some

background is due to the Earth's crust, it is not responsible for the majority of the effect. Pacini's next experiment investigated how the rate changed with depth. This measurement was the first of its kind in physics, and he found a substantial decrease compared to the surface [12]. Crucially, he also correctly interpreted his results, asserting that "ionization ... in the atmosphere ... is independent of the direct action of radioactive substances in the soil" [3], and also suggested an extraterrestrial origin. This development was a key milestone in the field, and is not widely appreciated by members in the community.

At roughly the same time Pacini was studying ionization vs. depth, Albert Gockel made three balloon flights to 3000 m and recorded ionization rates. He found that the rate did *not* decrease with height [13], a result which contradicts the terrestrial origin hypothesis, and also corroborated Pacini's results. However, skeptics argued that if one had both crust and extraterrestrial sources, then at this altitude the extraterrestrial source would compensate for the decline in ionization due to the crust. This criticism would be resolved by the thorough series of experiments carried out by Hess.

1.6 The decisive balloon flights

Following the work done by Pacini and Gockel, Hess sought to provide a systematic confirmation of the altitude dependence of ionization. He flew seven times in 1912, carrying three Wulf electrometers. Two of these had 3 mm thick zinc walls for sensitivity to highly penetrating radiation, and the third had 0.188 mm walls for

7. Fahrt (7. August 1912).

Ballon: „Böhmen“ (1680 cbm Wasserstoff). Führer: Hauptmann W. Hoffory.
Meteorolog. Beobachter: E. Wolf. Luftelektr. Beobachter: V. F. Hess.

Nr.	Zeit	Mittlere Höhe absolut m	Mittlere Höhe relativ m	Beobachtete Strahlung Apparat 1 q_1	Apparat 2 q_2	Apparat 3 q_3	Apparat 3 reduz. q_3	Temp.	Relat. Feucht. Proz.
1	15h 15—16h 15	156	0	17,3	12,9	—	—	—	—
2	16h 15—17h 15	156	0	15,9	11,0	18,4	18,4	1½ Tag vor dem Aufstiege (in Wien)	
3	17h 15—18h 15	156	0	15,8	11,2	17,5	17,5		
4	6h 45— 7h 45	1700	1400	15,8	11,4	21,1	25,3	+6,4°	60
5	7h 45— 8h 45	2750	2500	17,3	12,3	22,5	31,2	+1,4°	41
6	8h 45— 9h 45	3850	3600	19,8	16,5	21,3	35,2	−6,8°	64
7	9h 45—10h 45	4800 (4400—5350)	4700	40,7	31,8	—	—	−9,3°	40
8	10h 45—11h 15	4400	4200	28,1	22,7	—	—	—	—
9	11h 15—11h 45	1300	1200	(9,7)	11,5	—	—	—	—
10	11h 45—12h 10	250	150	11,9	10,7	—	—	+16,0°	68
11	12h 25—13h 12	140	0	15,0	11,6	—	—	(nach der Landung in Pieskow, Brandenburg)	

Figure 1.3: Raw data from Hess' 1912 balloon flight. All electrometers display an ionization rate $\propto e^z$ where z is altitude. Figure from [2].

higher sensitivity to β radiation [2]. The first six flights reached a maximum altitude of 2 km and Hess observed no change in ionization rate [14]. His seventh flight obtained an altitude of 5300 m and showed a distinct increase in the rate; a table of his results is provided in Figure 1.3. Hess' results were impressive for a variety of reasons. The intrepid nature of the experiment notwithstanding, by using Wulf electrometers he provided precise measurements. Furthermore, by ascending to several kilometers his investigation probed length scales many orders of magnitude greater than previous studies. Finally, he also offered a clear and falsifiable explanation for his observations

> "The results of the present observations seem to be most readily explained by the assumption that a radiation of a very high penetrating power enters our atmosphere from above, and still produces in the lowest layers a part of the ionization observed in closed vessels. The intensity of this radiation appears to be subject to transient variations, recognizable in hourly readings."

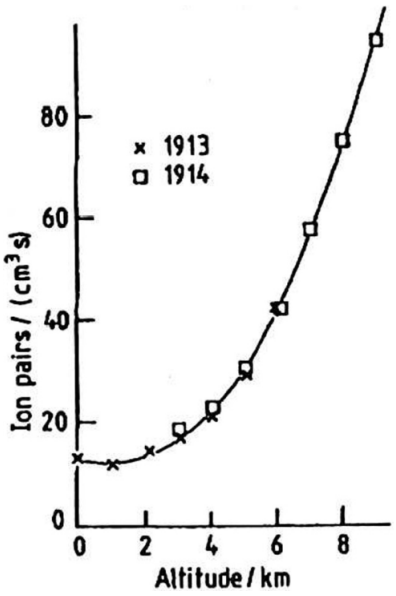


Figure 1.4: Data from Kolhörster's many flights. No error bars are shown, but systematic errors were likely on the order of 2-5 ions cm^{-3} s^{-1}. The exponential dependence on height is clearly visible. Figure from [16]

[15]. Hess' work was a significant advance and set the stage for more ambitious flights.

Werner Kolhörster was a young contemporary of Hess who made several ascents to extreme heights. In 1913 he was able to pilot his balloon, with the help of an oxygen mask, to a height of 6.2 km, roughly 1 km higher than the max altitude achieved by Hess. A year later he improved on this flight by ascending to 9.3 km using a 2200 m^3 hydrogen balloon. For reference, this volume is about 1.2 times the capacity of a Boeing 747 cargo freighter aircraft. His combined results are shown in Figure 1.4 from [16]. Kolhörster's data provides definitive evidence for the extraterrestrial origin of much of the ionization background. At roughly double the height reached by Hess, Kolhörster found the rate was almost an order of magnitude larger than at sea level. His data falsify the crust/radioactive decay hypothesis. Although the

outbreak of World War I stalled the field somewhat, rapid progress was made to further characterize the sub atomic particles which ionize the air.

1.7 The birth of particle physics

With the directionality of the source of background radiation resolved, researchers became increasingly interested in what was doing the ionizing. After the war, a majority of the 1920s was spent expanding the simplistic picture of fundamental particles[2], which included ionized hydrogen and electrons. Significant effort was spent investigating Millikan's claim that cosmic rays were gamma rays (the modern concept of the photon was still in development) formed by hydrogen fusion out in the larger cosmos. During this period many important advancements were made that allowed the experimental verification of Millikan's assertion.

After finishing his dissertation in 1906 on electrical cascades through gases, Hans Geiger began his career as an assistant under Ernest Rutherford at the University of Manchester. Rutherford had a great deal of expertise on radioactive decays and transmutation (for which he ultimately received the Nobel Prize in chemistry) and

[2]With the possible exception of some squabbling between Robert Millikan and Hess/Kolhörster. Millikan and his group took data using a shielded apparatus which contradicted the balloon flight results. He published a note, [17], refuting Kolhörster's results and suggested the background "is of local origin", primarily radioactive material that's been released into the atmosphere. He eventually reversed his position in 1926. Strangely, after he publicized this reversal, American scientific journals began referring to the ionizing particles as "Millikan Rays". Millikan is credited with coining the term 'cosmic rays' in 1928 at a lecture in the UK.

was interested in how charge was distributed in atoms. He set out to test the plum pudding model, developed by J. J. Thomson, which described the atom as a soup of uniform positive charge with negative electrons sprinkled throughout to ensure neutrality. Geiger, with the help of Ernest Marsden (Rutherford's undergraduate) conceived of a design that launched α particles at a thin Au sheet placed in front of an observing screen. If atoms were described by plum pudding, exiting particles should suffer minimal deflections, resulting in small deviation angles at the detector screen. Geiger and Marsden employed the scintillation technique, a particle physics workhorse discussed at great length later in the text, but needed an accurate way of recording the number of dim flashes of light in complete darkness. They attempted to do this using the naked eye at first, but quickly realized human observers could only provide about a minute of accurate data. Despite the limitation, data collection was "brute forced", and they observed large oblique deflections [18]. Rutherford's interpretation of the data posited the atom to be a compact positively charged nucleus surrounded by orbiting low mass electrons. The impact of this result on our understanding of the structure of fundamental particles and matter cannot be overstated.

The group had achieved an impressive result; however, Geiger was eager to find a less tedious approach for recording scintillation counts and in 1911 developed an instrument that could be used to count α particles in normal light conditions. The early Geiger counter obviated the need for human eyeballs by converting ionizing particles to large electric signals via electron cascades of inert gas. These signals could then be piped to a display or speaker to provide the researcher with scaler (counts) data. A diagram of a typical Geiger counter is shown in Figure 1.5. This

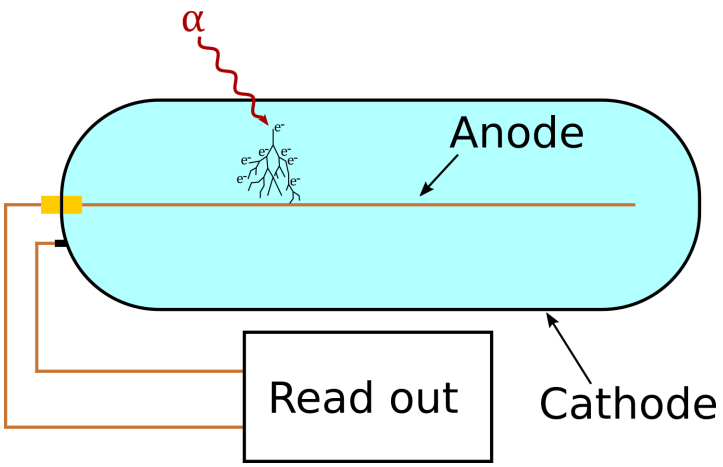


Figure 1.5: The counter (also referred to as a Geiger–Müller tube) design is based on two electrodes at high voltage separated by a gas (Helium, Argon or other inert elements are common). An ionizing particle penetrates the outer shell (cathode) and begins ejecting electrons off the molecular gas (cyan fill). This develops into a cascade of negative (electrons) and positive (molecules) particles. These ions result in a current due to the electric field between the anode and cathode, which is ultimately detected by the readout.

instrument had a significant impact on cosmic ray detection, and is still widely used today. For example, Bruno Rossi employed Geiger counters in a coincidencing circuit which investigated the penetrating power of cosmic rays. He found, using various combinations of lead and iron shields, that the primary[3] particle entering his apparatus encountered very little attenuation compared to secondaries produced inside the chamber [19]. This work was important for several reasons. The invention of the coincidencing circuit allows the investigator to reconstruct a particle's path with a high degree of confidence due to the elimination of accidental detections of background and other miscellaneous processes. Secondly, Rossi demonstrated that the type of secondaries produced depended on the atomic number and thickness of his shielding. This provides a hint about the cross section of the primary, but more importantly shows the interaction of the primary with other media is complex. The importance of this technology can't be understated–coincidencing is at the core of triggering systems in many particle physics experiments. An example is the muon spectrometer of the ATLAS experiment shown in Figure 1.6.

Throughout the 1930s researchers expanded on the works of Rossi, Kolhörster and others using careful analysis of cloud chamber photographs. Blackett et al. designed chambers that were triggered using arrays of Geiger counters which produced an impressive collection of photos, some of which showed a shower of tracks with negative and positive particles present [21]. Around the same time Carl Anderson and Robert

[3]To avoid confusion later, we note that Rossi's "primaries" are, in today's language, secondary particles (muons) generated by an air shower cascade from a primary incident in the upper atmosphere.

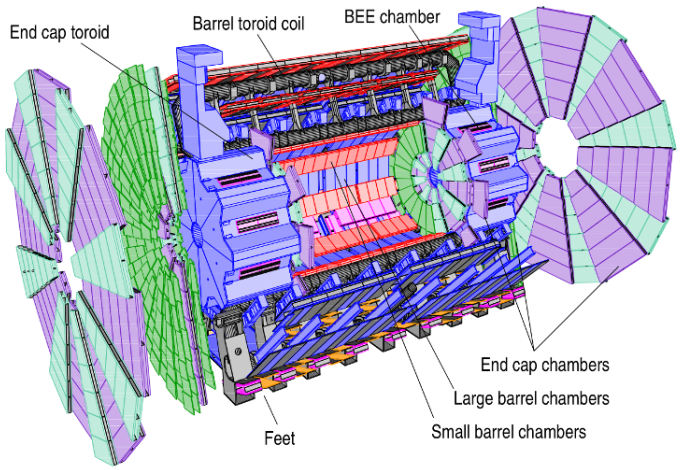


Figure 1.6: The ATLAS apparatus at the LHC. Coincidencing is performed on different chamber layers based on programmed geometrical windows to recover transverse muon momentum. Figure from [20]. While more complex than early Geiger counters, the coincidencing methodology is largely the same.

Millikan were also studying cloud chamber tracks. In one of their events they observed a track which appeared to have similar properties to the electron, but curved in the opposite direction. A photo of the event appears in Figure 1.7 This work resulted in the 1936 (3 years after publication) Nobel Prize being shared by Anderson for the discovery of the positron, and Hess for the pioneering work on cosmic rays which catalyzed the discovery. This approach continued through the 1930s leading to the discovery of numerous other particles. Among these was the mu meson or muon, which is important for air shower physics.

The 1930s was a highly productive era for cosmic ray physics. It saw the development of sophisticated detector techniques and hardware as well as the identification of secondary particle charge and mass. Moreover, the field significantly contributed to the development of subatomic and particle physics, especially in the following

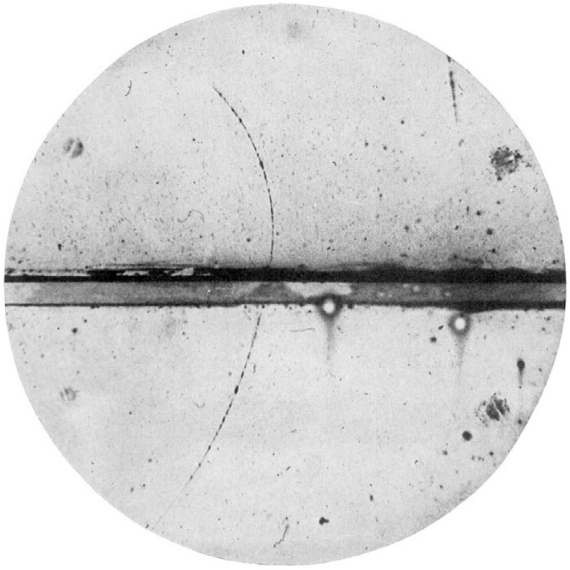


Figure 1.7: Image of positron track which appears in the famous Anderson paper [22]. In this experiment systematic uncertainties allowed for a mass of $< 20m_e$ and charge $< 2e$. While these errors are large by modern standards, the discovery is consistent with a positively charged electron.

decades. With the stage set, let us now transition to a high level review of cosmic ray experiments up to the present day.

1.8 A survey of cosmic ray experiments

Unlike modern colliders which are capable of measuring key properties of an interaction directly at the vertex, cosmic ray detectors rely on measuring subsequent generations of particles which appear well after the initial interaction point in the upper atmosphere. This phenomenon is known as an extensive air shower (EAS). Measurement of the properties of this shower allows one to infer, or "reconstruct", details about the primary interaction. A detailed technical discussion of air shower physics is discussed in a later chapter, but a tutorial schematic is shown in Figure

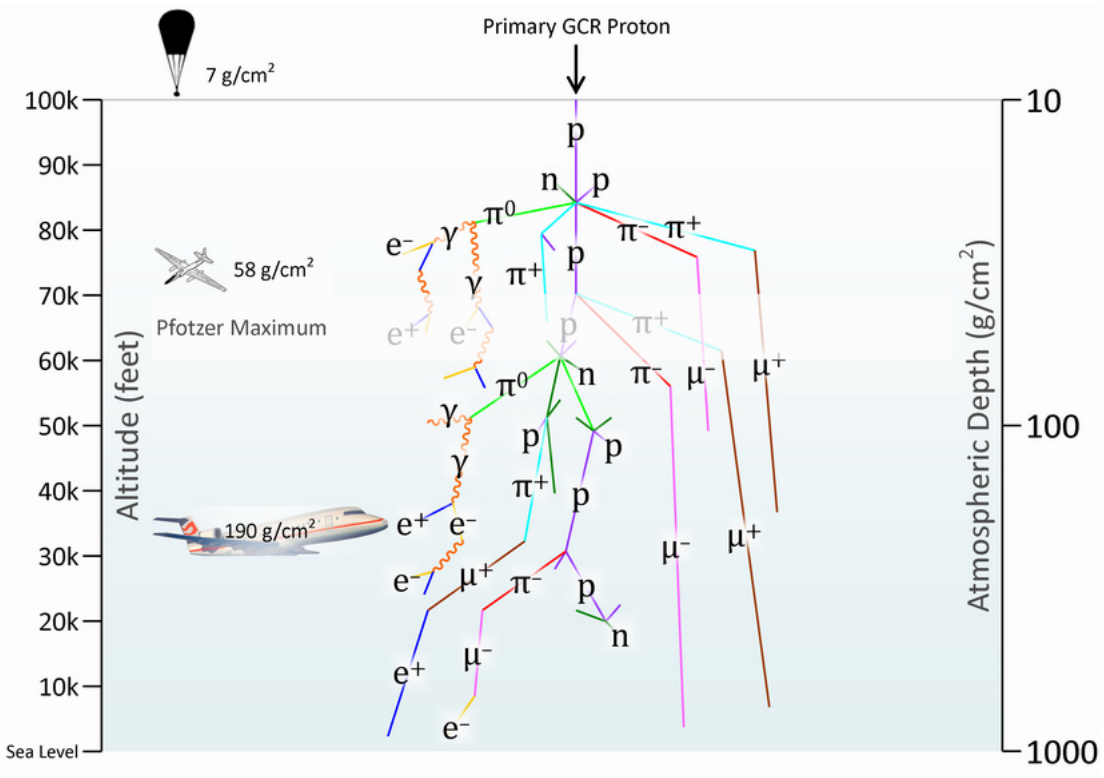


Figure 1.8: In this cartoon, taken from [23], an incident proton interacts with an atmospheric nucleus. The particle density grows, as new particles continue to cascade reaching a maximum at some atmospheric depth. This parameter is referred to as $X_{\max}$ or the Pfotzer-Regener maximum.

1.8 for a $\sim$ 1 GeV primary. The pioneering work characterizing particle cascades was performed by Rossi [19] and Regener [24] and later corroborated independently by Auger et al. [25]. In this section we discuss early incarnations of cosmic ray observatories and their influence on later designs.

1.8.1 The 1950s

After his initial work describing general features of EAS, Rossi narrowed his focus to properties of individual showers. At MIT he proposed to study airs showers with a grid of ionization chambers. Using electronic coincidence and photographic records

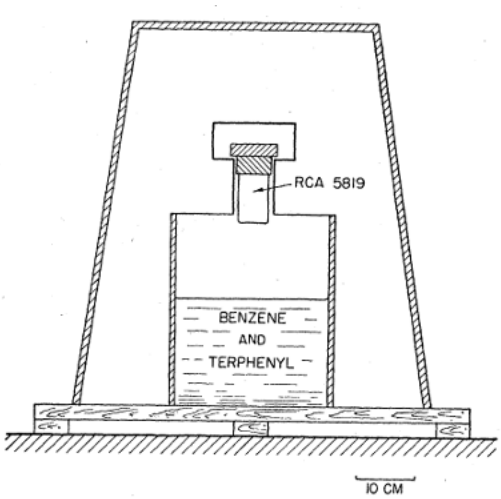


Figure 1.9: Counter used in Bassi et al.'s array [28]. The bottom was covered in aluminum foil to increase light into the RCA 5819 photomultiplier. The protective cover is an "inverted ash can."

of pulse heights his collaborator R. Williams was able to report absolute rates for showers with a certain geometry using 27 showers with $E > 10^{16}$ eV total energy [26]. This work is the first example of a prototype cosmic ray surface detector. Although timing resolution of the chambers was limited ($\sim$1 μs) and the spacing of his detectors was tight ($\sim$ 7 m) Williams was able to calculate the location of the shower core, using theoretical results available to him at the time, and assuming purely vertical showers [27]. The latter assumption being justified since 80% of his showers were at a zenith angle of $< 80°$ [26]. This is an example of a modern technique that is still in use, although timing resolution has increased by 100$\times$.

Bassi, another collaborator of Rossi's at MIT, developed a similar apparatus which used three large liquid scintillation counters (re-purposed five gallon drums that their benzene was delivered in). A diagram of the scintillation counter is shown in Figure 1.9. The group arranged the 3 detectors in various configurations: a vertical

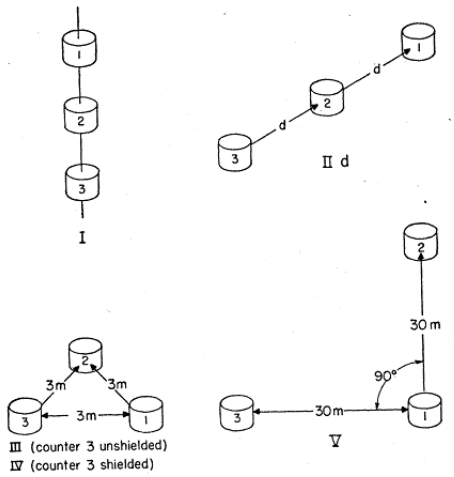


Figure 1.10: Multiple array configurations were used to study the radius of curvature and thickness of the shower disk. Figure from [28].

muon telescope, a linear string, an isosceles triangle, and a right triangle (see Figure 1.10). The vertical arrangement allowed the group to establish the characteristic pulse shape for downward muons. The remaining geometries were used for basic shower reconstruction. The authors present an extensive and complete treatment of the reconstruction theory in [28]. The fundamental insights of this work are still applied today with additional enhancements, such as correcting for the Earth's magnetic field, atmospheric layering, etc. Using the electronics and instruments available at the time, the group was able to put a lower limit on the curvature of (low energy) shower fronts of 1300 m [28]. Additionally, for the e^- component of the shower with $E \sim 20$ MeV they inferred a shower disk thickness of around 1-2 m, while more penetrating particles (presumably muons) reside in a disk with thickness of 2-3 m [28]. The group also upgraded their array by replacing liquid scintillator with solid panels of area 0.87 m^2 and increasing the number of detectors to 15.

During the same period an independent group in the UK (Culham, south east England) was also constructing detector arrays. As described by Watson in [27], the apparatus consisted of 16 Geiger counters (area 200 cm^2) in a square lattice covering about 0.6 km^2. Investigators determined the power law index derived from their observations agreed with results found by the MIT group, serving as a valuable cross check of detector technologies. Additionally, the group played an essential role in the development of both water and air Cherenkov detectors. Water Cherenkov detectors provide a large, economical detection volume–they will be encountered many times in following chapters.

Considerable work was also done by Soviet researchers during this period using Geiger counter arrays and hodoscopes. Emphasis was placed on understanding hadronic physics by developing a calorimeter based on ionization chambers. In addition to this contribution, the Soviet program discovered the "knee" feature of the cosmic ray spectrum above 2×10^{15} eV. The amount of progress made during this decade was impressive–particles that were discovered only decades prior were routinely being used to reconstruct extensive air showers and infer properties of the primary. The detection and reconstruction techniques developed during this time serve as the foundation for modern cosmic ray physics.

1.8.2 The 1960s

During this decade air showers arrays continued to be improved and enlarged. During the mid to late 1950s the MIT group built a new array at the Agassiz Station

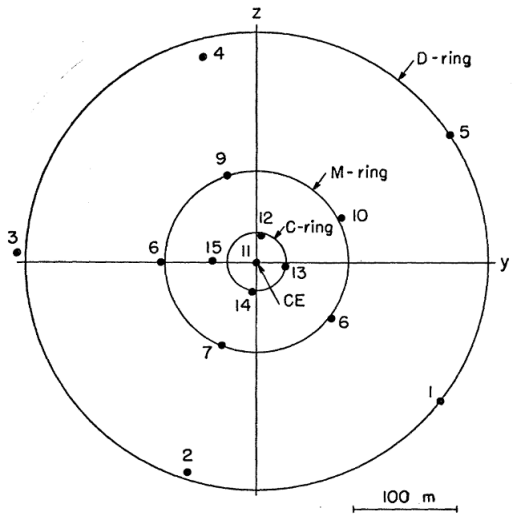


Figure 1.11: Concentric rings of different radii allowed the group to observe showers over many decades of particle densities. Figure from [29].

near the Harvard College Observatory. This experiment is included in this section since many of the important results were published in the early 1960s. Like earlier arrays, the Agassiz experiment was built to collect spectrum and arrival direction data with the ultimate goal of understanding the origin of cosmic rays while also providing the data to validate or falsify shower theories based on primary particle interactions [29].

The final design of the array employed plastic scintillators instead of liquid scintillator in a can used in the early prototype described in [28]. The detector consisted of a 42" diameter and 3" thick disk with a 5" PMT enclosed in a reflective can. The geometry of the setup allowed the investigators to observe a moderate range of secondary intensities (or primary energies, in modern parlance). A schematic is provided in Figure 1.11. In their results paper, the group discussed the use of different lateral distribution functions, most of which had just recently been published. For "smaller"

showers a basic exponential function was used

$$\Delta = \left(N/R_0^2\right) f\left(r\right) \tag{1.1}$$

where Δ gives the average density of particles, R is related to the perpendicular distance to the shower core, N is the density of the shower disk and $f\left(r\right)$ is the "structure function" which encodes the radial decay behavior of particle density. Their paper was among the first to include fits using the original Nishimura–Kamata (NK) structure function [30] as well as the Greisen approximation of that function (NKG) [31]. Using a particular parametrization of the NKG function, the group's largest shower was found to have a particle density of 3.49×10^9 arriving from $\alpha = 112°$, $\delta = 53°$ [29]. The energy of a primary for vertical showers at roughly 1 km above sea level can be approximated using

$$E_0 \sim 3.9 \times 10^6 \text{GeV} \left(N_e/10^6\right)^{0.9} \tag{1.2}$$

where N_e is the shower size and E_0 is the primary energy [32]. From this we see that the group's largest shower had a primary of $E_0 \sim 6.02$ EeV. This was among the first measurements of an ultra high energy cosmic ray, but wasn't the highest energy primary observed.

After the Agassiz Station experiment, the MIT group redeployed these detectors to form a larger ray at Volcano Ranch in New Mexico. The setup used 19 detectors spaced 442 m apart. A diagram taken from [33] appears in Figure 1.12. With this large scale detector the group was able to observe the extremely high energy events, both of which were published in PRL. The first was on the order of 10 EeV derived from an observed particle density of 5.5×10^9 [33]. This paper also explained why such a

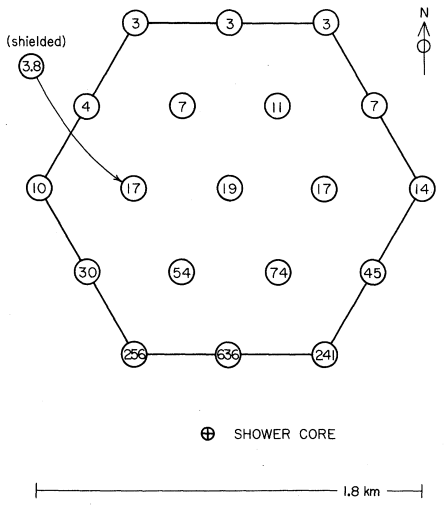


Figure 1.12: The physical footprint of the array was about 2 km^2, but the sensitive area was larger. The experiment ran for 180 days. Figure from [33].

primary couldn't originate within the galaxy and described how primary composition (proton or other nucleus) might affect shower density on the ground. For the second event a density of 5×10^{10} was observed, resulting in a primary energy of 1.0×10^{20} eV, or 100 EeV [34]. This paper also explains issues with fitting lateral distribution functions, but argues why the reported value is robust, and can at worst be an underestimate of primary energy.

In this decade the conventional surface detector array was built and operated. These instruments provided new insight into the characteristics of air showers while gathering evidence for extremely energetic primaries. Many of the theoretical tools for shower reconstruction were applied to data. In my estimation, these experiments serve as the foundation for modern cosmic ray observatories and collaboration, as later work will be technological enhancements to the methodologies presented in this section.

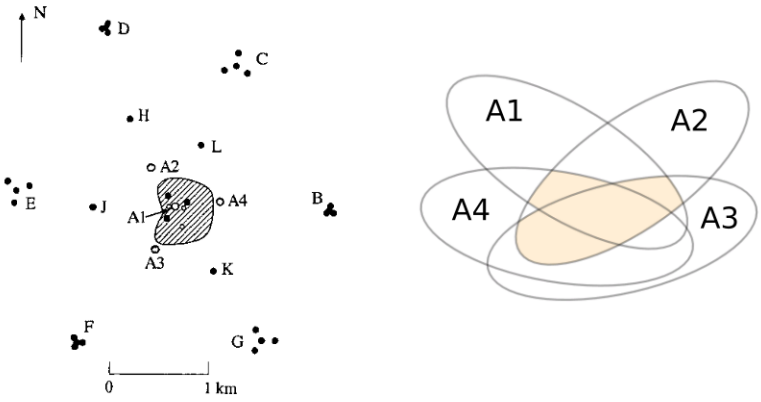


Figure 1.13: Left: scale drawing of Haverah Park array, taken from [35]. Open circles are triggering stations (A1–A4), stations H–L are 2.25 m^2 tanks and the rest are 13.5 m^2 sub arrays. The shaded area contains an infill array of 1 m^2 detectors. Right: logic Venn diagram for trigger condition.

1.8.3 The 1970s

During this decade the Haverah Park array construction was completed and collected data from 1967–1987 using water Cherenkov type detectors. This experiment introduced novel engineering techniques such as wireless communication to send commands to record oscilloscope traces of the PMTs [35]. The array's trigger condition was set to 10 vertical equivalent muons (VEM) traveling through a water depth of 1.2 m at the central A1 station along with any two of (A2,A3,A4) in coincidence within 4 μs. A diagram of the array geometry is presented in Figure 1.13. This experiment made many valuable contributions, notably, showing a hardening of the energy spectrum between 0.4 and 4 EeV and flattening above 10 EeV [35]. Their

data did not point to any spectral cutoff even upwards of 100 EeV, a violation of the GreisenZatsepinKuzmin (GZK) effect, a phenomenon which is discussed later. Intensity calculations showed at 100 EeV the expected particle flux was 1 per km^{-2} per steradian per century, an oft quoted figure in the community demonstrating the rarity of these events, and the need for large area arrays.

Another experiment built during this time was the Sydney University giant air shower array (SUGAR). It ran from 1968 to 1979 near the town of Narrabri, Australia. The array had 54 stations which consisted of two large liquid scintillators, each with area of 6 m^2, separated by 50 m. One PMT was used for both volumes, and these stations were buried. The array had an effective area of about 100 km^2. As in other experiments, the array recorded physics events when $\geq$ 3 stations registered local triggers coincident within 80 μs. After collecting 13727 events the Sydney group found evidence for the ankle at 10 EeV but did not observe flux suppression associated with the GZK cutoff [36], which agreed with Haverah Park data. The reported spectrum had a typical power law dependence with an index of 3.35 ± 0.01. Additionally, the group showed that arrival directions were consistent with isotropy given the array's angular resolution of 4°, ruling out nearby extragalactic sources [37].

In the late 60s a new detection technique was being investigated. Instead of detecting Cherenkov light produced by secondaries in a water volume, it observed the de-excitation of N_2 molecules resulting from the particle cascade. This approach was initially researched by Suga & Chudakov [38, 39]. Theoretical work was also performed by Bunner and Greisen at Cornell in the early 60s. The first experiment to observe the phenomenon was carried out in Japan by Hara et al. at the Dodaira

Observatory [40]. Their apparatus used a Fresnel lens to focus light onto a camera of 24 PMTs which generated a $23^\circ \times 20^\circ$ field of view. This method provides a near calorimetric measurement of the energy deposited by the shower. A noteworthy event from their 5 month run was a track length of 18.4° with duration of 1.9 μs, corresponding to a primary energy of $\sim 10^{19}$ eV.[4]

The first *definitive* observation and reconstruction of an air shower was carried out at Volcano Ranch by a Utah group in 1977. Their telescope used three 1.5 m diameter mirrors with 12 PMTs each (36 total) and observed the air volume directly above the surface detector (SD) array. It was run in coincidence with the SD array and is the first example of a *hybrid* cosmic ray observatory design. The group observed showers of primary energy ranging from 5×10^{16} to 2.5×10^{18} eV and found values about 10% larger compared to SD only reconstruction [42]. This project was a successful proof of concept, and would serve as an initial prototype for future designs of HiRes and the Telescope Array Project.

[4] An arXiv note by Bruce Dawson studies the plausibility of this value using modern fluorescence techniques adopted by the Pierre Auger Observatory; he finds this value to be reasonable, but warns it's probably an overestimate due to absolute light calibration issues [41]. Instead, an upper limit of 5×10^{18} eV should be used. Nevertheless, the work by Hara et al. is among, if not the earliest known fluorescence detection of cosmic rays. These measurements would be repeated by a different group a few years later.

1.8.4 The 80s and 90s

Following the pioneering experiments in the previous decades, more sophisticated arrays were commissioned to collect data over specific ranges of the CR spectrum. One example is the Akeno EAS Array, located ∼100 km west of Tokyo. This experiment started in the early 80s as a 1 km^2 array of proportional muon counters and scintillators. Eventually Akeno was expanded to 20 km^2 and reported arrival direction excesses in the direction of CygX-3 (a galactic compact X-ray source) and SS433 (eclipsing X-ray binary system discovered by CWRU astronomers) for energies $> 5 \times 10^{17}$ [43]. To boost the confidence of this result, the collaboration further expanded their surface array to 111 surface detectors and 27 muon counters covering an area of 100 km^2, creating the Akeno Giant Air Shower Array (AGASA). A map of the array and diagram of the scintillators, from [43], used is shown in Figure 1.14.The expected flux of the large detection area for primaries with $E > 10^{19}$ eV was 90 per year [43] and after 14 years of operation the complete data set had about 1000 of these events [44]. Interestingly, AGASA's spectrum continued beyond the predicted Greisen-Zatsepin-Kuz'min theoretical cutoff for extragalactic sources. This feature survived after the AGASA collaboration revisited the analysis in 2002 [45].

We will take a brief diversion from arrays studying the highest energy events to a recent project devoted to studying the composition in the spectral range of 1 to 100 PeV: KASCADE-Grande. As the name suggests, this was an upgrade to the original KASCADE array which used 252 stations on to precisely measure the EM and muonic components of showers. The ratio of these components is a reliable proxy

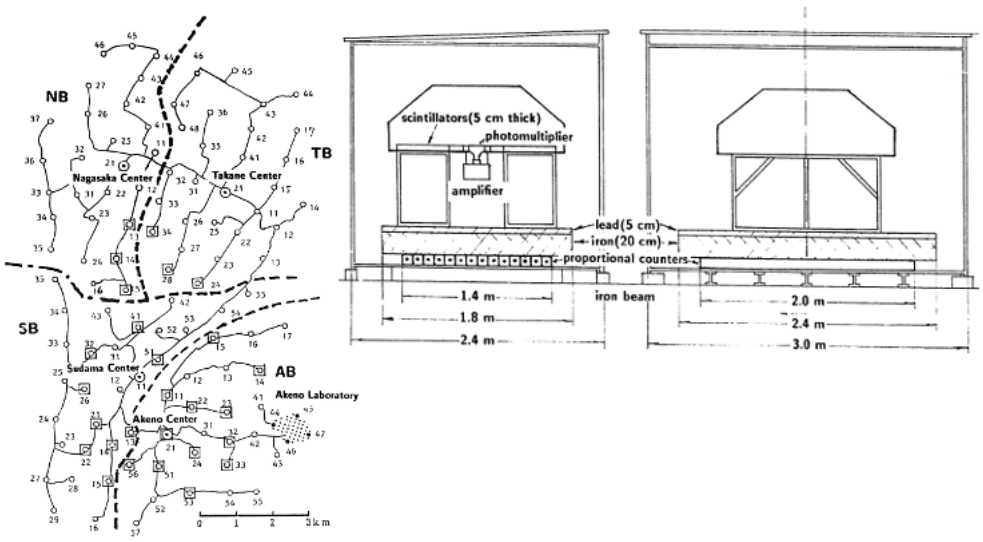


Figure 1.14: Left: A scale map of the AGASA array. Data was transferred through a fiber optic network (black lines) from scintillator detectors (open circles) and muon counters (open squares). Right: Schematic of the detector. Figure from [43].

for primary composition, and KASCADE observed that this changed from light to heavy at the 'knee' (around 3 PeV) of the spectrum. KASCADE-Grande was formed by re-purposing detectors from another decommissioned array (EAS-TOP), increasing the detection array to 700×700 m^2. This detector also uncovered an increase in flux at roughly 100 PeV, which is referred to as the second knee. It should be noted that compared to other energy domains, the data around this feature is somewhat sparse since a limited number of experiments can observe it and its a weak change in spectral index. Additionally, the KASCADE-Grande data suggest the second knee is associated with a transition toward heavier primaries [27]. We now return to experiments focusing on the highest energy primaries.

A prominent large surface detector array was built in Yakutsk, Russia and operated by the Institute of Cosmophysical Research and Aeronomy. It's goals were fairly

typical: measuring the flux above 10^{15} EeV with the GZK cutoff in mind, understanding chemical composition of primaries, studying high energy muon production and looking for anisotropy in arrival directions. The experiment consists of 64 scintillators, 6 of which are below ground, and 48 non-imaging atmospheric Cherenkov detectors atop some of the 58 above ground scintillators (2 panels) connected to a central computer with wires. This is an example of an early hybrid approach where the Cherenkov light is used as a calorimetric proxy for shower energy. The array trigger scheme is hierarchical, with station level events corresponding to particle densities of > 0.5 m^{-2} in two scintillators within 2 μs. One of the physics triggers, termed "trigger-500", happens when $\geq$3 stations are in coincidence within 40 μs. A second physics trigger, "trigger-1000", when the same condition is satisfied for stations separated by 1 km. The array geometry is shown in Figure 1.15 [46]. After about 20 years of operation the experiment observed the knee and ankle feature of the spectrum as well as GZK suppression. Compared to simulations, their depth of shower maximum data suggests light primaries from 1 to 10 PeV, becoming heavier up to 30 PeV, then light again around 100 PeV. In terms of arrival directions, no significant excess flux was found over their entire energy range. These results are summarized in [47].

In addition to many instances of ground arrays, fluorescence type campaigns were also underway. In the Utah desert, the University of Utah group operated Fly's Eye 1 and 2 from 1981 to 1992. Instead of observing a manufactured detection volume, these telescopes used 880 PMTs to collect fluorescent light from an ≈ 65 deg^2 portion of the sky [48]. By design the instrument was meant to collect data in "stereo" mode (where both telescopes see the same shower) for improved accuracy, but "monocular"

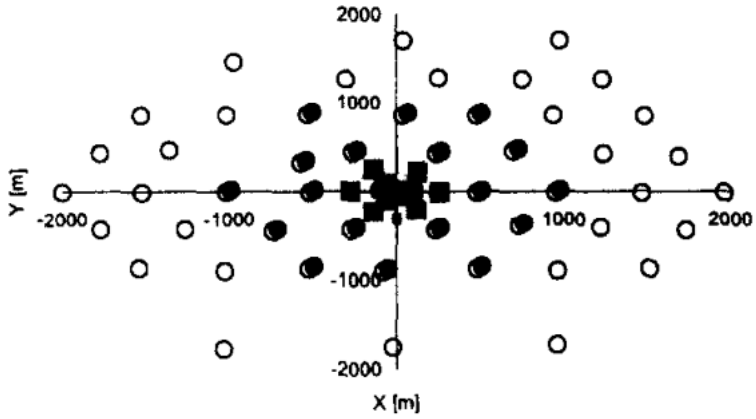


Figure 1.15: Diagram of the Yakutsk array. Open circles are scintillators, filled circles are Cherenkov detectors. Squares designate detectors situated in subarray installed in 1995 to study low energy showers. Figure from [46].

analysis was also possible (using data from one telescope). A bird's eye view of Fly's Eye 1 along with a close up of an individual telescope, which can be found in [48], is shown in Figure 1.16. The significant advantage of FD telescopes is the measurement of X_{max}, a parameter closely correlated with primary composition, for individual EASs. Interesting results of from this experiment include the detection of the 51 Joule "Oh my god" particle corresponding to a primary energy of 320^{+92}_{-94} EeV [49] of unknown composition. At this energy the GZK horizon is only about 30 Mpc and the reported arrival direction, in galactic coordinates, was $(9.6, 163.4)$ deg.

The Fly's Eye spectrum agreed with Haverah Park, Yakutsk and AGASA from $10^{17.5}$ to 10^{19} eV, but at $10^{19.5}$ eV (ankle region) there is a discrepancy of roughly a factor of 3 [50]. Following the success of the Fly's Eye campaign, the Utah group upgraded the instruments to use larger mirrors (4 m^2) and a greater number of PMTs

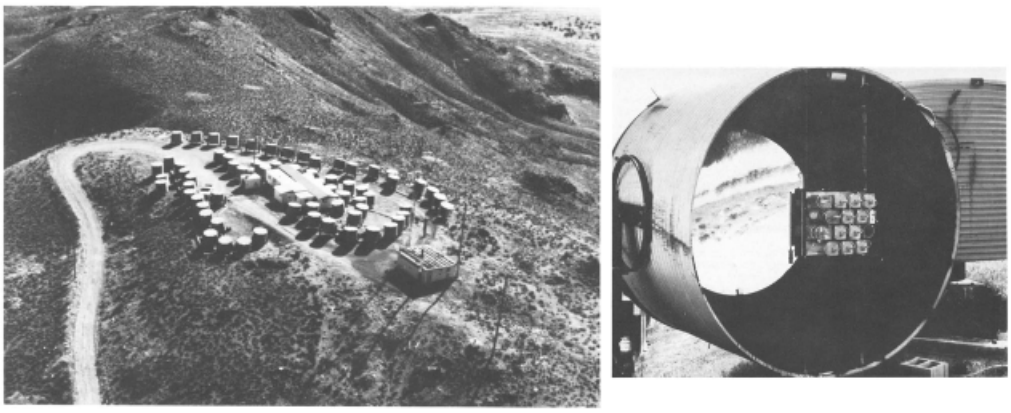


Figure 1.16: Left: the Fly's Eye 1 on the summit of Little Granite Mountain in Dugway, Utah. Right: closeup of telescope unit which uses a 62" mirror with 14 PMTs. Light is collected and focused by parabolic concentrators. Figure from [48].

(256), resulting in the High Resolution Fly's Eye (HiRes) which ran from 1997 to 2006. The HiRes detector found a 5σ level GZK suppression of flux as well as the ankle feature (a slight increase in flux) at 4 EeV. Additionally, an AGN correlation search was performed for objects within $z < 0.018$ but no signal was found [51]. In terms of composition, the data closely followed QGSJET-I simulations of pure proton primaries and placed an iron fraction upper limit of 0.1 at 90% confidence from 1 to about 100 EeV [52].

1.8.5 The 2000's and 2010's

Modern giant arrays share many features of the pioneering experiments explored in this chapter. The two largest active arrays in the world studying the highest energy particles are the Pierre Auger Observatory in Mendoza province, Argentina and the Telescope Array Project in the state of Utah, USA. Since both of these experiments are central to this dissertation, they will be described in detail in Chapter 3. In the next chapter we describe the main objectives of cosmic ray physics, outstanding problems and challenges, and how observables are used to make progress in the field.

1.9 Summary

In this chapter a comprehensive history of cosmic ray physics was presented, with an emphasis on experimentation. The origins of the field from spontaneous ionization of electroscopes through the advanced detection of extensive air showers initiated by cosmic ray primaries was described. Ultra-high energy cosmic rays are extreme phenomenon requiring precision timing, specialized particle detectors, and sophisticated data acquisition tools–not to mention an abundance of patience.

Chapter 2: The Nature and Origin of Cosmic Rays

2.1 Preamble

In this chapter we discuss the technical details of our current understanding of cosmic ray phenomenon. This includes the origin or "sources" of primary cosmic rays and how they might be accelerated to such an extreme fraction of the speed of light. The distance traveled by primaries will be extragalactic and/or galactic space and so we must consider propagation effects such as spallation and magnetic deflection during their journey. Finally when the primary arrives at Earth, it impacts an initial target nucleus in the atmosphere triggering a massive cascade of daughter particles. The energies involved in these interactions are at the limits of our understanding of hadronic physics. This cascade results in a plane of particles landing on the Earth's surface which are detected by a variety of techniques. The main data products of these techniques are an inference of (1) *energy* (2) *arrival direction* and (3) *atomic composition*. From these values a spectrum can be built, and we can attempt to extrapolate the primary chemical composition. All this information is used to study the origin and nature of ultra-high energy cosmic rays (UHECR).

2.2 Sources

Although considerable effort has been exerted in the field, there is still no global consensus about the origin of the highest energy particles. If we take the highest bird's eye view, the candidates can be divided into two sets: top down and bottom up.

2.2.1 Top down

Considered exotic by the mainstream community, top down scenarios appeal to new physics beyond the standard model. One production scenario for UHECR is the decay of super-heavy particles. While many papers have been written on the subject (a dated review appears in [50]), let us consider a representative example by Kachelriess *et al.* [53]. In this paper the authors first assume a mixture of relic (long-lived) particles lodged in the galactic halo. To reproduce energy spectra published in the literature the mass and lifetime are tuned such that $m_X > 10^{12}$ GeV and $\tau_X > 10^{12}$ years. How these particles were produced is also open to speculation: they could emerge from reheating effects following inflation or could be formed by the decay of topological defects (in essence, broken symmetries following a phase transition in the early universe, an example being a cosmic string which forms when axial symmetry is broken). In Kachelriess *et al.* [53] the particle is taken to be an SU(2)×U(1) (electro-weak) neutral fermion described by the Lagrangian

$$\mathcal{L} \sim \frac{1}{m_{pl}} \bar{\Psi} \nu \phi \phi e^{-S} \tag{2.1}$$

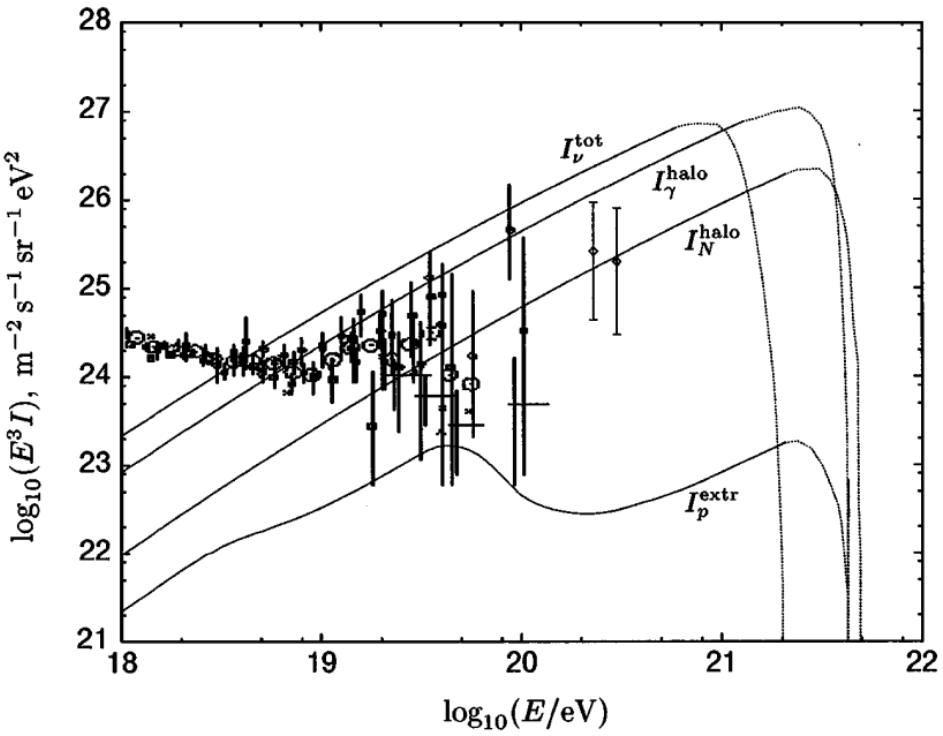


Figure 2.1: Energy spectrum prediction of X particle UHECRs for various decay scenarios: I_N^{halo} means nucleons (p,$\bar{p}$,n,$\bar{n}$) from the halo, I_p^{extr} are extragalactic protons, I_γ^{halo} are photons from the halo and I_ν^{tot} are extragalactic and galactic neutrinos. Figure from [53].

where m_{pl} is the Planck mass, $\bar{\Psi}$ is the X field, ϕ is SU(2) scalar with vacuum expectation energy 250 GeV and S is the action of a wormhole related to conserving R-parity.

With this framework in place and taking the dark matter halo radius to be 100 kpc, $m_X = 10^{13}$ GeV and $r_X = 5 \times 10^{-11}$, which is related to the lifetime and density of X particles, the authors produce the spectrum shown in Figure 2.1.

Instead of decaying heavy particles it has also been proposed that UHECRs could be produced from the interactions of extremely high energy cosmic neutrinos with relic neutrinos in an extended galactic halo. In this scenario relic neutrinos have a mass

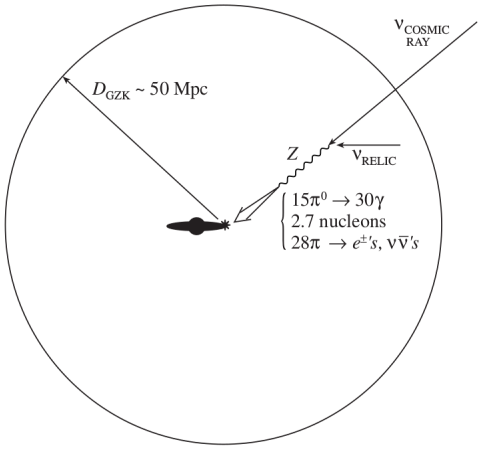


Figure 2.2: A diagram of the Z burst process showing how trans-GZK energies could be reached for primaries observed at Earth. Figure taken from [55].

of about 1 eV, which may not be unreasonable, see [54], and UHECRs are generated from a "Z-burst" annihilation event: $\nu_j + \bar{\nu_j} \rightarrow Z \rightarrow$ nucleons γ, where j indicates a flavor. This boson rapidly decays resulting in extremely high energy nucleons and photons. Since neutrino cross-sections are minuscule, this process would require vast distances or high densities to achieve non-trivial annihilation probabilities [55]. A schematic of the process is shown in Figure 2.2.

There are numerous types of topological defects, each leading to different observables for massive particle decay. While not an exhaustive list, these include superconducting cosmic strings, decaying vortons, magnetic monopoles, and cosmic necklaces. For a closer look at these see [56]. Beyond the decay of these phenomena into super heavy particles, it has also been suggested that they themselves are UHECR candidates. Currently many of these proposals are disfavored since the predictions disagree with the measured photon limits.

2.2.2 Bottom up

The bottom up paradigm assumes that Standard Model physics is enough to explain the energetics of UHECRs, meaning the primary particle is accelerated at astrophysical sites by a plausible electromagnetic mechanism—an Occam's Razor approach. However, even this paradigm has difficulties at the highest energies.

There is some agreement that lower energy CRs likely result from galactic sources. Let's start with lower energy sources and understand why they don't simply scale up. One of the most prominent features of a galaxy is its stellar population. It's well known that particles can be accelerated to GeV energies simply by solar flares (e.g. FERMI-LAT observations [57]). Given the right stellar conditions, perhaps it's possible to boost these accelerations by several orders of magnitude. In addition to flares, spacecraft experiments have also confirmed that particles are accelerated to keV and MeV energies by interplanetary shock waves due to the solar winds.

Stellar sources, more specifically the death of stars as type II supernovae (SN), are plausible sources for the bulk of higher energy CRs, possibly up to a few TeV. This can be seen from the power requirements to achieve measured average energy density for cosmic rays. The energy density of cosmic rays can be written as

$$\rho_E = \int \frac{4\pi E}{\beta c} \left(\frac{dN}{dE} \right) \, dE$$

. Using a canonical power spectrum of $\frac{dN}{dE} \propto E^{-2.8}$ for protons gives a value of roughly $\rho_E = 0.8 \text{ eV/cm}^3$. This can be used to compute the average power required to sustain

this energy, assuming the power source embedded in the galactic disk

$$P = \frac{V_D \rho_E}{\tau_\ell} \sim 1.5 \times 10^{42} \text{ erg/s}$$

assuming a galactic disk volume of $V_{gal} = \pi r^2 \Delta z = \pi \left(20 \text{ kpc}\right)^2 \ (1 \text{ kpc}) = 3.7 \times 10^{62}$ cm^3 and a galactic containment time of $\tau_\ell = 1 \times 10^6$ yrs. This power can now be compared to that generated by supernova ejecta, assuming a mass of 10 $\text{M}_\odot$ a velocity of $\upsilon = 5 \times 10^6$ m/s with a frequency of 20 years gives $P' = 2 \times 10^{42}$ erg/s. The crude calculation demonstrates, from the similarity of P and P', that supernovae are energetically plausible sources for lower energy CRs. Incidentally, the calculation also shows the mechanism must be efficient for the given choice of parameters. While encouraging for bulk, higher energy CRs, it's doubtful these sources could be 100% efficient, making them unlikely capable of powering $>$ 100 TeV CRs. Because of this we must look at other galactic processes, such as interactions between SN ejecta and interstellar medium (ISM) gas as follows.

Shock/statistical/Fermi acceleration is caused by the transfer of macroscopic kinetic energy of a moving plasma to charged particles, increasing their energies by orders of magnitude above that of thermal processes. The increase in energy is proportional to the number of encounters with this cloud, or shock front. Fermi's prescription, see [58], leads to a maximum energy of $E \leq E_0 \left(1 + \xi\right)^{t/T_{\text{cycle}}}$ (second order) where t refers to the time spent undergoing acceleration and ξ is the fractional energy gain per encounter $\xi = \frac{\Delta E}{E}$ and T_{cycle} is the time for one encounter (in and out of the shock front). Let's look at an example case. We can assume the energy increase is "moderately" efficient giving $\xi = 0.01$. A characteristic injection energy might be the

kinetic energy of a proton recently ejected from a type II SN, $E_0 = \gamma m_p v = 188$ keV with $v = 0.02c$. The cycle time depends on the probability of the particle interacting with a strong magnetic field deviation or magnetic mirror. Deriving the appropriate numbers for t and T_{cycle} requires a detailed treatment of the particle kinematics. This was originally done by Fermi, and is reviewed in (Davis, 1959)[59]. The result of the calculation is an exponent of roughly $t/T_{\text{cycle}} = 300$. Plugging in our numbers we find a maximum energy for our proton of $E \leq 3.7$ MeV. This process can be made more efficient if it occurs at a supernova blast wave (first order) instead of at a moving plasma cloud. Acceleration through this site is limited by the lifetime of the shock which can be approximated as the time required for the supernova ejecta to diffuse to a similar density to that of the ISM. Following Gaisser's formula, see [60], we have

$$\frac{4}{3}\pi \left(vt\right)^3 \rho = M_{\text{SNejecta}}$$

Using the same values for ρ, M and v as above gives $t = 1000$ yr. Gaisser also provides an equation for the maximum energy in this scenario:

$$E_{max} \leq \frac{3}{20}\frac{v}{c} ZeBvt$$

Taking $B = 1$ μG gives $E_{max} = 15$ TeV. Ultra-high energies are not achieved, but the calculation gives a plausible explanation for the very high energy (VHE) galactic CR component.

From the previous arguments we've seen how small scale diffusive acceleration fails to reach the highest energies. However, if one considers processes operating on galactic scales, the energies can increase dramatically. Consider a galactic wind forming a termination shock due to the stream of material from galactic supernovas

(for a modern treatment of this phenomenon, see Zirakashvili and Volk (2006) [61]). Cosmic rays produced in the disk that don't escape amplify Alfvén waves, coupling them to thermal gas, leading to a pressure gradient, which in turn drives a galactic wind. The flow becomes supersonic at around 30 kpc. Zirakashvili and Volk [61] argue that this radially outward flow should terminate in a shock at several hundred kpc, where out streaming particles are re-accelerated.

The primary difficulty with this picture is that it will cause a discontinuity between particles accelerated in the disk and at the termination shock. The authors argue this problem can be avoided by requiring the maximum energy to vary from $Z \times 10^{15}$ eV at low galactocentric latitudes to $Z \times 10^{17}$ at high galactic latitudes. Following their analysis they find

$$E_{max} \approx Z \sin\theta \left(\frac{\Omega(\theta)}{5 \times 10^{-16}\ \mathrm{s}^{-1}}\right)\left(\frac{B_g}{2\ \mu\mathrm{G}}\right)\left(\frac{R_g}{15\ \mathrm{kpc}}\right)^2 0.6 \times 10^{17}\ \mathrm{eV}$$

Here Ω refers to galactic angular velocity at latitude θ, R_g is a galactocentric radius and B_g is the poloidal field strength at that distance. Assuming a spherical termination shock at $R_S = 300$ kpc the authors find a max energy of $E_{max} = 10^{17}$ eV for $Z = 1$ and $E_{max} = 2.6 \times 10^{18}$ eV for $Z = 26$, achieving ultra-high energy levels.

From all the previous arguments it might be apparent that a dimensionality constraint plays a role in general acceleration process due to astrophysical collisionless magnetic shocks . This line of thinking was formalized by Anthony Hillas and is visually represented in the magnetic field vs. length phase space—an example plot is shown in 2.3. The Hillas plot can be used to rule out systems which do not have the combination of field strength and size necessary to reach observed energies. For exam-

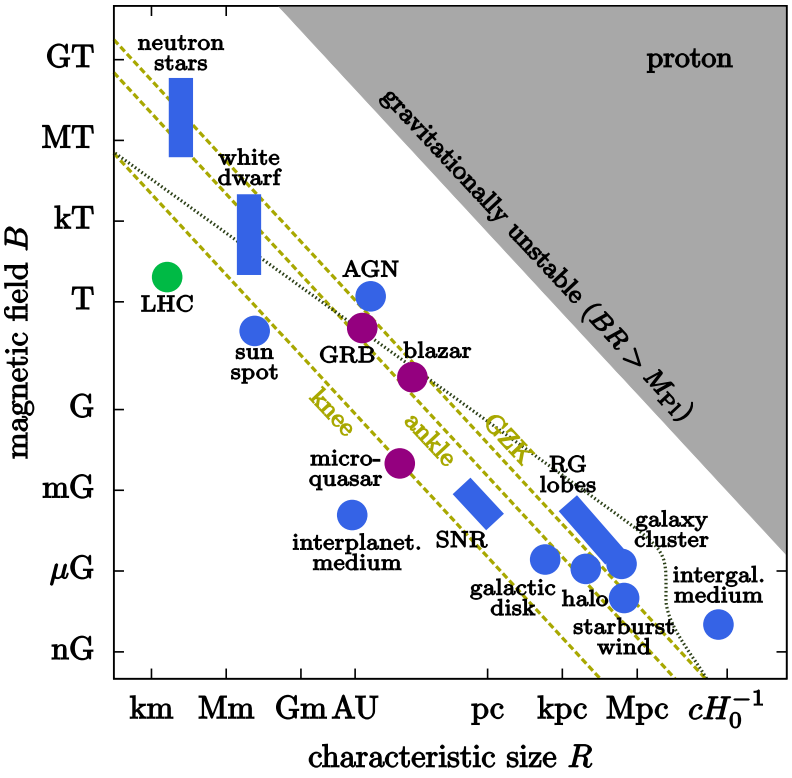


Figure 2.3: A typical Hillas diagram, taken from [62]. Independent measurements of B-field and length scale for various acceleration sites are shown and the dashed lines represent the maximum achievable energy per the Hillas criterion. For numerical reference the cosmic ray knee is ~ 3 PeV, the ankle ~ 5 EeV and the GZK cutoff about 60 EeV.

ple, the gyroradius of a charged particle in a constant B-field (normal to momentum) can be written as

$$r_L \sim \frac{E}{Z \times 10^{-3} EeV} \left(\frac{B}{10^{-6}\mathrm{G}} \right)^{-1} \text{ pc} \tag{2.2}$$

If we assume a 50 EeV proton and 1 μG field this gives $r_L = 50$ kpc, which is about twice the radius of the Milky Way. Thus the Milky Way, as an astrophysical source, fails the Hillas criterion and can be ruled out as an accelerator bottom up source for 50 EeV protons.

Neutron stars are potential sources due to their enormous rotational energy. A

typical neutron star with mass 1.4 $M_\odot$ and period $P_{10} \times 10$ ms (P_{10} is a pre-factor for 1 ms) gives an energy of 2×10^{50} erg/P_{10}^2. If these objects inject this energy into the cosmic ray bulk every 30 years it would be 5x larger than the power required to reach the local cosmic ray energy density.

Another tantalizing scenario is a neutron star inside a supernova remnant. The acceleration mechanism here, in addition to energy provided by rotational motion, might involve matter from the inside shell of the shock falling back on the star. One expects to find heavier nuclei at this part of the shell, which may explain mixed composition. In [60], the author estimates the power provided by rotation as

$$L_d = \frac{2}{3}\frac{\mu^2\Omega^4}{c^3}\sin^2\theta$$

where μ is the magnetic moment and θ is the angle between the magnetic and rotation axis and Ω is the angular frequency. However, a typical neutron star will generate $L_d \sim 2 \times 10^{39}$ erg/s which is sufficient to power cosmic rays with $E > 100$ TeV *only*. The maximum energy of particles accelerated at the interior of the shell due to the neutron star's surface field can be approximated as

$$E_{max} \sim \frac{eB_*R_*^3\Omega^2}{\sqrt{3}c^2}$$

where R_* is the radius of the star and B_* is the field magnitude at the surface. Assuming the pulsar has a 10 ms period with 10^{12} G field gives $E_{max} \sim 10^{17}$ eV, as shown in [60]. Modern estimates for millisecond pulsars place upper limits of the field at 10^{13} G [63]. These are known as soft gamma ray repeaters (SGRs) or anomalous X-ray pulsars (AXPs) and are classified as magnetars. Using this limit with a 10 ms period gives an E_{max} of 10^{18} eV. Because these are known as 'millisecond' pulsars it's

conceivable that the period might be 5 ms, or even 1 ms. Using a period of 5 ms results in factor of 4 increase in the max energy. If we assume a 1 ms period then we get a factor of 100 increase, giving a potential max energy of 10^{20} eV.

Another possible point source considered in [60] is a neutron star binary system. It can act as an accelerator when it accretes matter from its companion star or from a wind. Shocks in the accretion flow might be a natural candidate, however there are many features of the binary system that impede acceleration. The primary difficulty is the dynamical nature of the field deviations that scatter particles. In this environment they aren't frozen into the shock front which reduces the efficiency of energy transfer from the bulk motion of the cloud. Furthermore, because acceleration occurs in a region with very strong fields the probability of escape is very small due to a tight Larmor radius. This can be seen using $r_L = E/ZB$ kpc with B in μG and E in EeV. Assuming a 10^{18} EeV proton and typical field value of 10^8 G gives a Larmor radius of about 310 km, a factor of $\sim$100 the size of the neutron star which is much smaller than the shock size.

While shocks in the accretion disk may not be able to accelerate particles to the highest energy an alternative mechanism known as the "disk dynamo" can transfer energy to particles because of a difference in angular speed between the disk and neutron star. Particles are accelerated by a rotating electric field that is formed by the intersection of neutron star B field lines with the inner edge of accretion disk. The maximum energy can be approximated as

$$E \approx 3.5 \times 10^{14} \text{ eV } B_{12}^{-3/7} \; L_{38}^{5/7}$$

Here B is the field magnitude in units of 10^{12} G above the surface of the neutron star and L_{38} is the accretion luminosity in units of 10^{38} erg/s. From the equation above it's seen that lower field magnitudes lead to higher energies. If we take the luminosity to be Eddington, which depends on the neutron star mass and Thomson cross section (known constant) we have

$$L_{Edd} = 1.4 \times 10^{38} \ \frac{M}{M_{\odot}} \ \text{erg/s}$$

Using a typical pulsar mass of $M = 1.4 \ M_{\odot}$ and a lower bound of $B = 10^8$ G for the field of a millisecond pulsar [63] we find $E_{max} = 2.9 \times 10^{16}$ eV. If we consider that the field might be up to half as much as the assumed lower limit, this increases the max energy by a factor of 1.64, not much help. There is also little flexibility in changing the luminosity term as the Eddington limit for neutron stars is a ceiling based on the Chandrasekhar limit.

An intriguing recently proposed model describes how the inner bulge of the galaxy ejects superwinds reaching upwards of ~ 1600 km/s due to the increased star formation rate. This gas then collides with the halo gas a few kpc from the galactic center resulting in a strong termination shock where cosmic rays may be accelerated. For an observer outside the galaxy this phenomenon would appear as a large symmetric bubble about the galactic bulge. Their large extent makes these attractive candidates, satisfying this key aspect of the Hillas criteria. A paper outlining this scenario is [64]. The author derives a maximum energy of

$$E_p \lesssim 5 \ Z \ \text{EeV} \ \left(\frac{R_{max}}{1\text{kpc}}\right) \ \left(\frac{B}{5 \ \mu\text{G}}\right)$$

Taking a typical field of $B \sim 6 \ \mu$G and a max bubble radius of 3 kpc one finds

an upper limit of $E_p \lesssim 18\ Z$ EeV. Note that the small radius differentiates this mechanism from the mechanism discussed earlier that proposed a termination shock with the ISM, a distance of 100s of kpc away.

Finally, we briefly discuss active galactic nuclei and radio galaxies as potential accelerator sites. Nuclear regions are attractive candidates because of their compact size and potentially large B-fields. Reverberation mapping studies have shown the broad line region of the nucleus to be on the order of 0.01 pc [65] and the magnetic field can be estimated using jets of Fanaroff-Riley (FR) type II galaxies or an argument based on equipartition with electrons [66], giving a value of about 5 G. Using the dimensionality argument, protons can be accelerated up to 10 EeV. While AGNs are energetically capable accelerators, there are issues with the escape of the primary. Since the nuclear region is surrounded by an intense radiation field the primary will lose energy to this through a variety of processes and will limit the escaping energy to around 10^{16} eV. Another possibility is acceleration in the hot spots of radio lobes in FR galaxies where the density of photons is lower, allowing the primary to escape. In these regions the large magnetic field is extended over a larger region, permitting maximum energies of up to 10^{21} EeV. However, the caveat with this scenario is most FR objects are at cosmological distances, so the primary is unlikely to reach us without being degraded by intermediate matter.

Although we've discussed many source candidates, this section should not be taken as an exhaustive list. The theory of acceleration is an active field, and new models are continuously being developed and tested.

2.3 Propagation

After the primary leaves the acceleration site(s) it traverses the intermediate space, eventually being observed at Earth. Along this path the primary can interact with matter such as the intergalactic medium (IGM) or interstellar medium (ISM) or radiation fields, such as the cosmic microwave background (CMB). These processes will contribute to energy loss and possible fragmentation of the primary. Furthermore since primaries are nuclei (charged particles) they will be deflected by astrophysical magnetic fields. In this section I describe the dominant spallation effects and sources of magnetic deflection.

2.3.1 Radiation field interactions

Perhaps the most well known process associated with the highest energy primaries is the Greisen-Zatsepin-Kuzmin (GZK) cutoff [67, 68]. This results from the scattering of protons with the CMB via the delta resonance

$$p + \gamma \rightarrow \Delta^{+} \rightarrow p + \pi^{0} \tag{2.3}$$

$$p + \gamma \rightarrow \Delta^{+} \rightarrow n + \pi^{+} \tag{2.4}$$

since these photons are blue-shifted up to gamma rays in the primary's rest frame. This process is known as photo-pion production and it will "turn on" at roughly this threshold energy

$$E = \frac{m_{\pi}\left(m_{N} + m_{\pi}/2\right)}{\epsilon} \approx 7 \times 10^{16} \left(\frac{\epsilon}{\mathrm{eV}}\right) \quad \mathrm{eV} \tag{2.5}$$

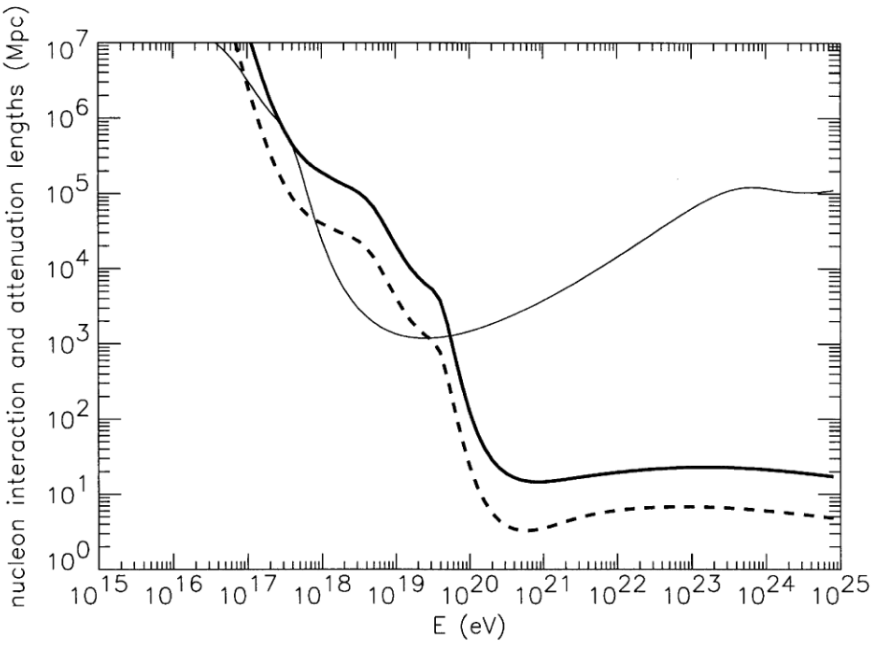


Figure 2.4: The solid line and dashed line is for proton photo-pion production attenuation and interaction, respectively. The thin solid line is for proton pair production channel. Primaries with $E = 100$ EeV originating beyond 50 Mpc are unlikely to be observed at Earth. Figure from [56].

where ϵ is the energy of CMB photon—roughly $\epsilon \approx 1$ meV. This gives a threshold energy of 70 EeV. As a result, most protons with $E \geq 70$ EeV have a high probability of reacting through this channel, continuing to lose energy through pion production, or conversion to neutrons which further accelerates energy loss. The consequence of this process is a distance horizon for the highest energy particles, since particles above the threshold energy will be attenuated. A useful plot (from [56]) showing the relevant length scales is shown in 2.4.

At lower primary energy photo-pion production is supplanted by the pair production resonance

$$p + \gamma \rightarrow p + e^+ + e^- \tag{2.6}$$

which occurs at the threshold of $E \approx 5 \times 10^{17}$ eV. Below this threshold, the most significant effect energy loss due to expansion of the universe (redshifting). While other interactions are possible, such as with molecules in the IGM, since it is a rarefied gas the cross-section is minimal and can be ignored.

In addition to protons and nuclei, photodisintegration is also the primary loss mechanism for ultra-high energy photon primaries, but in the form of pair production (where $'$ refers to primary photon)

$$\gamma' + \gamma \rightarrow e^+ + e^- \tag{2.7}$$

and with infrared background photons instead of the CMB around 150 TeV. The secondary electrons/positrons are also subject to inverse Compton scattering off the background photons. Unlike the CMB which is isotropic and has been very precisely measured, the universal radio background is not well constrained and is generated by galactic and extra-galactic sources.

For ultra-high energy neutrino primaries an analogous process to photodisintegration exists which is interaction with the relic neutrino background (RNB). The most likely energy loss channels are

$$\nu_i + \bar{\nu}_j \rightarrow \ell_i + \bar{\ell}_j \tag{2.8}$$

$$\nu_i + \bar{\nu}_i \rightarrow f + \bar{f} \tag{2.9}$$

$$\nu_i + \bar{\nu}_j \rightarrow \nu_i + \bar{\nu}_j \tag{2.10}$$

where i and j are the (not duplicated) flavor, ℓ is a non-neutrino lepton, and f is a charged fermion. Like the previously discussed primaries, degradation from interactions with nucleons (IGM or ISM) is negligible since the RNB density is much greater

than the baryon density. This effect is most important for "cosmogenic" neutrinos (resulting from pion decay from photodisintegration of GZK energy protons) with energies of a few EeV. It is important to note that presently there have been no observations of ultra-high energy neutrinos ($\gtrapprox$10 PeV) by the IceCube observatory [69]. This has important implications on the possible class of astrophysical accelerators for neutrinos, and possibly cosmic rays. Currently, IceCube observations are compatible with blazar and pulsar sources [69].

2.3.2 Magnetic deflection

Since UHECRs can be composed of charged elementary particles or nuclei they're trajectory can be modified by $\boldsymbol{B}$-fields. Since UHECRs likely originate in extragalactic sources, the intergalactic magnetic field (IGMF) must be considered in addition to the galactic magnetic field (GMF) as the primary makes its way to Earth.

At present the IGMF is poorly understood and its strength and morphology seriously under-constrained. No evidence exists for a general, organized IGMF; usually only approximations (or limits) of its average magnitude are available. Cosmological structure formation models predict values of $10^{-9} \leq B \leq 10^{-8}$ G along filaments [70]. For fields with coherence lengths equivalent to the Jeans' length, modern simulations tuned to Faraday rotation measures place an upper limit of $B < 1.7$ nG with 2σ confidence, and an upper limit for coherent fields across the universe of $B < 0.65$ nG [71]. For regions inside galaxy clusters the average field strength can be larger. For an average electron density of 10^{-3} cm^{-3} and length scale of about 1 Mpc the

expected field strength is on the order of 1 μG, which agrees with observed rotation measure dispersions of ~ 100 rad m^{-2} [72].

As one transitions to smaller length scales, like the size of a galaxy, the field strength and morphology become increasingly complex. The characteristics of magnetic fields in elliptical galaxies will be different compared to spiral galaxies which will differ from barred spirals. To some extent the field tends to follow the matter distribution, but not exclusively (in some situations the field is pitched up or down relative to the spiral arms). Furthermore, GMFs typically have multiple components: (1) the coherent/regular field which stays organized over dozens of kiloparsecs with strength ~ 5 μG, (2) the striated or ordered random field and (3) the turbulent or isotropic random field, both of which have $\langle B \rangle \sim 4$ μG. For a cartoon depiction, see Figure 2.5. The exact mixture and configuration of these fields in the Milky Way is not exactly known. Considerable effort has been spent on modeling the field, but a detailed description of every model is beyond the scope of this dissertation. Instead, I will present the basic building blocks of most models here, and a detailed description of a specific model in a later chapter.

For spiral or barred-spirals the two most common disk morphologies are axisymmetric (ASYM) and bisymmetric (BSYM), see Figure 2.6. In the ASYM case, the field lines follow the spiral pattern and either all point outwards or inwards. For the BSYM case, the fields are anti-symmetric relative to the spiral pattern/spin of the galaxy. This configuration allows for field reversals along a fixed azimuth. Symmetric fields are possible where the field has the same orientation in one half of the galaxy, relative to the zero meridian, and opposite in the other half which forms a "mirror"

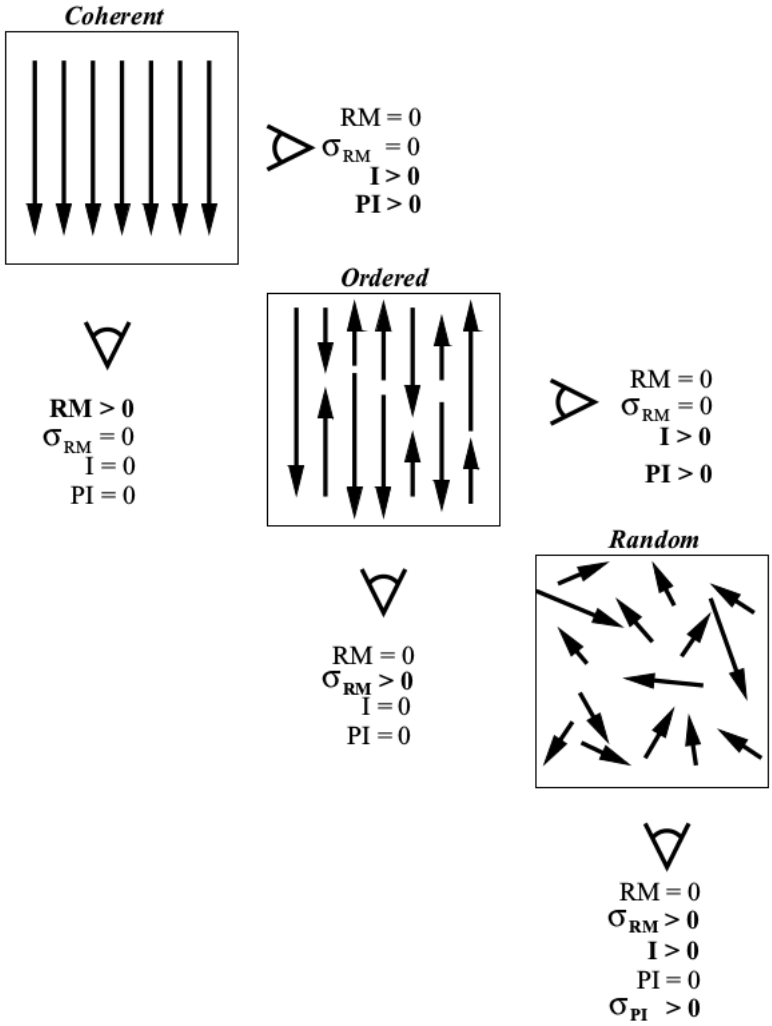


Figure 2.5: A simplified schematic of magnetic field lines for the three main configurations believed to be present in the Milky Way, reproduced from [73]. These are inferred from the presented observables, which depict the integrated line of sight. I denotes intensity and PI refers to polarized intensity.

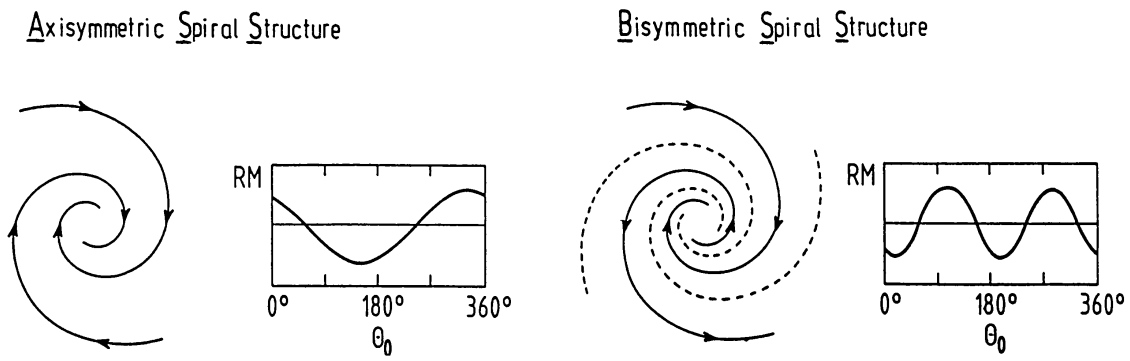


Figure 2.6: Diagram showing a basic axisymmetric (left) and bisymmetric (right). Figure from [74]. A key difference is the presence of multiple field reversals in the bisymmetric case.

symmetry. Anti-symmetric fields have the opposite configuration. Another configuration is a ring magnetic field, which follows the closed circular pattern of the host galaxy. Beyond the disk, it's also possible for the halo field to exhibit these types of symmetries, and even have a mixture of them.

The halo field, which is also referred to as "thick disk", is believed to extend roughly 3 kpc from the disk (but not necessarily having the same dorsal and ventral length) with a linearly decreasing field strength, tapering off to about 2.5 μG [75]. The structure in this region is purported to be uniform due to low variance of Faraday rotation measures [76]. However, it is possible for the northern and souther hemispheres to be have distinct properties. As noted previously, it's possible to have different symmetries in the disk and halo. For these situations a combined disk and halo configuration is possible. One example is the disk-even-halo-odd (DEHO) which exhibits a "butterfly pattern" in rotation measure skymaps.

Given these technical challenges, the UHECR community faces something of a paradox: theorists are presented with substantial difficulty in explaining, using known

(astro)physics, how primaries are accelerated, but experimentalists cannot simply come to the rescue due to the enormous complexities involved with the ballistic trajectories of the primaries. Nevertheless, experiments *can* place certain limits on physical scenarios which helps to point theorists in the right direction. In the next subsection we describe the impact of the primary with the atmosphere, and the foundations of cosmic ray indirect detection.

2.4 The extensive air shower (EAS)

After the primary has finished its tortuous journey through inter-and galactic magnetic fields, and if its trajectory happens to intersect with the Earth, it will collide with an atomic nucleus belonging to an atom of a molecule in the thermosphere: the mixture is 20% N_2 and 80% O_2 at about ~ 100 km above the surface. To get a sense for the violence of this event, let's compare the center of mass (CoM) energy of a 10 EeV proton primary fixed target collision with O_2 to head on collision of proton beams at the LHC. A typical proton beam at the LHC will have kinetic energy $E_p \approx 7$ TeV resulting in a CoM energy of $\sqrt{s} = 14$ TeV. If we assume the cosmic ray scatters off a proton in O_2 the CoM energy will be $\sqrt{s} \approx \sqrt{2E_p m_p} \approx 141$ TeV, where we've approximated the proton mass to $m_p \approx 1$ GeV/c^2.[1]

One of the most important rules of elementary particle physics can be thought

[1] This simple calculation shows the initial interaction probes physics beyond our current, and probably far-future, technological capabilities. Given this, it is entirely possible there are interactions taking place not predicted by the standard model. However, in this section I present a description grounded in standard model processes.

of as a Murphy's law equivalent: if a process can happen, it will happen. A more precise statement would be: if an interaction doesn't violate the laws of physics (causality, conservation rules, etc.) it will happen, at various probabilities. The first (and subsequent) interaction can be described by

$$p + p \rightarrow \text{e.m.} + \text{hadrons} + \text{muons} \tag{2.11}$$

where e.m. refers to electromagnetic particles usually γ and $e^{\pm}$ resulting from pair production. Hadrons is a catch-all and includes $p, \bar{p}, n, \pi^{0,\pm}, \eta, K^{\pm}, \ldots$. Muons are generated from the decay of pions—neutrinos are as well, but so little energy is transferred to these particles they can be safely ignored. After the first interaction there is still considerable energy to be partitioned into new particles, generating a cascade effect where the particle number continues to multiply, forming a localized and dense core of hadrons. The core continues to grow in size and density until it reaches a maximum number, at which point the average energy of the daughter particles is insufficient for creating new generations. A cartoon depiction of this process is shown in Figure 2.7 The slant depth at which this occurs in the atmosphere is called $X_{\max}$. The slant depth is used to reduce the physics to a 1-dimensional transport equation. For a specific energy, the flux can be calculated using

$$\frac{dH_i(E,X)}{dX} = -\frac{H_i(E,X)}{\lambda_i(E)} - \frac{1}{c\tau_i\gamma\rho}h_i(E,X) \tag{2.12}$$

where the subscript i refers to a species of hadron, λ is the mean free path, τ the lifetime, γ the Lorentz factor and ρ is the density of air. A further necessary step is to integrate over all possible energy values, which is not explicitly shown in this equation. In general an analytic solution is not possible without additional approximations.

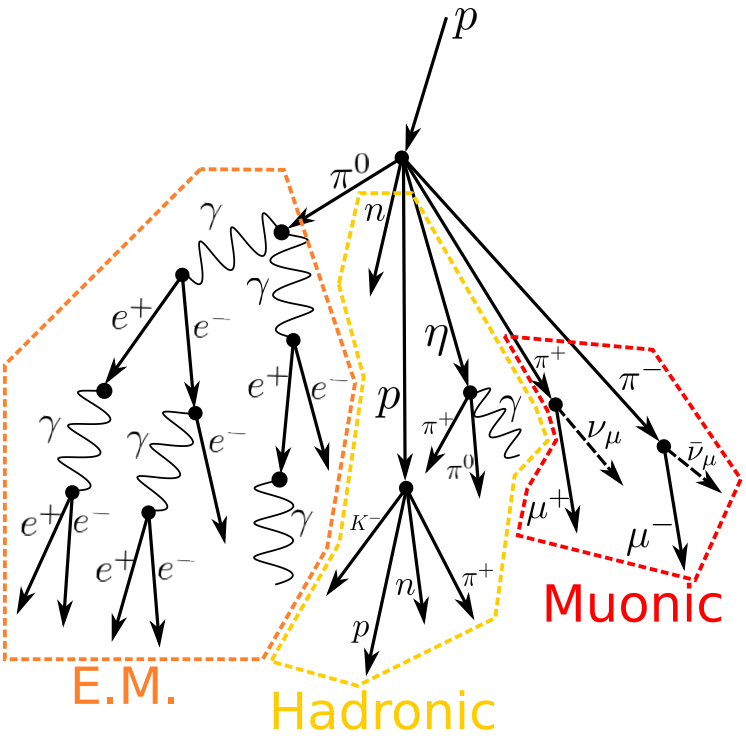


Figure 2.7: Initial interaction of primary and target atom. The first few generations of particles are shown, along with a grouping of the main classes of particles. Note: the diagram is for illustration and does not explicitly obey all conservation rules.

A common technique for solving the transport equation is numerical/Monte Carlo integration.

As the EAS travels toward the surface, the swarm of particles forms a disk/plane, with its highest density at $X_{\max}$. The over all density then continues to decrease until impact. In the shower plane, the density is extremely high in the core, roughly $r < 300$ m from the shower axis ($r = 0$) for a primary of ~ 1 EeV. The density decreases exponentially in the shower plane. However, technically speaking the shower front is not an ideal plane, but has some curvature: an approximately spherical or conical face with radius on the order of ~ 2000 m and a thickness of about 3–5 m. The shower development is rich with information about the primary's composition and initial energy.

To understand how primary energy is related to particle content in air showers we can consider the simple Heitler toy model [77] which follows the branching of the primary (energy E_0), which occurs after one collision length, λ. Branching continues until the depth of $X_{\max}$ at a critical energy E_c, after which nodes will be "absorbed" due to energy loss and other decays. In this model $X_{\max}$ is related to the primary energy by

$$X_{\max} = \lambda \frac{\ln\left(E_0/E_c\right)}{\ln 2} \tag{2.13}$$

The obvious limitation of the model is that one must know the critical energy parameter. If branching occurs purely in terms of electrons in air, this value would be $E_c \sim 80$ MeV, compared to muons in rock which is about 500 GeV [60]. The critical point is that $X_{\max} \propto \ln\left(E_0\right)$. The model can also be modified for nuclear primaries

which changes the dependence to

$$X_{\max} \propto \lambda \ln\left(\frac{E_0}{AE_c}\right) \tag{2.14}$$

where A is the number of primary nucleons. A more sophisticated treatment considers the spectrum of decaying hadrons into muons. Past work [78]on Monte-Carlo cascade equations gives an estimate for total number of muons at the ground;

$$\langle N_\mu \rangle \approx A \times \frac{14.5 \text{ GeV}}{E_\mu \cos\theta} \left(\frac{E_0}{AE_\mu}\right)^{0.757} \left(1 - \frac{AE_\mu}{E_0}\right)^{5.25} \tag{2.15}$$

where θ is the zenith angle of the primary and $E_\mu = 0.53 \text{ TeV } \times \left(e^{0.4X} - 1\right)$, where X in this equation is slant depth in kilometers of water equivalent, instead of air. This more complicated form accounts for higher energy primaries pushing deeper into the atmosphere before their first interaction, and the increased density will enhance pion reactions. In the limiting case where $E_0 \gg AE_\mu$ the equation simplifies to

$$N_\mu \propto A^{0.243} \tag{2.16}$$

demonstrating how the multiplicity of muons at a given depth is a direct tracer of the primary composition. Although there are more sophisticated modern forms of this equation which are optimized for specific air shower experiments, the underlying principles are the same. However, most detector types in use respond (in complicated ways) to multiple families of secondary particles (e.m., hadronic, and muonic) and so disentangling each component is crucial for determining the composition.

Since air shower detectors are spaced out on the ground and sample the shower plane at discrete points it is useful to measure the number of muons as a function of core distance. Here we use the multiplicity distribution derived from [79], but this

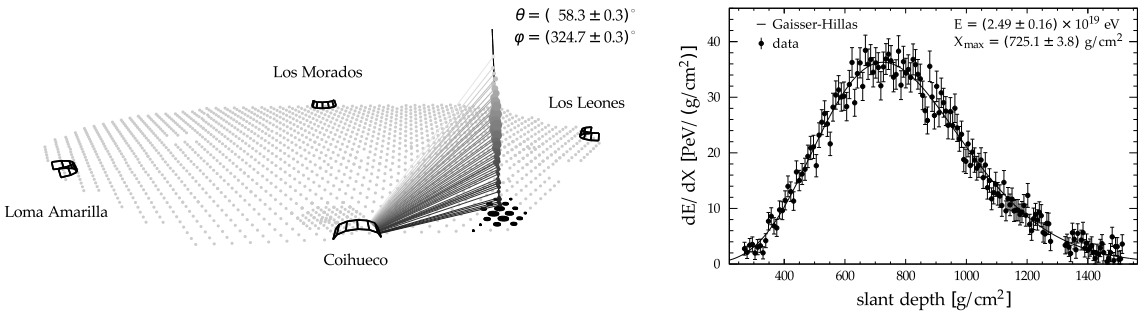


Figure 2.8: Left: diagram showing a telescope observing fluorescent light from an EAS. Right: plot showing the energy deposition as a function of slant depth. A fit to this data returns $X_{\max}$. Figure from [80].

can be modified depending on the experimental setup (see Chapter 3)

$$\frac{1}{N_\mu}\frac{dN_\mu}{dr} = C\frac{r}{(r+r_0)^\alpha} \tag{2.17}$$

where r is distance from the core, C is a normalization constant

$$C = (\alpha-1)(\alpha-2)/r_0^{2-\alpha}$$

with r_0 and α being free parameters. This is an example of a lateral distribution function (LDF).

In addition to inferring EAS properties from measuring particle densities on the ground, it is also possible to do this by measuring fluorescence photons generated by the shower as it propagates through the atmosphere (e.g. Fly's Eye experiment described in previous chapter). An example event at the Pierre Auger Observatory is displayed in Figure 2.8 The fluorescence photons are mostly generated by the e.m. cascade and are directly proportional to the primary energy. In other words, after accounting for losses due to multiple scatterings by aerosols, etc., measuring air fluorescence of an EAS gives an almost calorimetric energy. With this technique $X_{\max}$

is determined in a way similar to using densities at the ground—by fitting data to a function to recover the observable. An example is the Gaisser-Hillas function [81]

$$\frac{dE}{dX} = \frac{dE}{dX}\Big|_{X_{\max}} \left(\frac{X - X_0}{X_{\max} - X_0} \right)^{\frac{X_{\max} - X_0}{\lambda}} e^{\frac{X_{\max} - X}{\lambda}} \tag{2.18}$$

where X_0 and λ are "shape" parameters and are usually estimated by the experiment. As we will see in Chapter 3, it's possible to join the fluorescence and particle counting technique together to form a hybrid detector which benefits from the high duty cycle of surface detector arrays, and calorimetric measurements from fluorescence.

In keeping with the theme of the air shower development encoding all the important physics, an additional observable that can be reconstructed is the arrival direction: the angle offset from zenith on the ground, combined with the observation time (hour, minute, second). This results in a unit vector in zenith, θ, and azimuth, φ, (in local station coordinates) which can then be converted to right ascension (α) and declination (δ), or galactic longitude (ℓ) and latitude b.

This vector, or shower axis, is found using either fluorescence or surface array techniques. Since fluorescence telescopes are very fast multi-pixel cameras, they are able follow the shower from start to finish and record the orientation of the shower axis. With some geometrical transformations the track in elevation and azimuth can be expressed as an arrival direction since the distance to the shower axis from the detector is known. An example Auger event seen by a fluorescence telescope is shown in Figure 2.9.

Arrival direction reconstruction works in a similar way for ground base arrays, but instead of following the timing through the atmosphere, the timing follows the

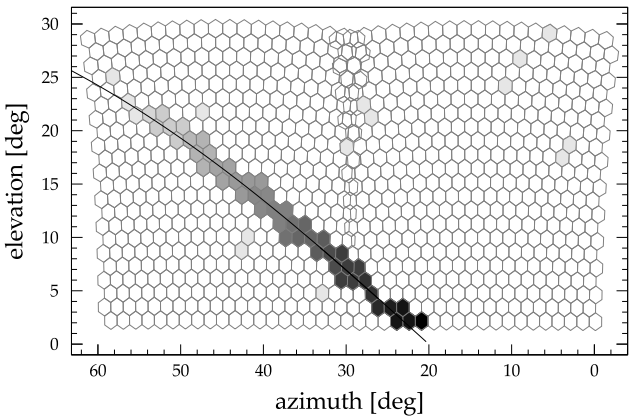


Figure 2.9: In this plot, lighter pixels occur at earlier times. The progression of the EAS can be clearly seen. The black line follows the direction of the shower axis, and is sometimes referred to as the "shower detector plane". Figure from [80].

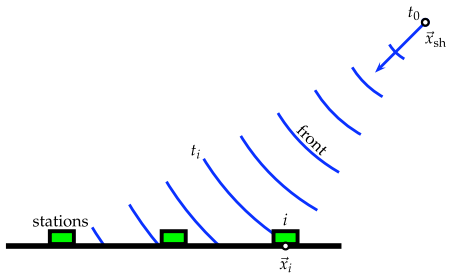


Figure 2.10: Diagram showing the shower progress and its intersection with a ground array detector plane. Figure from [82].

impact of the shower plane on the ground. If enough stations sample the EAS, the start times can be fit to an inflating sphere model, see Figure 2.10.

Since these techniques can vary from experiment to experiment, I will go into more specific details for the reconstruction process of the Pierre Auger Observatory in Chapter 3. From these reconstructed EAS observables, one can generate the two most significant and widely published data products: the energy spectrum and arrival direction anisotropy statistics. In the next chapter we describe the two largest cosmic

ray observatories in operation and also present and compare these important data products.

Chapter 3: The Pierre Auger and Telescope Array observatories

3.1 Preamble

In this chapter low-level technical details are presented about the design of the two largest hybrid cosmic ray detectors. I discuss their designs, how they translate particle counts to signals, along with reconstruction techniques. In many respects, the design philosophy and operation of Auger and Telescope Array (TA) are similar, but attention will be called to important differences such as sensitivity to air shower channels, surface detector array efficiency, and the reconstruction/analysis pipeline. This chapter will build on the basic physics concepts outlined in a previous chapter. I also note that while the importance of fluorescence detector can not be overstated, since they are not a focus on much of the work done in this dissertation, a higher level overview is given instead of dwelling on specific details.

3.2 Experiment sites

For hybrid observations a high desert climate is ideal to improve the seeing and duty cycle of the fluorescence telescopes. For TA, a site roughly 200 km south west of Salt Lake City, Utah, USA was chosen at 112.91° W, 39.3° N. The average elevation

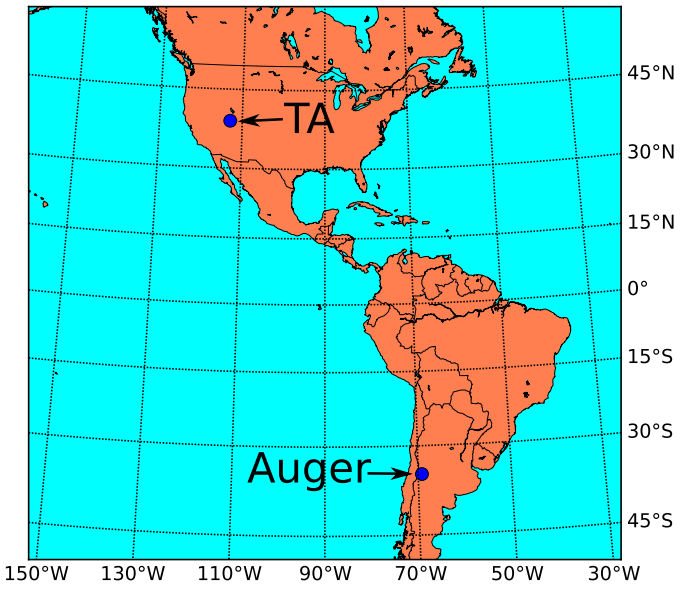


Figure 3.1: The locations of the two experiments using a Lambert-Conformal projection.

is 1382 m above mean sea level (a.m.s.l) and the Köppen climate classification is tropical and sub-tropical steppe climate (BSk). The public land is owned by the US federal government and managed by the Bureau of Land Management (BLM).

The Auger experiment is roughly 350 km south of Mendoza, Argentina at 69.2° W, 35.2° S with the headquarters in the town of Malargüe. There is limited industrial activity in the town resulting in little pollution, and high quality air. The site is mainly a flat plain with multiple dirt and gravel roads. The average elevation is 1400 m a.m.s.l in a dry-summer subtropical (a.k.a Mediterranean) climate (Csb). A map showing the locations appears in Figure 3.1. A majority of the land is owned by about 100 ranchers and is used mainly for animal grazing, with some also being owned by the province of Mendoza or the Argentina Natural Resources Administration. A small annual rent is paid per detector.

3.3 Collaborations

Because construction of massive hybrid observatories requires large capital investments the projects are usually composed of many institutions hailing from different parts of the world. The Telescope Array Project consists of 28 institutions across 5 countries (and 3 continents). There are about 130 TA members. Funding is likely derived from government agencies who support fundamental research in the 5 participating countries.

The Pierre Auger Observatory is formed by 82 institutions across 16 nations (4 continents) who are signatories to an international memorandum of understanding which has become a well known model for collaborations of this size. The project is funded by 2 EU international agreements as well as UNESCO, along with support from about 46 government agencies within the member countries and a few private research foundations. Presently there are 459 collaboration members. The project was initially conceived by Nobel laureate Jim Cronin and Alan Watson. Construction and early operation was managed by Fermilab, recently taken over by Karlsruhe Institute of Technology.

3.4 Surface detector arrays

The main data generator for Auger and TA are the arrays of surface detector (SD) stations covering large areas. The general design consists of a detector medium, a power source, photomultiplier tubes (PMTs) and electronics to record PMT signals,

and a radio communication system. The stations are light tight and self contained, and operate continuously (100% duty cycle) unless prevented by a fault or malfunction. In this section I described the Auger and TA SD array separately and conclude with a comparison highlighting key differences and similarities.

3.4.1 Auger SD station

Presently there are 1660 stations forming the array in a triangular grid. The main array uses a separation of 1500 m which reaches full efficiency for primary energies of $E > 3$ EeV. An infill array of 61 stations with 750 m is used to probe energies of $E > 100$ PeV between the 2nd knee and ankle spectral features—it covers an area of 23.5 km^2. It becomes fully efficient at $E > 300$ PeV. The infill is also equipped with scintillators, two 10 m^2 and two 5 m^2, buried 2.3 m below the surface. This is colloquially referred to as AMIGA: Auger Muon and Infill for the Ground Array, and in addition to low energy sensitivity, it can be used for composition studies.

The Auger SD station is a 3.6 m diameter 1.2 m high water Cherenkov detector (WCD) filled with 12,000 L of highly purified water. The water volume is contained within a highly reflective Tyvek® liner which has three windows for downward facing PMTs (Photonis XP1805/D1) at the top. The PMTs are symmetrically distributed and each is 1.2 m from the central axis. A diagram and photo is shown in Figure 3.2.

For each PMT, two channels are read out by the electronics: low-gain (AC coupled anode) and high-gain (8th dynode). A 5-pole Bessel filter with -3 dB attenuation at 20

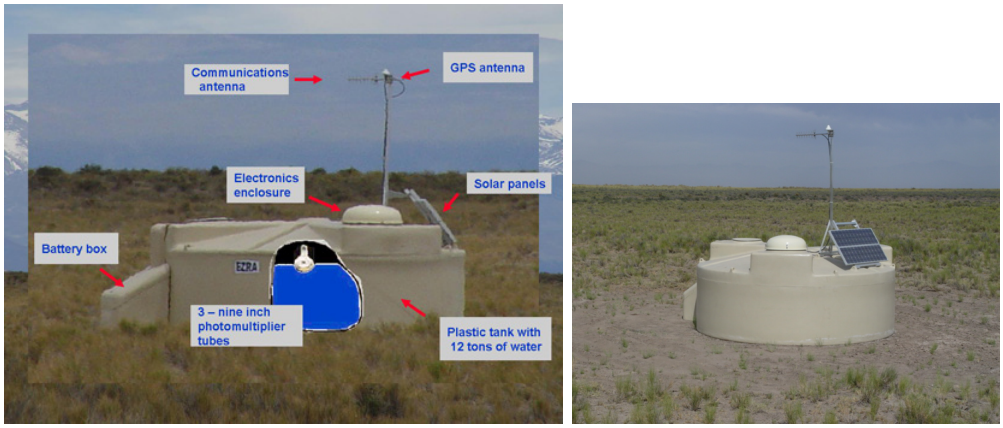


Figure 3.2: Left: the Auger WCD with cross-section showing internal tank. Right: photo of WCD operating in the Argentine pampa.

MHz is used in each channel to reduce noise without effecting the temporal features. The gain difference of the anode and last dynode is 32. Both channels are digitized by Analog Devices AD9203 10-bit 40 MHz semi-flash analog to digital converters (ADCs) with range 0-2 V. Digitizing both channels increases dynamic range (up to 15 bits) and allows sampling the EAS closer to the high density core. High voltage is provided to the PMT directly from its base board and linearity is verified using a programmable LED flasher setup.

The electronics motherboard (also called "unified board" or UB) is mounted inside a dome on the larger of the 3 PMT hatches. It's a semi-modular design, with a detachable front-end board, global positioning system (GPS) satellite receiver and radio modem. The UB has an IBM PowerPC 403 GCX 80 MHz central processing unit (CPU), 32 MB of system memory (DRAM) and runs the OS9 operating system.

The front-end contains the ADCs and Altera ACEX ®and Cyclone ®field programmable gate arrays (FPGA) which implement the triggering algorithms. When

a trigger occurs, digitized data stored in an external memory buffer, and eventually retrieved by the OS if a physics trigger occurs.

Detecting air shower requires synchronized, accurate and precise clocks in order to find coincidences between neighboring stations. A Motorola Oncore UT+ is used to receive 1 pulse per second (PPS) timing from rubidium or cesium atomic clocks onboard the satellites. An application specific integrated circuit (ASIC) running a 100 MHz 27-bit clock which latches to shower triggers. An additional 40 MHz ADC clock is latched to the PPS. These systems are used in tandem to provide timing accuracy for shower events with 10 ns root mean square (RMS) GPS synchronization precision.

To translate from ADC counts to physical units a continuously running online calibration system is used. Because of their omnipresence (rates of a few kHz) atmospheric muons are used for the calibration source. Specifically, the signal (Cherenkov light) from a central and vertically through going muon. This corresponds to a minimum energy of $E \approx 240$ MeV. Histograms are formed by slow control processes for a minute's worth of data at rate of 70 Hz, for events satisfying a "calibration event trigger", which is a threshold condition with a 3-fold coincidence component. The ADC bins at peak current, as well as the number of bins corresponding to the integrated signal are clearly visible in the histograms. The technique is verified by externally triggering on a vertically aligned muon telescope, where results are published in [83]. In Figure 3.3 the calibration histograms are displayed. In this plot, the peak due to vertical muons is clearly visible (second hump if looking at solid black lines) and is referred to as Q_{VEM}. Given the mean gain for PMTs is 3.4×10^5, this results in

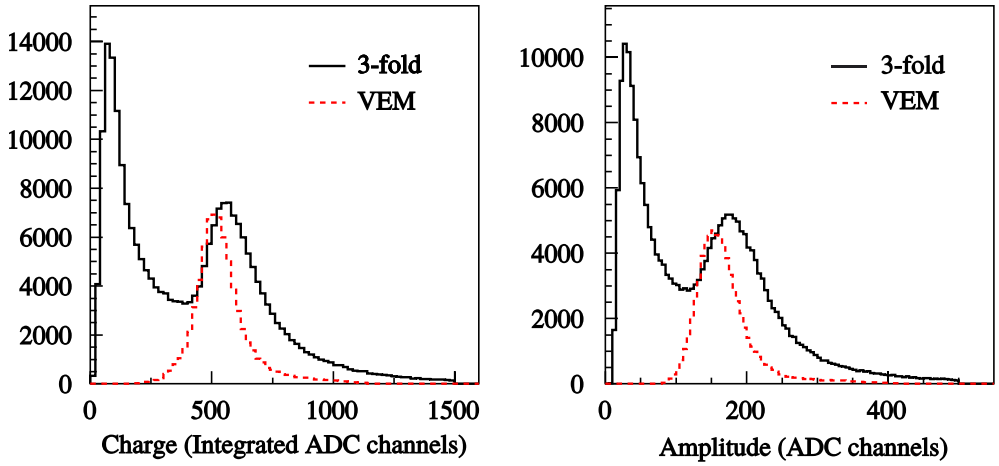


Figure 3.3: Left: charge (integrated signal) histogram. Right: pulse height/current (signal peak) histogram. Red-dashed line shows data generated using the muon telescope trigger, while the solid black line shows typical data from the calibration trigger. Figure from [83].

about 94 photoelectrons (p.e.) per VEM. The first hump in the data is due to a convolution of the calibration trigger on a sharply declining population of low-energy particles. For the charge histogram, the offset of the true vertical muon hump in real data is about 5-10 bins, corresponding to a systematic error of about 1-2%. The formal definition of the signal is

$$S_D \ [\text{VEM}] = \frac{\sum_{i=i_0}^{i_f} Q_i}{Q_{\text{VEM}}} \tag{3.1}$$

where S_D is the dynode (high gain) signal in VEM units, i_0 is first signal bin and i_f is the last signal bin, and Q_i is the ADC count in the 25 ns bin. The start and stop bin is dynamically found for each trace using an algorithm that compares moving statistics of the signal to the baseline. The anode signal is calculated in the same way, but includes the gain ratio

$$S_A \ [\text{VEM}] = \frac{\sum_{i=i_0}^{i_f} Q_i}{Q_{\text{VEM}}/32} \tag{3.2}$$

The station's radio modem, the subscriber unit (SU), is connected to the UB using a non-standard serial cable and uses a custom communication protocol to send data and receive commands from the central data acquisition system (CDAS), via a base station unit (BSU) which handles long distance communication. The data rate is about 1200 bit/s. In addition to data and commands, the station will also periodically transmit diagnostic information such as voltage levels, temperature, humidity, water level, scalers, etc. to monitor the performance of the detectors.

Although it uses a different detection medium, the general design of a Telescope Array SD station is similar.

3.4.2 Telescope Array SD station

Currently there are 507 stations arranged in a square grid using a separation of 1200 m covering an area of 700 km^2. The array is fully efficient for primaries $E > 10$ EeV with $\theta < 45^\circ$ with an aperture of 1100 km^2. Like Auger, an infill array is used in TA, the TA low energy extension (TALE) [84], to increase sensitivity to low energy showers. In TALE two spacings are used: 40 counters with 400 m separation for an efficiency of 10% for $E = 30$ PeV and about 60 stations with 600 m separation giving an efficiency of 10% at 100 PeV.

The TA SD station is an elevated rectangular scintillator about 2 m $\times$ 1.5 m. The detector medium is polyvinyl toluene plastic scintillator. Four slabs are used in total to form a two layer system (2 on top, 2 on the bottom). In each slab, the top side is embedded with grooves where 104 wavelength shifting fibers are used to guide

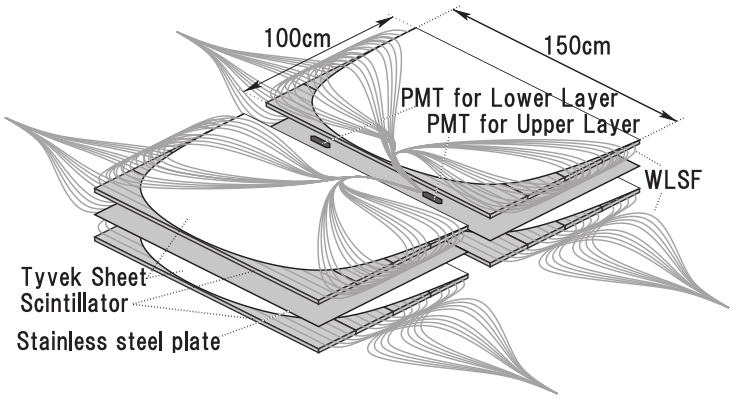


Figure 3.4: Exploded diagram of the TA SD station detector, from [85].

light to PMTs (Electron Tubes 9124SA), one for each layer located in the middle. The scintillators are wrapped in diffusive reflective material (Tyvek ®) and encased in a stainless steel box, under an iron housing for weather and sun protection. A schematic showing the inside view of the detector, from [85] is shown in Figure 3.4.

A single channel (anode) for each PMT is read out by the electronics and digitized by a 50 MHz 12-bit Analog Devices AD9235 flash analog digital converter (FADC). High voltage is provided by the PMT base. Linearity is maintained by an LED calibration system.

A modular electronics kit is situated inside a battery box, below the solar panel mount. The main board has a 266 MHz Renesas SH-4 processor and system memory which runs the TA data acquisition firmware. Data is connected to a front-end board where triggering algorithms run on a Xilinx SPARTAN 3 FPGA. When a trigger occurs, the waveform data is sent to main memory to be handled by the CPU. The FPGA also monitors the FADC baseline and adjusts threshold levels to maintain trigger rates.

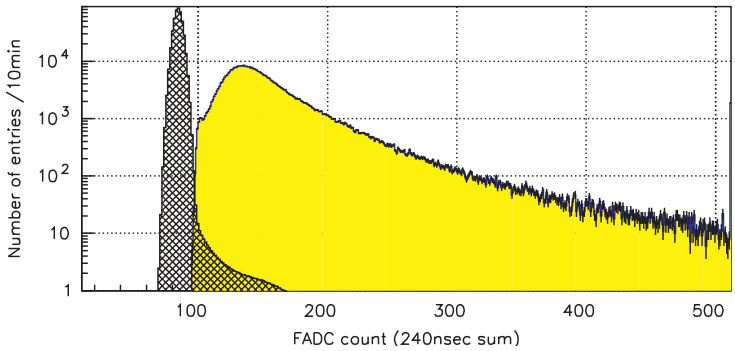


Figure 3.5: The hatched distribution is for the baseline or "pedestal" integrated counts, and the yellow distribution shows integrated Level-0 counts from the calibration system over a 10 minute period. Figure from [85].

The PPS is provided by a Motorola Oncore M12+ GPS receiver. A time stamp with 20 ns precision is provided by a 50 MHz clock component on the main board. In addition to the time stamp computed by the firmware, the number of clock counts after latching to the PPS are also saved and transmitted with trigger data.

To convert to physical units, minimum ionizing particles are used as a calibration source. According to Geant4 simulations of the detector volume, a vertical muon with kinetic energy 300 MeV is the lower bound for producing fluorescent light from the de-excitation of π orbitals in the plastic scintillator. The deposited energy spectrum for this muon peaks at around 2 MeV. The TA calibration system collects 240 ns worth of data from Level-0 triggers (signal $>$ 0.3 MIP), -4 bins ahead and +8 bins behind the trigger time. This is used to build a charge histogram where the peak corresponds to 1 MIP. An example of this data is shown in Figure 3.5, taken from [85]. The formal conversion from hardware to physical units is

$$S \ [\text{MIP}] = \sum^{12} S_0 - 1.5 \sum^{8} S_p \tag{3.3}$$

where S [MIP] is the integrated signal in MIP units, S_0 are baseline ADC counts and S_p are Level-0 trigger ADC counts and the factor of 1.5 is related to PMT gain.

Each e-kit has a communications board which allows the station to join a wireless local area network (WLAN). Packets are sent and received using a custom networking protocol to data acquisition (DAQ) computers deployed at the base of the three communications towers in the array. In addition to shower data, diagnostic information about PMT linearity, power, GPS and the environment are also sent and recorded.

3.4.3 Comparison summary of detector types

The fundamental role of SD stations is to count particles in the EAS, as these secondary particles allow inferences about the primary. Both collaborations use detector mediums sensitive to EAS components, but neither have the ability to accurately distinguish the types of particles depositing signals (methods exists for rough approximation). In the Auger WCD and the TA scintillator the following channels will contribute to the measurement: $\gamma, e^{\pm}, \mu^{\pm}$. The symmetry and aspect ratio of the WCD give it a mostly uniform response over a large span of zenith angles, while the scintillator slab will have a reduced cross-section at higher zenith angles. Both mediums are widely used in particle physics and are well tested and understood technologies.

Both experiments rely on PMTs to record fast and dim (macroscopically speaking) pulses of light. The collection area of the Auger PMTs (9" diameter) is larger which directly observes the water volume. The TA PMTs (1.13" diameter) are coupled to

a cookie which joins multiple WLS running through the scintillation volume.

Although the high level electronics design is similar in both experiments, the actual architecture and vendors used is quite different. Since TA was constructed a few years after Auger, it benefits from higher speed and denser processors. A detailed comparison of specific differences in how the electronics handle the analog waveform, through digitization, and processing of data bits is beyond the scope of this dissertation. However, it is fair to say that both designs handle the analog data in a way that is most efficient and compatible with the hardware design goals. A detailed description of the Auger front can be found in [86], and further details about the local station controller (LSC) and communications systems can be found in [82, 87]. Further details on the TA hardware can be found in [85, 88, 89].

For convenience, a summary highlighting the key features of each experiment is shown in Table 3.1. This should not be taken as an exhaustive list.

3.5 SD triggers

In this section I discuss the hierarchical scheme used to separate EAS signals from background. The format follows the last section where the experiment's design and implementation are described separately, and then compared. This time we start with the TA triggers, followed by Auger.

Item	Auger	TA
Station Geometry	cylindrical	rectangular cuboid
Dimensions	$r = 1.8$ m, $h = 1.2$ m	$l = 2$ m, $w = 1.5$ m, $h = 0.012$m
Detector medium	Ultra pure water (> 15 MΩ cm)	Polyvinyl toluene: C_9H_{10}
Grid type	triangular	rectangular
Main array separation	1500 m	1200 m
Infill name	AMIGA	TALE
Infill separation	750 m	400 m, 600 m
Main array exposure	≈ 5500 km^2 sr yr	≈ 890 km^2 sr yr
Main array efficiency	100% for $E > 3$ EeV	100% for $E > 10$ EeV
Upgrade plans	Auger prime	TAx4
Clock	GPS	GPS
Hemisphere	South	North
Calibration unit	VEM	MIP
Mean altitude (m.s.l)	1400 m	1402 m
Main array area	3000 km^2	700 km^2
Wireless bandwidth	1100 b/s	11 MB/s
Light sensor technology	PMT	PMT

Table 3.1: Some of the main features of the TA and Auger surface detector arrays.

3.5.1 TA SD array trigger system

The first stage is defined as Level-0 (L0), which continuously evaluates 128 FADC bins for both layers. When a sampled waveform is integrated and has ≥ 15 counts more than the baseline (which roughly corresponds to ≈ 0.3 MIP), and this condition is true for both layers simultaneously, this signal is promoted to a Level-0 trigger. These occur at roughly 750 Hz.

The next stage is defined as Level-1 (L1). This operates in the same way as L0, but the integrated threshold is ≥ 150 counts instead of ≥ 15, which corresponds to about 3 MIP. This occurs at roughly 18 Hz, see Figure 16 in [85]. Unlike the previous trigger, when a L1 occurs the time stamp is transmitted to the communication tower and received by the DAQ computer. Programs parse all incoming L1's, which in addition to time stamps have station identification (ID). The station ID provides a location. The program uses the timing and position information to determine if a space-time correlation exists which indicates an EAS has been observed. The geometrical requirement is 3 adjacent stations with each pair having a separation of $\leq$ 2400 m, and the temporal requirement is the absolute difference of all L1 permutations must be ≤ 8 μs. If these are satisfied, it is recorded as an Level-2 (L2) trigger event, which can be considered the "physics" trigger. Figure 3.6 shows the allowed combinations of stations for this trigger. After finding an L2, the DAQ computer records the first occurring L1 time stamp from the triad as the trigger time, and sends a request for FADC (128 bin length, or 2.5 μs) waveforms from all stations with $\Delta t < |32|$ μs of the trigger time. The waveforms are then combined with the 10

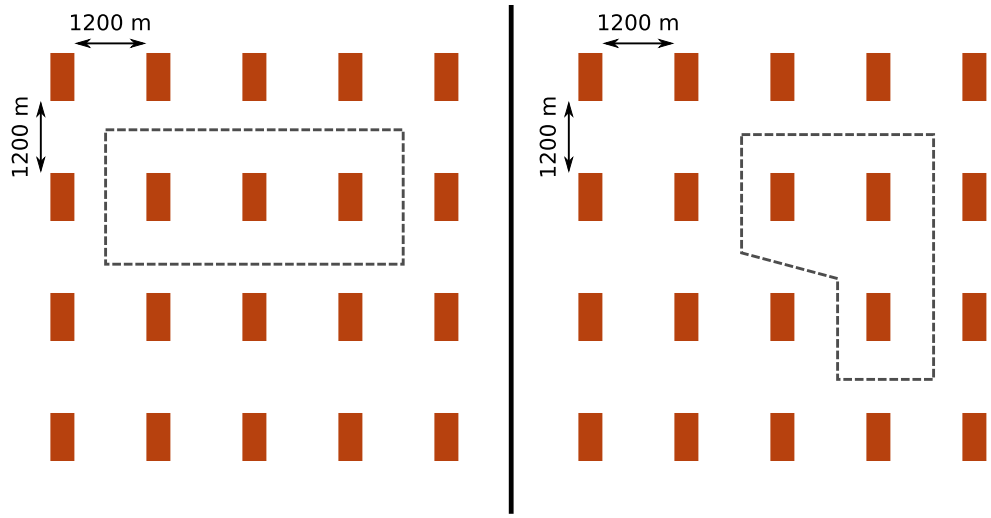


Figure 3.6: Left: an example L2 configuration where the maximum separation between any pair is 2400 m. A rotation of ±90° is also allowed. Right: the second L2 basis configuration. Here a rotation of any multiple of ±90° is allowed.

min. monitoring/calibration data to find the integrated signals during offline analysis and ultimately reconstruct the shower.

3.5.2 Auger SD array trigger

There are several local trigger algorithms running on the Auger SD station which can be broken up into 3 groups: the calibration (discussed in the last section), scaler, and EAS (physics) triggers. The scaler trigger is used for ancillary physics purposes, such as studying the impact of solar activity on galactic cosmic rays [90]. The first level of shower triggers, called "T1", are implemented using two separate algorithms. The first is a basic threshold trigger which requires the coincidence ($|\Delta t| < 1$ ns) of 3 PMTs with signal $> 1.75\ I_{\mathrm{VEM}}$, which is roughly 88 counts over the baseline. This T1 type, usually denoted T1-TH, is designed to find signals where features are not

spread out in time. In terms of an EAS, this is usually a shower at high zenith angle which must traverse more atmosphere, resulting in a shower plane richer in muons when it reaches the surface. This effectively reduces the background atmospheric rate from 3 kHz to $\sim$ 100 Hz. The second T1 algorithm is a time-over-threshold (ToT) which requires 13 time bins in a 3 μs (120 bins) gate to be larger than 0.2 I_{VEM} (10 counts above baseline), and this condition must be in coincidence for 2 out of 3 PMTs. This trigger is designed to capture vertical events (small zenith angle), and sample the EAS far from the core of high energy showers, in addition to cores of low energy showers. The T1-ToT occur at a rate of roughly 1.2 Hz.

The second layer in the trigger scheme is called T2, which places further constraints on the sampled T1 events. A T1-TH can be promoted to T2-TH if a coincidence of the 3 PMTs have a signal $>$ 3.2 I_{VEM} (about 160 counts over baseline). Any T1-ToT is automatically promoted to a T2, becoming T2-ToT. This layer gives a T2-TH rate of about 23 Hz while the T2-ToT is the same as T1-ToT.

Two additional triggers for the array have been implemented in the summer of 2013 called time-over-threshold-deconvolved (ToTd) and multiplicity-of-positive-steps (MoPS). These are designed to improve the array efficiency for lower energy primaries, but also for photon and neutrino initiated showers. Technical details of the algorithm is discussed in [82].

After each PPS cycle the SD station wirelessly transmits the list of T2 events (all types) to a central processor, the central data acquisition system (CDAS). Even though the radio link is not high-throughput, this stream consumes only about half of the available bandwidth. A data structure is instantiated for each second which

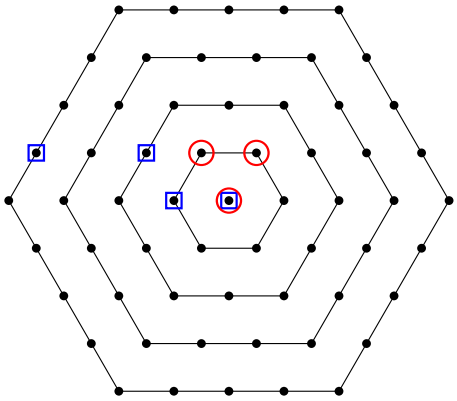


Figure 3.7: Stations are black points. An arbitrary station center is chosen. Red circles denote 3-fold event, while blue squares consider 4-fold event. Figure taken from [82].

contains a list of T2 events from all stations. The CDAS program searches this data structure for temporal clustering by examining T2 correlations within ± 25 μs. If a correlation is found with ≥ 3 members, the program will then test a spatial correlation using the station IDs.

The spatial criteria are as follows: a 3-fold coincidence where every T2 is T2-ToT and at least 1 station is in the first hexagon and the other is not further than the second hexagon (relative to a center station), a 4-fold coincidence that expands on the 3-fold rule but instead accepts any T2 type in addition to a 4th station not further than the 4th hexagon. A diagram, from [82], depicting example cases is shown in Figure 3.7. When a spatial coincidence is detected, the program then checks an additional timing constraint: each T2 timestamp should be within $(6 + 5n)$ μs of the center station, where n is the hexagon number associated with the T2. If all these conditions are satisfied, the program promotes the event to a T3 and sends readout requests for all stations within a 30 μs window and within 6 hexagons of the central

T2. The rate of T3 triggers is about 0.1 Hz.

In addition to physics triggers, a "random" T3 is also sent every 15 minutes which selects a random T2 time stamp. The waveform readout is collected and archived for T2 trigger monitoring purposes. A "2-fold coincidence" trigger also exists which examines T2 from a doublet setup (2 stations separated by 10 m). If there is a T2 coincidence within 1 μs a T3 is sent to the doublet. The rate is naturally 0.8 Hz, but the events are down-sampled to 0.0017 Hz due to bandwidth restrictions.

3.5.3 Comparison of SD triggers

From the previous sections we see that both experiments use standard air shower array coincidencing techniques to search for EAS events. In terms of the local triggers, both use an initial threshold trigger to thin out background events—Auger T1-TH and L0 for TA. However, Auger also implements the T1-ToT which inspects the temporal features of waveforms. This is a significant difference. Beyond the initial triggers, both experiments use a secondary threshold trigger to further separate low energy and high energy shower candidates: the TA Level-1 trigger and Auger T2-TH trigger. Recall that Auger T1-ToT are also promoted to T2-ToT.

For physics triggers, both arrays consider spatial and temporal coincidence, but the implementations are different. The coincidence window for a TA L2 is 8 μs compared to 30 μs for the Auger T3. The spatial 3-fold coincidence in TA is fairly tight, requiring the maximum distance between any pair to be $\leq$ 2400 m. The Auger 3-fold is very similar to TA, and allows for a maximum separation between pairs

Item	Auger	TA Notes
Base level local trigger	T1-TH,TOT	L0 (TH)
Base level rate	100 Hz, 1.2 Hz	750 Hz
2nd level local trigger	T2-TH, T1-ToT	L1 (TH)
2nd level rate	23 Hz, 1.2 Hz	18 Hz
Time window used in physics trigger	± 25 μs	± 8 μs
Max pair separation in physics trigger	3000 km	2400 km
Data collected in physics trigger	All neighbors within ± 30 μs	All neighbors within ± 32 μs

Table 3.2: Highlights of the main differences between Auger and TA SD array trigger systems. From the text, "TH" means a threshold type trigger, and "ToT" is time-over-threshold.

of 3000 m. One difference is that the Auger 3-fold permits a diagonal 'string' type geometry while TA does not. Also, the Auger 3-fold is generated using a different trigger algorithm (T2-ToT) compared to TA. Auger implements a permissive 4-fold trigger mode for which there is no TA analog.

Table 3.2 gives a summary comparison of the SD array triggers for the two experiments.

3.6 SD EAS reconstruction

In this section the shower reconstruction techniques are discussed and compared. While these will be similar to those presented in the last chapter, since the experimental designs are different, we will see how each collaboration has modified the canonical distributions to be optimal for the respective arrays.

3.6.1 Auger SD reconstruction

As described in the previous chapter, massive ground detector arrays permit sampling particle densities of the EAS. Since all stations subscribe to a synchronized clock, the directionality can also be inferred based on the intersection of the EAS with the ground. Beyond the local station level hierarchy, Auger also implements "offline" triggers. In this level, the first physics trigger is called T4 and this provides an additional quality cut on the T3 data to ensure that real showers and not spurious background events are propagated through the analysis pipeline.

The T4 cut employs an algorithm which compares FADC waveform features between the candidate T3 stations. It looks for temporal agreement in waveform features such as peak location and peak width. The T4 is designed to discard T3 triggers formed by accidental muons. For example, a low energy shower which triggers two adjacent stations, and random background muons which triggers another in an allowed hexagon. While a valid T3, the data from this event should not be included in any analysis since it is dominated by noise. The T4 prevents this situation. The implementation of the cuts keeps 99% of stations with actual shower signal [82].

The final level in the entire hierarchy is also a fiducial offline quality cut known as the T5, of which there are 3 variants. The strictest (highest quality) is the 6T5. This event is defined as a T3 with a central station and all neighbors in the first hexagon participating, with the central station reporting the highest signal. This is also referred to as a "working hexagon." This cut permits the most accurate reconstruction of the shower ground impact. The 6T5 is used to compute the aperture (or exposure) of the observatory, see [91], and is used exclusively in energy spectrum analysis, e.g. [92–94]. The two other variants are 5T5 and 4T5, where 5 and 4 stations from the nearest hexagon also participate in the T3 trigger, respectively. These are less strict than 6T5, but still provide somewhat precise ($\sim 1.2^\circ$) arrival directions and also benefit from increased statistics.

The arrival direction is computed using station location and EAS timing. The shower front is modeled as a spherical surface expanding at c described by

$$c\,(t_i - t_0) = |\vec{x}_{\rm sh} - \vec{x}_i| = r_{\perp,i} \tag{3.4}$$

where $\vec{x}_i$ are the ground locations of the station (e.g. northing, easting, elevation), $\vec{x}_{\rm sh}$ is a virtual origin and t_0 is the observation start time, giving Equation 3.4 four free parameters. An example plot, from [82], showing time difference vs. perpendicular distance from station to the shower plane appears in Figure 3.8. This fit, along with ground striking time of the core, is used to calculate the curvature of the expanding sphere.

The lateral distribution function (LDF) is a key shower property used to infer where the EAS lands on the ground, as described in Section 2.4. Sampling this func-

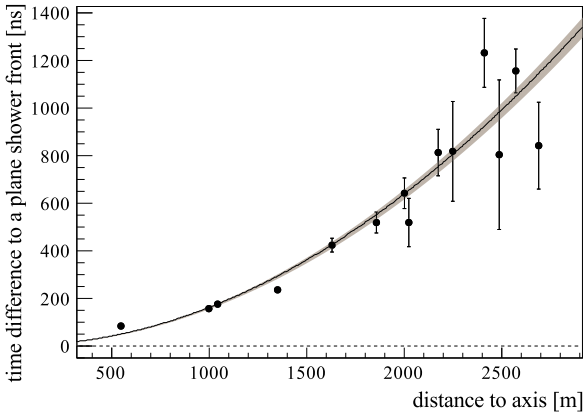


Figure 3.8: Time difference data and the best fit model for an example shower reconstruction, taken from [82]. Gray region is model uncertainty.

tion is challenging due to its exponentially decaying nature, which requires detectors with high dynamic range. While the low and high gain channels augment Auger's capability, stations near the core do become saturated and cannot be used reliably for a fit. A method does exist to attempt recovering the "true" signal [95], but this is an extrapolation. For Auger events, the Nishimura-Kamata-Greisen (NKG) function [30, 67] is modified to give the best results based on the array configuration. The signal as a function of distance in the shower plane is

$$S(r) = S(r_{\text{opt}}) \left(\frac{r}{r_{\text{opt}}}\right)^{\beta} \left(\frac{r + r_1}{r_{\text{opt}} + r_1}\right)^{\beta+\gamma} \tag{3.5}$$

with $r_1 = 700$ m, $r_{\text{opt}} = 1000$ m for the main array with 1.5 km spacing, and β and γ are fixed parameters determined from events with more than 4 stations. An example shower and LDF are shown in Figure 3.9. In Auger, the value of $S(1000)$ is used to set the energy scale. Its uncertainty, $\sigma_{S(1000)}$, consists of 3 sources

$$\sigma_{S(1000)} = \sqrt{\sigma^2_{\text{STAT}} + \sigma^2_{\text{SYS}} + \sigma^2_{\text{RAND}}} \tag{3.6}$$

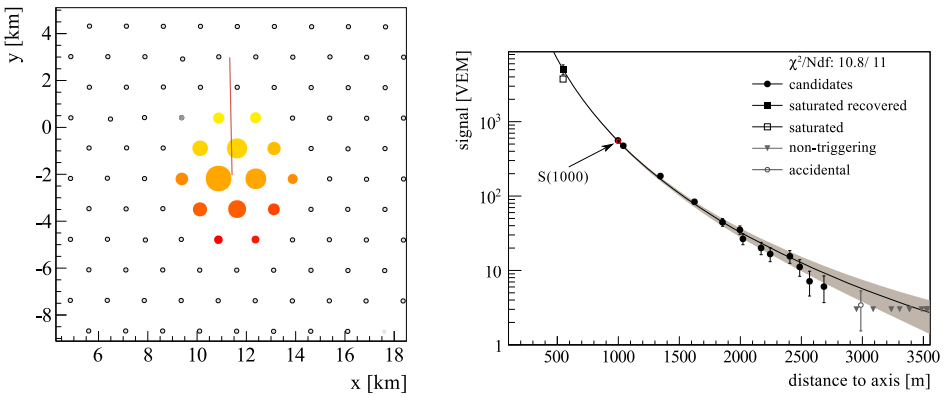


Figure 3.9: Left: plot showing a 6T5 high energy event hitting the SD array, from [82]. The colors show the time evolution where early times are warm colors, late times cooler colors. The circle size is the log of the signal and the line depicts the EAS arrival direction. Right: plot showing SD station signal vs. distance from shower core (shower plane), from [82]. The line is the best fit NKG function and shaded area the uncertainty.

where σ_{STAT} is statistical uncertainty from the limited dynamic range, σ_{SYS} is the systematic uncertainty from assumptions built into the LDF, and σ_{RAND} is natural shower variability. The rough contribution as a fraction of signal is 20%–6% (decreases as energy increases) for the first two, and 10% for the last term.

It is important to note that $S\,(1000)$ is not only energy dependent, but also zenith angle dependent due to attenuation of muons by the atmosphere. To account for these effects the Constant Intensity Cut (CIC) method [96] is used. The Auger CIC function is a third degree polynomial of $x = \cos^2\theta - cos^2\bar{\theta}$ with a reference angle of $\bar{\theta} = 38°$. The full function with coefficients is given in [82]. The second to last step for going from signals in the SD array to an energy estimator is computing

$$S_{38} = \frac{S\,(1000)}{f_{\mathrm{CIC}}\,(\theta)} \tag{3.7}$$

to give an equivalent signal in VEM to a $\theta = 38°$.

3.6.2 TA SD Reconstruction

Beyond the L2 trigger, TA also uses a higher level offline technique for selecting events. First the waveforms of the L2 and neighboring stations are checked for accidental muons, which manifest as a single pulse separated in time from the multiple pulses associated with the shower. These are isolated and removed. Next, the event is checked for spatial and temporal clustering: station pairs whose separation is less than $\sqrt{2} \times 1250$ m are "contiguous" in space, and a time difference $\leq d/c$, where d is station separation and c is the speed of light [89, 88]. After this step the shower arrival direction is calculated using a χ^2 minimization technique for a function involving station time stamps, pulse heights, shower position on the ground and the core position, which is calculated using a center of gravity approach. For the formal derivation of the equation see pages 113–118 of M. Allen's PhD thesis [88]. The result of this procedure is shown in Figure 3.10. This process is the final level of the TA selection hierarchy. After stations have satisfied these conditions they are used in a LDF fit.

An empirical LDF derived from [45, 97] is used to fit station data on the ground

$$\rho\left(s\right) = A\left(\frac{s}{r_M}\right)^{-1.2}\left(1+\frac{s}{r_M}\right)^{1.2-\eta(\theta)}\left(1+\frac{s^2}{10^6}\right)^{-0.6} \tag{3.8}$$

$$\eta\left(\theta\right) = 3.97 - 1.79\left(\sec\theta - 1\right) \tag{3.9}$$

where s is the perpendicular distance from shower core, $r_M = 91.6$ m is the Molière unit, θ is the zenith angle, and A is a scaling factor. See [88, 89] for more details.

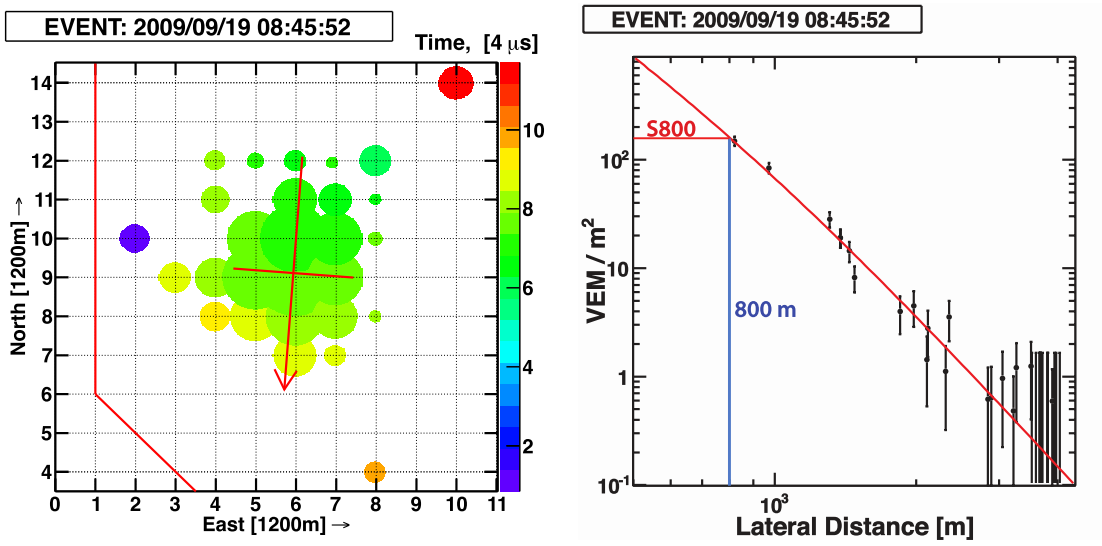


Figure 3.10: Display for a typical shower, taken from [88]. The shower core is at the middle of the cross-hairs, shower direction is indicated by the arrow and circle diameter is linearly proportional to signal size. Signals on the periphery which aren't contiguous are not included in further analysis, e.g. red, blue, orange circles.

The uncertainty used for this quantity is defined as

$$\sigma_\rho = \sqrt{0.56\rho + 6.3 \times 10^{-3}\rho^2} \tag{3.10}$$

which is used, along with $\rho(s)$ in a χ^2 minimization fit. Example data and a fit is shown in Figure 3.10. The fit to this curve allows the calculation of particle density at any given lateral distance from the core.

For the energy calibration the signal at 800 m was determined to be optimal for the TA reconstruction, where the methods described in [98] were used to determine that distance. The fit LDF is evaluated at this distance, $\rho(800)$, and is used with the reconstructed zenith angle for energy assignment from a large Monte-Carlo (MC) estimation table [89]. The table is produced by accurately simulating the response of

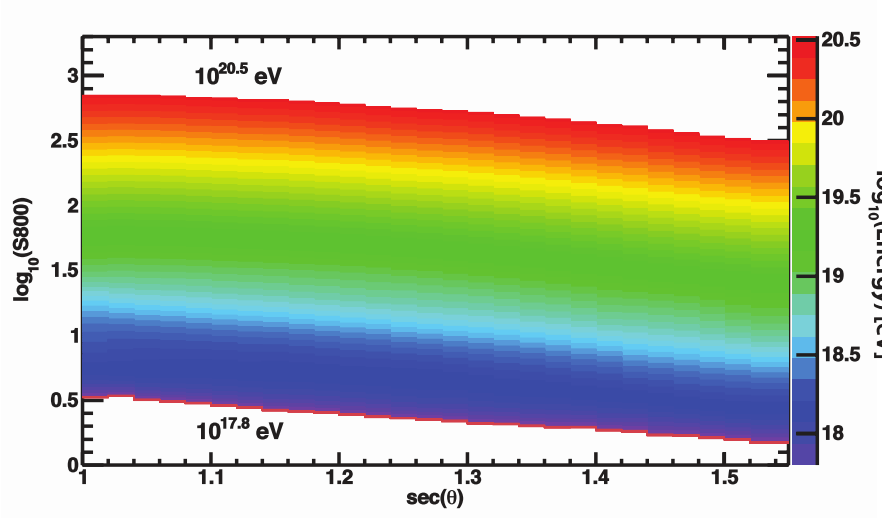


Figure 3.11: Contour plot showing the TA SD energy table, from [89].

the SD array, detector, and electronics and generating the shower using the CORSIKA code (version not specified) with QGSJET-II (version not specified) high energy event generator and FLUKA (version not specified) low energy event generator. The result provides the *simulation input* SD energy as a function of $S(800)$ and $\sec\theta$. The simulation coordinate which best resembles the data is then used to infer the SD energy scale. This table is depicted in Figure 3.11.

3.6.3 Comparison of reconstructions

In the previous two sections we examined how the SD array specifically can be used to reconstruct key shower parameters and make inferences about properties of the primary particle. Both experiments implement an additional layer(s) above the main physics trigger which further increases the quality of reconstructions which will be used in analysis. In Auger, random muons and spurious coincidences are filtered out by T4. The final layer, T5 ensures a high quality geometrical reconstruction. For

TA the final layer automatically removes spurious muons and checks for good EAS geometry by enforcing contiguous pair spacing and tight temporal correlations. This can be viewed as rolling the Auger T4 and T5 into one step.

We also saw how both experiments used timing and station ground location to determine core position and shower arrival direction using maximum likelihood/χ^2 minimization techniques. This data, along with station signals, is also used to find the best fitting LDF, which is essential for finding the primary's energy. Both experiments use a modified NKG function best suited to their array design. From this fit, the optimal signal value is found: $S(1000)$ in Auger and $\rho(800)$ for TA. Auger then converts this to an equivalent signal for an incident primary at 38° using the reconstructed zenith angle. In TA the SD energy is assigned by finding the closest match in an MC lookup table, which is model dependent.

At this point the final step in both experiments is to use the hybrid energy calibration to express these energies in terms of a near calorimetric measurement.

3.7 Fluorescence Detector Systems

In this section the FD telescopes of both arrays are described and compared. Since the focus of this dissertation is related mostly to SD systems, a brief, high-level treatment of the FD is given. Although this choice has been made, the importance of these detectors cannot be understated, since they are vital for an accurate energy estimator.

3.7.1 Telescope Array FD

Currently 3 fluorescence detectors (FD) are operated by TA: Middle Drum, Long Ridge and Black Rock Mesa. The FD is an essential component of hybrid detectors because they can provide near calorimetric measurements of EAS energy. The atmosphere is used as the detector medium providing a much larger exposure compared to the SD array, which is limited to a planar surface. However, unlike SD arrays, the FD cannot be operated during day-time and also requires cloudless and moonless nights. As a result, the duty cycle of FD is much lower than SD. Some technical specifications of the Middle Drum (MD), which are representative of the other two sites are considered below.

The TA Middle Drum (MD) fluorescence detector is refurbished equipment from the HiRes experiment comprised of 14 telescopes. After several years of operation, data (e.g. the spectrum) from the HiRes setup and MD are in agreement [88]. The MD is installed north of the SD array, 21 km from the Central Laser Facility (CLF). The individual telescopes use 5.1 m^2 (effective collection area of 3.72 m^2) spherical mirrors and are instrumented with 256 PMTs optimized for UV sensitivity. The field of view is 3° –31° in elevation and 112° in azimuth. The PMTs are run at a gain of about 10^5 and a threshold trigger rate of 200 Hz. The first level trigger condition requires clusters (> 3, two of which must be adjacent) of PMTs to be in time coincidence of 25 μs . Higher level trigger conditions involve patterns of these cluster triggers and further details can be found in [88], Chapter 3. A schematic of the instrument is shown in Figure 3.12.

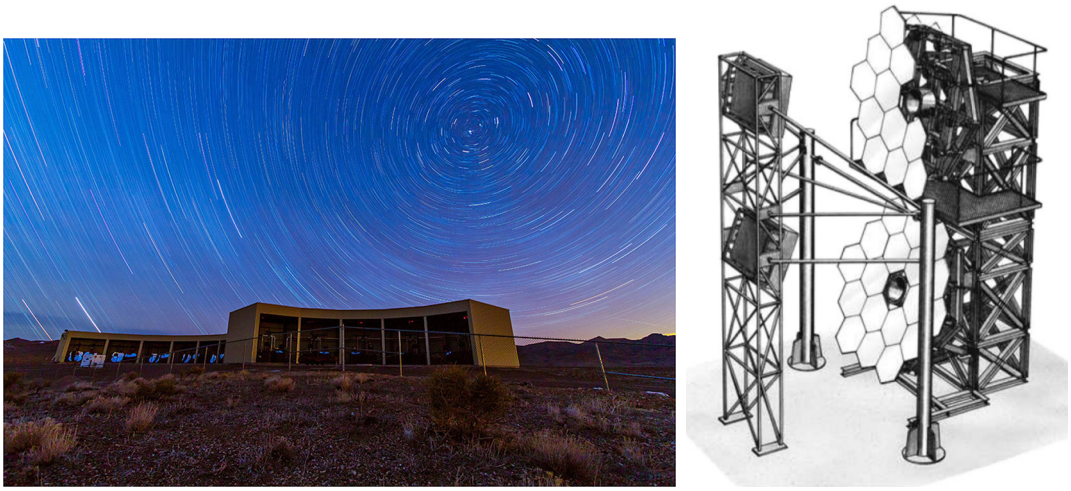


Figure 3.12: Left: Time lapse photo of an FD building, taken from [99]. Right: Diagram showing the FD telescope apparatus, taken from [100]. In the upper segment, mirrors have been omitted to show the mounting framework.

A photometric calibration is performed on each PMT using a xenon flasher system. This is a very stable source with known output photon statistics. Using calibration runs of 500 shots at a rate of 1.5 Hz, the response of each PMT can then be expressed as a function of the known source. Additionally, a UV flasher is used each night during data runs to monitor the gain of each PMT.

Another critical aspect of fluorescence detection is an accurate description of the atmosphere so that attenuation/scattering of UV fluorescent light can be accounted for. If one does not properly consider these effects, the energy estimation using the FD technique will be systematically biased. These issues are addressed using models of the atmospheric density profile and aerosol distribution and continuous monitoring programs. In TA the density profile is derived from a National Weather Service (NWS) station in Delta, Utah, which performs radiosonde (weather balloon) flights.

This data is limited from ground level to 30 km, where TA then extrapolates the US 1976 atmosphere model at higher altitudes. Aerosols are a contaminant produced from volcanic, desert dust, and anthropogenic sources which attenuate UV light via Rayleigh scattering. The aerosol conditions can also change on a time scale of hours or days. These effects can be subtracted by performing a calibration of the telescopes using precise laser sources. In TA LIDAR (LI Detection And Ranging) and 355 nm YAG laser (housed at the CLF) are used. The data is then used in a Monte Carlo process to infer aerosol concentration, which is later used to correct the final energy measurement.

During the EAS progression, most (90%) of the ionization of air molecules is due to $e^{\pm}$ particles at GeV energies, which results in the observed number of UV photons (fluorescence yield) being nearly linearly proportional to the number of ionizing $e^{\pm}$.

3.7.2 Auger FDs

Auger operates 24 FD telescopes at four sites (6 per site). Each telescope covers 30° in azimuth and elevation (total of 180° in azimuth). The camera consists of a 22×20 matrix block of XP3062 Photonis PMTs at the focal plane of either rectangular aluminum mirrors, or hexagonal glass mirrors. Each pixel also uses a Winston cone light concentrator to avoid dead zones between the PMT cathodes. The FD aperture radius is 1.1 m and includes a UV filter designed to reduce background light, prevents airborne contaminants and dust entering the facility, and also augments the climate control. In front of the filter an annular lens is used to eliminate coma and spherical

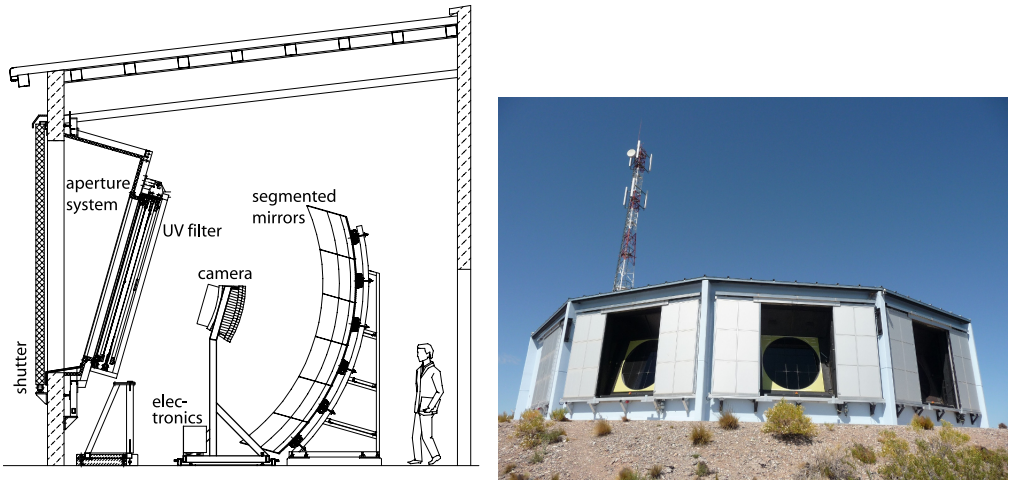


Figure 3.13: Left: Diagram showing the FD apparatus with person for scale, taken from [82]. Right: Photograph of the Los Leones FD site during the day, also from [82].

aberration. Figure 3.13 shows a schematic of the Auger FD.

A three stage trigger is used based on timing and geometry of neighboring pixels. The base rate/first level of each pixel is 100 Hz. The second level requires tracks in the camera of at least five pixels long. The rate for this level is between 0.1 Hz and 10 Hz. The final trigger is software based and discards spurious triggers caused by lightning, muons, or electronic noise. A computer analyzes data from the multiple cameras in a site, and if a coincidence is found, the data is read out and a hybrid trigger (FD-T3) is generated. These occur at roughly 0.012 Hz.

The FD performance is tested using lasers at CLF and eXtreme laser facility (XLF). Calibration is done using an end-to-end technique: a drum source equipped is used to fill the FD aperture with diffuse light generated by a UV LED pulser. This source is calibrated to a NIST reference photodiode within 3%. This procedure gives the translation from hardware units to physical units: 4.5 UV photons/ADC count.

Extensive atmospheric monitoring is performed at Auger. In fact, the data collected about the atmosphere dwarfs the collected data about cosmic rays. Commercial weather stations are used to record temperature, pressure, humidity, wind speed, etc. A balloon launching station was also used in the past to record these stations as a function of altitude, allowing the calculation of air density. This program has been discontinued in favor of the Global Data Assimilation System (GDAS) which provides the same information [101].

Aerosol optical depth profiles are built from CLF/XLF laser shots every hour during normal operation—this also includes information about the aerosol phase function. At each FD station an elastic LIDAR apparatus is installed which provides further data on aerosols. Infrared sensors, in addition to the LIDARs, are used to detect and monitor clouds. The presence of clouds will impact the FD aperture and by extension, the total exposure of the observatory.

3.7.3 Comparison of FD systems

The Auger and TA designs are similar with differences in hardware components, trigger implementation, and certain calibration techniques. Both are very sensitive to UV photons from 300-400 nm and observe large areas of the sky, which includes the entire footprint of the ground array. Both experiments also implement comprehensive atmospheric monitoring campaigns to account for systematic errors associated with using the atmosphere as a calorimeter. A summary comparing the main features of the observatories is given in Table 3.3.

Item	Auger	TA
# of buildings	4	3
pixels/camera	440	256
field of view per building	180° azimuth, 30° elevation	≈120° azimuth, 30° elevation
calibration method	end-to-end NIST source	end-to-end flasher and LINAC
aerosol density	LIDAR, 355 nm YAG laser	LIDAR, 355 nm YAG laser
atmospheric monitoring	GDAS, weather stations, radiosonde (deprecated)	local NWS radiosonde, 1976 atmosphere
up time	15%	

Table 3.3: Important parameters of the FD systems.

3.8 Hybrid Reconstruction and FD energy calibration

One of the key features of modern observatories is the ability to make near calorimetric measurements of the EAS to infer the primary energy. The technique relies on coupling the high accuracy FD observations with the high duty cycle of the SD array using a cross-calibration. This is done with high quality "golden hybrid events" which trigger the FD and SD simultaneously. In this section I briefly describe how Auger and TA handle this cross-calibration.

3.8.1 TA hybrid reconstruction & energy calibration

In TA the shower data set is searched for coincidences between the separate FD and SD reconstructions. If there are two events in a 2μs window these are treated as a single hybrid event, and added to this list [88]. The list is further analyzed by a separate reconstruction program which performs a combined chi-square minimization based on FD & SD timing and core position in the ground plane. This approach makes use of the additional data provided by both systems resulting in more robust fits to EAS reconstruction parameters.

The hybrid energy is calculated using an inverse Monte-Carlo approach based on the slant depth/shower profile where a Gaisser-Hillas function is adopted. Details of the procedure are outlined in Section 5.5 of [88]. An example of the timing and shower profile hybrid reconstruction is displayed in Figure 3.14.

The hybrid data set is used to express "SD only" events (e.g. daytime events) in terms of a calorimetric measurement. This approach helps to constrain the energy

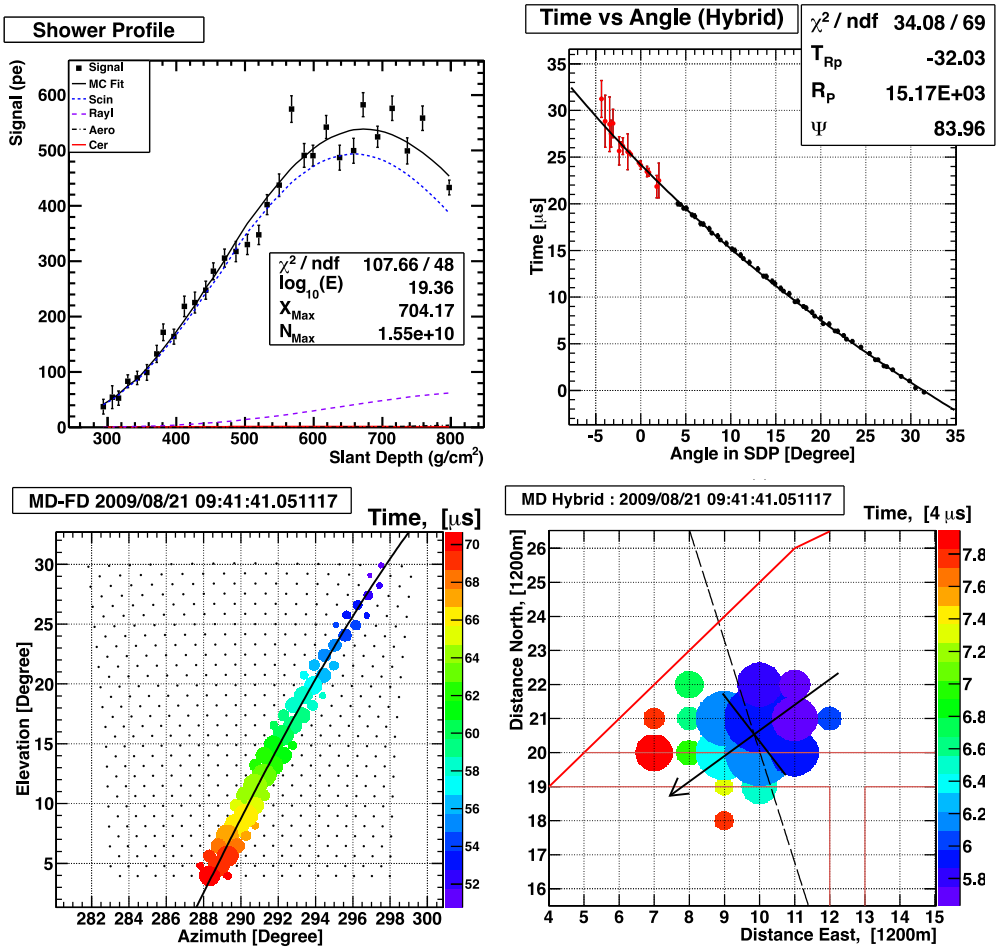


Figure 3.14: All plots from [88]. Top left: the FD reconstructed shower profile. Top right: combined timing reconstruction. Bottom left: signal track observed by FD. Bottom right: shower plane sampled by the SD. Important shower properties such as energy and $X_{\max}$ are also shown. This is the middle drum FD.

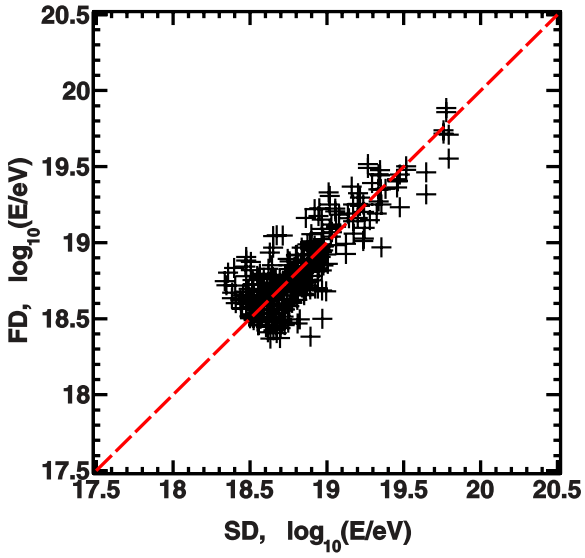


Figure 3.15: Energy cross-calibration for TA, taken from [102]. Data points are reconstructed energy for high quality hybrid events seen by the SD and 3 FD sites. The red line is the linear function with slope 1, $E_{FD} = E_{SD}$.

scale of the events, which would otherwise suffer from large uncertainties or biases in the MC simulations. In TA, the best agreement for hybrid event energies is achieved when the SD energy is reduced by 27%. With this scaling applied, the energy scales between the FD and SD appear to be linear. A plot showing the golden hybrid events demonstrating this correlation is given in Figure 3.15.

3.8.2 Auger hybrid reconstruction & energy calibration

The hybrid reconstruction uses FD data combined with SD timing information. Signals passing the trigger criteria are processed and pixels are subjected to an additional constraint of having SNR > 5. The FD timing and photoelectron counts are then used to infer the shower detector plane (SDP) using a χ^2 minimization. Once the

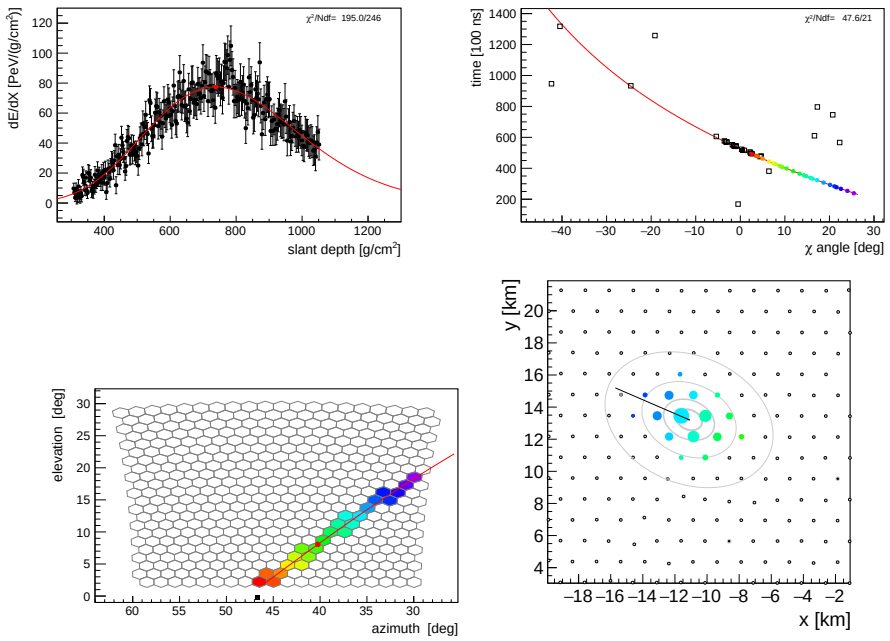


Figure 3.16: These plots were generated using Offline for an event recorded on 6/27/2011 with a reconstructed energy of about 50 EeV. Top left: shower profile for one FD site. Top right: combined timing fit, open squares are SD stations and colored circles are FD data points. Bottom left: shower track recorded by the camera. Bottom right: Shower observed on the ground by the SDs. The straight line shows the arrival direction, colors show timing and the size of the circle scales as the log of the signal.

SDP is established the timing information from the SD is added to further constrain the shower geometry. This is now used with timing information to infer the energy deposition as a function of slant depth, or the shower profile. After separating light generated by fluorescence, Cherenkov, and scattering, a Gaisser-Hillas function is fit to the shower profile curve, and integrated to give the calorimetric energy measurement, E_{FD}. Specific details of the procedure can be found in [82]. An example hybrid reconstruction is shown in Figure 3.16.

A hybrid cross-calibration curve is built from 1475 events between January 2004

and December 2012 with $E > 3$ EeV linking S_{38} (see 3.6.1) with E_{FD}. These events must pass strict quality cuts such as satisfying field of view criteria, accurate fit of the Gaisser-Hillas function, and an $X_{\max}$ accuracy within 40 g cm^{-2}. The curve is fit using a simulation-independent maximum likelihood approach which accounts for all known systematic energy uncertainties. The result is a power law described by

$$E_{FD} = A \left(\frac{S_{38}}{\text{VEM}} \right)^B \tag{3.11}$$

where $A = 1.90(5) \times 10^{17}$ eV and $B = 1.025(7)$. For this calibration the highest energy inlcuded is 79 EeV, meaning any SD derived energy above this value is an extrapolation.

These parameters can now be used to convert the SD signal to a calorimetric energy

$$E_{SD} = A \left(\frac{S(1000)}{\text{VEM}} \frac{1}{f_{\text{CIC}}(\theta)} \right)^B \tag{3.12}$$

The systematic uncertainty of the cross-calibration procedure is about 2%, while the statistical uncertainty of the SD energy estimator is around 16% for low energy and 12% for the highest energies. Data and a fit is shown in Figure 3.17.

3.8.3 Comparison and discussion

While Auger and TA may employ different algorithms and fitting procedures to the Gaisser-Hillas function and event timing curves, the approach to reconstructing FD data is largely the same, but optimized for the individual experiments. Both experiments employ restrictive quality cuts for their hybrid data set with the goal of minimizing the uncertainty for the energy cross-calibration.

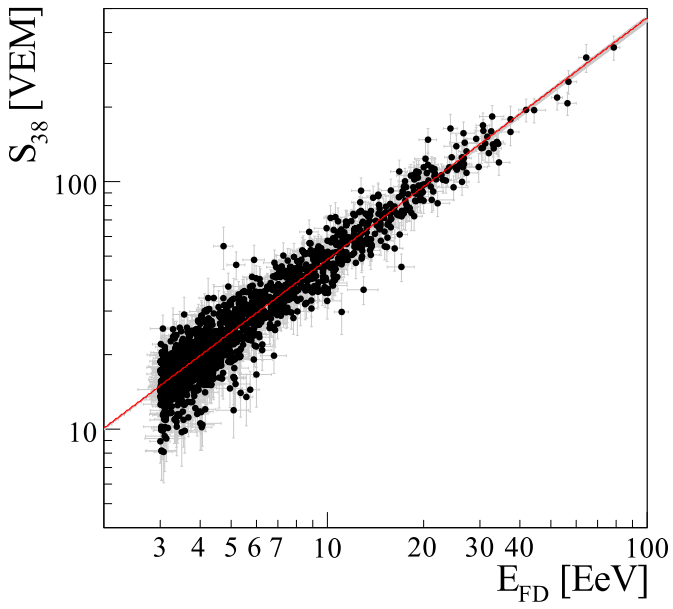


Figure 3.17: The golden hybrid events used to derive the cross-calibration function, taken from [82].

The main difference is the energy cross-calibration procedure. In TA, a comparison of the independently reconstructed golden FD and SD energies reveals general agreement if all SD energy data is reduced by 27%. Therefore, to represent the TA SD energy in terms of calorimetric measurement, one just applies the coefficient $0.73E_{SD}$. The SD energy reconstruction relies on high energy event generator models, which are themselves extrapolations of LHC data. In Auger the CIC method is used to first translate $S(1000)$ to an equivalent S_{38} measurement. This signal is then compared to the independently reconstructed FD energies of golden hybrid events. An empirical relation is derived using a likelihood fit, which is used to express SD events in terms of an FD equivalent.

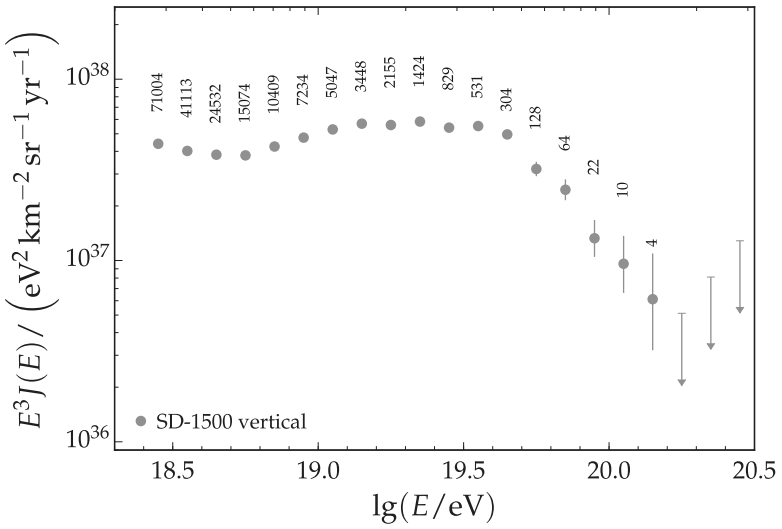


Figure 3.18: The latest Auger spectrum presented at the 35th ICRC [103]. Logarithmic binning is used and the number per bin appears above the data point. Error bars represent statistical uncertainty and the upper limit points are 84% confidence levels.

3.9 Energy spectrum, composition, and anisotropy results

In this section the most recent main results of each experiment are presented and discussed. Some interesting implications are also considered, but described in detail.

3.9.1 Auger results

First I discuss energy spectrum results. Since the array operates over a large span of energies, and also different modes (FD & SD), first we consider the spectrum formed purely from the main SD array with $\theta < 60°$. This is usually referred to as the SD 1500 "vertical" (low zenith angle) spectrum and includes 183,332 events above 3 EeV, see Figure 3.18. Features like the ankle and GZK cutoff are evident. To extend the spectrum to lower energies data from the infill array can be included,

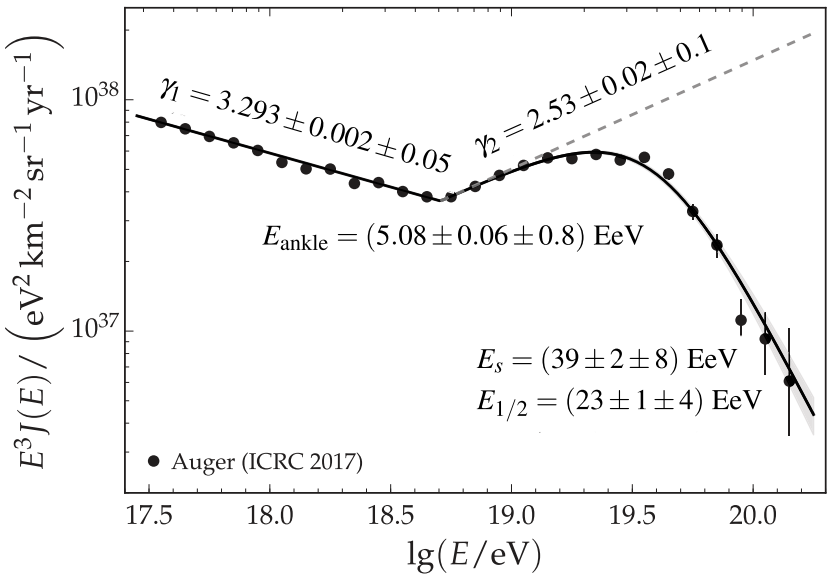


Figure 3.19: The most recent Auger combined spectrum along with the best fit parameters for the spectral features. The plot appears in [103].

and to further improve statistics data from inclined events ($\theta > 60°$) and hybrid showers are added to create a combined spectrum. A global maximum likelihood approach is used to determine the optimal way to combine and normalize the data. The combined spectrum is shown in Figure 3.19. After combination, the spectrum is fit with a broken power-law to obtain the spectral index, as well as the precise features of energies. The ankle is at $E = 5.08$ EeV and the flux suppression occurs at $E = 39$ EeV.

The Auger composition results have recently been updated for the FD, SD-1500 and SD-750 data sets. The FD set is for hybrid events covering the energy domain from 0.16 EeV to 79 EeV and includes 42,466 data points which survive the quality cuts. These data are used to calculate the elongation rate moments which can be compared to model predictions using various primary compositions. A description of

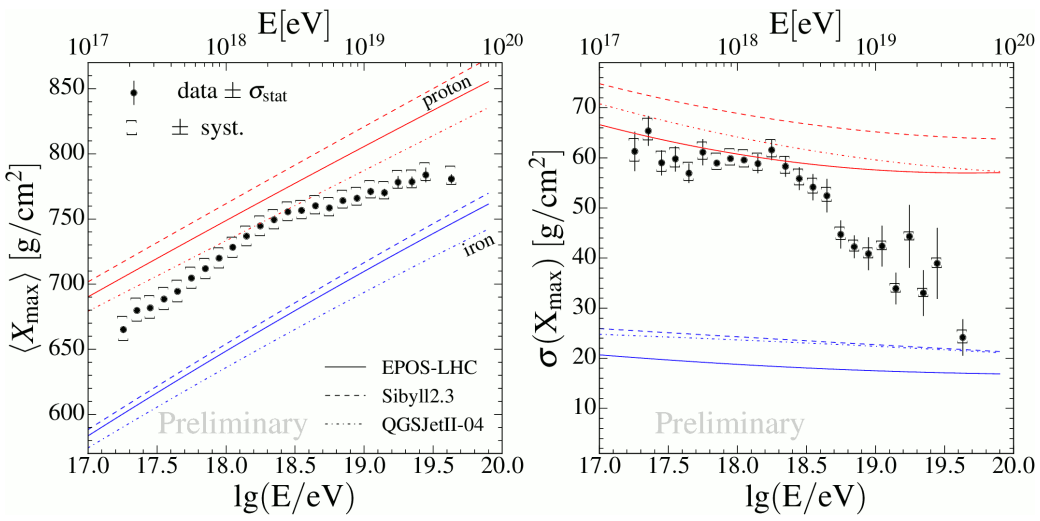


Figure 3.20: This plot is from recent results presented in [103]. The first moment (mean) is on the left while the 2nd moment (standard deviation) is on the right. The data is compared to three high energy event generator models, and considers light and heavy compositions.

the analysis is given in 3.1 of Auger ICRC proceedings [103]. The first and second moments appear in Figure 3.20. The SD data set uses a total of 81,575 from the infill and main array from January 2004 to December 2014 and considers only events with $\theta < 46.4°$ in the main array and $\theta < 39.7°$ in the infill. A description of the analysis is given in section 3.2 of [103]. The results are shown in Figure 3.21. The FD and SD $X_{\max}$ results both suggest an intermediate to light, to heavy composition with increasing energy. This trend is stronger for the second moment data. This picture is consistent with earlier results reported in [104].

In terms of arrival directions the Auger collaboration has reported a dipole anisotropy for energies $E > 8$ EeV in the direction of 95° right ascension [103]. Correlation studies have also been performed, and excesses have been observed around the Centaurus A, the most luminous AGNs in the Swift-BAT catalog, gamma-ray AGNs, and for

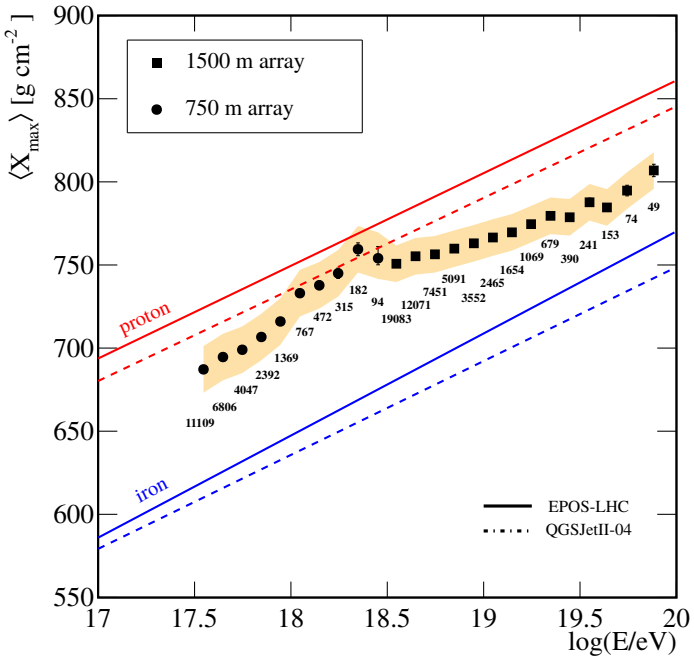


Figure 3.21: The mean $X_{\max}$ obtained from the SD data using the analysis procedure described in [103]. The tan band depicts the systematic uncertainty. Two event generator models, and heavy and light scenarios are considered.

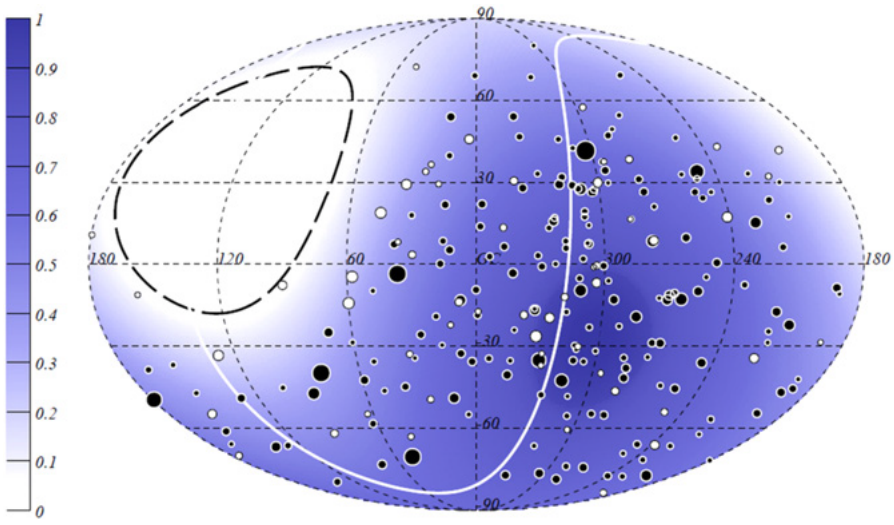


Figure 3.22: In this map, taken from [105], the relative exposure (color scale) is plotted with arrival directions. The size of the points is a function of energy. Black points refer to vertical, while white is inclined. The white line is the super-galactic plane and the dashed line denotes the field of view limit for the observatory.

star-forming galaxies for $E > 39$ EeV at an angular scale of 13° [103]. A representative recent sky map showing arrival directions for 231 $E \geq 52$ EeV events is shown in Figure 3.22. Many of the newest results reported in [103] are preliminary, and the refereed publications may appear after this dissertation has been submitted.

3.9.2 TA results

As a result of TALE and a bridging technique, TA can compile a spectrum ranging from 4 PeV to 100 EeV. While not the most recent result, the spectrum from [106] is discussed, as results have not significantly changed. Instead of a combined spectrum, TA presents data for the individual detection modes: monocular FD, SD and TALE. A broken powerlaw fit over the entire energy range captures four spectral features:

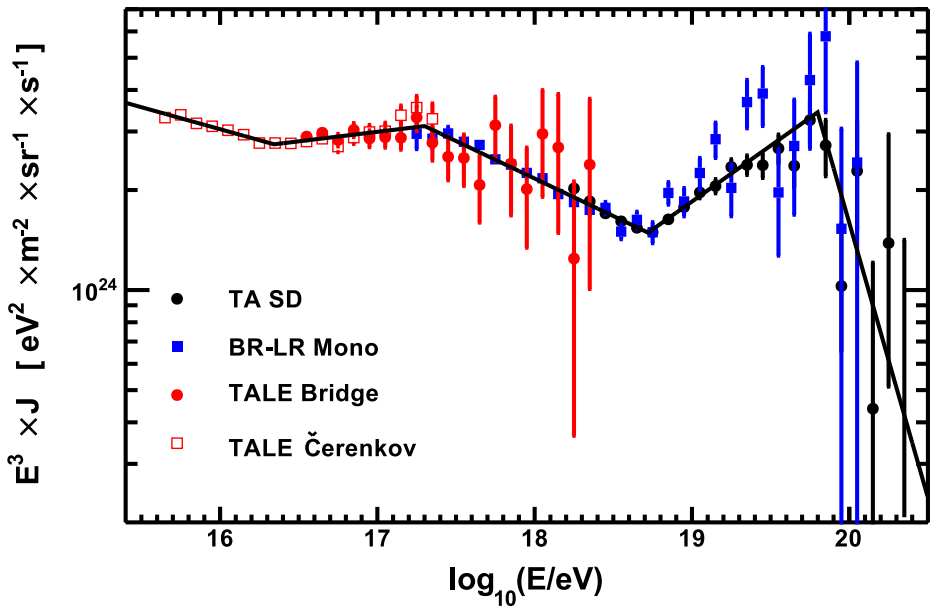


Figure 3.23: The full TA energy spectrum from [106]. The error bars are statistical errors, and the systematic error for SD and the FD-mono is estimated to be 20%.

flux suppression at 63 EeV, the ankle at 5 EeV, a second knee at 0.2 EeV and a low energy ankle at 22 PeV [106]. The spectral indexes as a function of increasing energy are: -3.132(1), -2.94(2), -3.226(7), and -2.66(2) before the break and -4.7(6) after. Figure 3.23 shows the data.

For TA composition studies one can examine the monocular FD data. I consider the results from the 2015 ICRC which includes the period of January 2008 to December 2014. The analysis compares the elongation rate to various high energy event generators. A plot of the first $X_{\max}$ moment is shown in Figure 3.24. The measurement is consistent with a light composition for most of the energy bins. The highest energy bin appears to favor mixed/heavy composition, however the reported $X_{\max}$ is only for 10 samples. Hybrid data is also presented in another proceeding [108]. This

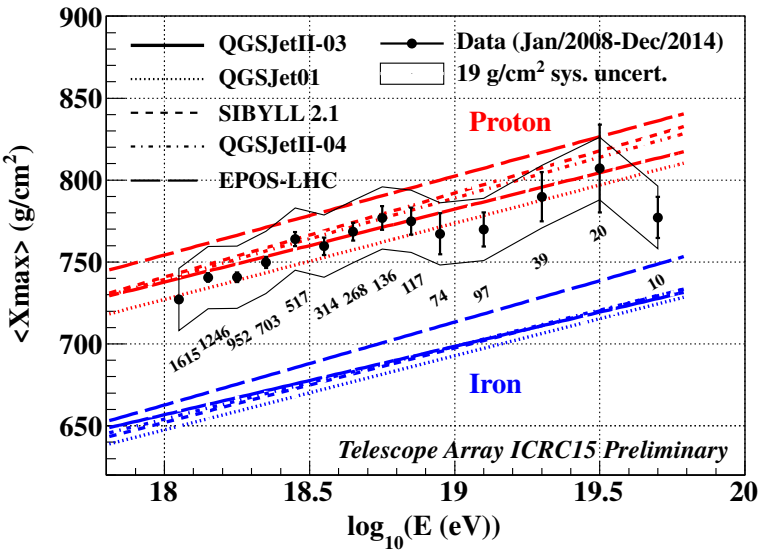


Figure 3.24: A recent composition measurement which appears in [107]. The number of data points is shown per bin, and the error bars represent statistical uncertainty, while the line band shows the systematic uncertainty of 19 g cm^{-2}.

analysis only considers the QGSJETII-03 model and is for 6 years of data, and also favors a light primary composition.

In terms of anisotropy an important recent result published by TA is the detection of a "hot-spot" for a sub-sample of events with $E > 57$ EeV centered at right ascension 146.7° and declination 43.2° [109], see Figure 3.25. The reported post-trial significance is 3.4σ and it is considered as a detection of intermediate-scale anisotropy. At the 2017 ICRC a summary proceedings, see [110], reported on a range of clustering tests: searching for correlations with large scale structure, medium and small scale clustering, correlations with AGN for energy ranges $E > 10$ EeV and $E > 40$ EeV. For these energy cuts, no significant deviation from isotropy was found [110], but the hot spot anisotropy persists. Another interesting result reported at the 2017 ICRC

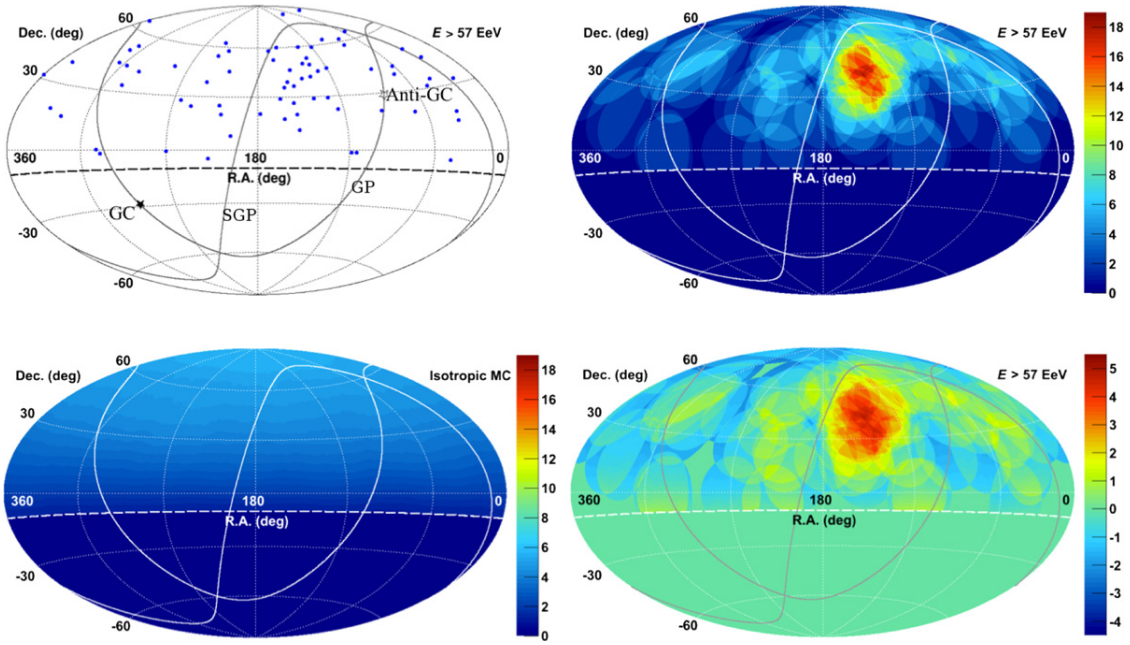


Figure 3.25: Plots showing the arrival direction data and simulation techniques for the hot spot analysis, taken from [109]. Top left: arrival direction data. Top right: number of events summed over 20° radius. Bottom left: expected background events using same summing technique. Bottom right: significance heat map formed from previous two plots.

is an energy-dependent anisotropy [111]. In this proceeding an analysis found that the hot-spot region also corresponds to a deficit of cosmic rays in the energy bin of $16 < E < 57$ EeV at a high confidence level which is indicative of galactic magnetic deflection effects.

3.9.3 Comparison

In this section I attempt to compare and contrast the most recent results of the world's largest observatories. In most cases a direct comparison of results must be done with extreme caution, and the same caution must be exercised when interpreting the *indirect* comparison, for a variety of reasons. Perhaps the most obvious is the fact that the observatories are situated in different hemispheres. This gives each observatory different sensitivity to astrophysics scenarios in the north and south hemisphere. This can manifest as a discrepancy in the common data products (spectrum, composition, anisotropy) which does not necessarily indicate problems with the data analysis methodologies. Another important distinction is the experimental design and characterization of uncertainties. Both experiments strive to provide a complete treatment of systematic errors, but the experiments are extremely complex, and these quantities continue to be refined and explored on both sides. When Auger and TA results are compared for a common measurement it is absolutely critical that one considers not necessarily the distance between data points, but the overlap of systematic and statistical errors.

First let's examine the energy spectra. Both experiments agree that the ankle

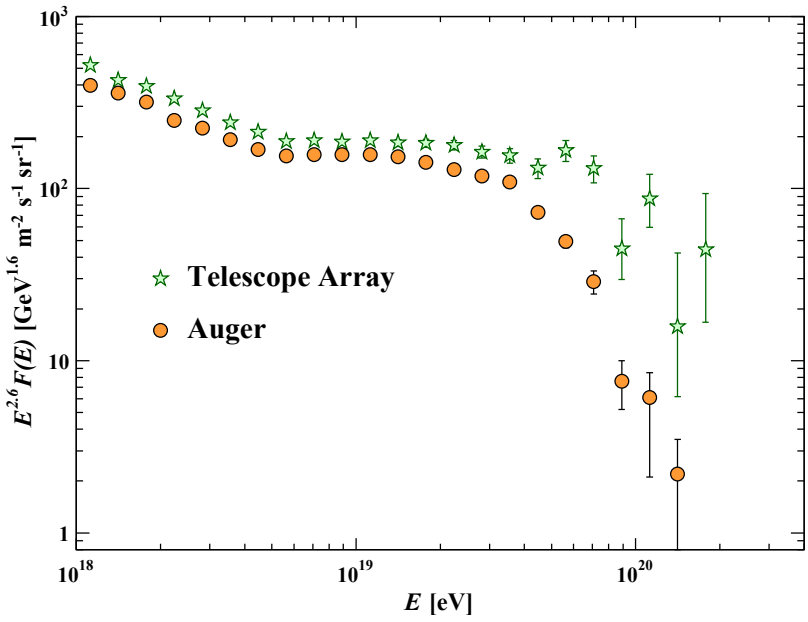


Figure 3.26: A comparison of Auger and TA data released at the 2015 ICRC. The plot is taken from [112]. Error bars are statistical errors, and systematic errors are not shown.

feature is prominent at a primary energy of ≈ 5 EeV. Auger reports a harder spectrum with flux suppression occurring at $E = 39(10)$ EeV compared to $E = 63(11)$ EeV reported by TA. This discrepancy is also present in $E_{1/2}$, the value at which the differential flux falls to half the value compared to the start of the intermediate energy range: 23(5) EeV for Auger compared to 52(11) EeV for TA. Figure 3.26 shows the Auger and TA spectra for the highest energy events. It is critical to note that the error bars are statistical errors and the y-axis has been scaled by E^3, thus systematic errors would appear as diagonals, but these have been omitted from the plot. In general the spectra agree within systematic uncertainties up to $E \approx 50$ EeV where they begin to diverge. This discrepancy can be explained by observation of different astrophysics scenarios, instrumental effects, uncertainty treatments, or statistical fluctuations since

the sample sizes at the highest energies are rather sparse. A joint Auger-TA working group was established in 2010 to investigate these differences. In results presented at the UHECR 2014 conference the group studied adjusting Auger data to be compatible with the TA fluorescence yield, as well as data in the various declination bands but were unable to attribute the discrepancy to these causes [113]. The group presented an update with plans to consider a comparison insensitive to anisotropy differences and also noted that the TA spectrum exhibits declination dependence [114].

The composition results have been directly compared in a few recent publications. In a joint proceeding of the 2015 ICRC it was shown that if Auger data is convolved with the TA middle drum FD acceptance (used in the TA calorimetric energy calibration) the first moment X_{max} data agree within systematic uncertainty [115]. Figure 3.27 shows a direct comparison between the data sets. During the 2017 ICRC these results were updated with a detailed statistical comparison of the X_{max} distributions and the results remain unchanged: when the observed Auger data is input into the TA reconstruction pipeline the results are consistent within systematic errors [116]. The indirect discrepancy is caused by acceptance biases and different analysis strategy trade-offs chosen by the collaborations. In Auger the reported X_{max} is that which minimizes detector and reconstruction bias, while in TA the approach is to reduce the influence as cuts which maximizes statistics.

For anisotropy and arrival directions TA reports an excess of the highest energy events with increasing confidence as more data are collected. This is contrasted with Auger where the highest energy arrival directions are compatible with isotropy. Auger has reported a variety of studies which seem to indicate excesses in the direction of

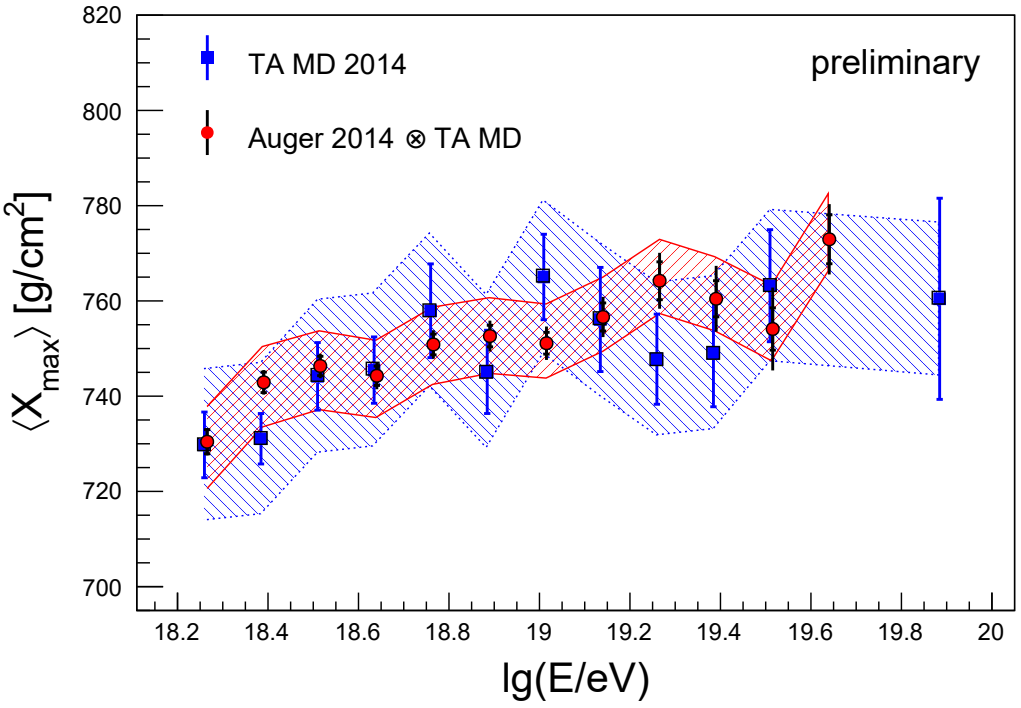


Figure 3.27: The first $X_{\max}$ moment using the direct comparison approach. The plot is taken from [115].

certain astrophysical objects, but it is difficult to draw a strong conclusion from these correlations as they may be spurious features of extragalactic + galactic magnetic lensing effects. In general the Auger arrival directions are correlated with nearby matter of the large scale structure. For lower energy regimes TA reports an agreement with isotropy and interestingly, a deficit of arrivals in the hot-spot region. In Auger for the energy range $E > 8$ EeV there are indications of a dipole anisotropy. A joint analysis performed by TA and Auger which benefits from full sky coverage and allows for multipole expansion also shows a dipole detection with high confidence [117]. A summary appears in Table 3.4.

Item	Auger	TA
Ankle	5 EeV	5 EeV
Flux suppression	39(10) EeV	63(11) EeV
$E_{1/2}$	23(5) EeV	60(7) EeV
Spectrum declination dependence	No	Yes
Dipole	$E > 8$ EeV	None
High energy anistropy	LSS	Hot-spot
Experiment composition	Light to mixed, possibly heavy at highest energies	Light

Table 3.4: Highlights of the measurements.

Chapter 4: Arrival direction probabilities

4.1 Preamble

In this section I describe the techniques used to backtrack CR primaries to potential sources. The candidate galactic magnetic field (GMF). A subset of high quality Auger events with $E > 50$ EeV is described. I discuss the software tools used and the relevant physical processes that are simulated. An initial Monte-Carlo sensitivity analysis framework is presented and the features of resulting arrival direction probability distributions are discussed. A second, updated analysis is described which uses a higher quality Auger data set and more powerful simulation tools. Additionally, this analysis includes the full JF12 field model with actual turbulent field realizations. A preliminary anisotropy analysis is performed using these distributions and compared to a traditional approach. Finally, a brief discussion on updates to the JF12 parameters, and the impact of heavier primaries on the arrival direction distribution is given.

4.2 The Lorentz force

Since UHECRs are ultra-relativistic particles their motion must be described by a covariant force law. The classical and well known Newton-Lorentz is

$$\boldsymbol{F} = q\left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right) \tag{4.1}$$

where q is the charge, $\boldsymbol{E}$ is electric field, $\boldsymbol{v}$ is particle velocity and $\boldsymbol{B}$ is the magnetic field and **bold** symbols represent vector quantities. The equation completely determines the particle dynamics and no assumption has been made of the time or spatial dependence of the electric or magnetic fields. The relativistic form can be written as

$$\frac{d\boldsymbol{p}}{dt} = q\left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right) \approx q\boldsymbol{v} \times \boldsymbol{B} \tag{4.2}$$

where $\boldsymbol{p} = \gamma m \boldsymbol{v}$, and $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$. For a full derivation in terms of vector potentials and the resulting Faraday tensor, the reader may consult Section 7.6 of [118]. In practice since $v \to c$, the B-field term dominates so the E-field term can be safely ignored. Additionally, the magnetostatic case is assumed to eliminate time dependence of the B-field. This is justified if the B-field dynamics are on the time scales of galaxy rotation speeds (200 km $\mathrm{s}^{-1} \ll c$). In this case the configuration of the field does not change appreciably before the UHECR exits the galaxy. For example, if we consider a "worst case scenario" where deflection causes the UHECR to cover 40 kpc of distance (entire galactic diameter), the transit time is about 130 kyr. In this time the galactic disk will rotate 0.08 degrees. Thus, the GMF can be treated as frozen in place. There are certain situations, such as shocks or blast waves in the ISM which are relativistic, where treating the B-field as constant in time is dubious, however

for particles with $E > 50$ EeV the transit time for such features (on the order of kpc) is small enough to ignore the time dependence. With these assumptions in place Equation 4.2 is a deterministic equation of motion and can be used to numerically integrate the particle's trajectory

$$\frac{d}{dt}\frac{\boldsymbol{v}(x,y,z)}{\sqrt{1-\frac{v^2(x,y,z)}{c^2}}} \approx \pm\frac{Ze}{m_0}\boldsymbol{v}(x,y,z)\times\boldsymbol{B}(x,y,z) \tag{4.3}$$

where Z is atomic number, m_0 is rest mass, e is elementary charge and $+$ is used for forward-tracking and $-$ used for backtracking.

4.2.1 Time reversibility and back propagation

When studying cosmic ray propagation one has the choice of starting with a source candidate and forwarding propagating the particle. The disadvantage to this approach, especially if complex magnetic field models are invoked, is the primary is not likely to hit the target (the Sun). To avoid these issues one can instead start with an observed arrival direction, e.g. measured by Auger or TA, and evolve the anti-primary forward in time. This is equivalent to evolving the original primary backward in time to its origin.

This technique is valid if the propagation process is known to conserve information; in other words, Liouville's theorem can be applied. For a detailed derivation of the theorem see [118]. In the context of CR propagation it is relevant for scattering and deflection processes. If for example the energy scale of magnetic turbulence or a magnetic field interface is on the order of the CR's kinetic energy the forces will be dissipative and information is lost. For $E > 50$ EeV particles, the interaction with

such features can be approximated as conservative. However, great care must be taken for lower energies ($\leq$ 10 TeV) where it has been shown that the conditions for Liouville's theorem no longer hold for environments like the heliosphere termination shock [119]. By using the highest energy particles we isolate ourselves from the onset of chaotic dynamics.

4.3 The Jansson-Farrar 2012 model

Understanding the galactic magnetic field (GMF) is not only of interest to cosmic ray researchers, but also astrophysicists interested in galaxy formation and evolution and dynamo theory. The GMF is typically modeled by using the spiral arm structure as B-field tracers and also looking at the structure of external galaxies for further guidance. A 3D parametrized model is developed and for various assumptions of the thermal and relativistic electron density, n_e, this model is fit to Faraday rotation measure (RM) and polarized synchrotron emission data. In this section I briefly describe the important features of the Jansson-Farrar 2012 (JF12) model [120], for examples of other popular models see [121–124]. These differ from JF12 primarily in the field morphology but use many of the same fitting procedures.

4.3.1 Model description

The JF12 model consists of three main large scale components: regular, random, and striated random. The regular field includes a disk, toroidal halo and an out-of-plane "X-field" component [120]. The disk has 8 logarithmic spirals and is defined

for $5 < r < 20$ kpc where $r = 0$ kpc is the galactic center position. This configuration is a generalization of [125]. The halo component is azimuthal with an exponentially decaying height envelope and separate north and south amplitudes. The "X-field" component is axisymmetric and poloidal and inspired by observations of edge on radio galaxies [126, 127]. The striated random field is included via a global scale factor to the regular field. The strategy of scaling the electron density introduces a degeneracy between the field and electron density, so the scale factor is written as a function of two additional parameters to break the degeneracy. These components are described by 23 free parameters.

The random field is an analytical extension of the regular field combined with the result of the fit for the striated component. The RMS strength is simply the disk and halo magnitude added in quadrature. The fitting procedure did not include explicit random field realizations. Technical details of the approach are explained in [128]. This component consists of 13 free parameters.

The overall model requires fitting 36 free parameters. A combination of RM (40,403 extragalactic measurements) and polarized synchrotron emission data (independent of RM) was used to constrain the model. A cleaning, foreground subtraction, and binning procedure was devised for the RM data (see [120]) resulting in a masked sky map of resolution $\sim 4^\circ$. The WMAP7 [129] Stokes Q and U synchrotron data is used where the original 1° pixels are degraded to $\sim 4^\circ$ and nearby sources are masked out to prevent contamination. Although the data can be considered independent, both are integrated line-of-sight observables which require a description of the galactic relativistic electron density. The authors adopt the simulation based GALPROP

model [130] since this leads to a better overall fit.

The model optimization/fitting is done with an adaptive Markov Chain Monte-Carlo (MCMC) using 100k iterations. The simulated skymap is compared to data is generated using the `HAMMURABI` program [131]. The authors assert the chain has converged since the Gelman-Rubin statistic for the 36 parameters is $\hat{R} < 1.03$ [120]. The best fit parameters are provided in [120] and achieve $\chi^2/\text{ndf} = 1.096$ with a Bayesian information criterion (BIC) of 7401. The reduced chi-square and BIC is compared to [121] and [132] and found to be superior.

4.3.2 MCMC convergence

The MCMC sampling technique is very useful for estimating unknown parameters for complicated models. However, it's important to keep in mind the theory only guarantees convergence for infinite iterations. In practice a large number of iterations (for a simple model) does asymptotically approach the target distribution, but there are numerous counter examples where the same number of iterations produces an obvious non-converged chain. A successful MCMC optimization is a delicate balance between number of iterations, model parametrization, proposal distributions, initial conditions and step sizes.

In the JF12 the convergence condition used was for each parameter to satisfy $\hat{R} < 1.03$ [120]. In general if $\hat{R} < 1.1$ it's an indication that the chain is well mixed and converged [133], however this is a rule of thumb. The $\hat{R}$ diagnostic is not a hypothesis test. In addition to $\hat{R}$ other diagnostics can be performed:

- Compare chains with different initial conditions: yield similar posterior distributions?
- Visual inspection of trace plots
- Visual inspection of joint densities
- Correlation coefficient for parameters
- Autocorrelation of parameter trace plots
- Geweke diagnostic [134]
- Generalized $\hat{R}$ diagnostics [135]

The JF12 optimization resulted in 3 posteriors with problematic shapes: w_h, r_s and z_0. This circumstance alone is enough to bring the chain convergence into question. However, it's also possible the initial conditions are too far from the target density and the discovered optimization is a local minimum.

The point of the sensitivity analysis is to take the ansatz that the true minimum has been found or is at least within the 36-ball formed by the parameter uncertainties, and determine the allowed arrival directions for the model. It can be considered the worst best case scenario. The (*) worst case scenario (the * means wildcard) would be a false minimum optimization which would render model predictions meaningless.

4.3.3 Parameter distributions

All parameters, save three, are assumed to follow normal distributions. As described in [120], the parameters for w_h, r_s and z_0 are the only values observed to

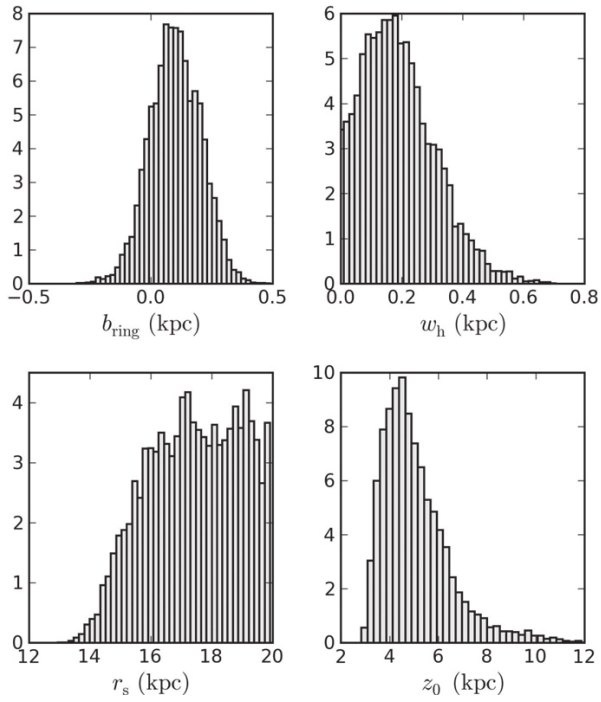


Figure 4.1: Histograms from Figure 7 in [120] which depart from normality. The top left plot shows a representative histogram which follows a normal distribution. The top right, bottom left and bottom right are special cases. The y-axis units are 10^3 counts.

deviate from normality. To account for this a truncated normal distribution is used for w_h and r_s. Although z_0 for the *regular* field is slightly skewed, it is assumed to be normal. For safety I also use truncated normal functions for r_0 and z_0 for the *random* field, otherwise sampling the default distributions could result in spurious parameters, i.e. $r_0 > 20$ kpc which is beyond the field's domain, or $z_0 < 0$ which would lead to a divergent halo field strength. The distribution parameters are estimated by eye using Figure 7 in [120], which is provided in Figure 4.1 in this document for convenience.

The truncated normal distribution is defined as

$$f\left(x;\mu,\sigma,a,b\right) = \frac{\varphi\left(\frac{x-\mu}{\sigma}\right)}{\sigma\left(\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right)} \tag{4.4}$$

Parameter	μ	σ	a	b
r_s [kpc]	18.25	2.75	12.6	19.5
w_h [kpc]	0.2	0.12	0	1.0
r_0 [kpc]	10.97	3.8	0	20
z_0 [kpc]	2.84	1.30	0	8

Table 4.1: Parameters chosen to form the truncated normal distributions to sample regular field w_h and r_s and random field r_0 and z_0.

where φ is the probability density function (PDF) of the standardized normal distribution and Φ is its cumulative distribution function (CDF). Table 4.1 lists the truncated normal parameters, and the PDFs are shown in Figure 4.2.

4.4 Precursor sensitivity analysis

In this section I briefly describe previous work which appeared in an Auger GAP note, and which serves as the foundation of the extended analysis presented in the following sections. This analysis does *not* make use of the updated parameters, nor does it include the effects of turbulent fields. It should be viewed as a helpful primer.

A collection of Monte Carlo simulations which explores a small subset of the total available JF12 galactic magnetic field model parameter space is presented. The goal is to investigate the extent to which published uncertainties in JF12 parameters might lead to model based systematic errors in cosmic ray source directions inferred

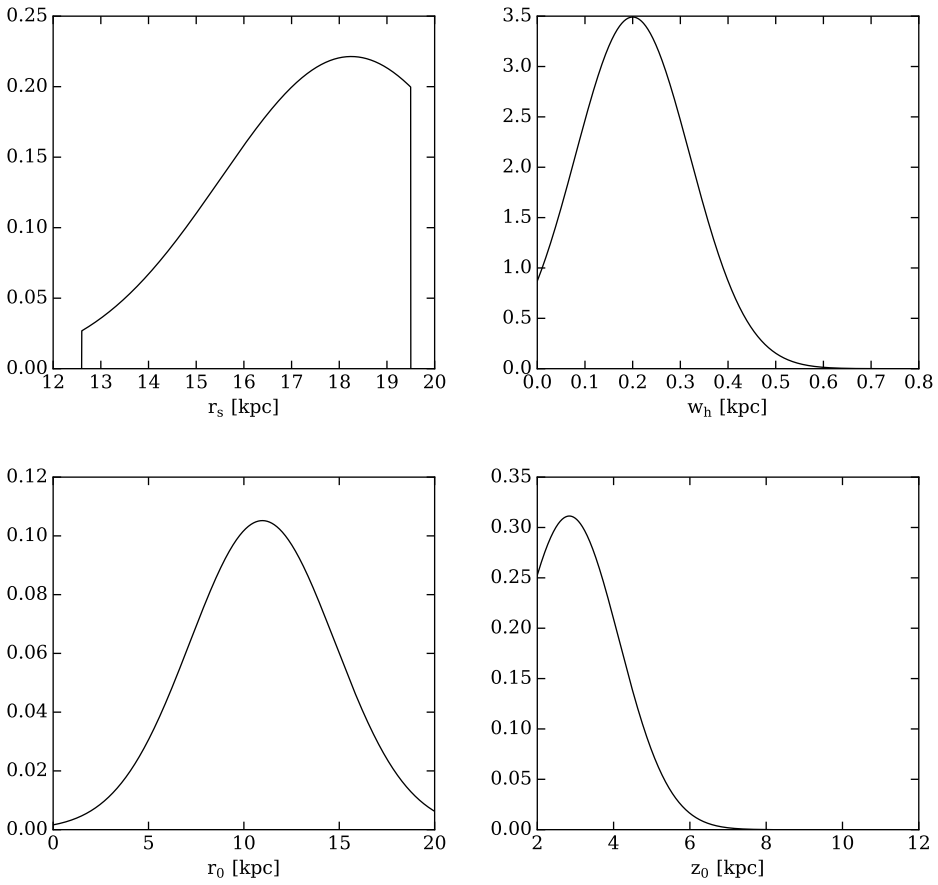


Figure 4.2: The PDFs of the truncated normal distributions. The top two plots are for the regular field while the bottom two are for the random component.

by backtracking for the highest energy Auger events, assuming extra-galactic source origins.

An older set of Auger Herald events with $E > 50$ EeV (of which there are 202) were selected as a sample for this study. A modified version of CRT 3.0.3 [136] was used to generate a sample of 200k backtracked events where parameters are pseudo-randomly sampled based on the published μ and σ value. This resulting set of (ℓ, b) galactic coordinates represents an approximation to the distribution of galactic entry points for each Herald event. We found that on average, the solid angular extent is 16.6 deg^2, many distributions are non-normal and exhibit a variety of morphologies. We also found a small number of multi-modal distributions. To verify our samples, while sparse, are representative of the true underlying distributions, we have also performed "deep" runs using a sample size 2×10^6 for select events.

Our results suggest that systematic errors on estimates for magnetic deflection due to uncertainties in the parameters of JF12 model are comparatively large with a typical mis-position of the backtracked galactic arrival heading of 5.3 ± 1.6 deg.

Since the term sensitivity analysis (SA) isn't widely used in astro-particle physics, we offer the definition provided by [137]

> *Sensitivity analysis studies the relationships between information flowing in and out of the model.*

Physical models are constructed to decode information in a way that helps us understand the world. However, it's unlikely that even the most carefully conceived model perfectly decodes natural processes. This imperfection is understood as uncertainty

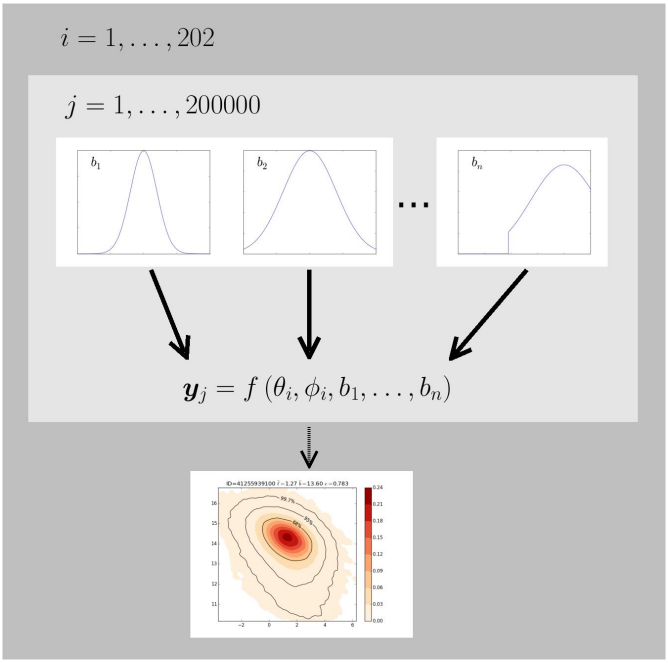


Figure 4.3: The dark gray set refers to the 202 Herald arrival directions which are converted to galactic coordinates and input into the JF12 model, depicted as f as constant. The light gray set represents the MC sampling step. For each iteration pseudorandom parameter values are drawn from their given probability density function, and input into f, which results in a CRT call that back tracks the particle to the galactic entry point, $\boldsymbol{y}$. After 200000 iterations, a sample output distribution is formed, which is depicted by the contour plot below.

in information provided by the model and will result as uncertainty associated with the encoding process [137], which in our case will be physical observables. A diagram of the process is provided in Figure 4.3. SA is used in a wide range of disciplines such as chemistry, engineering and economics. Use cases highlighted by [137] include: determining the dominant factors that contribute to output variability, determining which parameters are insignificant and can be eliminated from the final model, etc. For this study, we use a MC sampling SA technique to probe how the precision of

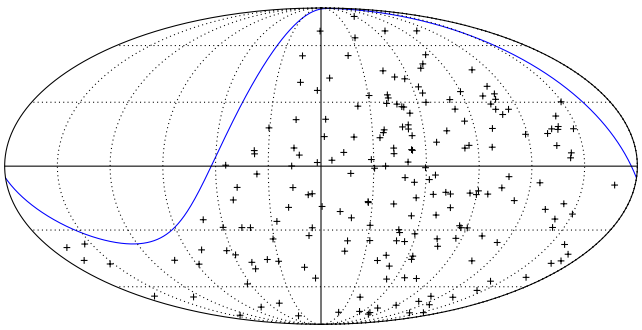


Figure 4.4: A Mollweide projection of the observed arrival directions (black 'plus' sign) and Auger field of view (blue line) in galactic coordinates. Galactic longitude increases to the left.

JF12 model parameters influences UHECR backtracking predictions: (ℓ, b).

I note these results should *not* be interpreted as a complete description of the JF12 galactic entry parameter space. At best, our samples can be interpreted as a highly sparse representation of the joint latitude, longitude marginal distribution. However, the analysis gives some sense of how the model input uncertainties impact the uncertainty of outputs, even though I do not claim absolute convergence of the output distribution.

The analysis uses the CDAS (v5r2) Herald reconstructed events with $E > 50$ EeV. It is important to note that energy and arrival direction observables are subject to some change (within reported uncertainties) as the Auger tasks are constantly refining the data analysis procedures used in Auger. A sky-map of our sample appears in Figure 4.4. The only post processing done on this data is transforming RA and dec to galactic coordinates for input into backtracking software.

Software used in this analysis includes a parallelized front end written in Python3

and GNU Bash which automates the process of building samples and feeding inputs into the back end program, CRT, which backtracks the particles. A thorough description of the CRT code is described in [136]. I make use of version `CRT.3.0.3`. The effects of Kolmogorov turbulent fields are not considered. We use the default numerical integrator and tolerance levels. We've made the following modifications to this version

- The `JF2012` class has been changed to accept user defined parameters for the JF12 model. The default design has these fixed in a header file
- The main `CRT` program has been edited to accept all input from `stdin`, in addition to the default mode of reading a steering file. This eliminates a potential IO bottleneck, and allows for more efficient communication with the front end

A complete record of revision history, along with the source for our front end, and installation instructions, is available at https://github.com/seanpquinn/crt-fork. This page also includes links to download a tarball of the MC samples for each event, as well as the deep MC samples. For portability, the samples remain in the Numpy `.npy` binary format, and can be accessed using Numpy's `load` function. Installation and user guides for Numpy and Python in general are available at http://www.numpy.org/ and https://www.python.org/, respectively.

The front-end loads arrival direction data and generates pseudo-random parameter values based on the distributions.

4.4.1 CRT backtracking call

The set of parameter values are handed to the back-end. The event is then backtracked through this JF12 field realization until CRT successfully exits, or a time out condition is met.

In some cases, the combination of parameters will generate a field that is numerically unstable for CRT. To avoid hangs, we use a timeout limit of 3 seconds (via a piped call to `GNU timeout`). If the integrator hasn't converged after 3 seconds, CRT is terminated, the galactic entry position is recorded as $(0, 0)$, and the iteration is flagged. In our simulations the event, 83371620600, had the most timeouts, 90. This represents about 0.05% of MC samples. The average time out rate for all events is 0.005%.

For each of the 200k iterations the front-end records the galactic entry points, and when finished, saves a binary file of the sample.

4.4.2 Analysis

The MC simulation generates a sample of (ℓ, b) coordinate pairs, or a $p = 2$ dimension random vector with $n = 2 \times 10^5$ samples, for $m = 202$ events. It's technically incorrect to treat these data as 2 dimensional (since they're on the unit sphere), I note that most of the distributions cover small solid angles and argue the flat sky approximation holds. I'm interested in characterizing the resulting bivariate distributions to understand the possible galactic entry locations. We make no assumptions about an underlying bivariate probability distribution, and analyze the MC discretely sampled

distribution.

4.4.3 Outliers

Certain combinations of JF12 parameters yield results that aren't technically forbidden by the model, but are likely to be physically spurious and prevent a meaningful characterization of the primary sample cluster.

We define an outlier to be

$$d_j > \bar{d} + 3\sigma_d \tag{4.5}$$

where d_j is the Mahalanobis distance of the jth element, $\bar{d}$ and σ_d are the mean and standard deviation of the sample's Mahalanobis distances. This is a statistical distance first introduced by [138], and given by

$$d_j = \sqrt{(\boldsymbol{y}_j - \boldsymbol{\mu})^T S^{-1} (\boldsymbol{y}_j - \boldsymbol{\mu})} \tag{4.6}$$

where $\boldsymbol{y}_j$ is the jth galactic entry position, $\boldsymbol{\mu}$ is the sample mean entry position and S is the sample covariance.

In general this approach isolates the bulk feature of the sample and makes further processing tenable. Example quality cuts are shown in Figure 4.5

4.4.4 Kernel density estimation

The majority of MC samples for the events are non-normal and exhibit complex morphologies. We do not attempt to fit a bivariate normal to the data, but instead approximate the underlying PDF using kernel density estimation (KDE), with a Gaussian kernel. Bandwidth selection is done using Scott's "rule of thumb approach" and

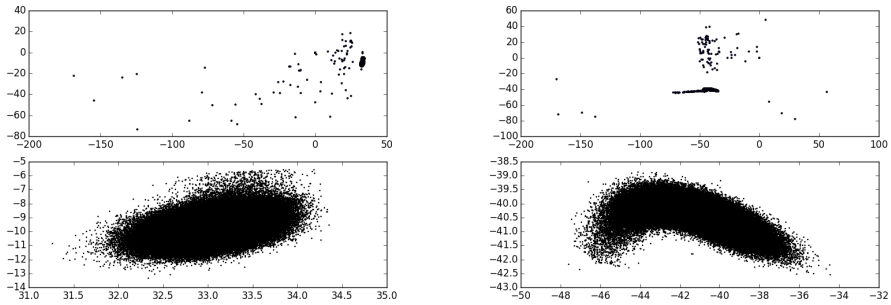


Figure 4.5: Entry point scatter plots for events 82685012000 (left) and 100802228200 (right). Abscissa is galactic longitude and ordinate is galactic latitude. The top panel displays the raw samples, the bottom shows the filtered sample.

calculating the Scott's factor [139] for every event's sample.

4.4.5 Probability contour calculation

To generate the $68\%, 95\%, 99.7\%$ contours we've implemented the following procedure:

- A 150×150 grid is created using the rectangle defined by (l_{min}, b_{min}), (l_{min}, b_{max}), (l_{max}, b_{max}), (l_{max}, b_{min}). This resolution is chosen for computational efficiency.
- The KDE is used to calculate the PDF at each grid point. For a 150^2 grid this will be on the order of 8 minutes.
- Points with a PDF of $< 1 \times 10^{-5}$ are considered to be a numerical artifact and are discarded. If left unfiltered, such points would introduce spurious tail structures to the PDF in regions where data is not present.

- The CDF is then calculated from the grid points and the contours are found by solving $\Phi\left(l, b\right) - \left(0.68, 0.95, 0.997\right) = 0$ numerically.

4.4.6 Sample statistics

To quantify the angular extent of the contour we calculate its Euclidean area using the polygon formed by the contour. This area is given by the "Shoelace" equation

$$\left|\frac{1}{2}\Big(\left(\ell_1 b_2 - b_1 \ell_2\right) + \left(\ell_2 b_3 - b_2 \ell_3\right) + \cdots + \left(\ell_{n-1} b_n - b_{n-1} \ell_n\right)\Big)\right| = \frac{1}{2}\left|\sum_{i=0}^{n-1} \ell_i b_{i+1} - \ell_{i+1} b_i\right| \tag{4.7}$$

where ℓ_i and b_j are the i^{th} and j^{th} galactic longitude and latitude in the set of coordinates that form the 99.7% contour.

To test for multivariate normality we look at the distribution of distances. By the law of large numbers ($n - p$ large where n is the number of samples and p is the number of variables) the sample of statistical (Mahalanobis) distances for a multivariate normal distribution will be χ^2_p distributed: $d^2_j \sim \chi^2_p$. Therefore if distance quantiles vs. χ^2_2 theoretical quantiles do not follow a straight line trend, we can reject the bivariate normal assumption.

Two sample QQ plots are shown in Figure 4.6 which are representative of the remaining plots.

4.4.7 Deep run

Since this analysis performs MC simulation on a high dimensional parameter space I do not claim complete convergence for the resulting joint longitude, latitude dis-

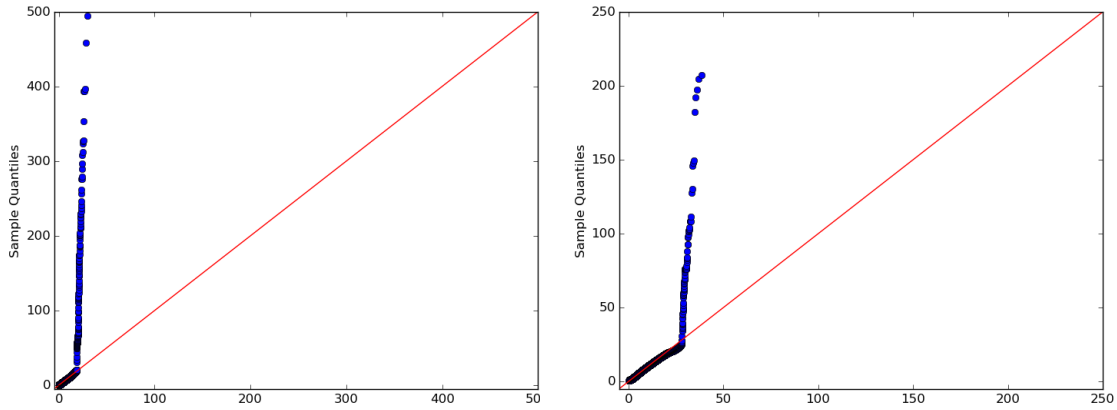


Figure 4.6: Sample distance quantiles stemming from a bivariate normal distribution would closely follow the theoretical quantiles of the χ^2_2 distribution. The QQ plots for these events show non-normal behavior, especially toward the tails of the distribution.

tribution. We reiterate that our results, *at best*, represent a sparse sample of the true joint distribution. It's possible that the shape, or statistics of our samples could change as $n \to \infty$. We address this criticism by performing 'deep runs' (sample size increased $\times 10$) on a subset of events with irregular distribution shapes. If the resulting distributions do not change significantly I argue for a converged regime.

Deep runs were performed for these events: 90078442400, 102414692000, 103648222300, 152677849300, 42391987800, 60354407800, 62967626400, 70844935900, 81183160400, 83286880300, which is about 5% of our data set.

I test the difference in underlying distributions in two ways. First, I check that the proportion of filtered outliers hasn't changed. For example, if a larger fraction of points are considered outliers in the deep run, this indicates the additional parameter space is contributing new components to the $p\,(l, b)$ distribution. A table of results

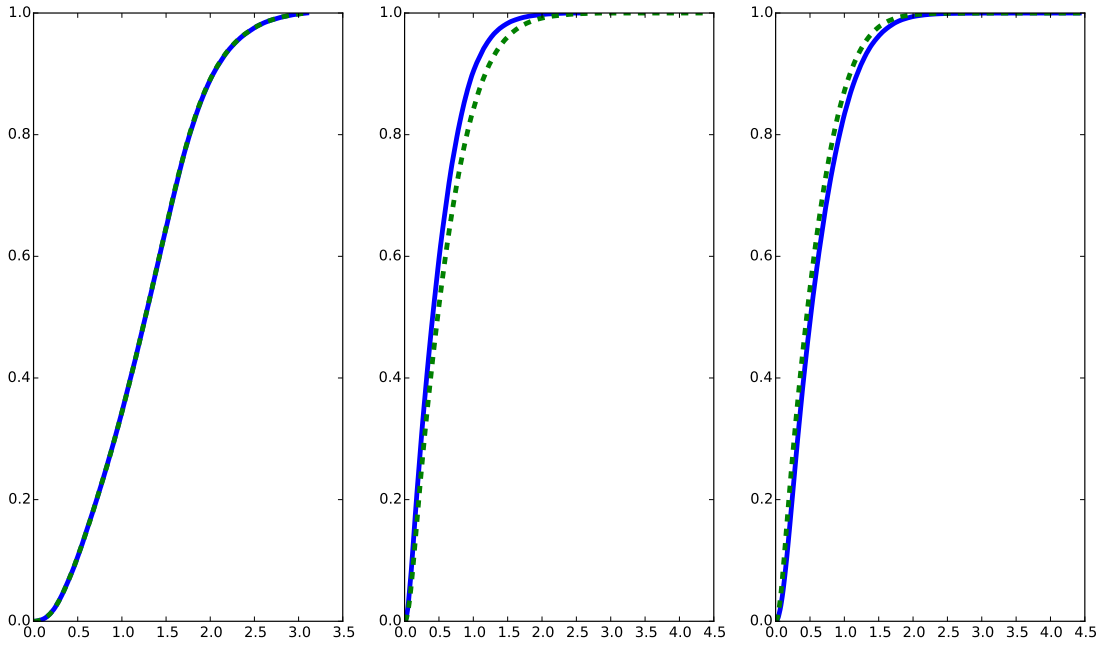


Figure 4.7: A sample of 3 empirical CDFs for events 102414692000, 60354407800, 83286880300 (left, middle, right). The standard sample size is the solid blue line and the deep run is the dashed green line.

appears in C. The fraction of outliers is mostly unchanged for the two samples: less than 0.1% for the 10 selected events.

Next, I compare the distribution distances using an empirical CDF plot of the sample distances. This plot is a useful diagnostic for detecting differences in distribution shape. A majority of the 10 events have overlapping CDF plots. For some, as shown in the sample plot in Figure 4.7 there are slight differences. These cases suggest larger sample sizes might be necessary.

4.4.8 Results

A majority of events have galactic entry distributions that are skewed and asymmetric while a small minority, 5 out of 202, can be considered bivariate normal, assuming a liberal inspection of the QQ plots. We find an average angular area of 16.57 deg^2 for the 99.7% confidence contours. In addition to these summary statistics, we highlight the shape and characteristics of a few example distributions, and also examine one event which exhibits a bimodal galactic entry distribution.

4.4.9 Distribution morphology, descriptive statistics

The event PDF profiles span a large range of shapes. A sample of six PDF contours which highlight the range of features are shown in Figure 4.8.

4.4.10 Multi-modal events

Two events (83371620600,102414692000) exhibit multiple PDF peaks. Both events are near the galactic plane ($|b| < 20$ deg) and are close to the origin meridian ($|\ell| < 20$) deg. For these events there are multiple locations at the galactic boundary where a primary has equal probability of arrival. This is more pronounced in 83371620600 which shows two well defined, seemingly symmetric, profiles whose peaks are separated by $|\Delta b| = 23$ deg. Event 102414692000 has two prominent side lobes separated by $|\Delta b| \approx 8$ deg. from the central peak. These results suggest that events near these galactic coordinates should be treated with extra caution, as there are multiple possible regions of entry. This is especially true for 83371620600, whose bimodal dis-

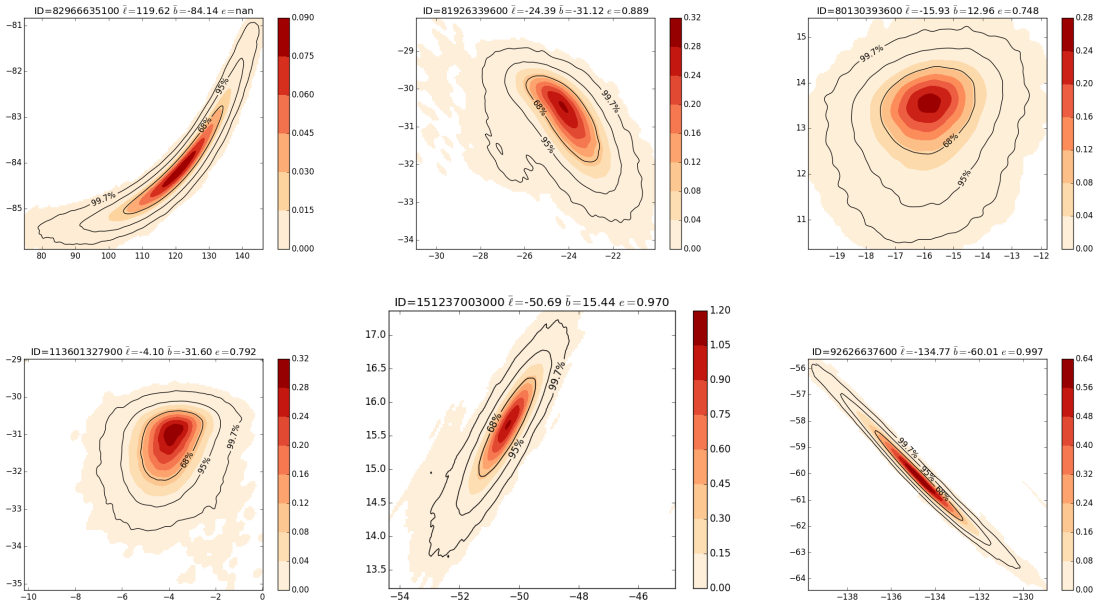


Figure 4.8: Contour plots which include identifying information in the title. The color scale refers to the probability density and it isn't normalized.

tribution covers a large angular area. Contours for these events are shown in Figure 4.9.

4.4.11 Objects in contours

While not technically appropriate, I treat the 99.7% contour as a "3σ" confidence interval and use this to constrain possible source candidates. I assume zero deflection by extragalactic fields for convenience and note there's evidence this field is non-zero [140, 141]. We adopt the half million quasar catalog for source candidates [142]. A skyplot of this catalog's coverage is show in Figure 4.10 A sample of three events with source candidates contained in the 99.7% contour are presented in Figure 4.11.

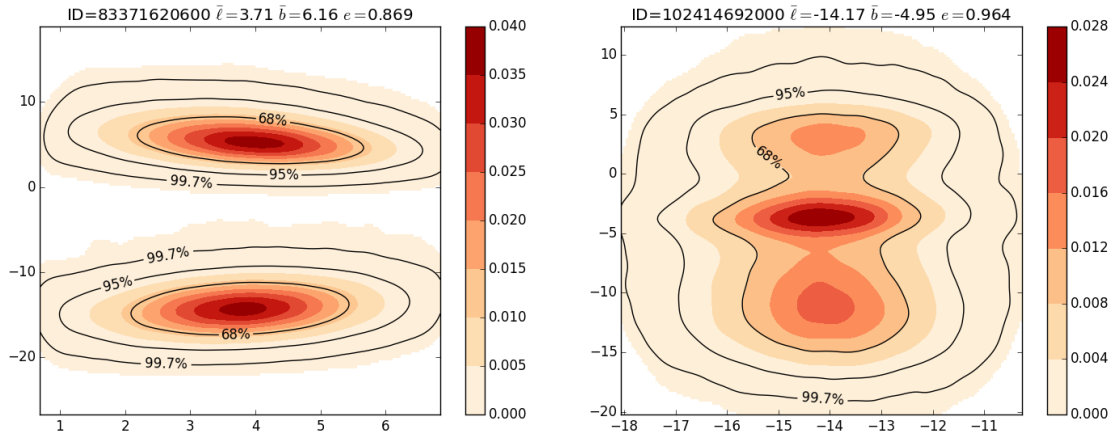


Figure 4.9: Similar to an earlier plot, abscissa and ordinate are galactic longitude and latitude, respectively. Since event 83371620600 has cleanly separated peaks, the analysis code was able to compute probability contours for each peak independently.

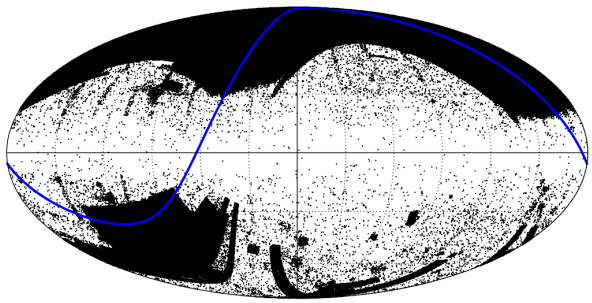


Figure 4.10: The HMQ catalog in Mollweide projection in galactic coordinates with Auger exposure (blue line). Coverage is better at higher absolute galactic latitudes.

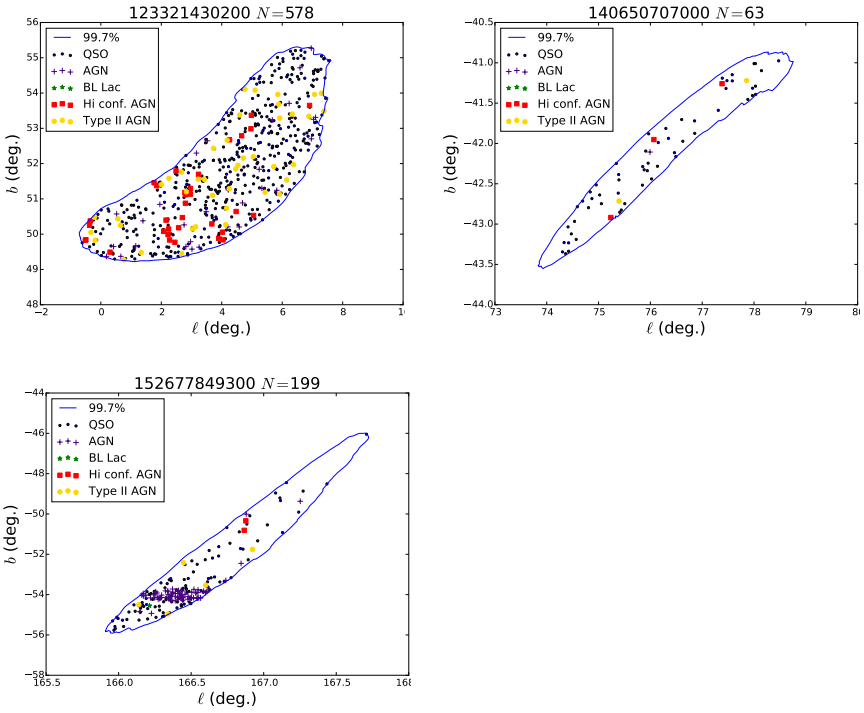


Figure 4.11: Flat sky plots of potential source candidates.

4.5 Discussion and conclusions

In this section I presented a preliminary analysis of a large Monte Carlo simulation of back tracked Auger events at the highest energy. The results represent a sparse sample of the joint $p(\ell, b)$ PDF from the larger 22 dimension JF12 parameter space which are used to construct "3σ" angular contours. We find that these contours are non-normal, exhibit a variety of directionally dependent morphologies, and cover an average angular area of about a dozen square degrees. Moreover, for certain galactic trajectories the JF12 parameter space allows for multi-modal galactic entry positions.

I find that a majority of the 99.7% contours do not contain an excess of source candidates. This is most likely due to the dearth of catalog coverage at low latitudes.

The analysis demonstrated that uncertainties in parameter estimates of the JF12 model, when simultaneously accounted for, produce non-trivial galactic arrival probability distributions. Furthermore, the angular size of the highest confidence regions pose a significant challenge for charged particle astronomy: given the present precision of this model it isn't possible to resolve individual sources for most events. This highlights the need for accurate and precise descriptions of astrophysical magnetic fields in order to successfully identify primary sources. A complete description of this early study can be found in GAP note 2016-012 [143].

4.6 New data set

A new data set is adopted for the improved SA. The energy cut is maintained ($E >$ 50 EeV), but uses the Auger 6T5 cut. These events have the smallest uncertainty in arrival direction of all trigger types measured by the array. It uses the Herald reconstruction v6r0. As of 5/31/2017 there are 224 events with $E > 50$ EeV. A sub-sample of 102 events is analyzed due to limited computer time. A histogram of the energy distribution is shown in Figure 4.12 and a skymap of reconstructed arrival directions is shown in Figure 4.13 along with the back tracked positions (described later).

4.7 The turbulent magnetic field

Galactic magnetic turbulence in the large scale interstellar medium (ISM) is currently poorly constrained, compared to the small scale ISM . The main way to study this phenomenon is using the variance of Faraday RM measurements. These are used to produce structure functions which can be used to study the spectral features of the turbulence. As more RM data becomes available and sky coverage increases the picture will become clearer, but a position dependent and high resolution map of the turbulent field's strength and length scale presently does not exist.

The most common form of turbulence considered is that of an isotropic stochastic incompressible fluid described by the Kolmogorov 5/3 law [144]. Many studies of the electron density power spectrum and solar wind phenomena are compatible with an

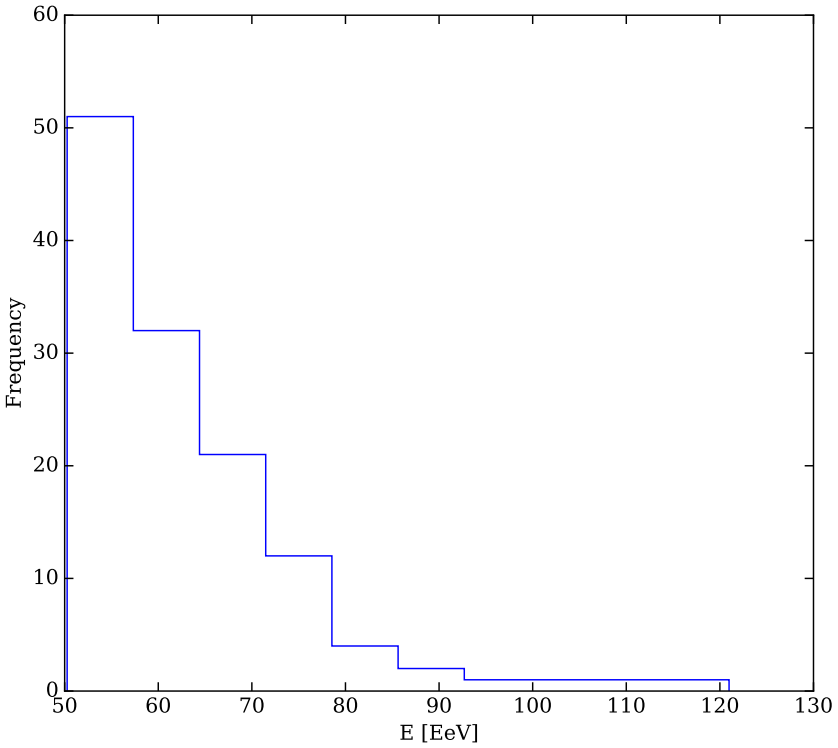


Figure 4.12: Energy histogram for events in the updated sensitivity analysis. The inescapable power-law nature of the statistics is clearly apparent.

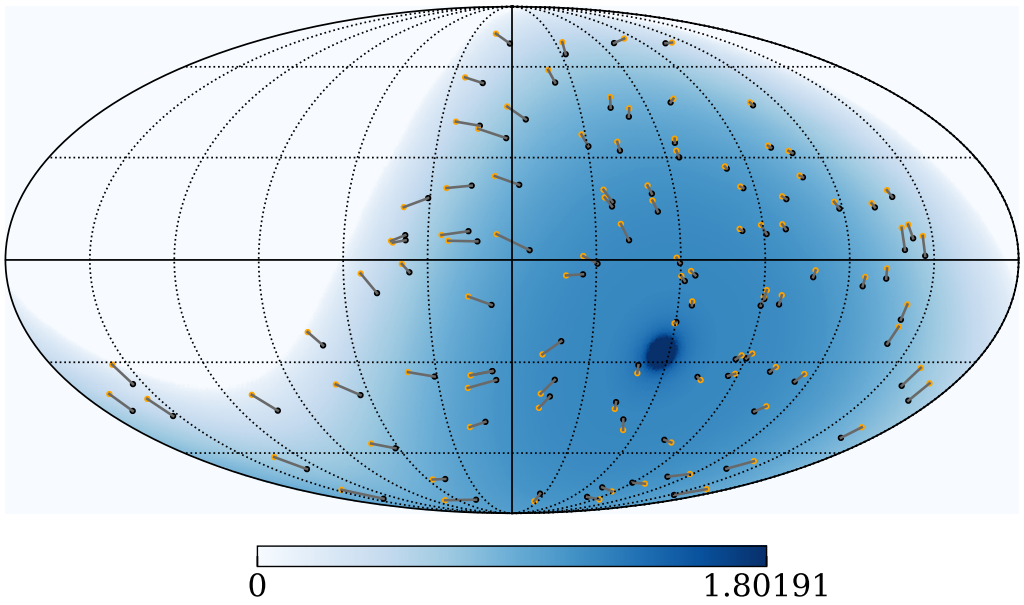


Figure 4.13: Skymap of observed 6T5 events (black) used in analysis and their deflected trajectories (orange). The color map shows the Auger relative exposure (arbitrary units). Galactic coordinates and a Mollweide projection are used with longitude increasing to the left.

index of $\alpha = 5/3$ [145–150]. An alternative model of turbulence is that developed by Kraichnan which predicts an $\alpha = 3/2$ scaling [151]. Anisotropic models also exist, such as that of Galtier et al. which predict an index of 2 for modes perpendicular to the mean B-field and -1/2 for parallel modes [152] or that of Goldriech & Sridhar which leads to an anisotropic Kolmogorov type spectrum [153]. The source of ISM turbulence is thought to be predominantly injected by supernovae remnants or super bubbles.

In an early study in the late 1960s, the rms field strength was found to be 1.5 μG with a correlation, or coherence, length of 250 pc [154]. Incidentally, while the authors only had limited data and used a simple model of the galaxy, these are consistent within two orders of magnitude (in some cases much closer) to current values found in the literature. A more recent study which considered variations between spiral arms and interarm regions using RM structure functions and depolarization modeling: it found outer scales on the order of 100 pc for interarm regions, but 2-10 pc within the spiral arms [155]. This result agrees with other studies which used independent observations [156, 157]. Additionally, other papers present similar scales in the spiral region of a few pc [158]. Some of the latest results point to an outer scale of 16-29 pc based on theoretical estimates using LOFAR diffuse synchrotron emission data [159]. This paper also provides a convenient diagram showing past estimates of the outer scale, which is reproduced in Figure 4.14.

The discrepancies are not necessarily contradictory, since different regions of the galaxy were probed, and the outer turbulence scale might be heavily dependent on local injection sources. Given the difficulty of fully specifying the coherence length

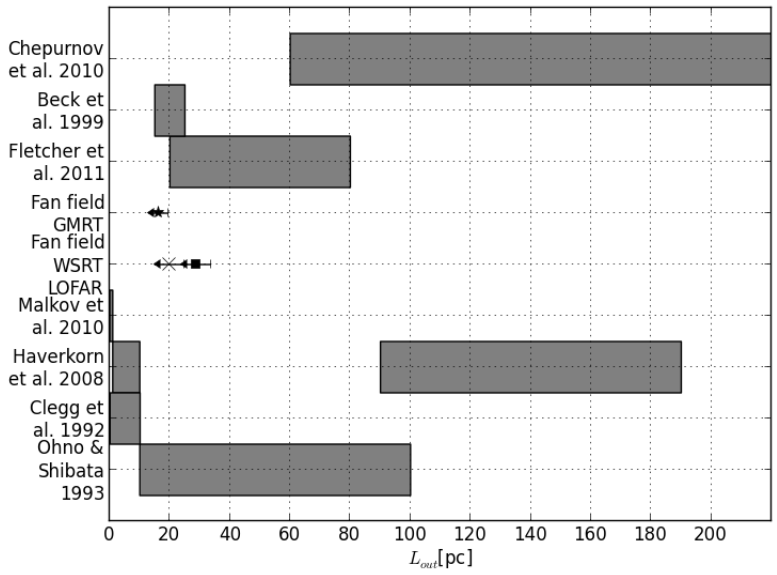


Figure 4.14: A summary of outer scales of turbulence found in different analyses. This is Figure 10 from [159]. Most studies show a large range of possible values while others show upper limits. Works cited can be found in [156, 160–162, 158, 157].

everywhere in the galaxy, a compromise is adopted: the turbulent random field follows a Kolmogorov power spectrum with constant $\ell_c = 60$ pc.

4.8 CRPropa3 and the new SA pipeline

The CRPropa3 code was used for the updated propagation study. Details of its design and physics capabilities can be found in [163]. While not as fast as CRT, it offers other benefits such as a Python interface, a full treatment of spallation and energy loss processes, periodic turbulent field grid scaled to the regular field, cosmological energy losses, and many convenient utility functions. The work-flow is almost identical to the previous study:

1. Auger energy (E) and arrival direction (Galactic longitude and latitude (ℓ, b))

and the Sun's position are input as initial conditions

2. The JF12 parameters are pseudo-randomly sampled from their distributions

3. A turbulent field realization (with a random seed) is generated based on parameters drawn in the previous step

4. The particle is propagated until: a timeout condition of 35 s is reached or the particle reaches the 20 kpc spherical boundary (the edge of the JF12 domain)

5. The particle's final energy, position, momentum, redshift and timeout condition flag are saved to a binary data file

A diagram showing this process is available in Figure 4.15. The starting location is taken as $(x, y, z) = (-8.5, 0, 0)$ kpc (within uncertainty of most up-to-date values [164]) in galactocentric coordinates. The initial direction $\hat{p}$ is found using by converting the observed longitude and latitude with the `dir2vec` Healpix function with *lonlat* flag enabled [165]. The particle's initial state is now fully specified and ready to be back-tracked. The next stage is the start of the 100k loop block. At the start of the iteration a B-field realization is created using the parameter distributions. The parameters specify the regular field and are also used for the turbulent component creation. Since a new turbulent field is generated for each realization, this also folds in the variance associated with different turbulent field morphologies. Technically speaking, to fully account for this variability one would need to produce a large sample of turbulent fields for each of the 100k realizations. However, even for a small sample such as 1000, this would be prohibitively expensive to compute since now 100

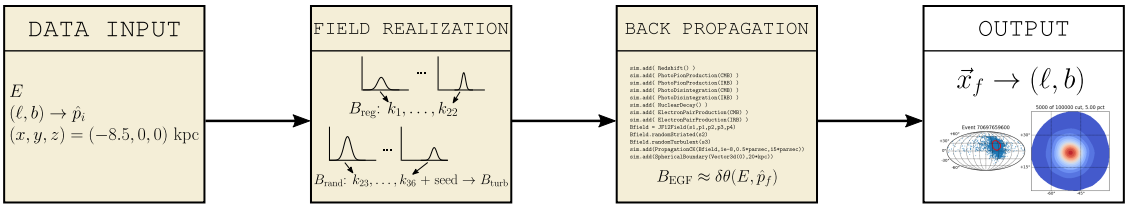


Figure 4.15: A flow chart showing the analysis framework. The process begins with Auger data starting at the Sun, generates a field realization, back tracks to the edge of the galaxy, and is propagated to the GZK horizon using an approximation of the turbulent EGF. The final output stage is a collection of 100k coordinates.

million trajectories must be backtracked. Creating a new turbulent field with each parameter realization attempts to fold their effects into the analysis.

With the field fully specified, the code is then instructed to account for the following physics: energy loss due to redshift (negligible in the galaxy), photo-pion production due to the cosmic microwave and infrared background, photodisintegration due to those backgrounds, nuclear decay, and electron pair production due to cosmic backgrounds. A Cash-Karp integrator is used with an enforced error tolerance of 10^{-8} with a minimum step of 0.5 pc and maximum step of 15 pc. The simulation is run using the GNU timeout utility with a threshold of 35 seconds. When complete the program saves the final position $\vec{x}_f$, final direction $\hat{p}_f$, final energy E and a flag indicating time out (0 if successful return).

An approximation function is used to account for deflections due to the turbulent extragalactic magnetic field (EGF), which is described in the next sub-section. This is a significant difference compared to the initial study and other improvements have been made in terms of source candidate selections.

4.9 Extragalactic turbulent field

CRPropa3 is fully capable of simulating the effects of EGF however this comes at the cost of computation time. A full treatment of extragalactic propagation would be a dissertation by itself since one must account for fairly strong and potentially ordered field strengths in nearby clusters and superclusters on the scale of 1 Mpc [166]. Since these are the injection scales it is reasonable to take 1 Mpc as the correlation length. With this condition and the assumption of a stochastic EGF and $N \propto E^{-2}$ generating spectrum and no energy loss, it's possible to derive an angular smearing function, see [167, 56], for a given propagation length

$$\delta\theta = \frac{e\sqrt{2d\overline{B^2\ell_c}/9}}{E} = 0.8^\circ Z \left(\frac{E}{10^{20}\text{ eV}}\right)\sqrt{\frac{D}{10\text{ Mpc}}}\sqrt{\frac{\ell_c}{1\text{ Mpc}}}\left(\frac{B_{\text{rms}}}{10^{-9}\text{ G}}\right) \tag{4.8}$$

where E is energy, d is propagation distance, ℓ_c is EGF correlation length, B_{rms} is the rms field strength and Z is atomic number. As mentioned previously, I use $\ell_c = 1$ Mpc and adopt $B_{\text{rms}} = 0.9$ nG based on the limits established in [71]. Protons are also assumed.

Smearing doesn't account for energy loss processes, so it is instead embedded into the propagation distance using an energy dependent GZK horizon radius, found in Figure 2 of [168], reproduced as Figure 4.16 in this document. This distance is then taken as the total extragalactic propagation distance d. A spline interpolation is used to calculate horizon distances not provided by the figure. Additionally, for energies $E > 110$ EeV, the horizon for $E = 110$ EeV is assumed. The same approach is used for $E < 55$ EeV.

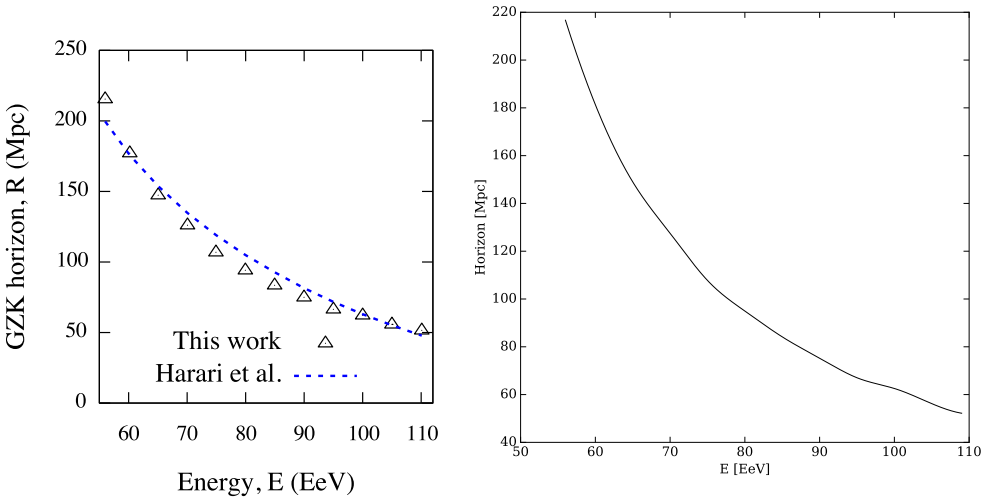


Figure 4.16: Left: Figure 2 in [168] which provides the horizon distance as a function of energy. Right: Interpolation of the left plot using a cubic spline.

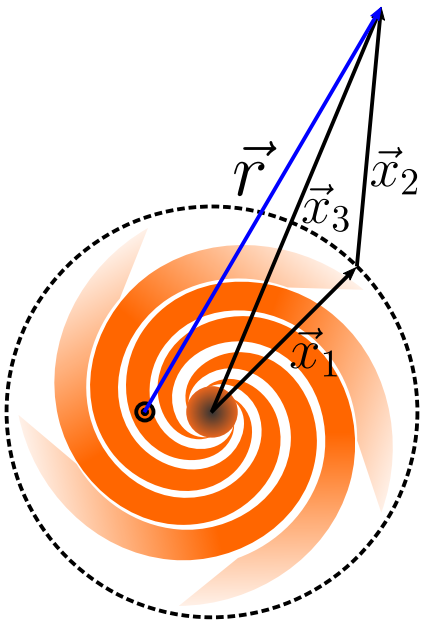


Figure 4.17: Diagram showing the vectors involved in propagation. After scaling to the GZK horizon a final galactocentric direction vector is constructed, and then transformed to galactic coordinates using the `astropy` *Python library.*

To complete the EGF approximation, the final primary coordinates for each of the 100k iterations are determined using the following approach

1. $\delta\theta$ is calculated using Equation 4.8.

2. The normalized momentum vector $\hat{p}_f$ is used as the mean direction for EGF smearing.

3. The `randVectorAroundMean` function from the CRPRopa3 `Random` module is used to generate a "smeared" version of $\hat{p}_f$.

4. The smeared $\hat{p}_f$ from step 2 is scaled by d to give the final position of the primary out to the GZK horizon from the galactic exit point. This is shown as $\vec{x}_2$ in Figure 4.17.

5. The resultant vector in galactocentric coordinates, $\vec{x}_3$, is found based on the exit location $\vec{x}_1$ and horizon position $\vec{x}_2$.

6. The direction of $\vec{x}_3$ is transformed into galactic coordinates using coordinate tools in the `astropy` Python library. The distance to the galactic center is identical to what is used in the propagation studies, 8.5 kpc, and the sun is taken to be at $z = 0$ pc in the galactic plane.

4.10 Data cleaning and outlier removal

A new method was devised to handle outliers. In the previous study cuts were made based on the difference from the mean of the Mahalanobis distance distribution. This does provide a good measure of the 'statistical deviation' for many points, but in

certain cases it does not properly handle the spherical nature of the data and leads to strange artificial boundaries. To address this concern, I've devised a new algorithm using robust statistics. First the spherical median is computed for the normalized final position vector, $\vec{r}$ in Figure 4.17, using the `mediandir` function from the `Directional` R package which is based on the definition found in [169]. This summary statistic is robust against spherical outliers and will more accurately determine the main core of the directional distribution.

With the spherical 'center' of the distribution determined, the angular distance from this point and each of the 100k iterations is calculated. The median of the resulting position angle distribution is calculated and points beyond this median, $\Delta\theta{+}3\sigma$ are labeled as outliers and cut. Figure 4.18 illustrates this process for a typical event. To handle any outliers which may have survived the first cut, an additional layered is used: points beyond the 95^{th} percentile are excluded from further analysis. An example of the raw data for each of the field realizations (where no timeouts occurred) along with the resulting cut is shown in Figure 4.19.

4.11 Arrival direction probabilities

After the quality cuts we are left with a sample of ≈ 95000 points. From this, a probability density function is found using a Gaussian kernel density estimate. Since the sample size is large, FFT methods are employed using the *fastkde* Python library [170, 171]. The resulting PDF estimate is then sampled on a 513×513 planar grid to generate a contour. A flat sky approximation is assumed and justified since the

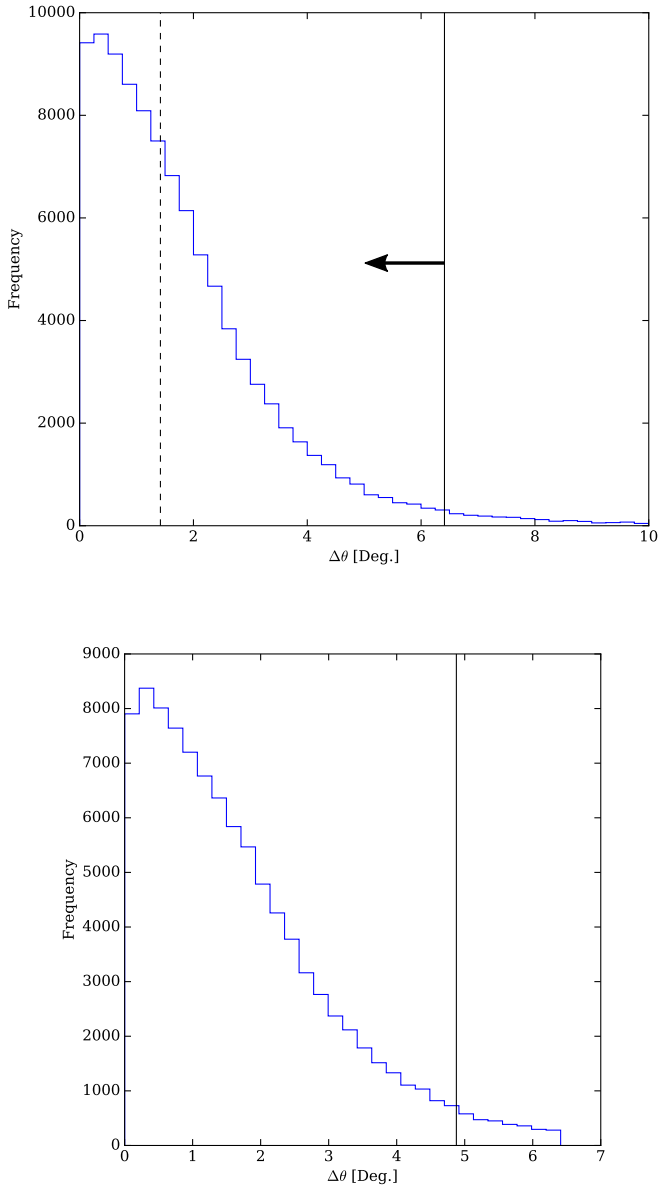


Figure 4.18: Left: Position angle distribution (from spherical median) for Event 82481424901. The median position angle is denoted by the dashed vertical line, and the $\tilde{\Delta\theta} + 3\sigma$ is the solid vertical line. Coordinates which have position angles smaller than this threshold are retained for further analysis, and those beyond are cut from the sample. Right: Position angle distribution after the cut on the left plot is applied. The solid vertical line is the 95^{th} percentile.

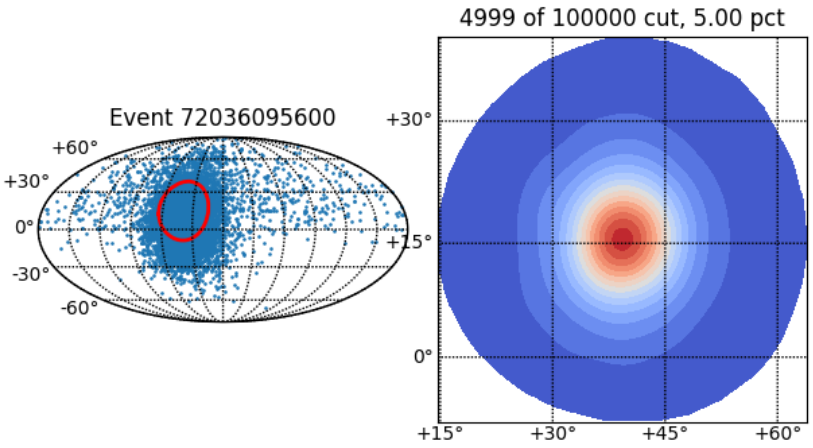


Figure 4.19: Left: Skymap in Mollweide projection showing all simulated arrival directions (teal) along with the quality cut boundary (red). Right: Kernel density estimate of the remaining sample.

angular scales of these distributions are small enough such that little information is lost. See Figure 1 in [172] for the impact of flat sky on various power spectra studies—for a multipole of $\ell \geq 6$ corresponding to angular scales of $\leq 60°$ the discrepancy is small. Given most contours are on the order of 40° the angular resolution of the PDF is $\approx 0.08°$. The distribution is confined to be within the cut region.

4.11.1 Characterization

Of primary interest is the morphology (shape) and size of the distribution. Morphology is important because in conventional anisotropy studies magnetic deflection is assumed to be simple Gaussian smearing, or the effects is expressed as a circle with some angular radius (usually about 10°). These approaches are not physically moti-

vated compared to using a GMF model. The community generally views magnetic field models as unreliable and instead resorts to the mentioned ad-hoc methods. This is problematic for several reasons. Uniform smearing over the sky is unrealistic and will lead to biased results: given the Milky Way is either a disk or barred spiral [173] one expects different total deflection as a function of galactic latitude. A 75 EeV primary arriving at $b = 75^\circ$ experience much less net deflection than one arrival at $b = 3^\circ$. The uncertainty is much lower for the former compared to the latter, even for an unknown GMF model, since its path length through a magnetized medium is considerable smaller. Assigning the same uncertainties to two such events is wrong, but this practice is common in the literature. Furthermore, the study discussed in this chapter is physically motivated: the GMF is at least partially constrained by astronomical observables. My study shows that in many cases, even after accounting for the inherent statistical error of the model due to the MCMC, the error contour can have a radius smaller than the typical spread found in the literature (e.g. 20° in [109] or 3.4° in [174]) for certain sky positions. There are cases where the contour can also have a larger radius, especially toward the galactic plane and in certain longitude bands due to the disk geometry. Moreover, in most instances the contour is not Gaussian or circular in shape, and the shape can change depending on location. In rare cases, the contours are multimodal and cannot be described by the widely used ad-hoc treatment. The main point is the conventional position uncertainty weighting does not properly reflect the actual statistical uncertainty of a physically motivated magnetic field model, and it introduces potentially spurious biases into an anisotropy analysis.

A valid counter-argument is that my analysis does not offer any treatment for a possible systematic error in the magnetic field model. However, this point was addressed in an earlier section and corresponds to the worst-worst case scenario where the MCMC has converged on a false minimum. The only way to contend with this issue is a running a longer chain to continue probing the parameter space, and one can only be truly certain about convergence after infinite time.

The sensitive analysis provides a novel way of including the model uncertainty. While the distributions generated with this approach can be complicated and unwieldy, they are at least physically motivated and give an accurate representation of the longitude and latitude joint probability density for the 35 parameter JF12 model. Now let's examine the main results for the "new and improved" framework using CRPropa3.

The raw data, cuts, and density estimates are given in Appendix A. The density estimates, 1σ & 2σ contours, and deflection maps are given in Appendix B. For this sample the mean deflection out to the GZK horizon was found to be $\delta\theta = 4.51 \pm 2.82$ degrees (uncertainty is standard deviation). The deflected end point uses the most probable value (MPV) of the final arrival direction distribution, and the deflection angle is calculated between this and the observed coordinate (i.e. the position angle). A histogram showing the deflection values appears in Figure 4.20. It shows that for a majority of events the total deflection is a few degrees, but a fair number also experience relatively large $> 8^\circ$ deflections. These ranges are the best case scenario since this value will grow as charge increases. We can also investigate how this magnitude changes vs. galactic longitude and latitude. Figure 4.21 shows plots which

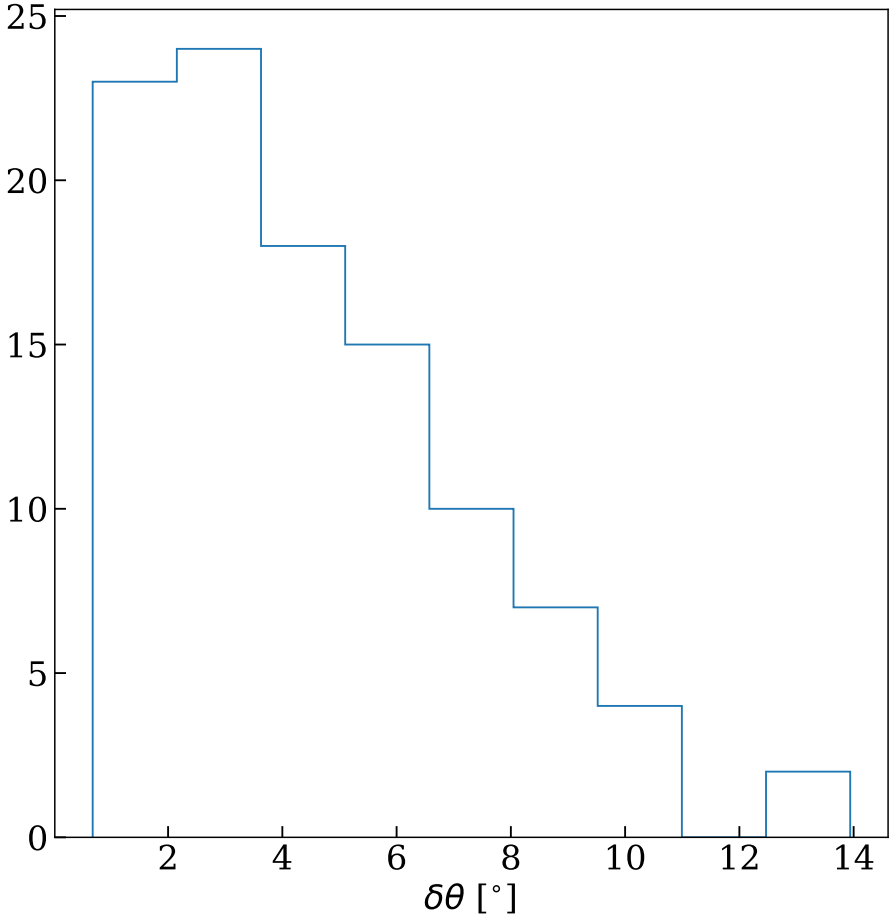


Figure 4.20: Histogram of deflection values for events.

examine this dependence. It's clear that toward observed arrival directions with headings toward the galactic center are severely deflected. As one moves to higher or lower latitudes it will decrease, with the exception of around $-150°$. Of the 102 events, it appears a small interval from -60 to -130 degrees has the lowest average deflection. The picture for galactic latitudes is somewhat similar, with larger deflections present at low latitudes. One important difference is the variance in this data which make it very difficult to isolate an interval where the average deflection is low. There is an obvious trend of a decreasing mean deflection toward the southern galactic pole, but unfortunately it approaches a mean value of around $4°$ instead of 0. There isn't enough data near the north galactic pole to make an inference about behavior in that region, although one could consult a deflection map generated in the original JF12 paper [120]. With these results, one could speculate that a possible sweet spot which minimizes deflection contamination is a heading toward the south galactic pole around longitude $-100°$.

Next we can look at the model uncertainty behavior. It's found that the average area of the 1σ contour, which is taken as the model uncertainty is 316.64 deg^2. For reference this is about the size of the constellation Lupus. The variance is about the same as the mean. In Figure 4.22 I show the histogram of 1σ angular areas. The distribution is fairly peaked, but also exhibits a heavy tail. The tail results from areas of high uncertainty mostly stemming from the galactic disk, which can be seen in Figure 4.23. We see that toward the galactic plane the 1σ contours become large, which is a consequence of the uncertainties associated with the spiral arm field parameters. This can also be seen in the pseudo-periodic structure in the longitude

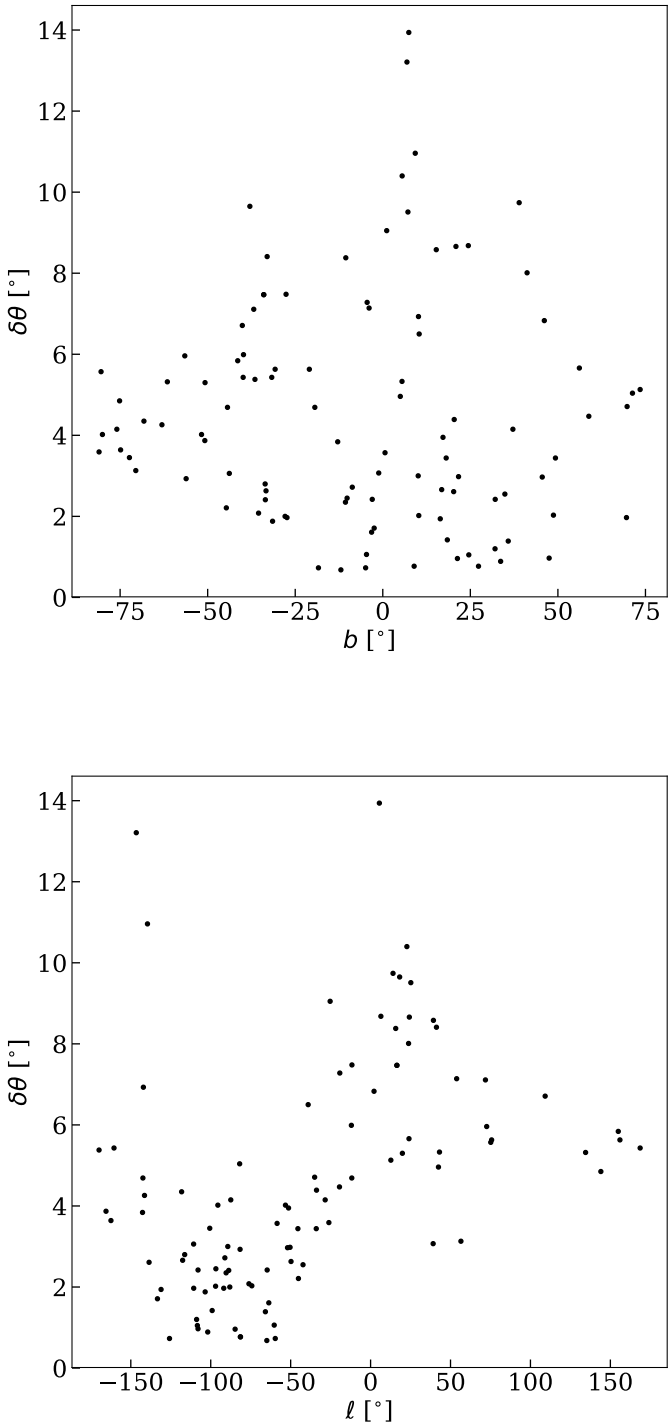


Figure 4.21: Left: deflection vs. galactic latitude using [-90,90]. Right: deflection vs. galactic longitude using [-180,180].

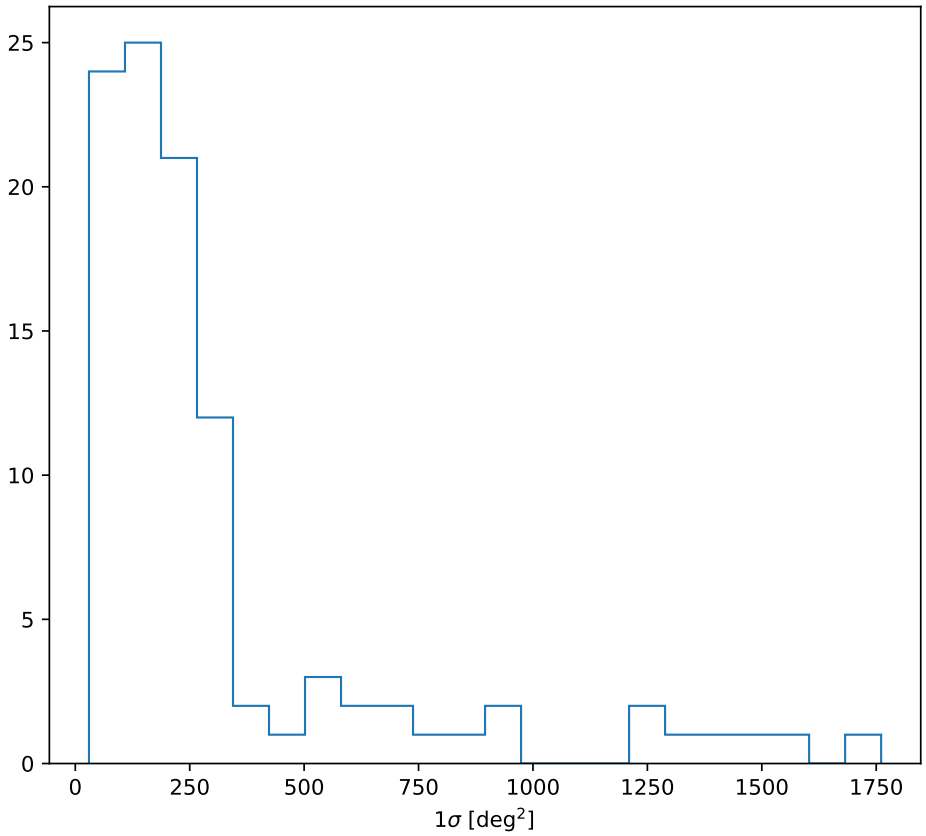


Figure 4.22: Histogram of JF12 uncertainty angular size. The distribution is mostly contained in the range of < 400 deg^2.

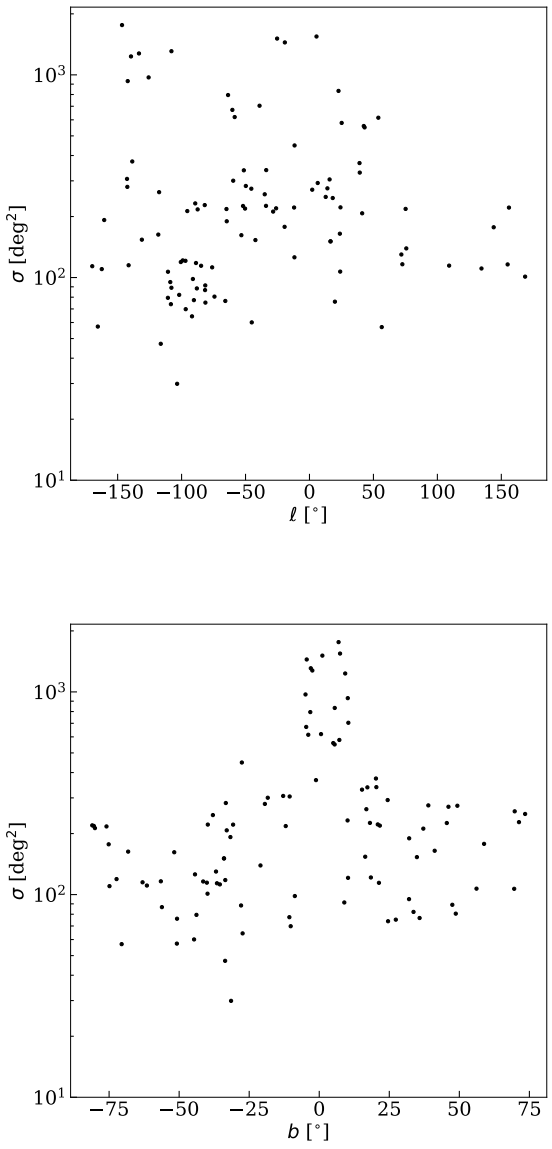


Figure 4.23: Top: model uncertainty area as a function of galactic longitude. Bottom: model uncertainty as a function of galactic latitude. Note logarithmic scales used in both plots.

plot: for intervals where the line of sight intersects a spiral arm, the 1σ area becomes large. These contours are the physical foundation for object candidate searches. They provide a weighted region in the sky out to the event's respective GZK horizon where particles could originate. In the next section I present a study about potential sources for the 102 $E > 50$ EeV Auger events considered in this dissertation.

4.12 Source candidates

This analysis uses the same catalog as the precursor: the half million quasar catalog [142]. Before analysis, a cut is made to select only events within the maximum allowed GZK horizon (for 55 EeV this is about 210 Mpc): objects with $z > 0.05$ are discarded. This leaves a total of 3613 objects. A histogram showing the composition of this cut catalog is in Figure The object keys are as follows

- Q: QSO from literature with broad line unresolved
- A: AGN including extended, Seyfert and low-luminosity
- B: BL Lacertae
- q: High-confidence photometric QSO from SDSS NBCKDE/SDQSO
- K: Type II AGN

which follow the HMQ definitions [142]. These classifications can have further properties listed below

- R: Radio loud association

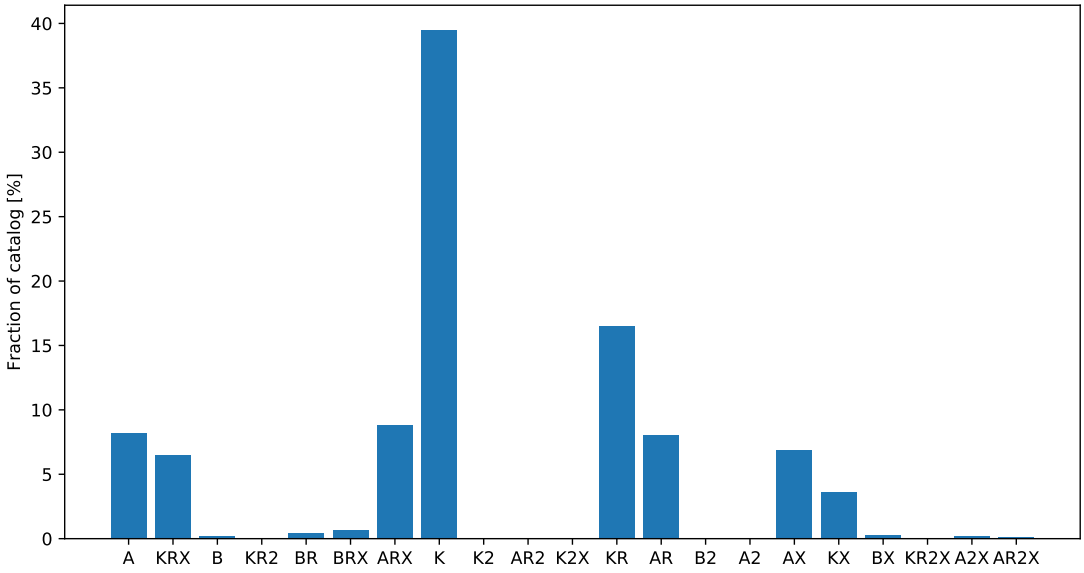


Figure 4.24: Histogram showing the fraction of object classes composing the 3613 objects with $z < 0.05$. *The class codes are defined in Table*

- X: X-ray association
- 2: Double radio lobes

. Therefore the interpretation of ARX in Figure 4.24 would be: X-ray and radio loud active galaxy.

The prescription for identifying source objects is given below:

1. The 1σ contours in galactic coordinates are extracted from the kernel density estimate described in an earlier section. The contour is found numerically.
2. Objects with galactic coordinates embedded in this contour are recorded.
3. The co-moving distance, $d_C\,(z)$ is calculated for each recorded object. I adopt the Planck 2015 ΛCDM cosmology [175].

4. Objects with $d_C(z) < d_{\mathrm{GZK}}$ are kept and others are discarded.

5. These and the 1σ contour are then plot on a flat sky and used for analysis

The entire collection for 102 $E > 50$ EeV events are provided in Appendix C. An analysis of the objects contained in the contours leads to an interesting result. Assuming that the catalog is volume complete (a large caveat, no doubt) and immune to other types of observational biases, naively we would expect the contours to be populated by mostly variants of Type II AGNs and variants "unresolved" AGNs if the contour locations were being generated by a random process. However, we find that BL lacs are associated with 14 of the 102 events. Of the 14, one BL lac object is associated with 4 unique events. Given that the fraction of BL lacs is 1.61% (58/3613) it's extremely interesting that the error contours contain 10 (17%) of them. This might not constitute a "smoking gun" discovery, it is suggestive that BL lacs are plausible sources and might contribute to some clustering, given the magnetic field model uncertainties.

These objects are a special highly variable AGN class which have their relativistic jet aimed in our direction. The jets themselves are poorly understood, but are powered in some way by the central super-massive black hole such as through the Blandford-Znajek process [176] or Penrose mechanism [177]. The extragalactic termination shocks they form are energetically capable of producing 100 EeV particles and meet many of the stringent efficiency requirements for particles to escape [178]. Some example plots of contours associated with BL lacs are shown in Figure 4.25.

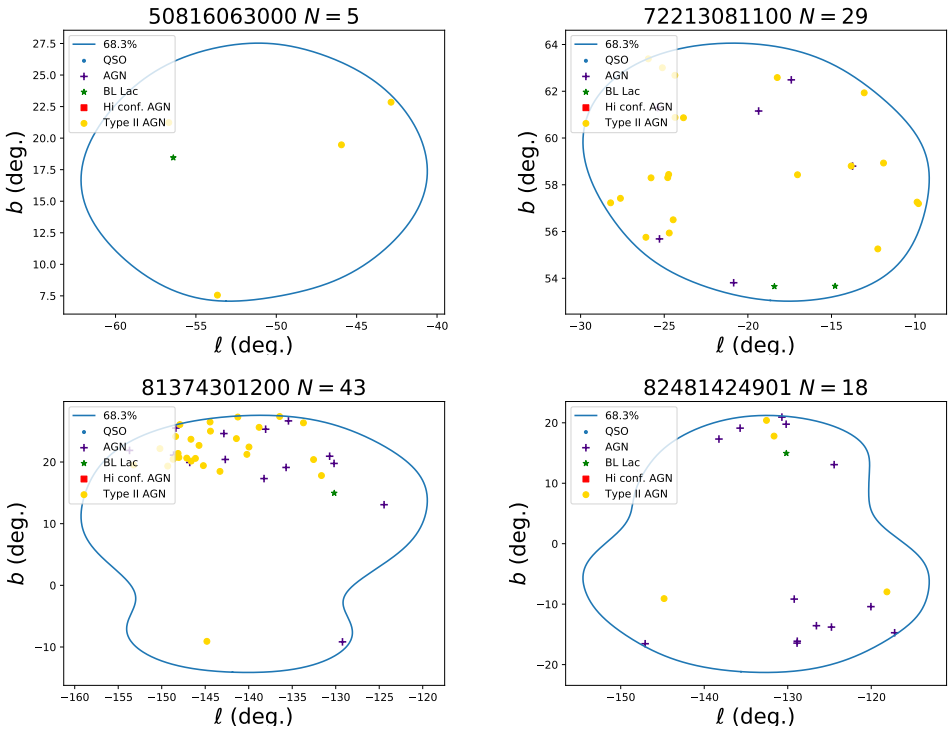


Figure 4.25: Sample of objects associated with BL lacs.

4.13 Anisotropy implications

One of the traditional ways to search for anisotropies in arrival directions is to characterize excess clustering. This usually involves assuming the observed heading is smeared in some way. Examples of this might be using a Gaussian profile with some characteristic angular scale, or drawing a circle around the arrival direction. The purpose is to encode magnetic deflection uncertainties into the clustering analysis, otherwise any signal could simply be an artifact of the deflection physics, rather than source distribution in the sky. The updated sensitivity analysis presented in this section is a major improvement over these ad-hoc methods since a distribution function is provided containing all the uncertainty information associated with the

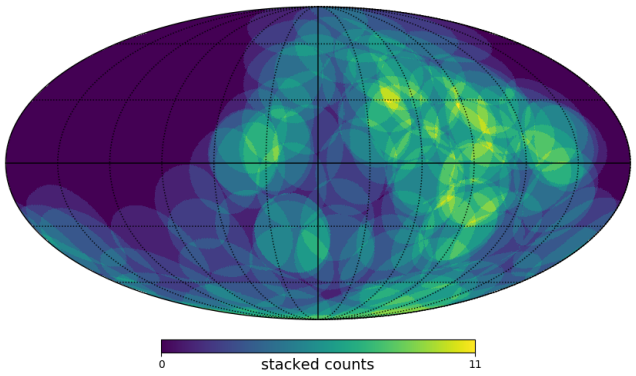


Figure 4.26: Observed arrival directions (at center of circle) with a uniform uncertainty of $r = 20°$ using a Mollweiede projection in galactic coordinates.

deflection which itself is physically motivated by other independent astrophysical observations.

As a test case, I compare the overlap of the probability distributions from this analysis to the overlap using a uniform circular uncertainty. A standard technique used in clustering analysis is testing for Li-Ma significance [179]. In this test, if any counts overlap, the count in the overlap region also adds. An example plot in Figure 4.26 has been created for our data set and I've used a radius of $r = 20°$.

Visually inspecting this figure and specifically looking far large excess compared to the average number of counts, there appears to be considerable signal at longitude -90° between -30° and 30° latitude. Additionally, there is considerable signal in this map near the south galactic pole. This result can be compared directly to the distributions shown in Appendix C.

When constructing the "stacked" map using distributions, I normalize each event using the most probable value. This results in a small region where the probability

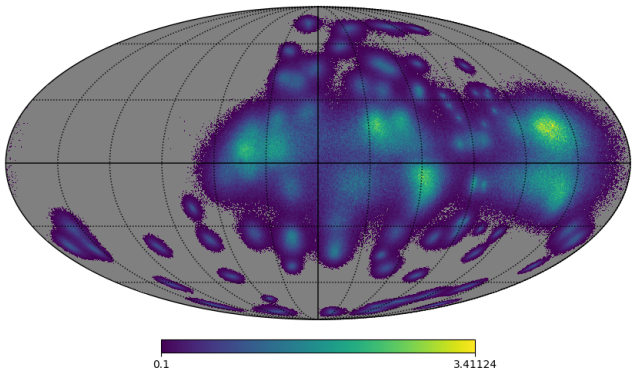


Figure 4.27: Stacked arrival directions using distributions. Note the color scale is not the same as Figure 4.26, if it were, the features would not be visible. The color scale is in units of counts. A Mollweiede projection is used in galactic coordinates.

of finding the deflected particle is 100%, according to the best case scenario for the JF12 model. The rest of the distribution contributes to a fraction of an event in a weighted sense: there is little reason to consider a whole count near the tails if the event is not likely to originate from that direction, unlike the circular method which assigns a probability of 1 in this region. I show the equivalent stacked map using the weighted distributions in Figure 4.27.

The contrast between the two approaches is rather stark. Firstly, there is a significant difference in the dynamic range: the circle method predicts very large signals compared to an average count, while the distribution method signal is not so exaggerated. Secondly, the location of features are completely different. In the area around -90° a large excess is shown, but in fact the uncertainty from the field model is relatively tame there, and it's not appropriate to co-add the events together. A similar situation happens near the galactic south pole where the field model error is

further reduced, and where the predictions can be used with higher confidence. In that region there is relatively little clustering, but the circles yield large signal.

The power in the distribution method is the more reasonable weight assigned to the directional spread of the uncertainty compared to ad-hoc methods.

4.14 Planck data updates and parameter compromise

Since the initial publication of [120] many improved measurements of the galactic environment have been made. Chief among them is the recent synchrotron and dust emission data released by the Planck collaboration in 2015. In an "intermediate results" paper the collaboration provides *preliminary* updates to three popular magnetic field models to make their optimized parameters consistent with Planck data [180]. This work also benefits from an updated relativistic electron model, see [181]. One important caveat is that the `Commander` component separation results in large uncertainties in amplitude and morphology of emission [180]. This approach is subtly different from that in [120] where the analysis was subject to large uncertainties in WMAP component separation in the galactic plane, i.e. resulting in potential large systematic uncertainties in random disk field parameters [180]. The Planck paper parsimoniously adjusts the original JF12 parameters until longitudinal emission profiles align with the Planck data [180]. The authors very plainly state their new parameters are not the best fit and this would require significant future work.

The authors strongly challenge the random component parameter value of JF12, arguing they are too large and likely a result of fitting to anomalous microwave

emission in the WMAP data based on the JF12 model's over-prediction of the galactic variance of synchrotron polarization compared to Planck data [180]. The dominating field strength in a sole spiral arm is also scrutinized [180]. However, the authors do caution that the updated parameters might underestimate the modified random field strength. They state that the original JF12 parameters and the new ones should be treated as lower and upper bounds bracketing the real parameter space [180].

Given this situation I use a compromise between the updates (Jansson 12b in [180]) and originals: Planck values will receive 60% weight and original values 40% weight. The true best fit parameters need not be the midpoint at all, but this is the ansatz adopted. Table 4.2 shows the original, updated and weighted values used in the arrival direction study.

New parameters were used along with an increased turbulent correlation length of $\ell_c = 138$ pc motivated by [180, 182] to backtrack four events, limited by computer time. The analysis procedure described in earlier sections was followed for these events. One event is near the south galactic pole while the other three are near the disk. An example is shown in Figure 4.14. I compare the area of the 1σ contour to the original JF12 parameters in Table 4.3. The rescaling of the parameters according to the compromise weighting results in rather dramatic differences in the model uncertainty. The substantial reduction in the random component strength leads to drastically smaller error contours, especially in the disk region. For the four cases the deflection angle to the distribution MPV is similar. Although this is an extremely small sample size, it indicates that one might expect an order of magnitude difference in model uncertainty for the original vs. compromise JF12 parameters. The compro-

Parameter	Updated value	Original value	Weighted value	Uncertainty
$b_6^{\rm reg}$ [μG]	-3.5	-4.2	3.78	0.5
B_X [μG]	1.8	4.6	2.92	0.3
β	10	1.36	6.544	0.36
$b_1^{\rm rand}$ [μG]	3.12	10.81	6.196	2.33
$b_2^{\rm rand}$ [μG]	6.24	6.96	6.528	1.58
$b_3^{\rm rand}$ [μG]	3.12	9.59	5.708	1.10
$b_4^{\rm rand}$ [μG]	6.24	6.96	6.528	0.87
$b_5^{\rm rand}$ [μG]	3.12	1.96	2.66	1.32
$b_6^{\rm rand}$ [μG]	6.24	16.34	10.28	2.53
$b_7^{\rm rand}$ [μG]	3.12	37.29	16.788	2.39
$b_8^{\rm rand}$ [μG]	6.24	10.35	7.884	4.43
$b_{\rm int}$ [μG]	3.9	7.63	5.392	1.39
B_0 [μG]	7.332	4.68	6.2712	1.39

Table 4.2: Mean parameters used for the arrival direction study. The parameters are assumed to be normal, and the uncertainty is taken as σ. The uncertainty from [120, 128] is adopted.

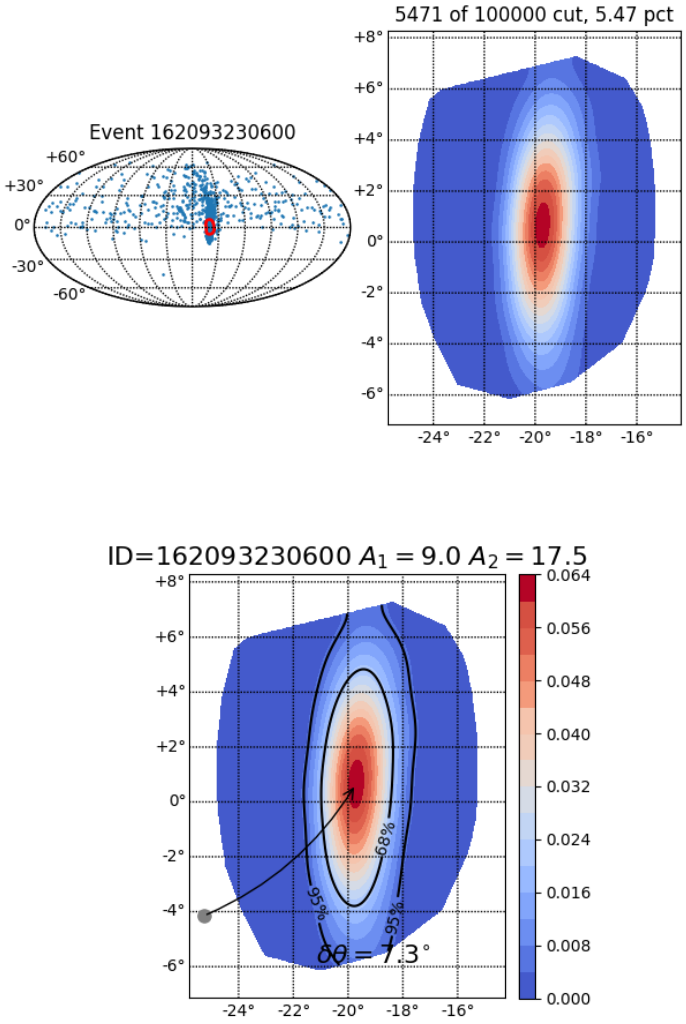


Figure 4.28: Top: Raw 100k backtracks along with the quality cut region. Bottom: Kernel density estimate of the quality cut region and the resulting 1σ and 2σ contours.

Event ID	Og. JF12 1σ [deg^2]	Comp. parameters 1σ [deg^2]	Og. JF12 $\delta\theta$ [deg]	Comp. parameters $\delta\theta$ [deg]
50543756300	212.8	13.2	4.0	4.2
81374301200	1232.9	0.7	11.0	4.0
82481424901	1275.3	0.9	1.7	3.8
162093230600	1445.5	17.5	7.3	7.3

Table 4.3: Comparison of model uncertainty and deflection for a few select events. The updated field strength parameters clearly have a significant impact on the model uncertainty. Original are Og. and compromise are comp.

mise parameters are designed to include information from Planck and WMAP, which likely bracket the true strength of the field components. A 60/40 weighting was used as an ad-hoc starting point. An accurate determination of the weighting, which might also include non-uniform mixtures of the parameters, e.g. 50/50 for b_1 and 80/20 for $b_3^{\rm rand}$ would require extensive simulation work, and re-optimization of the field model using the Planck data. This is well beyond the scope of this dissertation.

The results emphasize that the arrival directions are very sensitive to the field model parameter strengths. Even when the uncertainties on the individual parameters were kept the same, significant reduction in certain field strength parameters had large impacts on the arrival direction distributions. We see that while parameter uncertainty plays an important role, the correct assignment of parameter means is also paramount, especially if the correct mean is *not* covered by a spurious mean and its dispersion. In the next section I consider the effect of heavier primaries on the probability distributions using original parameters and the compromise updates.

4.15 Heavier primaries

This section briefly explores the model response for nitrogen and iron primaries. These are then compared to the proton result. The GZK horizon for heavier elements such as iron is roughly comparable to that of protons, but scaled down due to the increased photodisintegration cross-section of the heavier nucleus [183]. The reduced horizon distance would displace the final resultant vector by 1° or so, which is within the model uncertainty distribution, especially for higher charges.

Due to limited computer time, two events were simulated for nitrogen, and one event for iron. Although this is a tiny sample size, these are considered representative for the overall population. In Figure 4.29.

Nitrogen primaries increase the 1σ contour area by a factor of ≈40 for the two events. This is a significant fraction of the entire sky (41,253 deg^2.). The overall deflection angle increases by ≈10×.

Similar features are found for the iron event, shown in Figure 4.30. These results suggest that primaries significantly heavier than protons make it extremely difficult to isolate sources as the probable arrival directions cover a large portion of the sky using the original JF12 parameters.

We also considered heavier primaries using the "compromise" Planck parameters. Since these significantly reduce the random component field strength the resulting 1σ area has decreased compared to the original JF12 parameters. As before the uncertainty area is highly direction-dependent, but for the four events we see areas of a few hundred deg^2. and in the worst case scenario an area of ≈1000 deg^2. for

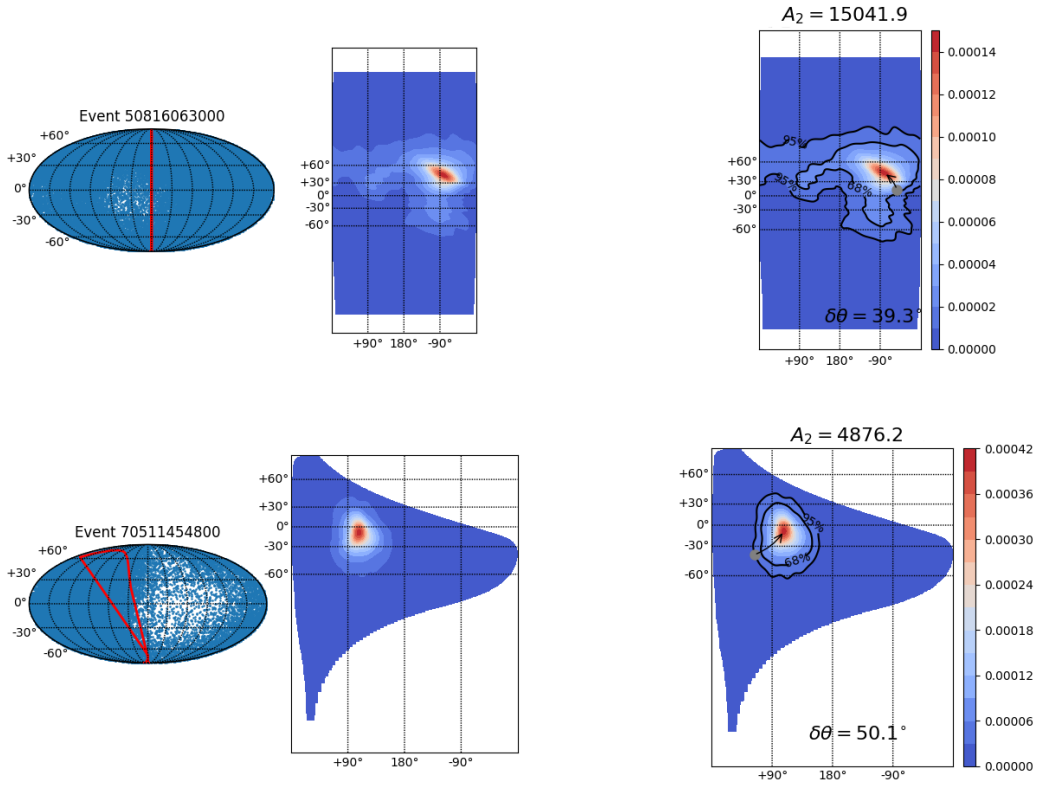


Figure 4.29: Two events where the primary was assumed to be a nitrogen nucleus. The arrival distributions show heavy tails and the distribution core is significantly expanded compared to assuming proton for the same event.

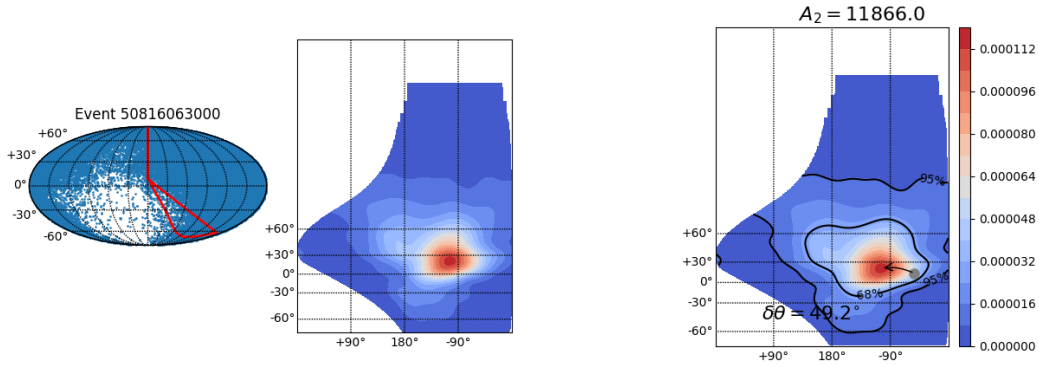


Figure 4.30: An event where iron is assumed as the primary composition. The distribution has a significant tail and the 1σ contour covers a large portion of the sky.

iron. This is a dramatic difference compared to the uncertainties for the original JF12 parameters iron run. Plots for nitrogen are shown in Figure 4.31. Iron plots are shown in Figure 4.32.

4.16 Discussion and conclusion

This chapter detailed a deep dive into cosmic ray arrival directions. Since these are charged particles traversing a magnetized environment, their trajectories will not be rectilinear. Therefore, in order to backtrack the to their source, an accurate description of the intervening magnetic fields is required. I performed an initial study using the regular component of the JF12 model and publicly available tools. To understand the inherent uncertainty of the large parameter space, a Monte-Carlo sensitivity analysis framework was devised. An initial set of Auger arrival directions with $E > 50$ EeV was used in an effort to minimize overall deflection. I found that this uncertainty was larger than the systematic error of observations, and in some directions can be rather severe. In rare cases, the JF12 parameter space allowed for multimodal distributions which could make source identification problematic for these events. With this study completed, I embarked on a more sophisticated approach which used CRPropa3, a more powerful simulation tool, and included all components of the JF12 model: regular, random, striated random.

In this analysis a new set of Auger events with $E > 50$ EeV was used which also passed the 6T5 cut. Like in the first study, I assumed a proton primary, which represents the best case scenario for deflection effects (which scale as Z). This study

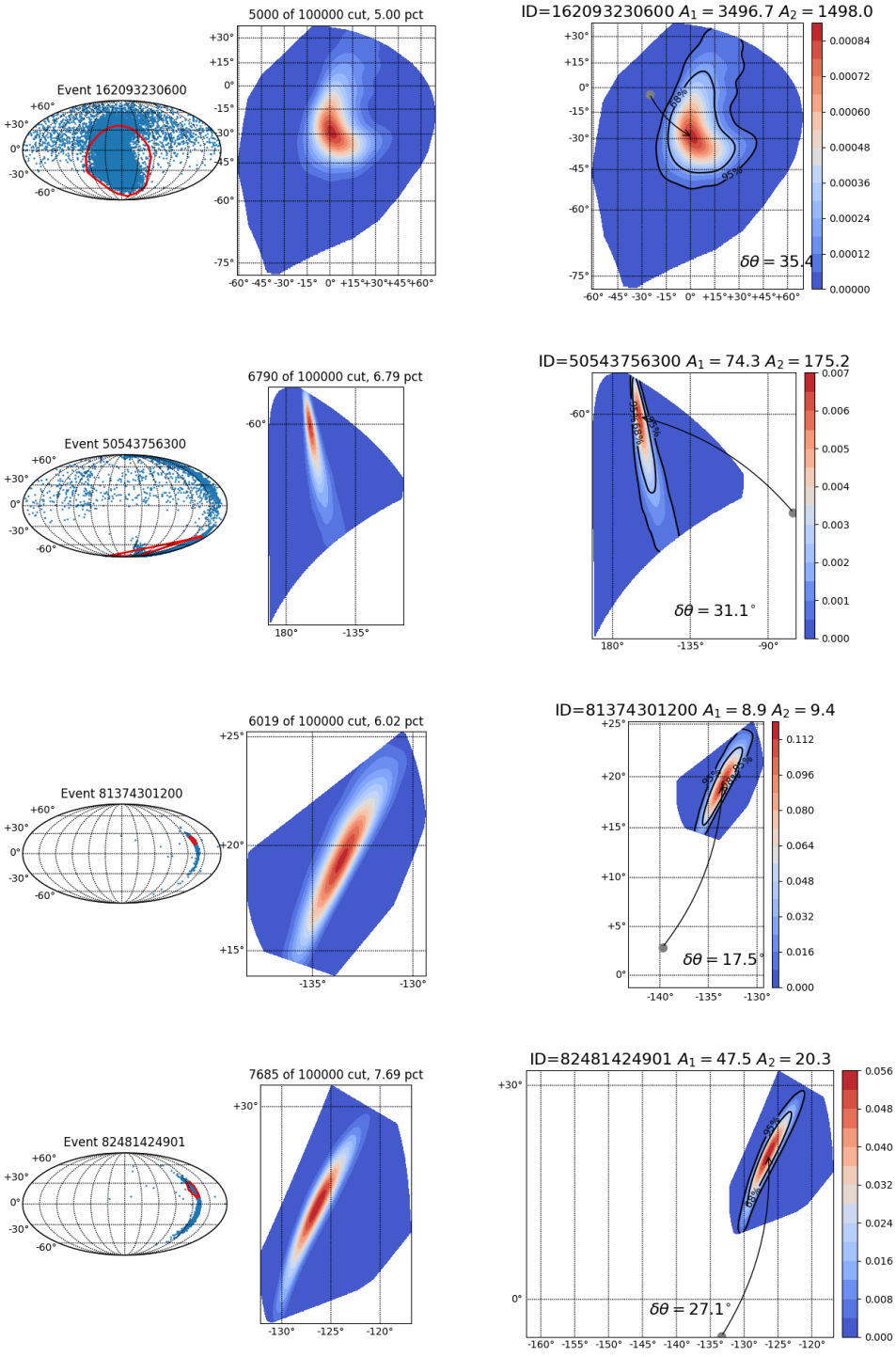


Figure 4.31: Distributions for nitrogen using updated parameters on four selected events.

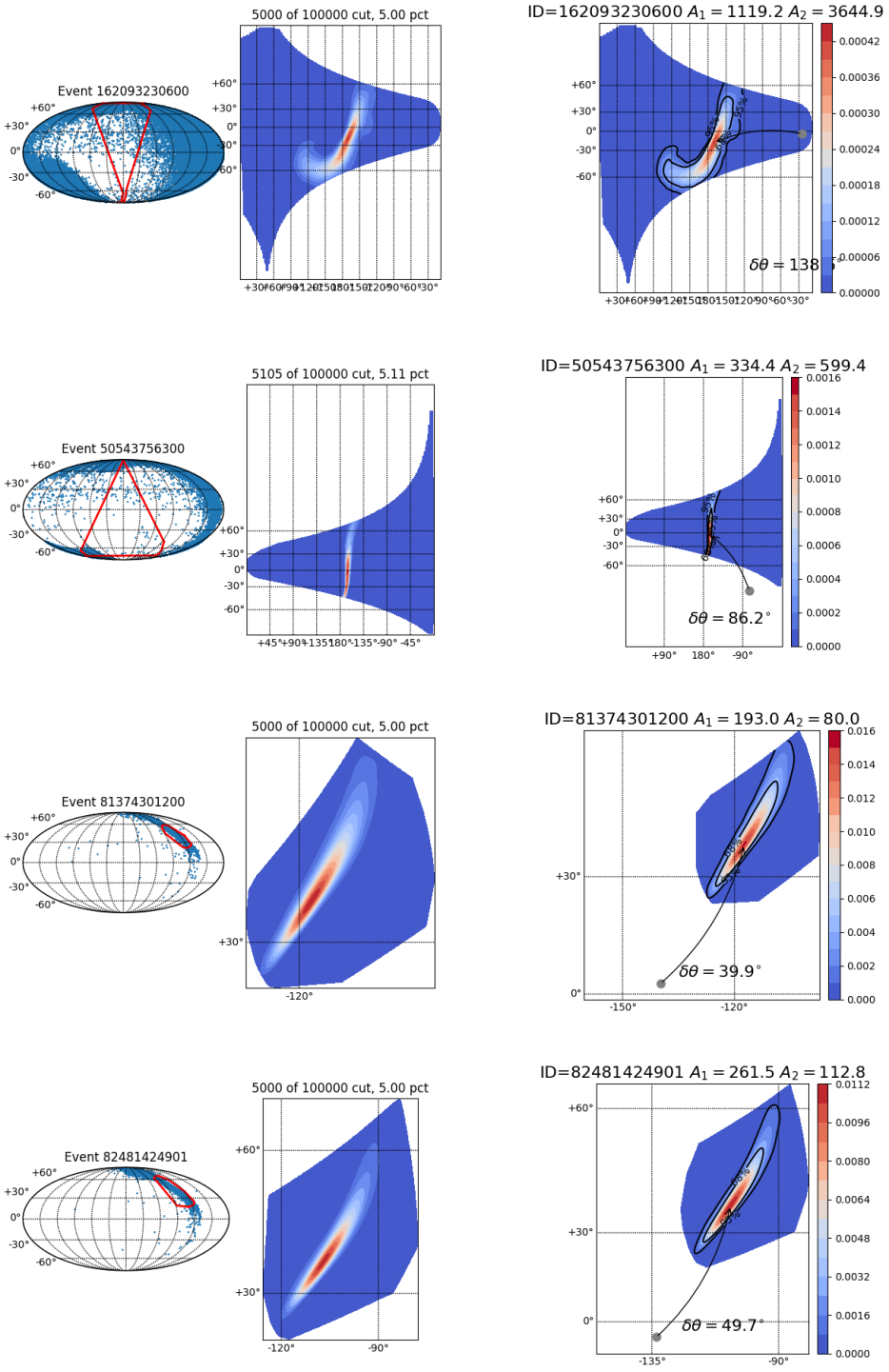


Figure 4.32: Distributions for iron using updated parameters on four selected events.

also accounted for the small effects of extragalactic turbulent fields using an analytic approximation formula. I found that the arrival direction probability areas increased significantly, to the order of hundreds of square degrees. Like the precursor study, there were some events which also exhibit multi-modal features. It was found that the area of the 1σ contours were dependent on direction. The JF12 model is more accurate for larger absolute latitudes and in certain longitudinal intervals.

The 1σ contours were then used to isolate potential sources using a literature complete catalog of quasar objects. An interesting excess of BL lac objects was noted, but this is ultimately *not* a conclusive science result. Finally, the anisotropy implications for this treatment of arrival direction uncertainty were considered. It was found that the distribution predicted signal was strikingly different than the uniform radius approach. Ultimately, this would lead to a difference in clustering significance when performing a complete Li-Ma analysis.

The overall size of contours was reduced substantially when considering the compromise parameters based on the recent Planck results. However, it should be cautioned that the compromise uses a 60/40 weighting and the true weighting could be different. It was merely an attempt to balance the results of JF12 and the recent Planck data.

Heavier primaries were shown to make source identification much more difficult due to larger 1σ contours which likely encompass many objects. For the original JF12 parameters and iron primaries it would be virtually impossible to say anything meaningful about potential sources: the model uncertainty covers large fractions of the sky. If the true magnetic field follows the JF12 model, and its parameters are bracketed

by the WMAP and Planck data, then even with iron primaries it's conceivable that sources could be identified. Model optimization uncertainty is extremely important for heavy composition, and it's reduction is critical for charged particle astronomy to be tenable.

Scripts and code used in this study will be made available at https://github.com/seanpquinn/arrival_dists.

Chapter 5: The Auger@TA Project: introduction and hardware

5.1 Preamble

In this Chapter I present the motivation for pursuing an *in-situ* cross-calibration of Pierre Auger Observatory and Telescope Array surface detector stations and some open questions that we hope to resolve. I discuss the details of the experimental program designed to achieve this goal, the instruments and hardware that were deployed to the field and the supporting infrastructure which enables data collection and analysis.

5.2 Compatibility of Auger and TA results

There are many advantages of combining the data sets of the two experiments. Among these, the most important are the increased statistical power that comes with larger sample sizes for spectrum and composition analysis in addition to full sky coverage which permits more sophisticated anisotropy analyses in terms of multipole expansions and other techniques. As mentioned in Chapter 3, an intensive study has been undertaken to perform a statistically robust compatibility check between the $X_{\rm max}$ distributions which accounts for different analysis techniques and detector re-

sponses *based on simulations* found the reported composition results to be compatible within systematic errors [115, 116]. A similar campaign is underway to understand the differences in the energy spectrum, see Section 3.9.3 of Chapter 3, but to my knowledge an analysis framework similar to the composition studies remains to be developed. It's very likely that this comparison will also use simulation strategies. Therefore both strategies rely on the accuracy and fidelity of simulations since it was not practical to relocate either experiment for a temporary cross-calibration.

However, there are many aspects of actual detectors that simulations can't completely capture, such as aging effects, hardware biases and other unexpected sources of discrepancy. To verify the simulation results and provide additional input for future studies with Auger and TA combined data sets, it's necessary to conduct a direct experimental comparison of detector responses. While this is prohibitively difficult for FD stations since they are large and complex systems, it is achievable for SD stations since these are designed to be placed in the field to begin with.

The Auger@TA project strives to improve on previous studies by performing an empirical comparison of detectors. This has the potential to reveal potential systematic biases in low level signal handling and analysis which can then be accounted for in higher level data comparisons. If no biases are discovered it adds confidence and validation to the results of past comparisons and encourages future studies based on established methodologies. Finally, it addresses a key unknown of the high energy spectrum discrepancy: is this caused by different astrophysical scenarios in the northern and southern sky, or is it a result of energy dependent experimental effects? Since the majority of data is produced by the SD in both experiments, ruling out the ex-

perimental effect would be a major step forward, and that is the primary objective of the Auger@TA program. Examples of such effects would be an unexpected response based on zenith angle dependence, a bias in the sensitivity to secondary air shower particles, especially $e^{\pm}$ and γ (electromagnetic) versus $\mu^{\pm}$ (muonic) components, or biases related to shower age, which is a function of core distance and detection time. Our data might also prove useful for the groups working on the planned Auger Prime upgrade [184].

5.3 Phase I

The project can be broken down into two phases. The first phase, which is the main focus of this dissertation (phase II will be described in a later chapter), involves the installation of an Auger SD station doublet and a TA SD station doublet at the TA central laser facility (CLF). The Auger doublet is comprised of a standard Auger water Cherenkov detector (WCD) that is identical to stations currently operating in Argentina. It's equipped with the same electronics board, PMTs, subscriber unit, etc. It is totally possible to integrate this station into the main Auger array and begin to collect data. I refer to this clone as Cosmo or Auger south (AS). The other member of the Auger doublet is a prototype Auger WCD which was designed and constructed for the northern hemisphere Pierre Auger Observatory. This extension was ultimately not funded and subsequently canceled, however, several prototypes were constructed and deployed into a research and development array (RDA) near Lamar, Colorado, see [185] for more details. I refer to this prototype WCD as Auger

north (AN).

The Auger north WCD has the same physical dimensions for the detector volume as AS. It uses a single, central, downward facing 9" PMT compared to the three symmetrically distributed AS design. The electronics board is similar to AS but has a smaller form factor since one PMT is digitized. An important difference is a new 10-bit 100 MHz FADC which handles four channels to improve the dynamic range. A newer lightweight Linux based operating system (Debian) is used which has many convenient applications compared to OS9000 on AS. The anode signal is split into $0.1\times$, $1\times$, and $30x$ channels and instead of the 8th dynode (used by AS) the 5th is used. The Auger doublet is separated by roughly 7.6 m center-to-center and is located roughly 20 m from the CLF structure, in the north-east corner of the enclosed CLF zone.

The TA doublet is comprised of one standard TA station which has been instrumented with electronics, PMTs and a parabolic radio antenna for long range communications. Unlike other stations in the array grid, the broadcast L1 triggers from this station are ignored by DAQ computers, so it's not a member of the grid which forms TA physics triggers. It operates in a standard data taking mode, constantly sending L1 trigger lists (which are ignored) and also listening for data requests. This station is located roughly 24 m from the CLF, in the south-west corner of the CLF zone. It's roughly 44 m from the Auger doublet. This station is called DET2421. The second TA station is equipped with PMTs and a battery box, but no electronics. Data is collected from this station differently, and is explained in detail in a later section. It's located about 18 m from the CLF, in the north-west of the CLF zone. It's roughly

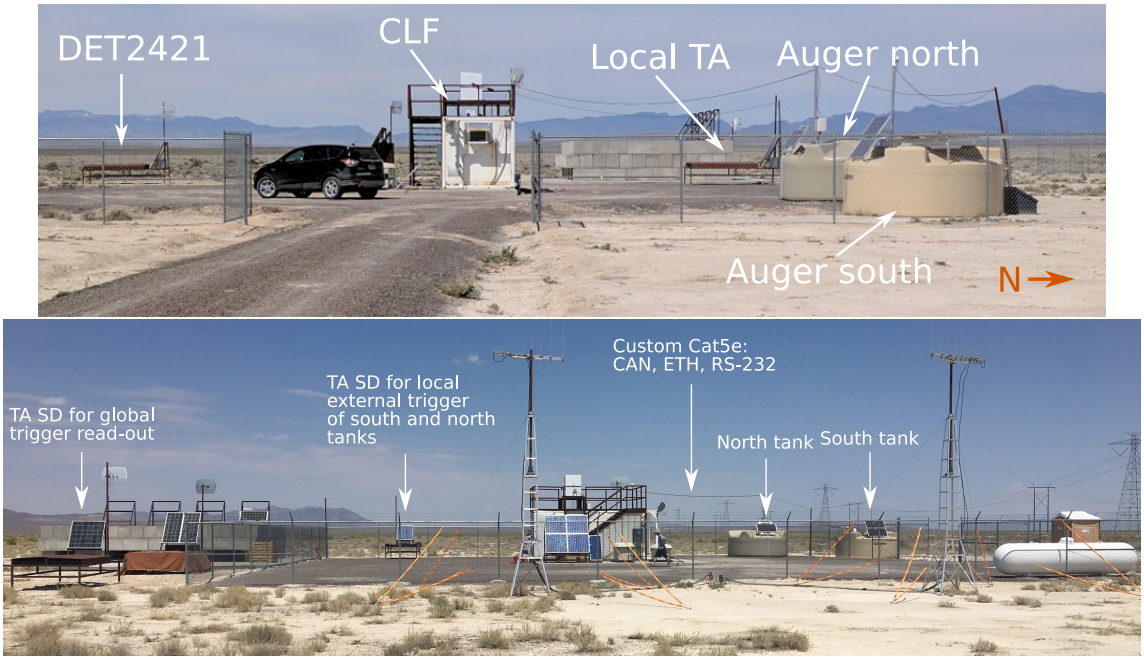


Figure 5.1: Top: Photo of the Auger@TA setup during summer 2017, north is to the right of the page. Bottom:Photo of the setup during the spring of 2016, north is into the page.

29 m from the Auger doublet, and is referred to as "TA local." Two photos of the setup taken at different angles are shown in Figure 5.1.

There are two critical goals of this phase:

1. Build a cross-calibration curve based on data collected with AS and DET2421 to understand the signal responses of the WCD and scintillator detectors for a large sample of showers which fire the TA physics trigger. Additional analysis can be done with this signals using shower reconstruction parameters which are available for a subset of events which pass additional quality cuts outlined in [102].

2. Verify that the AN and AS signal response are compatible and tightly correlated. This is to validate the design and functionality of the AN detector, but to also

account for any possible discrepancies by constructing a mapping function to translate between AS and AN. Completion of this step is needed before we proceed with phase II.

5.4 Auger@TA project infrastructure

To build a cross-calibration curve based on Auger and TA signals the detectors must be read out for the same extensive air shower (EAS). The AS and AN stations are designed to operate wirelessly and the TA physics trigger information cannot be simply transmitted over their local area network (LAN). This creates several challenges which need to be overcome. For the first point, we have supplanted the wireless communications system with physical wires. The AS is designed to send and receive digital data from the main board or "local station controller" (LSC) to the SU module. The SU then pipes this data into a radio modem and is transmitted. We have bypassed the SU and pipe the LSC serial output to a DB-9↔CAT5 interconnect coupled to about 40 m of CAT5 cable. This is secured to a guy wire which runs from the AS mast into the CLF structure. This is the only modification made to the AS detector.

The AN station is modified in a similar way. Instead of a serial bitstream, starting from the LSC, the CAN protocol is used, which is routed to the tank power control board (TPCB) housed in a weatherproof enclosure. An CAT5 cable is then connected to the "local radio" (LR) which transmits data to the antenna. We bypass the LR and instead send the CAN bitstream over a 40 m CAT5 cable into the CLF structure. Two

additional CAT5 cables are brought up from the LSC into the weather enclosure: one is for Ethernet networking and the other provides access to a Linux terminal trough serial. These are also coupled to 40 m CAT5 cables brought into the CLF structure. The guy wire therefore carries four cables total from the doublet into the CLF.

As mentioned in the previous section, since DET2421 is operated in normal data acquisition (DAQ) mode, traces and timing information is retrieved in the normal way over radio communications, and then extracted in post analysis. No modifications are necessary on this station.

Next, we need a way to collect trace data from the Auger doublet when an EAS occurs at or near the CLF. This is accomplished by using a TA electronics kit (e-kit) as a radio receiver. A parabolic antenna is mounted on the CLF and directed to the Smelter Knolls communications tower, allowing it to receive physics trigger messages (L2 triggers). Normally these messages are parsed internally, but we need a way to access and read the time stamps. Our TA colleagues were extremely accommodating, and Toshiyuki Nonaka-san modified the firmware to forward the universal coordinated time (UTC) L2 triggers over a serial line in ASCII format. This is connected to a TRENDnet TU-S9 (Prolific PL2303 chipset) serial to USB converter which is attached to an Auger single board computer (SBC).

5.5 Auger SBC

The Auger SBC was deployed in three incarnations. First we tried using a Technologic Systems TS-7250v2 for its industrial design and CANbus support during the

initial deployment. This setup operated for about 6 months and was replaced in spring 2016 with a new SBC: a Raspberry Pi Model B. This has proved as reliable as the TS-7250v2, and offers other benefits such as reduced power consumption and easier CANbus configuration. An upgraded version (Raspberry Pi 2 Model B) was installed in the summer of 2016 which included increased storage and a more sophisticated power supply. The import specifications and peripherals are listed below:

- 3× serial to USB converters
- 128 GB USB flash drive
- 1× CANbus adapter board (Canberry v2.0)
- Electromagnetic interference (EMI) protection circuit
- 12 V to 5 V DCDC converter (CUI Inc. PYB10-Q24-S5-T)

A photo of the second incarnation is provided in Figure 5.2.

5.6 Auger south global trigger software stack

The software used to gather and parse data can be broken down into two levels:

1. Radio and central data acquisition emulators
2. Data stream monitoring and trigger decision

For the AS station a `C` program which handles the Auger custom radio protocol, T2 lists and command inputs was provided by Ricardo Sato and is available at

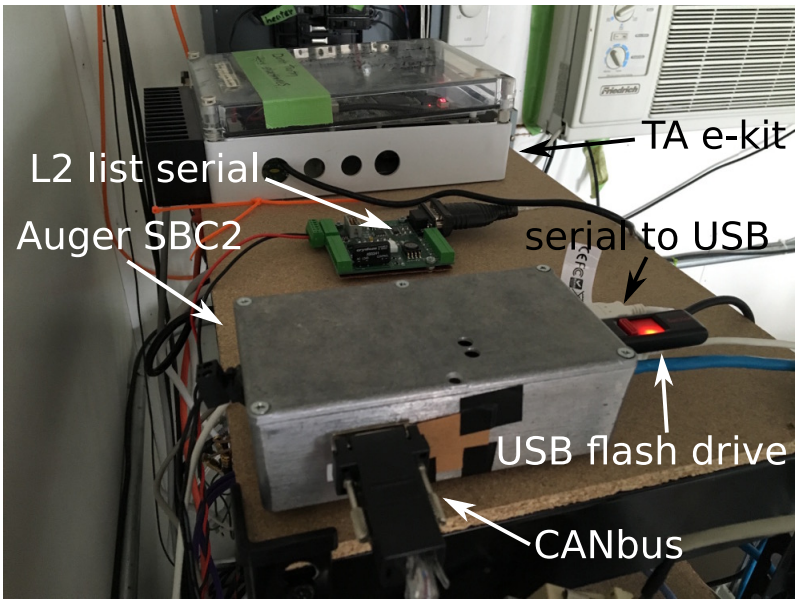


Figure 5.2: The second version of the Auger SBC with attached peripherals next to the TA e-kit which relays the physics trigger list.

https://imogen.phys.cwru.edu/auger_ta_svn/rsato_su_emu/. I made a small modification to the file writing behavior for the T2 list: it is opened in unbuffered append mode to avoid overwriting previous data and to have access in real time. Starting and stopping data collection is fairly straight forward. The following procedures can be used from a cold start.

1. Mount the USB flash drive. Verify the device first, usually `/dev/sda1`, then mount using: `sudo mount /dev/sda1 data/`

2. Identify the Auger south serial device. The most recent settings are listed in the `tty_devices.txt` file in the home directory. However, if the system has been rebooted it will be necessary to identify these manually using the `/var/log/kern.log` file, which will list the `tty` device assignments[1]. Search

[1]Since Auger south and the local trigger use the same chipset it is not possible to tell these

this file for "prolific" as this is the chipset used by the converter.

3. Navigate to `/home/pi/data/new_sbc`

4. Start the `cdas_su_emu` server using: `./cdas_su_emu /dev/ttyUSB2 /dev/ttyUSB2 3000 >> su_log.txt &` where `ttyUSB2` should be replaced with the correct ID.

5. Begin triggering. Issue the following command: `./cl s`

6. A new file called `T2_list.out` will appear if the previous command was successful, and data will be written in real time.

7. Execute the parsing program to express `T2_list.out` entries as GPS seconds: `python parse_auger_south.py > south_parse.log 2>&1 &`

8. Execute global trigger and AS time stamp coincidence search program: `python tag_as_trigger_dec2016.py > TAG.log 2>&1 &`

A diagram representing this process is shown in Figure 5.3.

5.7 Global trigger decision

The global trigger (synonymous with TA physics trigger) is a key component of the first phase, so I will present further details about the architecture of the

apart without taking a look at the datastream. I generally do this using `picocom -b 115200 /dev/ttyUSB1` if I think `ttyUSB1` is the local trigger. If you can see readable text scrolling by with GPS timestamps, that means `ttyUSB1` is the local trigger and the other Prolific device is the global trigger

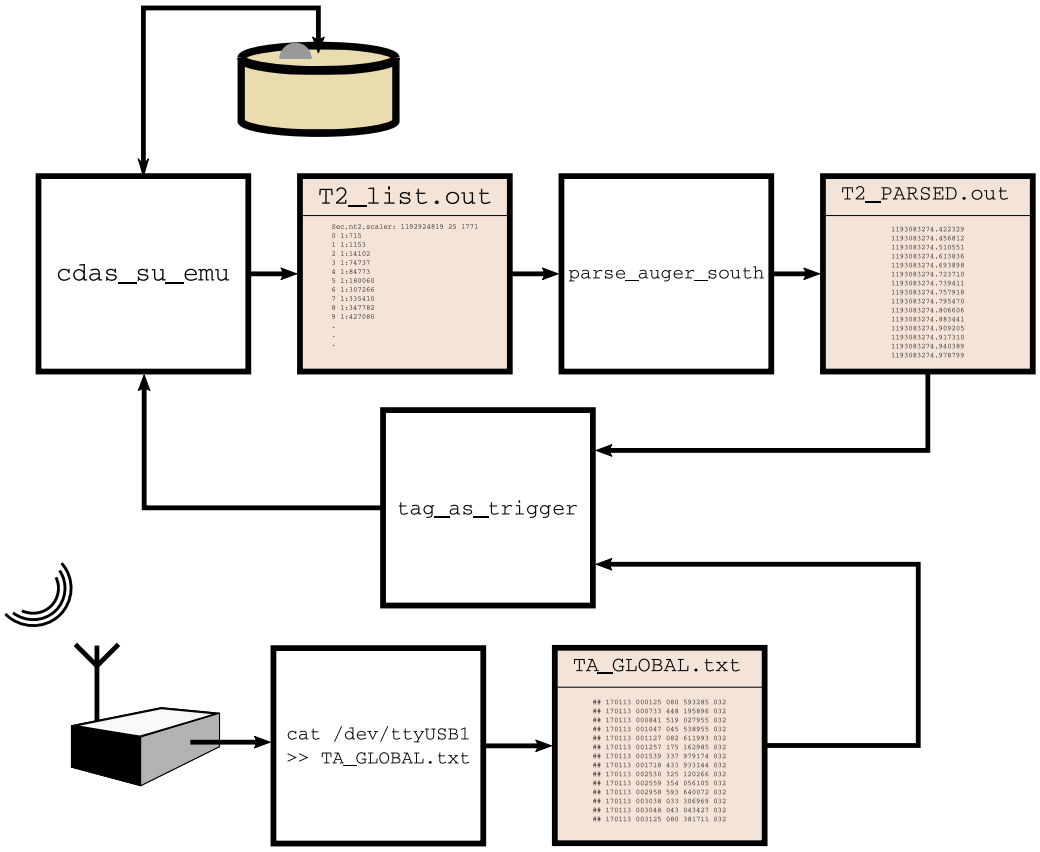


Figure 5.3: White boxes represent programs and dark tan boxes are files on disk. Arrows depict the communication direction. `cdas_su_emu` *is a binary compiled from* `C` *source. The user can interact with this server using the* `cl` *program (also a compiled binary) to send commands.* `parse_auger_south` *and* `tag_as_trigger` *are Python 2 programs.*

`tag_as_trigger` program. The design philosophy is rather simple: *look for all coincidences of TA physics triggers (L2) and AS T2 triggers in a window of 100 microseconds*. The L2 timestamps are given in UTC time and must be converted to GPS time in order for `tag_as_trigger` to perform the coincidencing. The `gpsFromUTC` function from the LIGO `glue` library is used. It can be accessed at http://software.ligo.org/docs/glue/glue.gpstime-pysrc.html#gpsFromUTC. The correct leap year must be manually specified by the user. As of October 26, 2017 there are 18 leapseconds for GPS compared to UTC, see http://tycho.usno.navy.mil/leapsec.html and https://www.iers.org/SharedDocs/News/EN/BulletinC.html.

The program also uses the `mpmath` Python library for GPS time arithmetic. This is required since GPS seconds with microsecond or sub-microsecond precision can't be represented as a 64-bit floating point number. For all arithmetic operations (mostly subtraction) time stamps are instead converted to a $22^{3.33} = 29833$-bit floating point equivalent. This ensures that no rounding errors contaminate any arithmetic. A subprocess is started using the the GNU `tail` utility, which reads the unbuffered file contents of `TA_GLOBAL.txt`, and its output is piped into a Python subprocess data structure. The main body of the program occurs in the `while true` loop. The first line of the loop uses a `readline` member function which is *blocking*. This means if no new data is written to the global trigger file, the loop doesn't iterate. This conserves a large amount of CPU cycles since the global trigger period is about 10 minutes (≈ 2 mHz).

When a global event occurs the subprocess object outputs a string. The program then sleeps for 2 seconds, to compensate in any possible delay due to the slow serial

line. After this a new subprocess starts which reads the last 120 lines from the parsed T2 file using the `tail` utility. This is transformed to a Python list of formatted strings. This approach guarantees that any possible Auger T2 is comfortably between the start and end points of the recent T2 history. Next additional testing and error handling is performed on the global trigger string, a conversion is made from UTC to GPS seconds, and it's cast as a string. Next, the absolute difference is calculated between the TA global timestamp and every element of the 120 length T2 list with 29833-bit precision. If a difference is found to be < 100 μs the *original* TA global trigger and parsed Auger T2 strings are appended to a file, along with a string of the time difference. A new subprocess is spawned which sends a T3 request to the AS station: `./cl T 1193083274 422329 30` where the first integer is the GPS second, the second integer is the microsecond (following the decimal place), and 30 is used as the window (in microseconds). This instructs the AS station to search its FADC memory buffer and if a trace with a timestamp within the window is found, it's returned over the serial line as an ASCII hex dump.

The choice of 100 μs is to ensure no events are missed. In terms of light travel distance, this would encompass an area of 30 km which one could argue is too large of a window. On the other hand, we also want to account for any possible GPS clock or other electronics offsets, which could be on the order of a few microseconds. Although it is a large window, most of the time if the shower occurs elsewhere in the TA array there will be no resulting coincidence with an Auger T2, and so the loop will continue to its initial waiting state.

5.8 Auger north software stack

Unlike the AS station which is incapable of storing data locally on disk, the AN station was built with a significantly more powerful processor and versatile OS. With the expert help of Laurent Guglielmi, a software engineer involved in AN project, programs were developed enabling storing every T2 level FADC trace. In this mode the detector will generate about 20 GB of data per day. It's extremely useful for trouble shooting and debugging, and is also a viable data collection mode. Custom software was written to accommodate this mode, and is described in a later subsection.

Like the AS station, AN can transmit data and receive commands using its radio. The CAN protocol is used for this communication. In this mode, I've implemented programs similar to the ones described in Section 5.7 which handle the real time DAQ. These are described in the next subsection.

5.8.1 Local storage mode

The station can be configured to save data to an external USB drive (usually 500 GB or 1 TB). Connect to the LSC on the CLF LAN and execute the following commands

1. Start the acquisition server: `das -v start` Exit code 0 means successful start
2. Tell the LSC to save data to disk: `ssdstart`
3. Start triggering: `stop -12345 control`

Binary data is saved to `/data/Events`, `/data/Monitor`, `/data/Muons`. In these directories a new folder is generated each day (UTC) in the format `YYYYMMDD`. A shell script was written to automatically transfer monitoring, calibration data (files in the Muons folder) and T2 timestamps each day. File are staged on `imogen`. The previous day's list of local and global trigger timestamps is transferred from the SBC using an `rsync` cronjob to `clarke`, a desktop machine on our lab's network, via `imogen`. The AN T2 lists are parsed into text via `testevt` program written by Laurent Guglielmi and compared to the local and global trigger lists. When a coincidence is found within 20 (100 for global) μs the AN T2 event index appended to a list. The list is then converted to a Bash shell script and transferred to the AN station. Half an hour later a cronjob runs this script extracting the specific FADC traces from the large data files containing roughly 20 minutes worth of traces. Individual trace files are relatively small (few kB) and easily transferred back to `imogen`after extraction. This technique is a large improvement over previous methods where the entire `/data/Events` directory would be transferred, requiring about 12 hours and significant bandwidth. After transfer from AN to `imogen`the binary FADC files are sent to `clarke`for analysis. A simplified network topography diagram is shown in Figure 5.4. Both trigger type data is archived on `imogen`and `clarke`and backed up on a daily rotation to a CWRU Google Drive account.

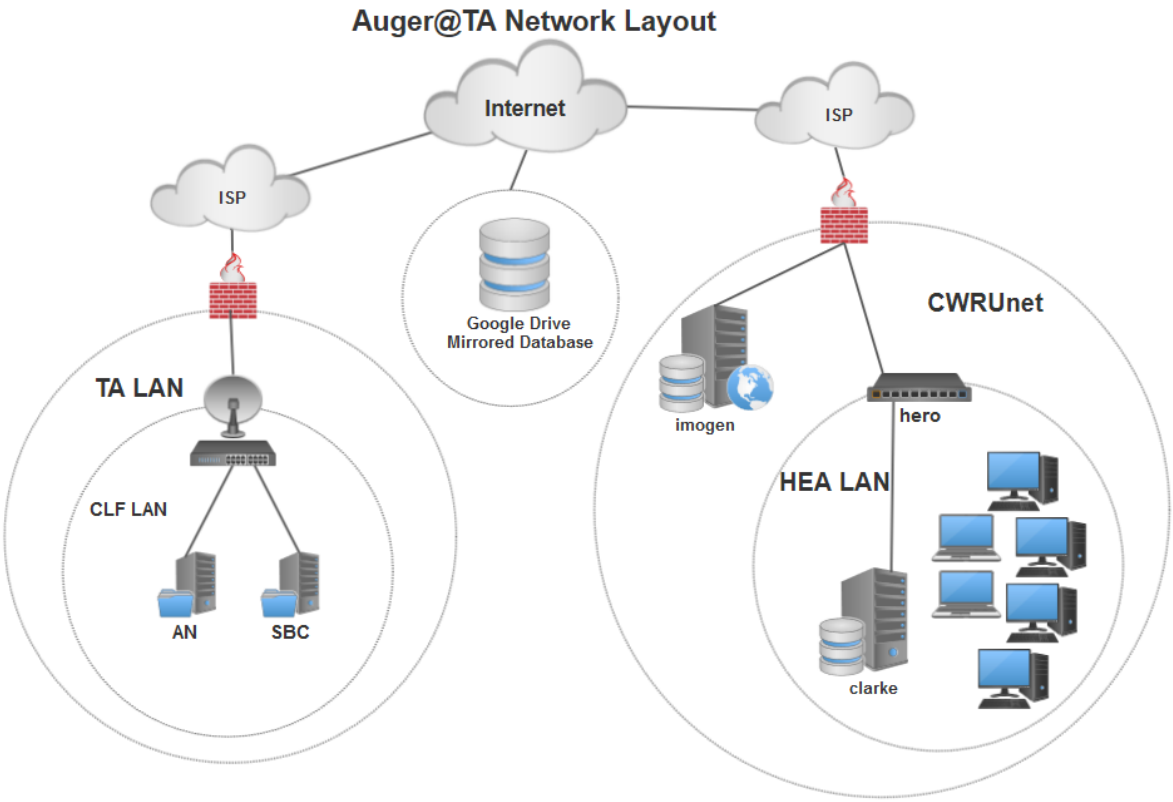


Figure 5.4: A high level overview of the network configuration. Machines directly involved in data handling and transfer are connected by lines. Dashed circles depict various subnets. The AN detector and SBC are data sources and share a long range radio link to the main TA network. The primary handler and staging area for data is `imogen`*, and local trigger analysis jobs are delegated to* `clarke`*.*

5.8.2 CANbus mode

To enable the radio link from a cold start follow the same steps as local storage mode, but without running `ssdstart`:

1. Start the acquisition server: `das -v start` Exit code 0 means successful start
2. Start triggering: `stop -12345 control`

To stop DAQ from a previously running local storage state (i.e. the `das` server is already up), use the following sequence:

1. Stop triggering: `stop -12346 control`
2. Stop storage: `ssdstop`
3. Start triggering: `stop -12345 control`

When triggering starts, a T2 stream is sent following every pulse per second (PPS) from the GPS clock. Every few minutes monitoring data is also sent. The software package which handles the CAN protocol and emulation services is `LSC_das`written by Lauren Guglielmi. It's a collection of `C` programs very similar to `cdas_su_emu`, but for AN. The SBC is configured to send and receive CANbus data by adding the following entries to `/boot/config.txt`

```
dtparam=spi=on

dtoverlay=mcp2515-can0,oscillator=16000000,interrupt=25

dtoverlay=spi-bcm2835-overlay
```

The SBC operating in the field as of October 29, 2017 already has these modifications. Assuming the CAN peripheral is attached, the link can be brought online by running `sudo /sbin/ip link set can0 up type can bitrate 125000` or executing `open_can.sh` in the home folder. The following sequence of commands will start the `LSC_das`package:

1. Navigate to `/home/pi/data/new_sbc/north`
2. Initialize the radio protocol process: `./csradio > csr.out &`
3. Initialize concentrator: `./cs > cs.out &`
4. Initialize postmaster and parsing program: `python parse_auger_north.py &`

If the programs return successfully the `AN_T2.txt` file will be updated with live T2 events. To send T3 requests to AN using the `LSC_das`package, uncomment line 116 in `tag_as_trigger_dec2016.py` before running that program. Timing offsets between AS and AN stations were found to be on the order of 5-10 μs. As a result it isn't necessary to monitor the AN T2 in addition to the AS T2, since a window of 30 μs is used for T3 requests to both stations. It was discovered in the field that sending a remote T3 request to the LSC would crash the `das` server resulting from insufficient memory allocation for the FADC trace. This problem was fixed in consultation with Laurent Gugliemli and T3s can be successfully transmitted over the wire.

5.8.3 LSC configuration

A nice feature of the AN station is that the configuration can be changed remotely. The command to change settings is `acqconfig`. All settings can be viewed using `acqconfig -L`. Note that `das` and triggering should be stopped when making changes. To change the PMT HV for example use this command: `acqconfig v1=1400`. This sets the HV to 1400 V. As of November, 2017 the VEM threshold is 50 and the HV setting is 1600 V. Running `saveconfig` will save the files currently loaded in the ram disk into the home directory of the LSC. More details on commands and setting up the station can be found in the lsc˙userman.pdf file.

5.9 Local trigger hardware

The local trigger was created to address the second critical issue for phase I: verification of equivalent performance of AN and AS stations. While it's possible to use the global trigger system for this study it will take a long time to generate large statistics for the comparison due to the low frequency of TA physics triggers. It was decided that we would devise a separate trigger exposed to lower energy showers to rapidly build a large sample. The northwest "local" TA station described in an earlier section was installed for this purpose. Since an e-kit was not available we also devised a primitive trigger based on a threshold comparator circuit assembled by CWRU electrical engineer Robert Sobin, a schematic is available in Appendix D. The threshold is set to -90 mV using a trimmer pot to yield a 2-fold coincidence rate of 2-5 Hz. The PMT reference voltage is nominally 1.2 V (set with a potentiometer) which

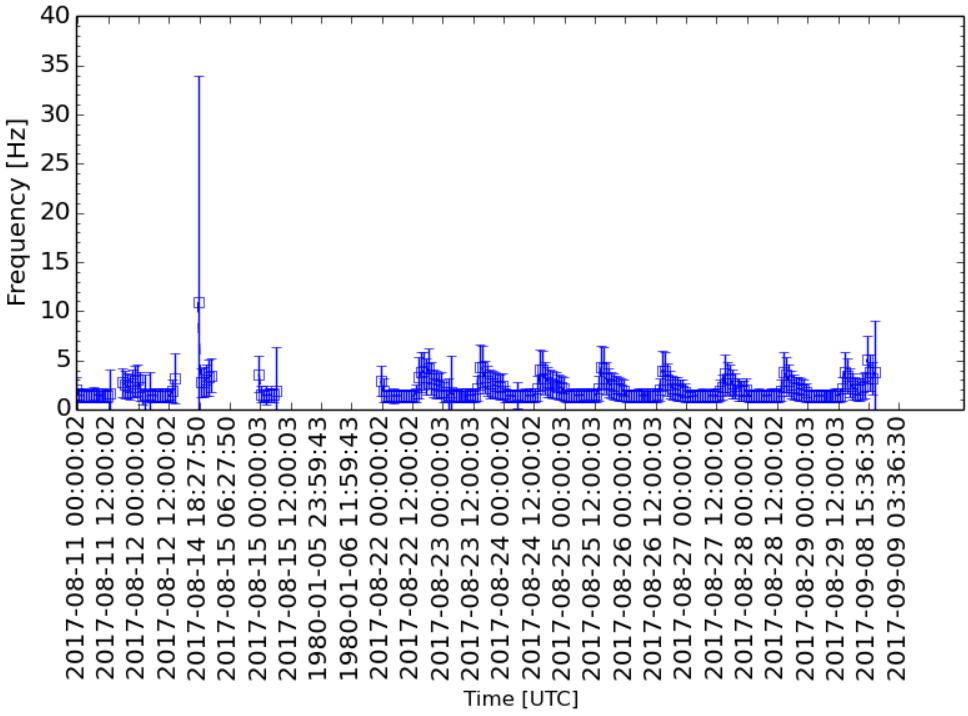


Figure 5.5: Local trigger rates over 12 hour intervals. The mean is used and error bars are one standard deviation. A slight periodic trend is observed corresponding to the day-night cycle.

is converted to an HV value of 1200 V. An example of recent rate data can be found in Figure 5.5. Discriminators are connected to both TA PMTs and their output is sent to an AND gate (Texas Instruments SN74HC08N). The output for this gate is then sent to a 50 Ω NAND line driver (Texas Instruments SN74S140N) and then propagated over the 40 m of RG58 coax into the CLF. The coax is terminated with at 50 Ω tee and the signal is passed into an octal bus (Texas Instruments SN74LVC245AN) which translates the voltage to 3.3 V. This is then sent to a MicroZed development board with carrier card to be time stamped relative to the GPS reference clock. The time tagging firmware was graciously provided by Robert Halliday, which required porting the code from a conventional Zed board. The firmware provides an easy to use ASCII

encoded bitstream over a serial line, which is connected to the SBC via an adapter. The MicroZed is capable of timetagging pulses at rates from 0-30 Hz. Higher rates are limited by deadtime due to parsing of certain GPS receiver message. This can be lifted by disabling that message in the GPS receiver, but this would require a major software modification in the local trigger parsing program.

5.10 Local trigger software stack

The local trigger programs are very similar to their global trigger counterparts. First a parsing program is started which converts the ASCII timing data from the MicroZed to GPS second timestamps and a second program searches for coincidences. The MicroZed sends continuous PPS messages with the UTC time and position information. A subprocess opens the serial device using the `cat` utility in unbuffered mode. This data structure is used to retrieve strings in blocking mode in the main loop. The position information is ignored and a UTC date is formed from this string. The same LIGO `gpsFromUTC` function is used to convert this date to GPS seconds. The program is listening for messages containing the `TESTevent` string followed by an integer giving the number of counts after the PPS for a 32-bit 750 MHz counter. These are arrivals of logic pulses from the local trigger hardware. Each counter value is converted to a GPS second with microsecond precision and appended to the `TA_LOCAL.txt` file.

The local trigger script watches for time coincidences between AS and the local trigger. Like the global trigger program, the `tail` utility is called as a blocking Python subprocess object and reads the `TA_LOCAL.txt` in the main loop. When a

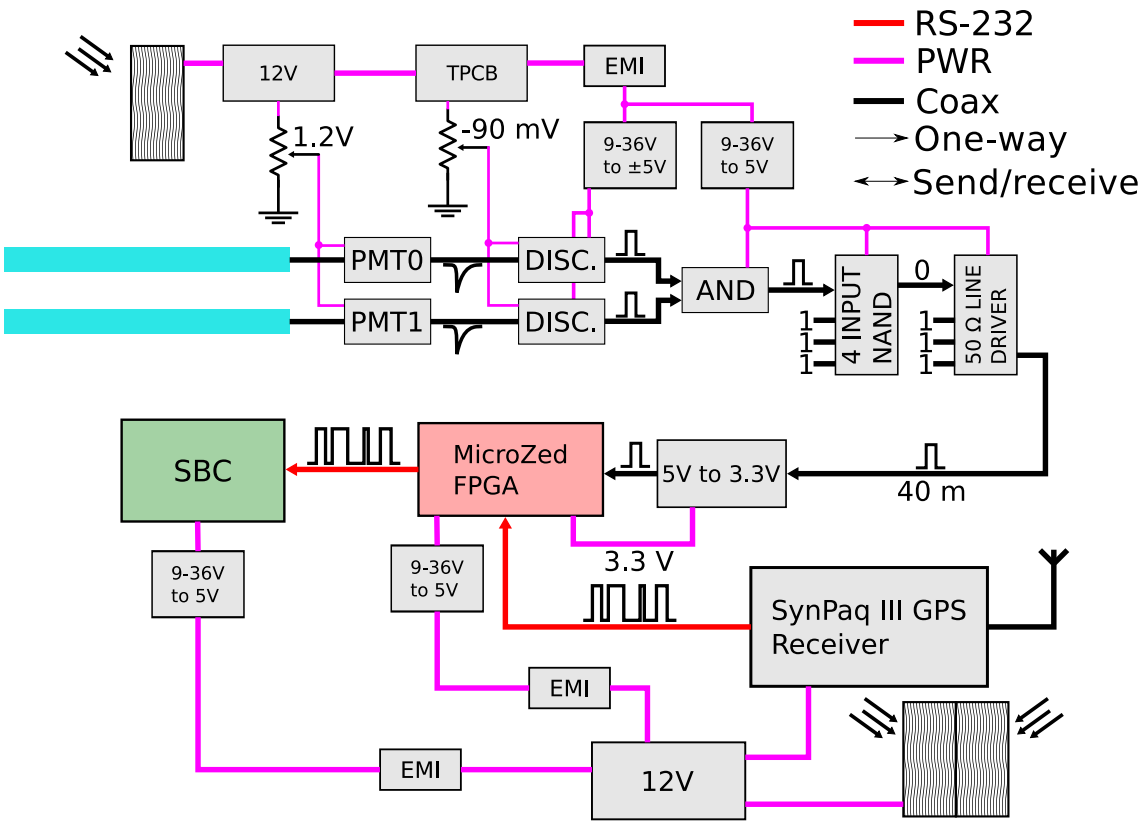


Figure 5.6: Key components of the local trigger design showing a mix of circuit elements and high level device blocks. The TA scintillators are depicted as aqua slabs. Components connected to the TA detector preceding the 40 m coax cable are housed in a weather proof battery box. Components following the 40 m cable are installed in separate enclosures inside the CLF structure. The local trigger discriminators and gates are powered by a TA solar panel (125 W) and 12 V battery. The MicroZed and its supporting electronics are powered by two Auger north (80 W) solar panels and 3 12V batteries.

new local trigger is received the program sleeps for 1 s and then reads the last 70 lines of the `T2_PARSED.out` file. This is done to account for any delays and to ensure any coincidence is padded from edges. The absolute difference for the new local timestamp and all 70 AS T2s is computed using 29833-bit precision. If a $|\Delta t| < 30$ μs is found a T3 request is sent to AS and AN using the AS T2 timestamp. The local trigger, AS and time difference is then saved to `TAL_AS_COINCIDENCE.txt`. The parsing script and trigger decision code is provided in Appendix E.

To begin local triggering from a cold start follow this sequence:

1. Start the parsing program `python parse_local_ta.py > PARSE_TAL.log 2>&1 &`. Verify the device ID on line 68.

2. Start the local trigger searching program: `python tal_as_trigger_dec2016.py > TAL.log 2>&1 &`

Chapter 6: The Auger@TA Project: data analysis and results

6.1 Preamble

In this chapter I discuss the methods used for unpacking the raw data for AS and AN. The procedure for finding the calibration values are discussed as well as trace construction and integration. Supporting services such as the publicly accessible monitoring website, which also serves as a data access point are also covered. The memorandum of understanding (MOU) entered with the Telescope Array collaboration is presented along with a discussion for specific materials we receive from TA. I describe TA trace construction and integration methods. Cross calibration curves for AS, DET2421 and AN are shown and fit. A simulation campaign which reproduces the observed reconstructed showers using Corsika and Auger $\overline{\text{Off}}\underline{\text{line}}$ is presented, and these results are compared to the data. Finally, I wrap up with a discussion about expectations for several detector performance scenarios and how these compare to the actual cross-calibration curves.

6.2 Auger south data extraction

As described in Chapter 5 local and global triggers are appended to the `T3.out`. This file is continually updated throughout the day, and at 00:00 UTC it is archived and transferred to `imogen`, an Auger@TA staging and analysis computer on the CWRU network. This machine handles decompression, decoding and trace analysis for AS data. Later, at 14:00 UTC the "unhex and sort" program will process the previous day's `T3.out` file and separate the global and local trigger types. It's a monolithic script that has a lot of moving parts, so I will describe it in some detail here.

A Python script was created to perform the following actions on the `T3.out` file which happen automatically in the Auger CDAS package. This involves stepping through the decompressed binary file and converting byte sequences to various precision integers:

1. Decompress the file using Python library `bz2`
2. Retrieve all bytes associated with the online calibration process such as VEM Peak, D/A ratio, 70 Hz rate, etc., using the `save_calib` function
3. Retrieve all calibration histograms generated by the online calibration such as Baseline, Pulse Height, Charge, and Shape using the `save_mon` function
4. Retrieve all FADC traces using the `x2` program written by Ricardo Sato
5. Retrieve GPS second and other timing information integers

This package makes heavy use of the Python `struct` library to handle big-endian data. The source can be found in Appendix F. Global and local triggers are separated by comparing the event GPS second timestamp to the local and global T2 archive files which contain the timestamps in coincidence. Additional functions found in this script are described in sections below.

6.3 AS calibration

Although it's possible to use the integer values returned by the online calibration process the recommended and more accurate approach, and what is used in the standard Auger framework, is to determine the 1-VEM hump in the charge or peak histograms [83, 186]. The "unhex and sort" program does this with the `find_vem` function which does the following for each of the 3 AS PMTs:

1. Load the histogram and subtract off the FADC channel offset, which is the first bin entry. This is illustrated in Figure 6.1 which shows the raw charge histograms and the start bins.

2. Filter the y-axis/histogram counts using a 1-D Savitzky-Golay filter with window length 45 and polynomial order 3 from the Python `scipy.signal` library [187], version 0.13.3.

3. Locate and record the pedestal peak in the first 60 bins of the filtered/smoothed histogram

4. Local and record the VEM hump after the first 60 bins of the smoothed his-

togram. The result of this process is also shown in Figure 6.1.

The charge peak bin value is the conversion factor, Q_{VEM} from Equations 3.1 3.2, needed to translate an integrated FADC trace to physical VEM units.

6.4 AS trace building and integration

The `x2` program extracts a text file containing the time and amplitude bins for the 6 PMT channel FADC traces: 3× dynode and 3× anode. The text file is loaded by the "unhex and sort" program which does the following:

1. Determine the baseline. The method is closely based on the one described in [188], but uses slightly modified thresholds. The algorithm uses moving statistics and thresholds to identify peaks in the traces. These segments are marked as 'signal' and are not included in the baseline calculation. Quiescent segments are averaged together to find the best estimate of the baseline. If signals are between baseline regions, the missing part is linearly interpolated to find the baseline.

2. Establish integration limits. The method is again based on [188] with minor adjustments to thresholds and edge trimming. In general, the beginning of the signal is easy to identify since it's a rapidly growing peak. Finding the end of the signal involves calculating moving statistics and comparing these to baseline values. In general, if the dynode is **not** saturated the same integration limits are applied for the anode signal calculation.

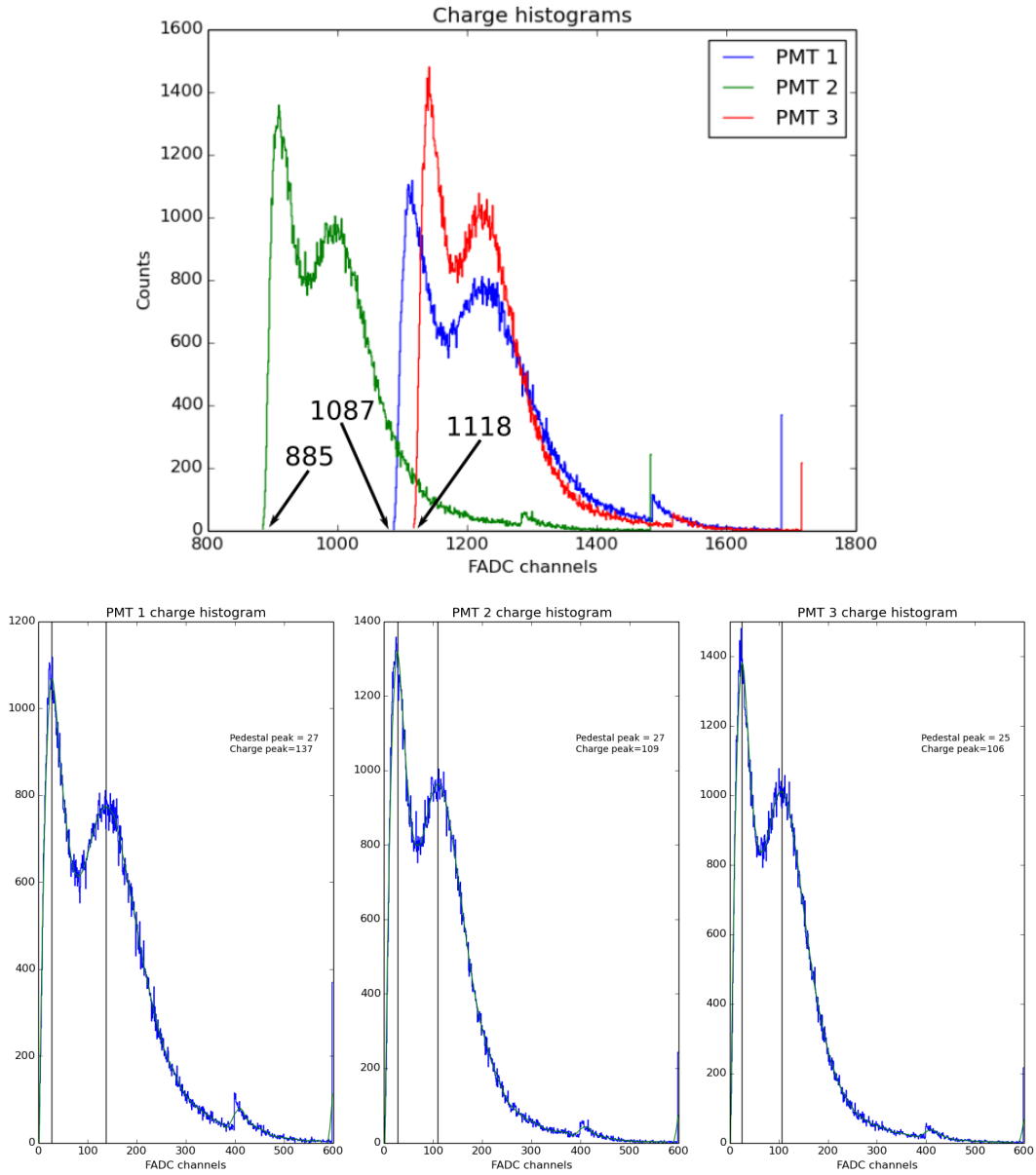


Figure 6.1: Top: Typical raw charge histogram after decompression and decoding. The starting bin offsets are subtracted for further analysis. Bottom: Smoothed histograms showing the pedestal and VEM hump bin location. Data from a global trigger event on October 10 2017 at 20:20:14 UTC.

3. Subtract the baseline and integrate the trace using the previously found limits and Q_{VEM} values.

The signals found from this procedure are archived in the files `south_global_signal.txt` and `south_global_signal.txt` sorted by GPS second. There are also 6 boolean flags indicating saturation for the channel (where false means unsaturated). Separate files `south_global_signal_extra.txt` and `south_local_signal_extra.txt` contain the same data, but also provide GPS second with microsecond precision, the UTC date with microseconds and the T2 type (9=TOT and 1=TH). All these files are in the `/home/augta/web_monitor` directory of `imogen`. Additionally, the unpacked integer data (FADC traces, calibration histograms, signal trace in VEM units, etc.) is available for all events on website hosted by `imogen`, which is described in a later section.

6.5 AN data extraction

The Auger north data is handled in a similar to way to AS with some minor differences in the binary decompression/decoding and VEM calibration procedure. The main decoding program is called "north daily events" which runs at 18:00 UTC everyday on the `clarke` host. The main steps are summarized below

1. Download previous day's binary event, calibration and monitoring files from `imogen`.

2. Process all muon files using Laurent's `anamu` program. If AN is being operated in local storage mode there will be about 1 GB worth of calibration data: 17

minute periods over 24 hours, or 84 files. Each 17 minute period is taken as a "calibration run" and the ADC to VEM value is calculated. Details are provided in the next section. If AN is running in radio mode there will be no calibration files, and `anamu` should instead be used directly on the event binary file.

3. Decompress and decode events using Laurent's `testevt` program and store output as a text file.

4. Get GPS time from the event file. Use this time to search through the calibration runs. Find the run which brackets the event timestamp and save this as the calibration value to be used.

and these are performed for global and local events. Trace data for 4 PMT channels and calibration data for 3 PMT channels is archived. The source is provided in Appendix F.

6.6 AN calibration

For AN the ADC to VEM conversion must be done using histograms. The histograms are not transferred in a binned state as in AS and must be built offline in order to find the VEM hump. For local storage mode the function `get_calibs` from the "north daily events" program is used to create histograms from the 17 minute calibration run using the A×30 channel only

1. `anamu` is used to dump all muon buffers to a text file.

2. The A×30 data is loaded into a `numpy` array.

3. The array is checked for integrity and clipping, then divided into $n \mod 64$ muon traces, where n is the number of buffers.

4. Two empty lists are created for charge and peak charge.

5. The muon traces are looped over and the baseline value of 511.5 is subtracted from the trace. A constant baseline is used since in the vast majority of studied traces the AN baseline is very stable. The summed trace is appended to the charge list and the max value of the trace is appended to the peak charge list.

6. Histograms of the charge and peak charge lists are made. Additionally, a Gaussian kernel density estimate is generated for these lists using the Python `scipy` library and Silverman's rule [189]. This technique is found to smooth the data without losing information about the muon hump.

7. The ADC to VEM conversion for the charge histogram is easily determined by finding the peak in the density estimate, after ignoring the first 150 points which contribute to the pedestal (the density is sample over 1200 bins).

8. The conversion factor (for charge and peak histograms), along with the bandwidth (which contributes to peak estimate uncertainty) used for density estimation is saved to a file for each muon data run.

A typical charge histogram and density estimate along with the charge peak histogram is shown in Figure 6.2. respectively.

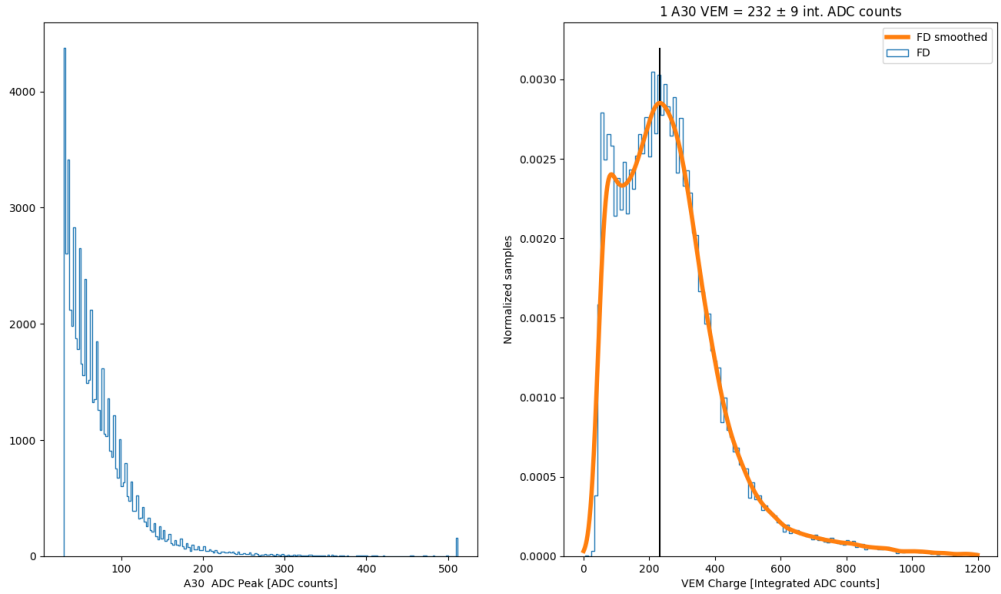


Figure 6.2: A×30 calibration run data for period April 22, 2017 from 05:03:10 UTC to 05:19:55 UTC. Left: Peak charge histogram. Right: Charge histogram. Here the algorithm produces an integrated ADC to VEM conversion factor of $Q_{\mathrm{VEM}} = 232 \pm 9$.

6.7 AN trace building and integration

After the calibration values are determined the "north daily events" program will unpack the binary T3 events into text files and integrate the trace to find the signal in physical VEM units. The baseline is subtracted from each anode channel and the bin associated with the trace peak is identified. The event signal is bracketed by taking the previous 30 (-300 ns) and subsequent 60 (600 ns) bins as the integration limits.

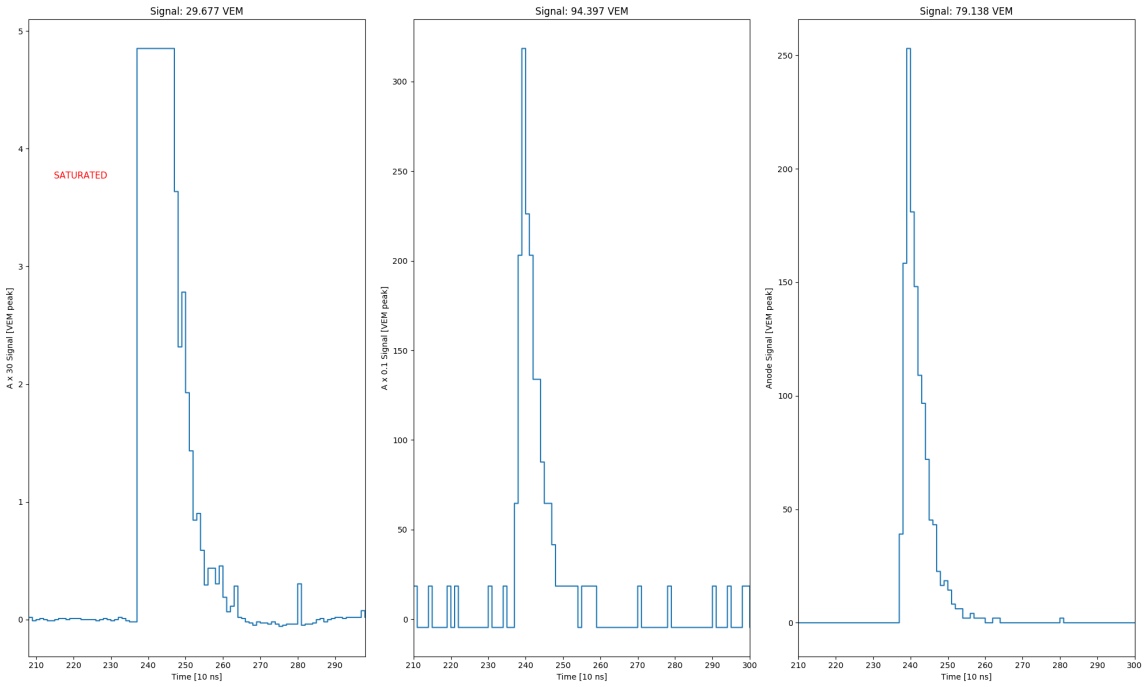


Figure 6.3: PMT traces and integrated signals for (left to right) $\mathrm{A}\times 30$, $\mathrm{A}\times 0.1$ *and* $\mathrm{A}\times 1$ *channels for an April 22 2017 event at 05:03:12.50149249 UTC. This is an example where the high gain* $\mathrm{A}\times 30$ *has saturated. Signal values appear at the top of the panels.*

The signal for the three anode channels ($\times 1$, $\times 0.1$, $\times 30$) is calculated using

$$S_{\mathrm{A}\times 1} \ [\mathrm{VEM}] = \frac{\sum_{i=i_0}^{i_f} Q_i}{Q_{\mathrm{VEM}}/30} \tag{6.1}$$

$$S_{\mathrm{A}\times 0.1} \ [\mathrm{VEM}] = \frac{\sum_{i=i_0}^{i_f} Q_i}{Q_{\mathrm{VEM}}/300} \tag{6.2}$$

$$S_{\mathrm{A}\times 30} \ [\mathrm{VEM}] = \frac{\sum_{i=i_0}^{i_f} Q_i}{Q_{\mathrm{VEM}}} \tag{6.3}$$

where Q_{VEM} is the value obtained from the $\mathrm{A}\times 30$ calibration histogram. An example of this process for the three anode channels is shown in Figure 6.3. The uncertainty in the signal is dominated by the calibration algorithm (bin width used) which results in a signal uncertainty on the order of 3–5 %.

6.8 AN-AS comparison

Validation of the AN stations is an important part of the Auger@TA project effort. Even if the micro-array is successfully installed in phase II, we require a way to understand the AN signals in terms of the AS response, since these are ultimately used in the Auger analysis pipeline. The AN-AS doublet provides an empirical mapping between the stations (should signals not be perfectly matched) and also allows for a performance assessment relative to AS. In this section I discuss the following topics: T2 coincidences and time tagging performance and integrated signal for local and global events.

6.8.1 2-fold coincidences

To test time tagging and the T2 trigger algorithm I consider four daily T2 lists from the 2017 period and examine the entire distribution of time differences. To do this an algorithm starts with a timestamp in one list (e.g. AS), finds the corresponding GPS second in the other list. In addition to that GPS second, the previous and future second are also considered. The minimum time difference is between the starting timestamp and a member of this sub-list. This process is then repeated for all elements in the first list. If the detectors were recording random arrival times one expects the time differences to follow a Poisson distribution. Time differences for several 2 day intervals during the 2016 and 2017 periods are shown in Figure 6.4. The histograms clearly show a sharp peak centered at $\Delta t = 0$ s which would be absent for a Poisson process. The peak bin is populated by 2-fold coincidences between the AN-AS doublet

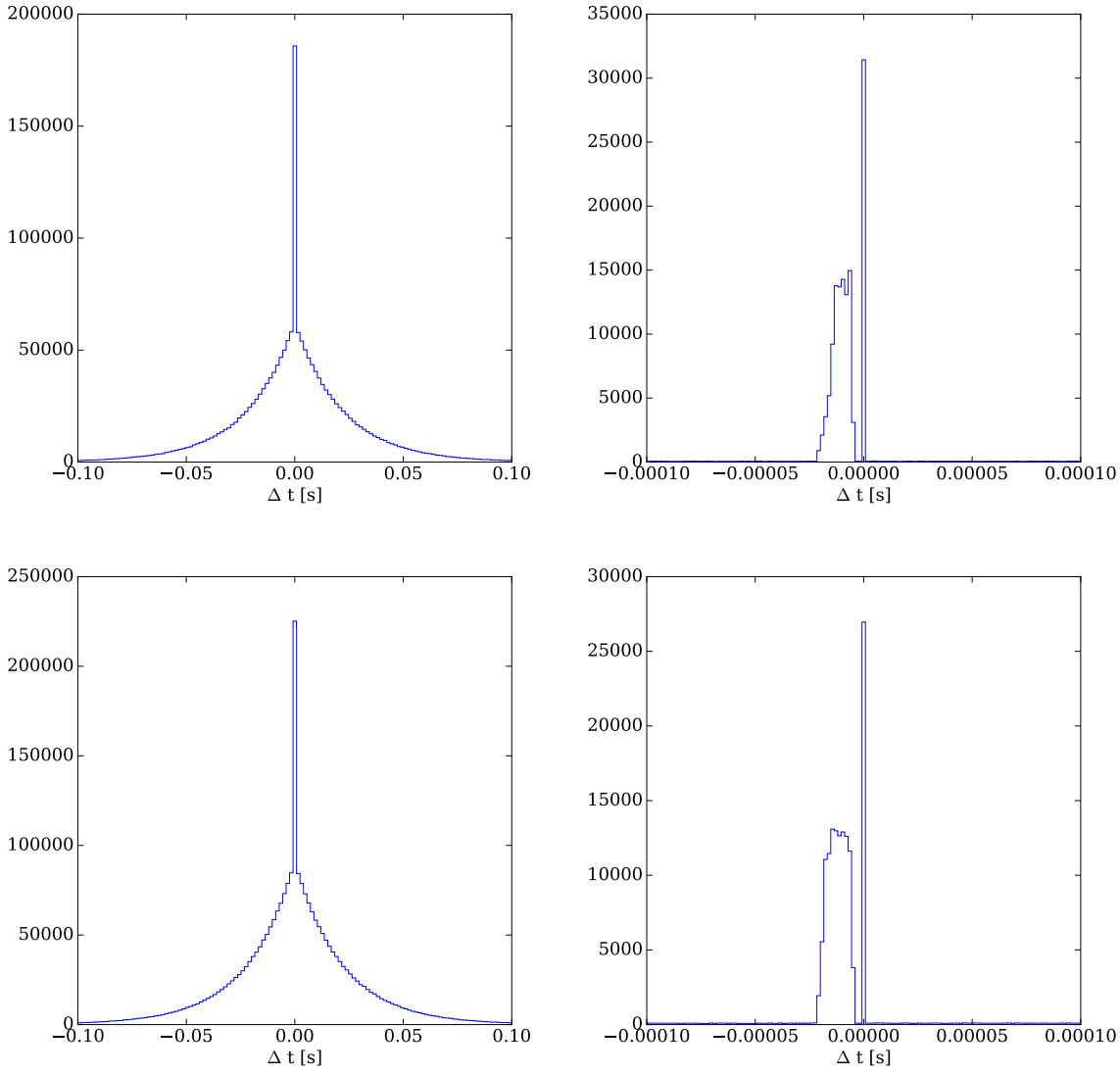


Figure 6.4: Time difference distributions for 48 hours during different months of the 2017 period.

Year	Dates	N	Average 2-fold rate (Hz)
2017	Feb 17,20	124719	0.72
2017	April 25,26	125242	0.72
2017	June 1,2	136626	0.79
2017	August 22,23	133319	0.77

Table 6.1: Samples from 48 hour periods during normal operations. N is the total number of observed two-fold coincidences over 48 hours.

and correspond to EAS candidates. The small offset peak is unusual, but likely the result of a small systematic offset due to electronic or firmware effects. The average rates for each of the samples and other useful information is shown in Table

6.8.2 T2 rate history

In general the AS T2 rate is stable at a rate of around 22-23 Hz. Stability is observed over different seasons throughout the year. During initial operation in 2015 the AN station exhibited a walking T2 rate which seemed to increase during winter months, see Figure 6.5. The AN stability has improved since this time. When in local operation mode the performance is improved if the LSC is configured with an appropriate HV. Reducing the rate to around 16 Hz has shown the best performance. An example plot is given in Figure 6.6.

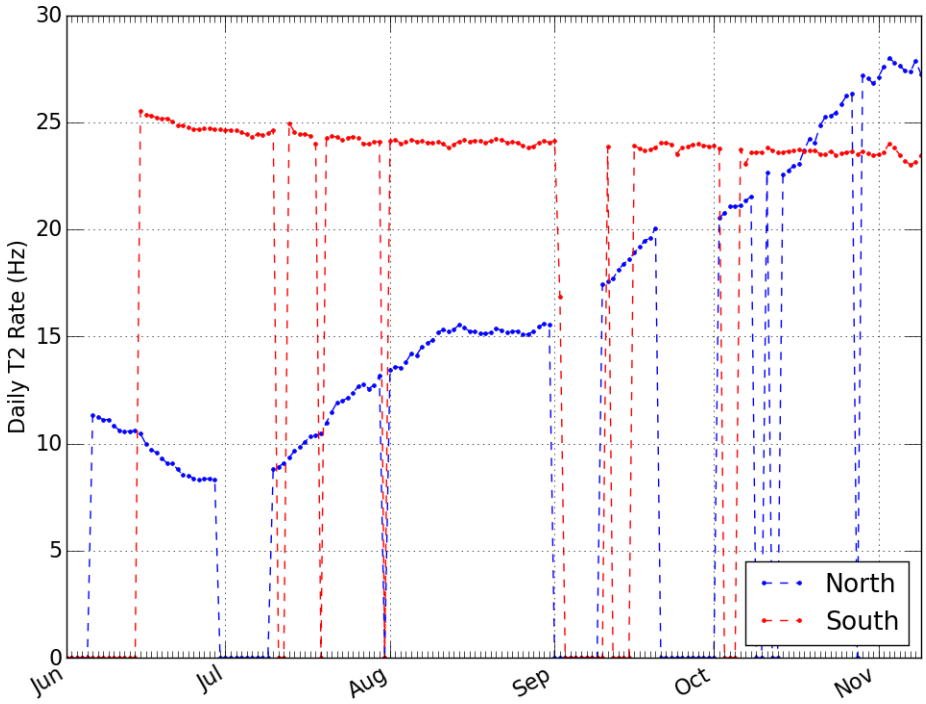


Figure 6.5: Daily T2 rate for several months of operation in 2015. A rate of 0 indicates the station was not collecting data for some portion of that day.

6.8.3 Integrated signal comparison

A comparison of the integrated signal is a paramount component of the project. The ideal case corresponds to identical detectors reporting the same integrated signal for every event. In practice this type of response is unlikely, even for doublets using a similar tank design, e.g. [190], and some scatter in the cross-calibration curve is expected.

For our comparison I use data from the May-September period of 2017 with 3823 local trigger and 15 global trigger coincidences. The data set and fits are shown in Figure 6.7. Note that no quality cuts have been made—this is raw data.

The two sets are fit independently using orthogonal distance regression (ODR)

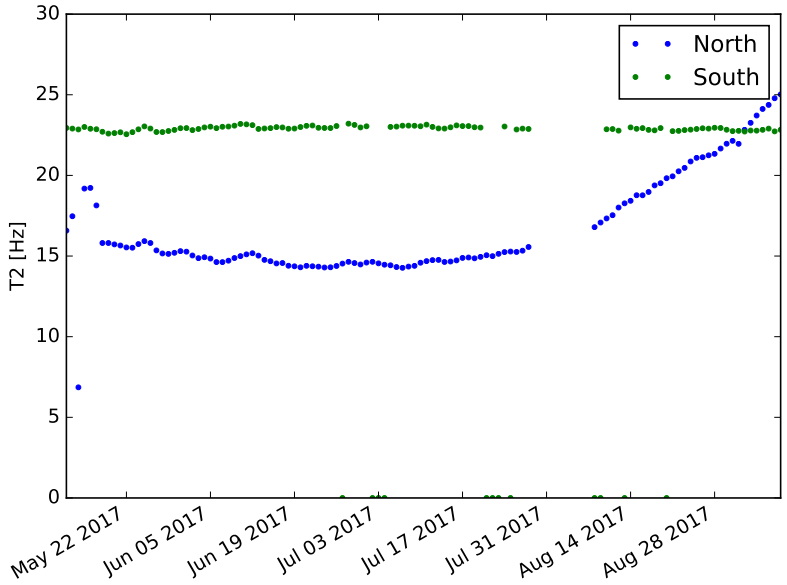


Figure 6.6: AN and AS T2 history over the summer months of 2017. The rate was quite stable during the summer months, but began walking in mid August following a reboot.

[191]. The local trigger set is fit with a powerlaw function

$$y = ax^b \tag{6.4}$$

where the optimal parameters are: $a = 0.8^{+0.2}_{-0.1}$ and $b = 1.1 \pm 0.1$. The global trigger is fit with a linear function

$$y = ax + b \tag{6.5}$$

with parameters: $a = 1.010 \pm 0.002$ and $b = 11.7 \pm 0.4$. For display purposes, the intercept of the global fit is not included in Figure 6.7.

The global fit indicates satisfactory performance with a slope close to unity, especially when the signal is produced by a known cosmic ray shower. The non-zero intercept is likely due to fluctuations in the low signal regime. The local fit gives

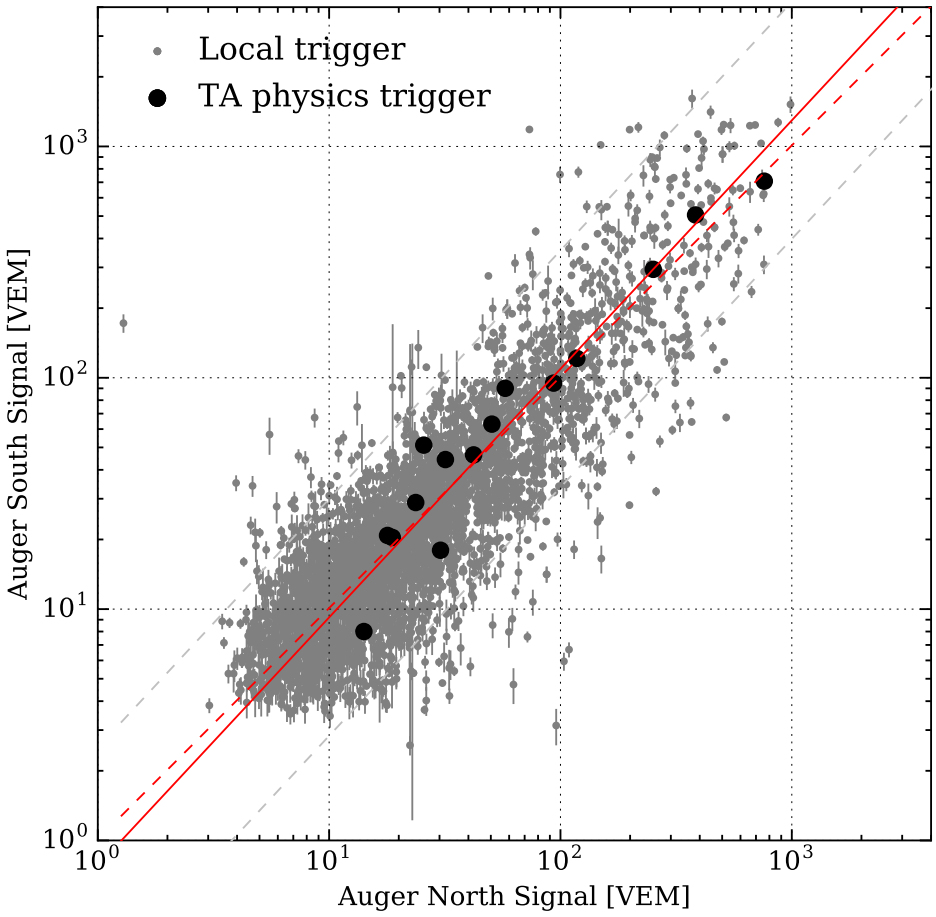


Figure 6.7: Plot of the 3823 local trigger and 15 global trigger integrated signals for the Auger doublet. The solid red line is the fit to local trigger data, while the dashed red line is the fit to global trigger points.

a smaller slope, but agrees with unity within 1σ. The exponent is also high, but agrees with unity within 1σ. The slight discrepancy in the local trigger results is not a complete surprise: unlike a traditional Auger doublet, the local trigger is not initiated by coincident T2 events from these stations. This is a consequence of the AN station being in local storage mode during the operation period. The local events are generated by coincidence with the threshold comparator circuit time stamps, which is highly biased toward fast large amplitude pulses (muon rich older showers). Additionally, the close separation means the local events are dominated by low energy showers which have uniform particle densities on much smaller length scales compared to > 30 EeV showers, for example. It should also be noted that the local trigger has no way of adjusting the HV on the PMT to maintain a background rate: the HV is nominally kept at 1200 V. It has been observed that the local rate fluctuates with the diurnal cycle. We believe the combination of all these factors serve to increase the AS-AN signal spread in the cross-calibration curve. Nonetheless, if the detectors are working properly, *on average*, we expect a consistent response. Visual inspection of the plot shows the average trend appears consistent with unity slope. Additionally, the data spread is not lopsided nor does there appear to be strong deviations away from the centroid in any particular signal range. Given these observations, the AN signal response appears to be consistent with the AS counterpart.

6.9 Monitoring website and data staging area

As described in earlier sections the `imogen` machine is the main point of contact for data transfer and AS analysis. Additionally, a website was built which shows local weather conditions (updated every 15 mins) as well as astronomy information. It provides T2 rates for AN and AS as well as pseudo-T2 rates for the local trigger system, updated every 24 hours. A small section also shows some news items and other important advisories. More detailed diagnostic data for AN is provided on the "Data" page which shows recent calibration, environmental and battery data. This page also shows the daily and weekly local trigger rates. A screen-shot of the landing page is shown in Figure 6.8. A wiki page is also hosted (using MediaWiki) and includes some helpful information about the project and provides a place for collaborators to contribute and edit articles.

This website also provides a link to the main data warehouse where users can access raw ADC traces and histograms, as well as processed traces with integrated signals by following https://imogen.phys.cwru.edu/monitor/data/. Each of the four links lead to a new page where each day is presented as its own folder. The internal directory structure for raw data is shown below

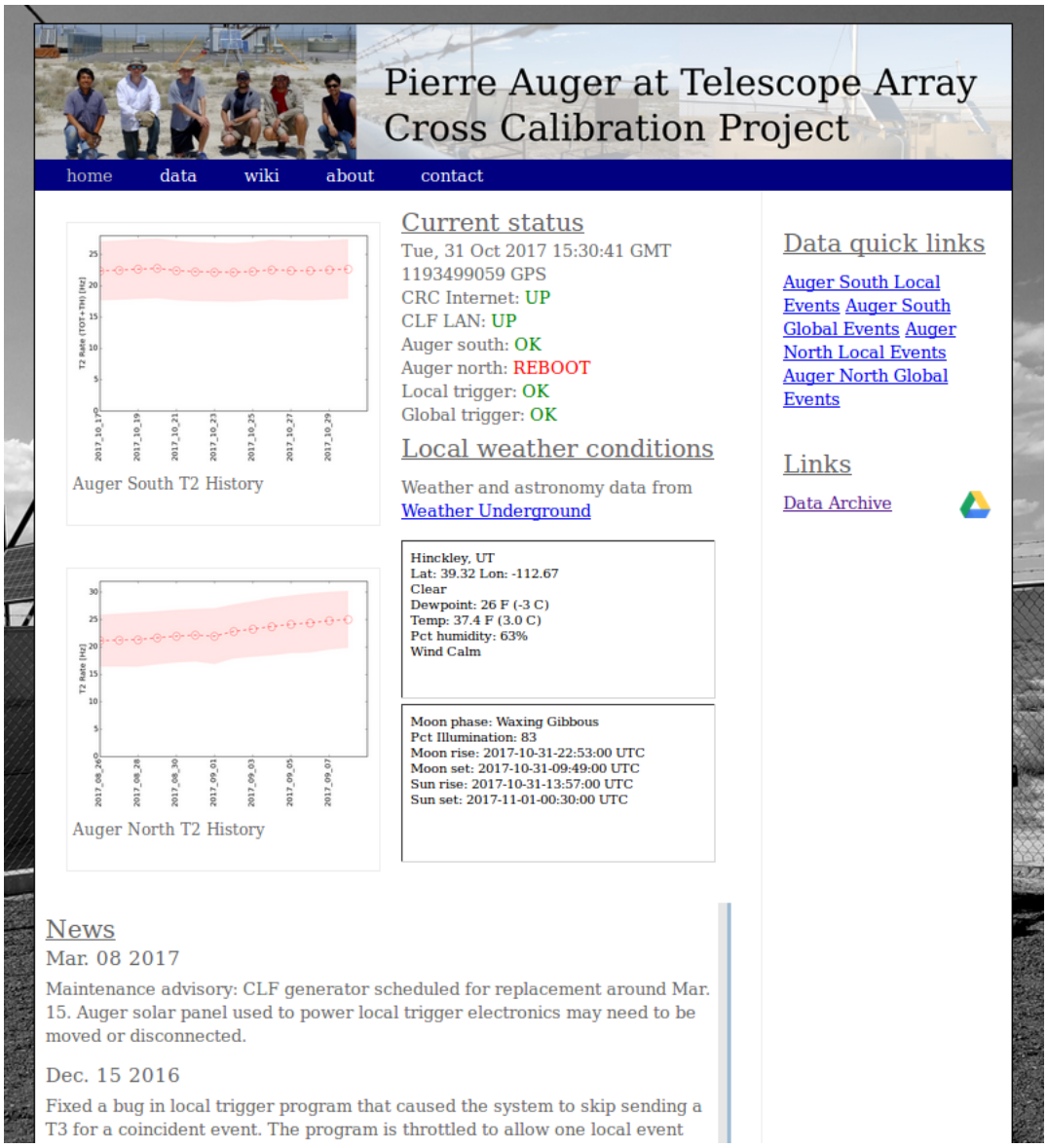
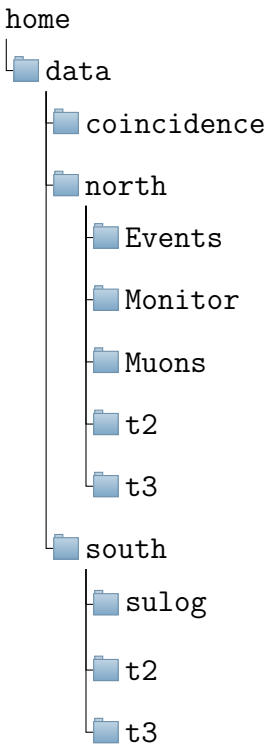


Figure 6.8: The home page of the monitoring website. The T2 rates of AS and AN are shown, along with the status of important components. Links to data files are on the right sight.



6.10 Memorandum of understanding

Since this project requires the cooperation of two large international collaborations it was necessary to formally define what materials are to be shared between Auger and TA. These details have been outlined in a MOU written by Fred Sarazin with feedback provided by CWRU which appears in Appendix G. As of summer 2017 the agreement has been approved by the Auger technical board and Telescope Array and it remains in effect. The main highlights can be summarized as

- All raw Auger data is stored on premises at TA in addition to being transferred to CWRU servers. Processed data is also available on publicly accessible web page.

- Waveforms and calibration data for DET2421 are shared which allows the integration of signals for this station. This point directly enables a cross-calibration of SD stations. It also enables a comparison study of PMT traces, discussed in a later section.

- Timing and integrated signals for neighboring stations: DET1216, DET1217, DET1316, DET1317 for custom reconstructions, or studying particle density vs. core distance.

- Access to TA reconstruction parameters for global events which pass the spectrum quality cut described in [102].

Currently we try to maintain a quarterly data transfer schedule for these materials.

6.11 TA trace construction and integration

Data for the DET2421 station is received from TA collaborators in a zipped text file. As of October 31, 2017 we have received two quarters worth of data [192, 193]. A Python script was developed, "build master", to parse the text file:

1. Layers are first separated

2. Baseline is subtracted using a cluster weight algorithm: the normal 128-bin length trace is reduced to 40 bins and a weighted average of the two most populous bins is used for the baseline estimate[1].

[1] A previous version of the program used a method similar to the one used for AS. In general this performed poorly for the TA traces since there are more peak features, and interpolation between signal regions proved to be problematic.

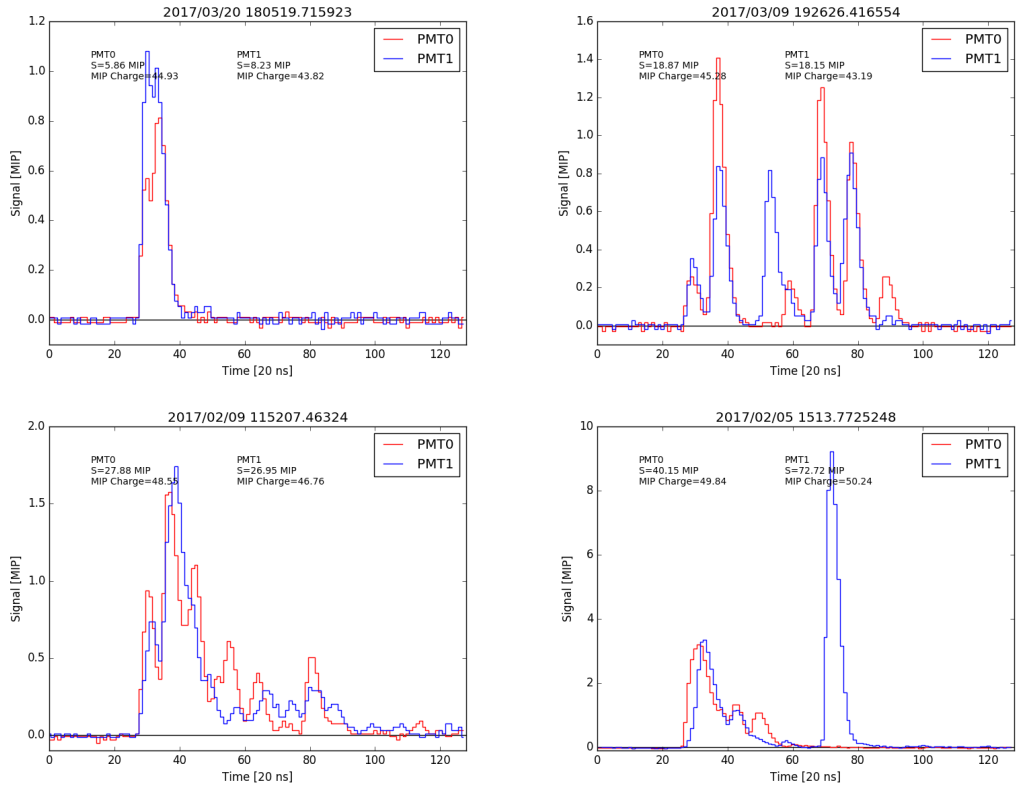


Figure 6.9: Four global event traces built with the MOU materials received from TA colleagues. Timing information appears on the top of the panels and integrated signal values are shown on the plots.

3. Event is flagged for saturation if either layer has one or more FADC entries of 4095.

4. Trace is integrated using the ADC to MIP conversion factor provided by TA.

The source can be found in Appendix H. A few representative examples of this process are shown for global events in Figure 6.9.

6.12 The Auger-TA SD station cross-calibration curve

With all the previously described hardware and software infrastructure in place it's now possible to perform an empirical comparison of the AS, AN and TA SD station level responses to the same EAS. The processed data can be compared in a variety of ways, but we have historically presented results using a MIP vs. VEM curve [194]. This provides an easy way to visually spot interesting features, such as departure from linearity, spread in the data or intercept offsets. As of October 31, 2017 we have 100 global events with data available for DET2421 and AS. Of these, AN has data available for 11. Three of the 100 events resulted in saturated detectors. The TA PMTs were saturated for all of these, and the Auger low gain PMTs were saturated for one. Of the 100 raw global events 13 pass the TA spectrum quality cut. The raw data for these results is displayed in Figure 6.10. The same plot can also be made but using Auger north data. This appears in Figure 6.11. Less data is available for AN due to detector duty cycle issues: there were extended periods where it can get stuck at boot and must be manually restarted at the CLF. In addition to looking at the MIP-VEM behavior we can also compact the data to a single dimension by calculating the ratio of S(TA)/S(Auger). This allows for a statistical characterization of the correlation between the station responses, but it can't be used to study subtle effects, e.g. a changing ratio as a function of signal strength, so it is somewhat limited. For thoroughness I include a variety of binning schemes in addition to a Gaussian kernel density estimate, since the choice of binning method does impact the visual representation of the data. The histograms are shown in Figure 6.12. The data

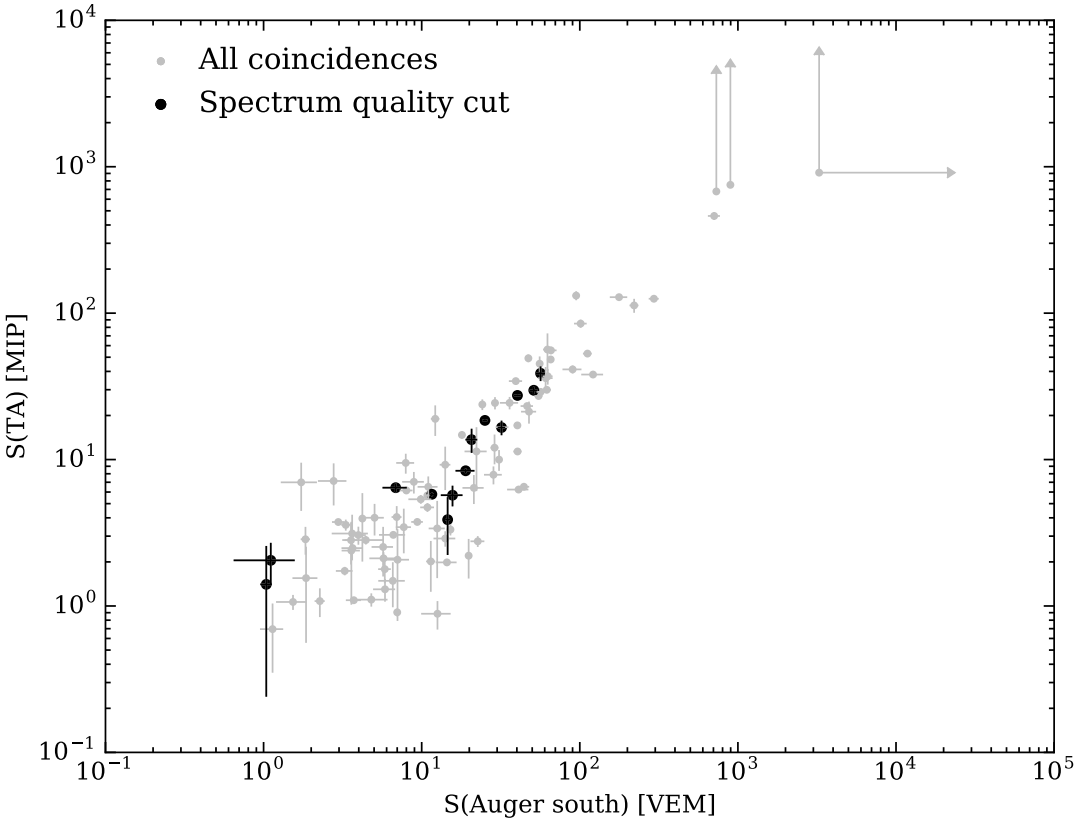


Figure 6.10: The cross-calibration curve for raw TA global trigger data. Data is for the first two quarters (Jan.-Jun.) of 2017. Error bars show the systematic and statistical uncertainty for the measurements. Saturation events represent lower limits; however, since

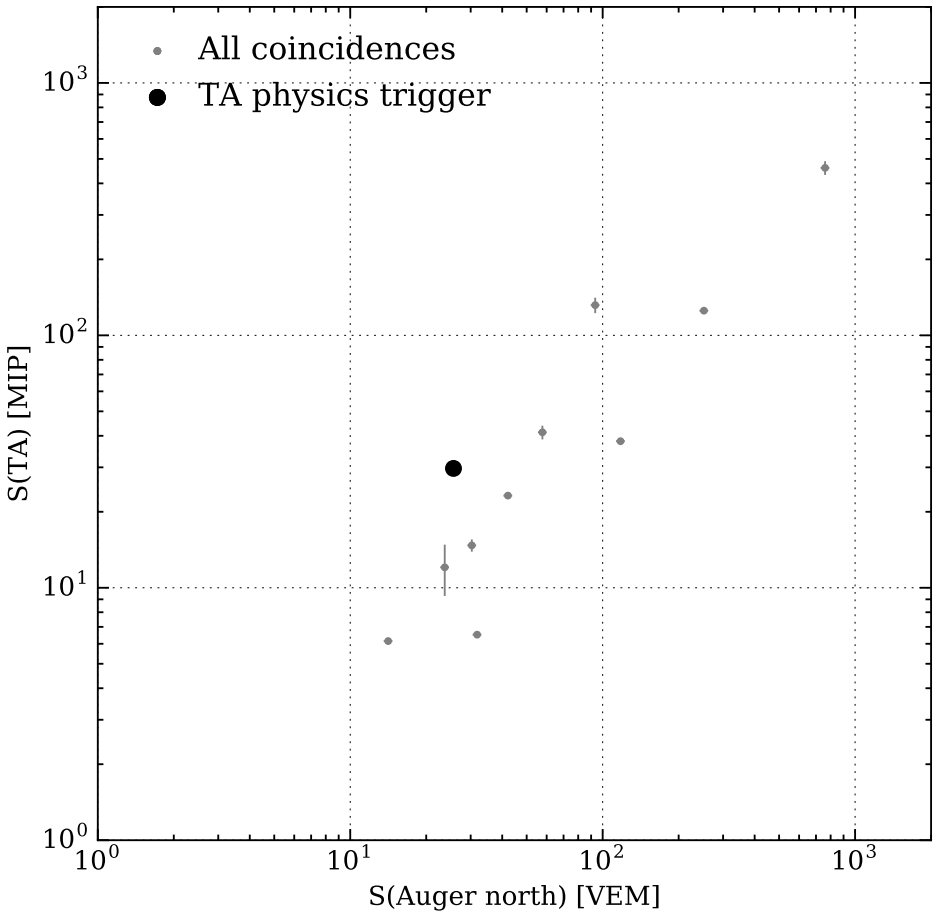


Figure 6.11: Cross-calibration curve using AN data.

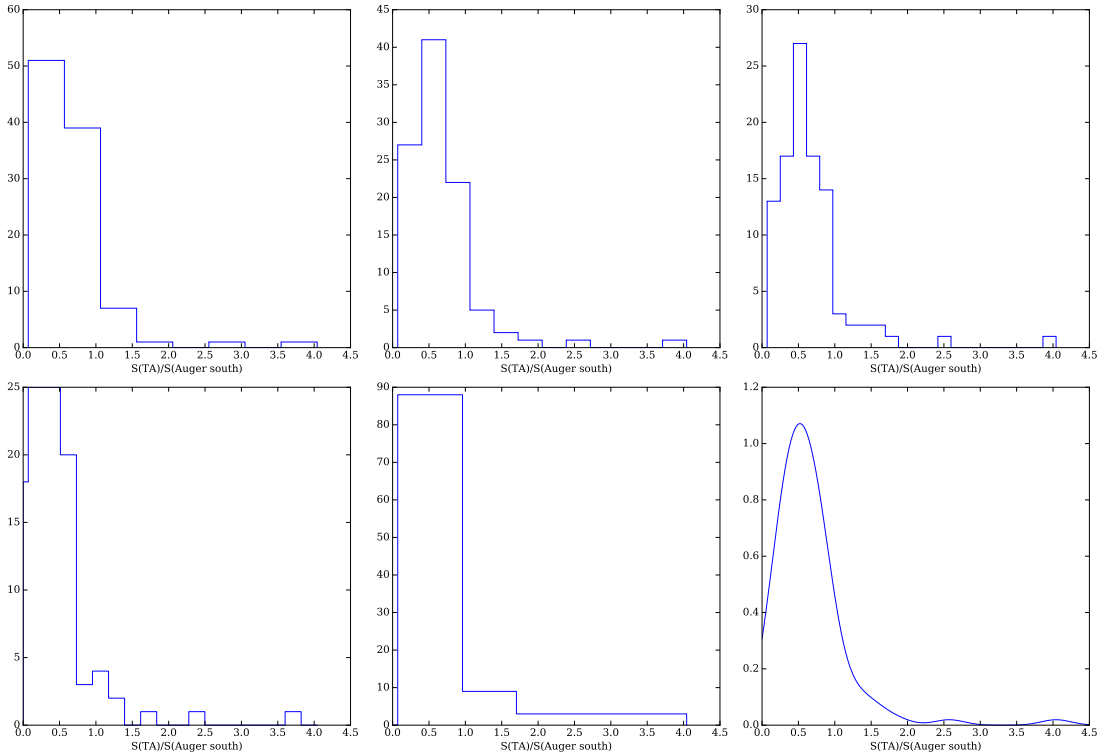


Figure 6.12: Histograms using six different binning methods. From top to bottom and left to right: Freedman-Diaconis, Doane, Sturges, Knuth, Bayesian blocks and a Gaussian kernel density estimation. Although the profile differs slightly between the methods, it is clear the data are mostly in the range of 0.4-0.7 with some spread.

Method	MPV	$\bar{x}^*$	σ^*	References
Sturges	0.07-0.567	0.663	0.508	[195]
Doane	0.401-0.732	0.669	0.501	[196]
Freedman-Diaconis	0.431-0.612	0.663	0.508	[197]
Knuth	0.291-0.732	0.655	0.511	[198]
Bayesian blocks	0.07-0.961	0.66	0.454	[199]
Gaussian KDE	0.521	0.686	0.542	[139]

Table 6.2: Summary statistics for the signal ratio data. The most probable value appears in the MPV column. The weighted mean using the histogram binning is shown in the $\bar{x}^*$ column along with the weighted standard deviation in σ^*.

distribution shows a mean around 0.6 with a fairly large dispersion of around 0.45. Some numerical values with summary statistics are listed in Table 6.2.

6.13 Expected station response

In addition to studying the empirical relation, we can contemplate the expected behavior for idealized situations. In this section I present an example case adopting the following assumptions

1. Both detectors have identical efficiencies for all shower components
2. Both detectors count all incident particles perfectly
3. Showers are planar

With these in place, the only remaining factor which leads to a different response is the detector geometry. The TA scintillator is further assumed to be a flat rectangle since the thickness is 1.2 cm, which is a small fraction of the length dimension. The shower flux through a TA detector is

$$\Phi = \sigma \ell w \cos\theta = 3\sigma \cos\theta \tag{6.6}$$

where Φ is the flux, $\ell = 2$ m, $w = 1.5$ m, θ is the shower zenith angle and σ is the shower density. The effective area as a function of zenith angle is therefore

$$A_{\text{eff}} = 3\cos\theta \quad [\text{m}^2] \tag{6.7}$$

For the Auger detector I do not assume a flat geometry since length, width, and height of this detector are of the same scale. I do approximate the tank as a perfect cylinder. First I consider the flux through the cylindrical wall. For a shower axis described by $\vec{x} = \sigma\hat{r}$ the flux can be found using

$$\Phi = \int \vec{x} \cdot d\vec{A} \tag{6.8}$$

A planar shower with zenith angle θ can be represented by the Cartesian vector $\vec{x} = \sin\theta\ \hat{x} + \cos\theta\ \hat{z}$. The area element in cylindrical coordinates will be $d\vec{A} = (\cos\phi\ \hat{x} + \sin\phi\ \hat{y})\, r\ d\phi dz$. Equation 6.8 can now be written as

$$\Phi = \sigma \int_0^h \int_{-\pi/2}^{\pi/2} r \sin\theta \cos\phi\ d\phi dz = 2rz\sigma \sin\theta = 4.32\sigma \sin\theta \tag{6.9}$$

where $z = 1.2$ m is the tank height and $r = 1.8$ m is the tank radius. The top of the tank is treated as a flat circle so its effective area is simply $A_{\text{eff}} = \pi r^2 \cos\theta = 10.18\cos\theta$ m^2. The total effective area of the Auger station is then

$$A_{\text{eff}} = 4.32\sin\theta + 10.18\cos\theta \quad [\text{m}^2] \tag{6.10}$$

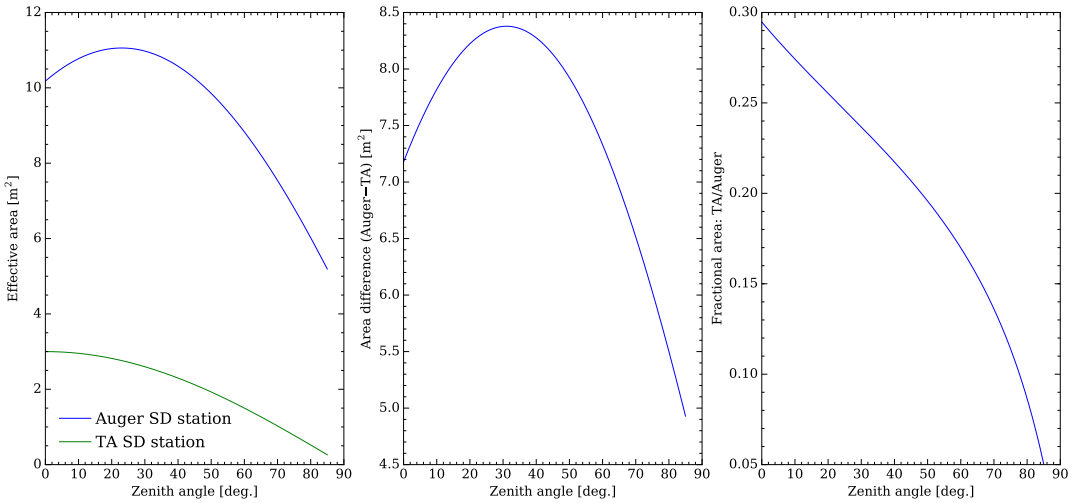


Figure 6.13: Left: Direct comparison of Auger and TA effective area as a function of zenith angle. Middle: The difference in effective area, which has a maximum value of 8.4 m^2 at 31.1°. Right: Ratio of effective areas.

A plot showing each detector's effective area, their difference and ratio is shown in Figure 6.13. For a range of zenith angles the effective area of Auger increases because the cylindrical wall becomes increasingly visible and adds to the circular top piece. Since the TA design is 2-dimensional it exposes an increasingly smaller cross-section, especially for $\theta > 50°$. This simplistic model predicts slope values uniformly ranging from 0.207 to 0.295 for $\theta = 0°$ to $\theta = 45°$, respectively, for a MIP-VEM cross calibration curve. The model prediction is compared to actual signal ratios which have zenith angle data available, shown in Figure 6.14. Information about core distance and shower arrival times, important for effects like rise time azimuthal asymmetry related to composition studies [200], are also included. At this time is difficult to draw robust conclusions due to the small sample size. It's fairly clear the toy model disagrees with the observed data. One glaring issue is assumption 1, where identical

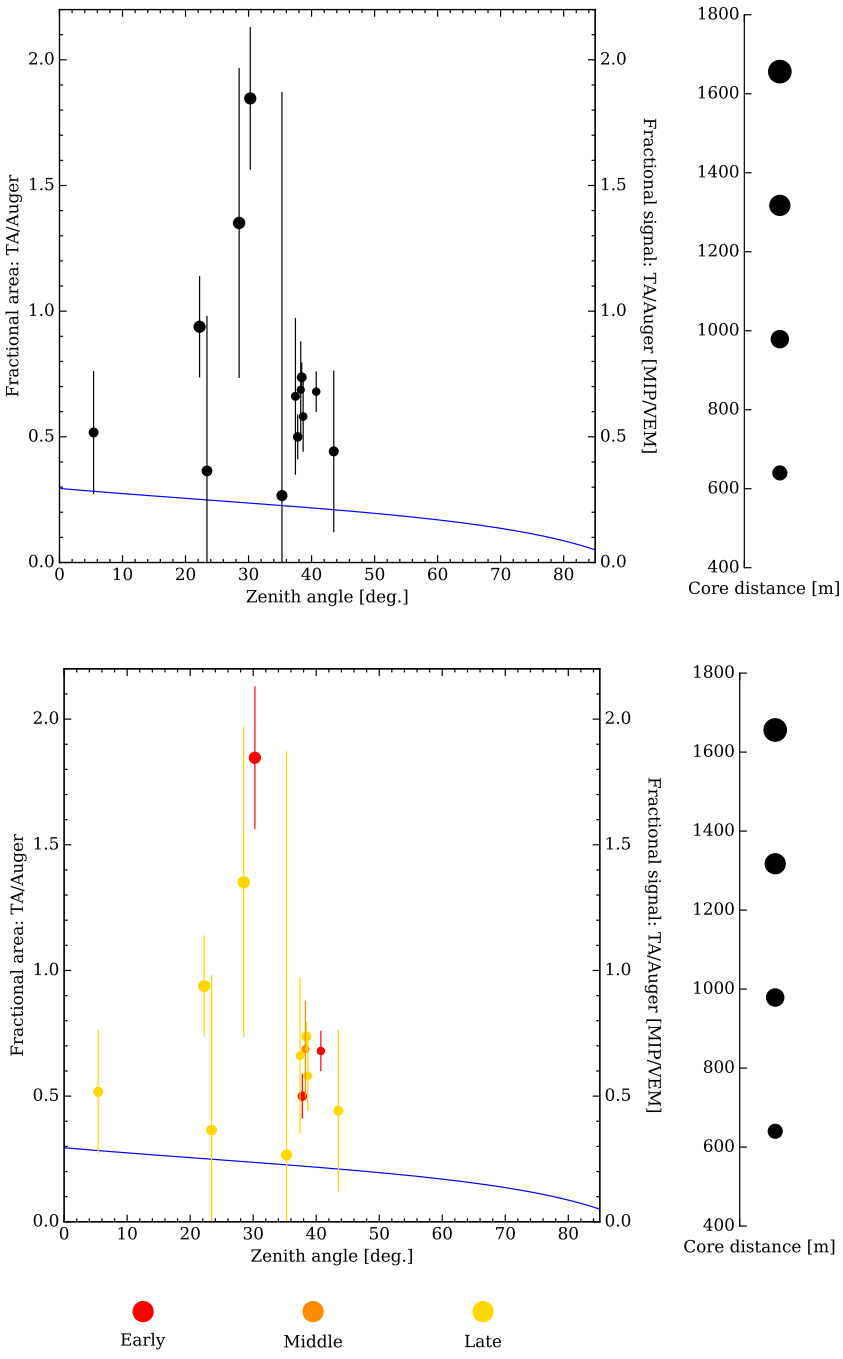


Figure 6.14: Top: The effective area ratio model is compared to global data with reconstructions. Marker size depicts distance from the core. Bottom: Same as left, but shower arrival time information has been added. Early means the station is among the first to sample the shower plane, middle means the station was roughly near the core (middle of impact), while late means the station was among the last stations to sample the shower plane.

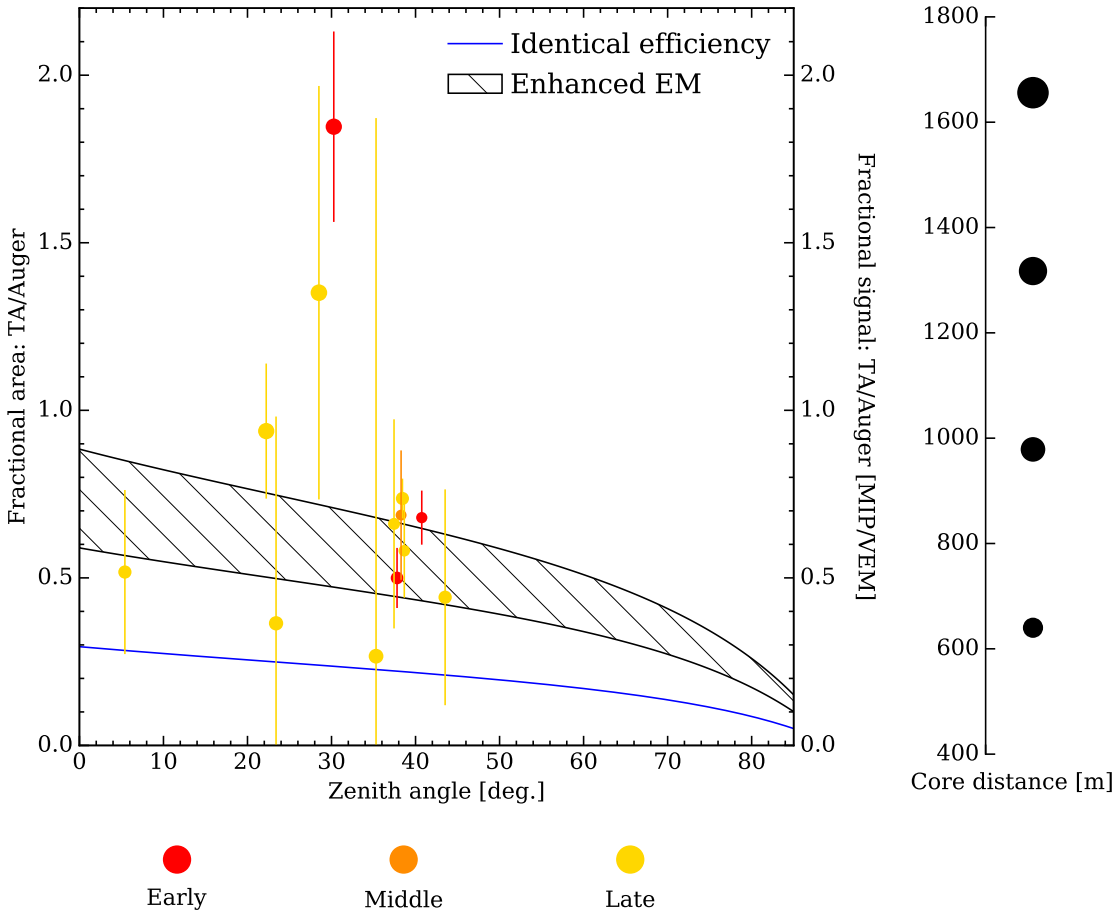


Figure 6.15: The same plot as Figure 6.14 but with a new model that accounts for the enhanced response of scintillator to the EM shower component.

efficiencies are assumed. This is known to be incorrect based on Auger upgrade (SSD) simulations which show that the ratios between integrated EM and muon component signals is roughly between 2-3 for a scintillator compared to a WCD for a large range of core distances [184]. I use this fact to present an updated toy model which includes a new region accounting for this difference in efficiency in Figure 6.15. This update to the toy model improves agreement with data. This is an important effect which should be considered when doing joint analyses in the future.

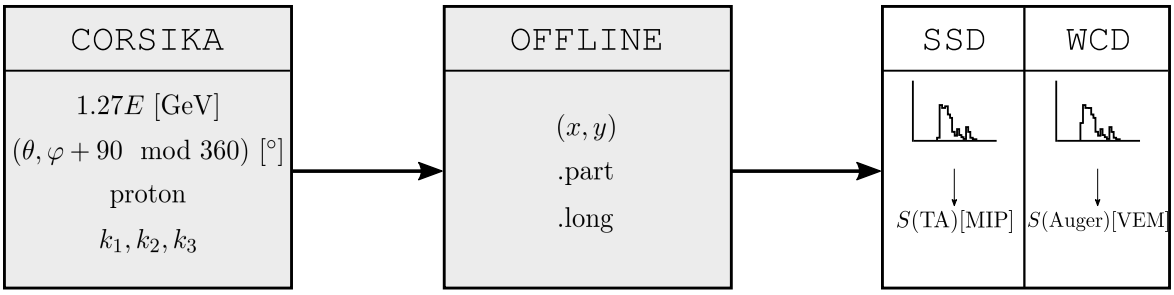


Figure 6.16: The simulation pipeline. Simulation data is generated in the grey modules and the list of parameters (in expected units) are the inputs. In the `CORSIKA` *box the seed is denoted with* k_i*. The final white box is an output stage where signals are integrated.*

6.14 Simulated response

In addition to comparing data to toy models, we can also look at simulated signals. The pipeline involves creating a custom `CORSIKA` [201] air shower based on TA reconstruction parameters and reconstructing this simulation using a custom geometry in Auger $\overline{\text{Off}}\underline{\text{line}}$ [202] which reproduces our setup in the field at the CLF. Figure 6.16 shows a diagram summarizing the work-flow. The primary energy, zenith angle and azimuth are entered into the `CORSIKA` steering card. A recent version, 75700 (current version 75800 as of November 5, 2017) is used and compiled with the following non-default options: `THIN`, `FLUKA` (version 2011.2c.6) [203, 204], `QGSJETII-04` [205]. The QGSJET interaction modeled is preferred due to faster run times. A direct comparison was done for 3 proton events and very little difference in the predicted MIP-VEM values was observed.

A random seed is used to determine initial interaction height and these range from mostly 25-50 km and a median depth of 13.5 g cm^{-2} and 100 showers are generated

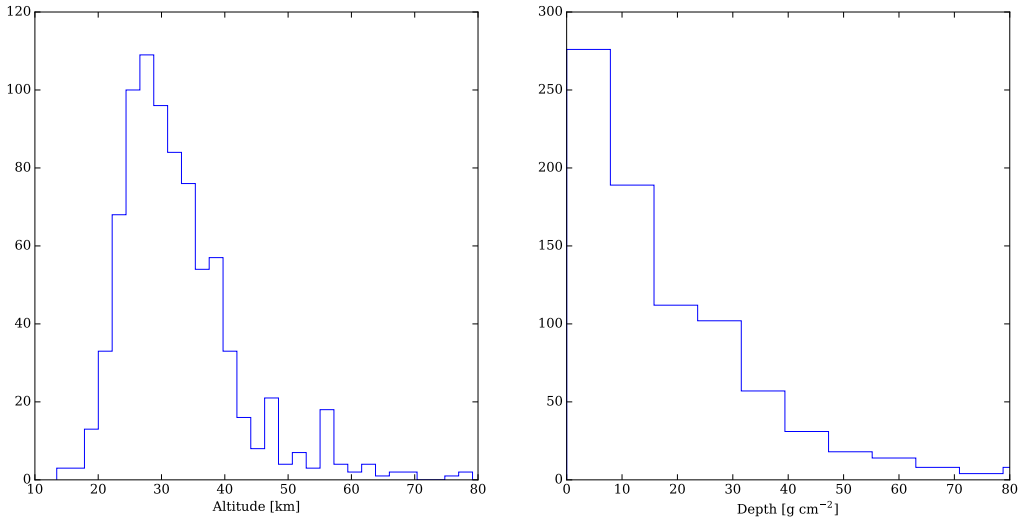


Figure 6.17: Left: Altitude of first interaction. Right: Depth of first interaction. This parameter is assigned by Monte-Carlo since it's not available in reconstruction data.

per event. The distribution for a sub-sample (800) of the total shower population is shown in Figure 6.17. To increase throughput, the 100 samples are split among multiple cores across two servers running Ubuntu Server 16.10. In terms of hardware, one is a dual socket quadcore Intel Xeon E5520 machine with 48 GB memory and the other is a dual socket hexcore Intel Xeon X5675 also with 48 GB of memory. The geomagnetic field is reproduced at the CLF using the IGRF-12 model [206] and the coordinates of the CLF: 112.908723° W and 39.296919° N. A useful tool for finding the field parameters is available at `https://www.ngdc.noaa.gov/geomag-web/?model=igrf#igrfwmm`. The field value is $B_x = 21.8787$ μT and $B_z = 45.8686$ μT where the mean is taken over the first few months of the year 2017 and the altitude of the CLF, 1391 m, is used. Other key parameters, such as thinning fractions, energy cuts, and flags can be found in an example steering card in Appendix I. The only parameters which vary according to the event are `ERANGE, SEED, THETAP, PHIP,`

`NSHOW, RUNNR, THIN`, all others remain fixed. The first two `THIN` parameters are fixed, and the last is a core cut parameter which is changed based on the core distance. Particles within the core cut radius are not written to disk, which saves on storage space. Proton, helium and nitrogen primaries are used; heavier species are avoided since they're not currently favored in these energy regimes.

The observed TA azimuth is defined with respect to due west. This is converted to `CORSIKA` azimuth using $\varphi + 90^\circ \mod 360$. Additionally, the provided energy is in terms of the TA FD calorimetric scale where a factor of 1/1.27 has been applied. We convert back to the SD energy scale by multiplying by the coefficient $1.27E$ per a TA collaborator's recommendation [207].

The trunk development version (rev 28752) of Auger $\overline{\text{Off}}\underline{\text{line}}$ is used which has scintillator surface detector (SSD) functionality, in addition to standard Auger WCDs. A note describing the implementation of the new detector module was presented at the 2017 ICRC [208]. To simulate the doublet setup at the CLF a reference tank is chosen in the Auger array, `iTank` (ideal tank) 5048 with coordinates (480629 m, 6091844 m) at altitude 1391 m. This serves as the Auger south tank coordinate. The core position is considered relative to this coordinate. The TA SD is instantiated as `iTank 5040` and given the location (480594 m, 6091805 m) to approximate the separation of the detectors in the field. The core is referenced to the midpoint (e.g. the CLF building) of this pair: (480612 m, 6091820 m). The SSD has a surface area of 3.84 m^2 and the TA SD has area 3 m^2. Therefore to express the integrated MIP signal of the SSD into a TA equivalent MIP, we multiply the SSD value by 0.78125.

The core position is placed at a fixed location on the *ground* specified by the

coordinates provided in the MOU materials: $(x, y) = (480612 + x', 6091820 + y')$ where x', y' are the observed easting and northing ground coordinates. An example of the $\overline{\text{Off}}\underline{\text{line}}$ configuration files EventGenerator and bootstrap XML files are shown in Appendix J.

To read out the FADC traces I use a program provided by David Schmidt: "DataWriter" [209]. The user can input the maximum distance from the core position for which to save trace information. Additionally, the sum of baseline subtracted FADC counts for high gain and low channels are provided for the SSD and WCD, as well as a flag indicating saturation. The simulated FADC sums are converted to particle counts using the following calibration values: 4735.1 Σ(ADC)/VEM for the WCD and 197.005 Σ(ADC)/MIP for the SSD [210]. An example reconstruction is shown in Figure 6.18 along with the equivalent TA event view plot.

The data are compared to simulations using the MIP vs. VEM plane. For the 100 simulation points a probability density function is estimated using a Gaussian kernel density estimator where the bandwidth is determined using Scott's rule. See Figure 6.19 for events 2–5, Figure 6.20 for events 7, 9, 10, 11 and Figure 6.21 for 12, 13, 14, 16. Error bars for simulation data are one standard deviation. For observed VEM data the error is found by adding systematic and statistical error in quadrature. Statistical error is defined as $\delta S_{\text{stat}} = \text{std}\,(S_1, S_2, S_3)$ where S_i is the signal in PMT i. Systematic error is defined as $\delta S_{\text{sys}} = 0.01\sqrt{S_1^2 + S_2^2 + S_3^2}$ which accounts for the 1% error in the online VEM calibration procedure. For observed MIP data, the error is purely statistical and uses the standard deviation. The mean of PMT signals is used for the observed value coordinates.

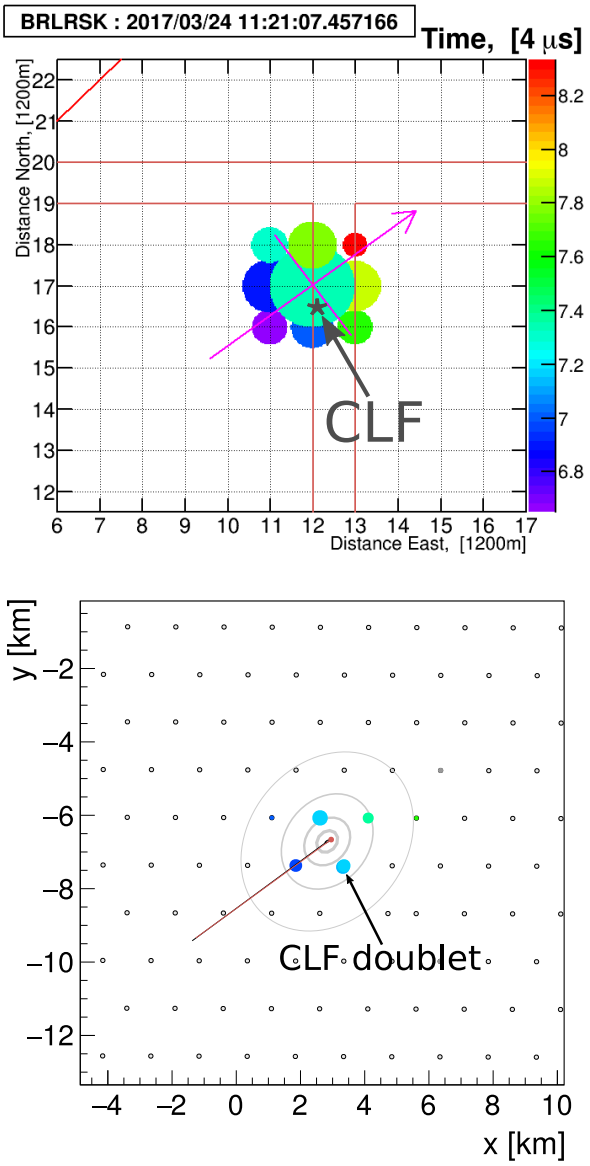


Figure 6.18: Left: Display for observed event (#9) provided by TA colleagues. Circle size is proportional to the integrated MIP signal and color reflects timing. Right: Offline reconstruction event display for one of the 100 CORSIKA simulations for event #9. Circle size corresponds to the log of the integrated VEM signal and color depicts timing data. One can distinguish the close separation of detectors by zooming in.

In Table 6.3 I provide the discrepancy between the proton simulations and observations. The quantities $\Delta_{m,v} = S_{\text{obs}} - S_{\text{sim}}$ where the subscript m refers a MIP measurement and v refers to a VEM measurement. It is the *actual* difference. The relative difference is also provided $\Delta^r_{m,v} = \frac{|S_{\text{obs}} - S_{\text{sim}}|}{(S_{\text{obs}} + S_{\text{sim}})/2}$ which is expressed as a percentage. The average of observed and simulated signals is used to give each equal weight in the reference term. The uncertainty is added in quadrature using standard error propagation methods.

Detector distance from the core, ρ, is calculated using a planar shower geometry

$$\rho = x_g \cos\theta \tag{6.11}$$

where x_g is the ground distance to the shower axis. See Figure 6.22 for a visual depiction.

The same procedure was repeated for helium and nitrogen primaries. The helium results are comparable to proton, showing minor improvements in discrepancy. For nitrogen there is a noticeable discrepancy reduction, but for a majority of events it remains large. The nitrogen results are provided in Table 6.4. The improvement to VEM based measurements does seem larger than the MIP based counterpart. In some cases the discrepancy for the MIP signal worsens for nitrogen simulations.

6.14.1 Energy scaling and future simulations

This study used a scaling of $1.27E$ to normalize to the TA SD energy scale. A more precise conversion will take into account several factors: (1) there is a known scaling offset between the TA and Auger energy spectra of around 10.4% [211] (2)

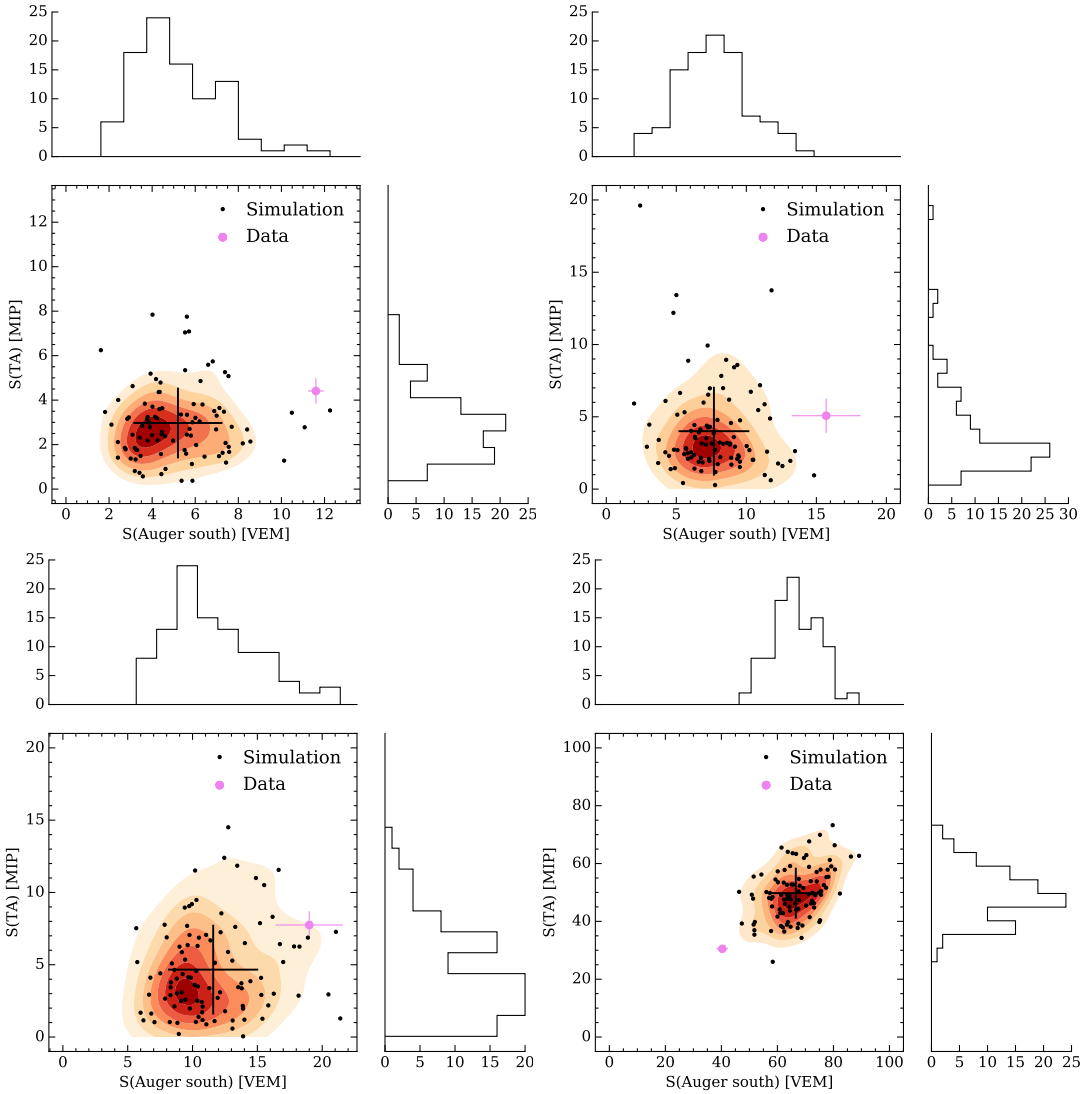


Figure 6.19: The MIP vs VEM plane for simulated and observed integrated signals. Top left: event 2, top right: event 3, bottom left: event 4, bottom right: event 5. The MIP and VEM simulation populations are also shown in the histograms. The energies are given in terms of the TA FD scale, and the azimuth follows the TA convention.

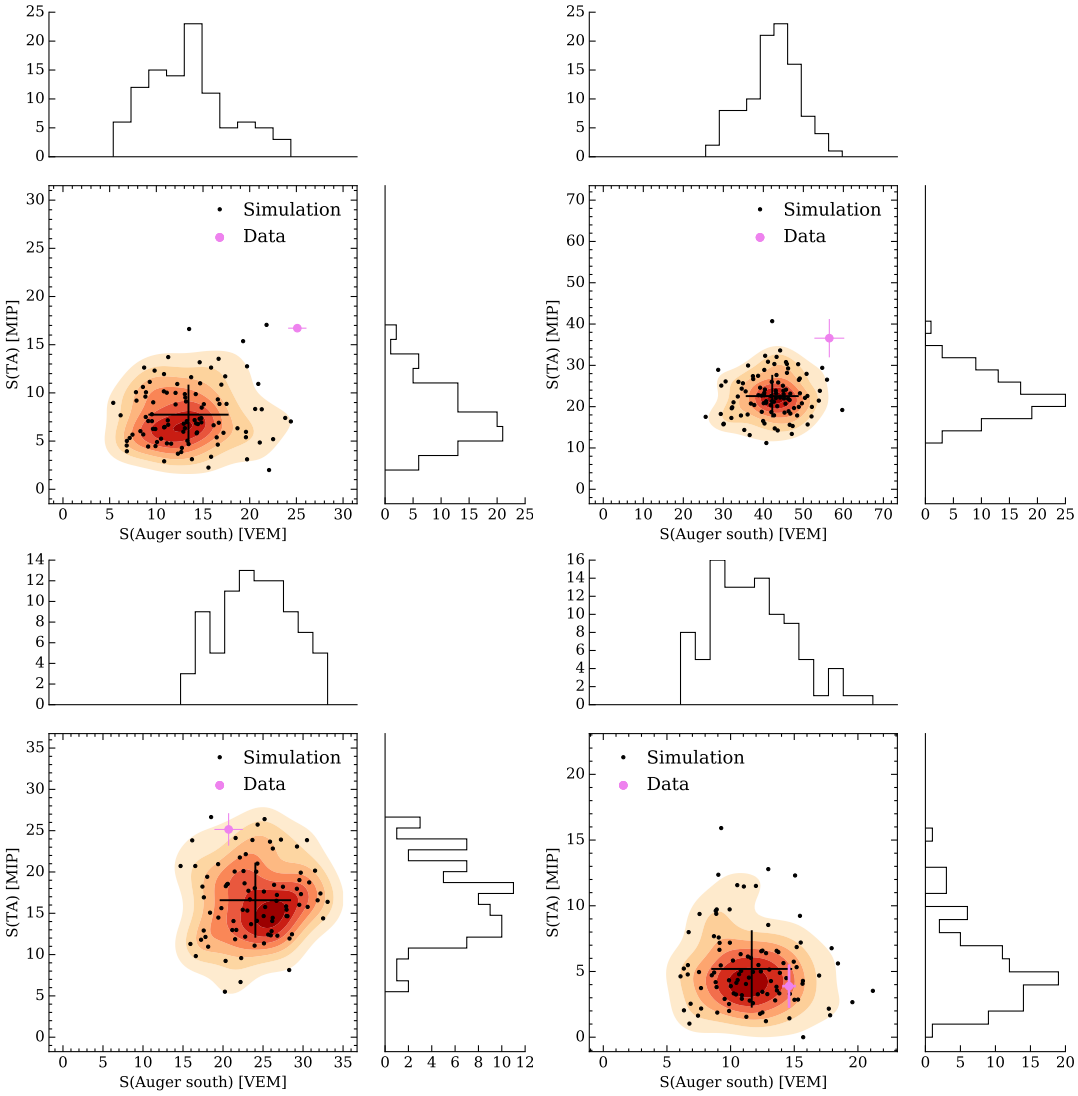


Figure 6.20: Extension of Figure 6.19 for remaining events. Top left: event 7, top right: event 9, bottom left: event 10, bottom right: event 11.

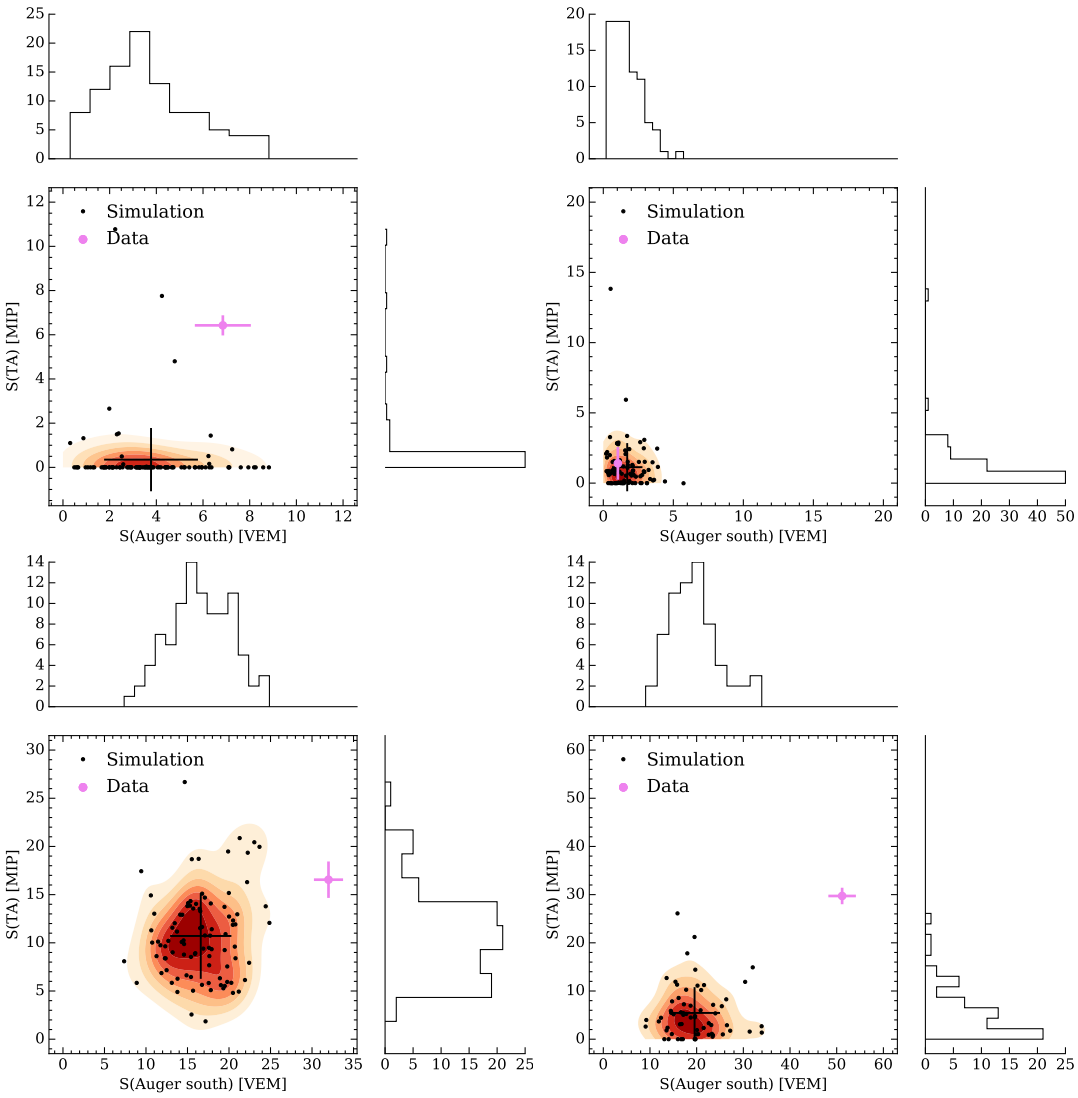


Figure 6.21: Extension of Figure 6.20 for remaining events. Top left: event 12, top right: event 13, bottom left: event 14, bottom right: event 16.

ID	E [EeV]	θ [°]	φ [°]	ρ [m]	Δ_v [VEM]	Δ_v^r [%]	Δ_m [MIP]	Δ_m^r [%]
1	1.52	29.62	194.68	851	▽	▽	▽	▽
2	1.08	37.80	64.58	906	6(2)	80(30)	1(2)	40(50)
3	3.42	23.40	154.75	1193	8(4)	70(30)	1(3)	20(80)
4	4.25	43.52	173.85	1014	7(4)	50(30)	3(3)	50(60)
5	3.61	40.74	87.15	687	-26(10)	50(10)	-19(10)	70(20)
6	4.46	16.54	174.16	202	•	•	•	•
7	4.57	38.44	222.90	985	12(0)	60(30)	9(3)	70(30)
8	3.60	30.27	233.02	1597	×	×	×	×
9	4.58	38.28	216.69	640	14(10)	30(20)	14(10)	50(20)
10	2.05	37.44	78.46	792	-3(5)	30(20)	9(5)	20(30)
11	10.18	35.30	38.73	1276	3(3)	20(30)	-1(3)	30(70)
12	7.57	22.23	102.83	1588	3(2)	60(50)	6(1)	200(100)
13	4.58	28.50	25.66	1656	-1(1)	50(60)	0(2)	20(200)
14	3.30	5.41	202.20	1001	15(0)	60(20)	5(4)	40(40)
15	1.23	30.59	55.96	1561	▽	▽	▽	▽
16	2.22	38.66	304.97	734	30(10)	90(20)	20(10)	140(50)

Table 6.3: Observed and simulated data (proton) for two quarters of project operation. Absent data is described by the following symbols: • for PMT saturation, □ no observed Auger data, ▽ no observed TA data, × simulation reconstruction fails to trigger. Total uncertainty is given in parentheses.

ID	E [EeV]	θ [°]	φ [°]	ρ [m]	Δ_v [VEM]	Δ_v^r [%]	Δ_m [MIP]	Δ_m^r [%]
1	1.52	29.62	194.68	851	▽	▽	▽	▽
2	1.08	37.80	64.58	906	6(2)	70(30)	3(2)	70(60)
3	3.42	23.40	154.75	1193	2(4)	10(30)	-2(3)	30(40)
4	4.25	43.52	173.85	1014	6(4)	40(30)	3(3)	50(50)
5	3.61	40.74	87.15	687	-20(10)	40(10)	-10(10)	40(20)
6	4.46	16.54	174.16	202	•	•	•	•
7	4.57	38.44	222.90	985	11(0)	50(30)	11(4)	90(40)
8	3.60	30.27	233.02	1597	-1(1)	70(60)	1(1)	60(90)
9	4.58	38.28	216.69	640	7(8)	10(20)	14(10)	50(30)
10	2.05	37.44	78.46	792	-1(6)	0(30)	5(3)	50(40)
11	10.18	35.30	38.73	1276	1(4)	10(30)	-2(4)	40(70)
12	7.57	22.23	102.83	1588	1(3)	10(40)	4(2)	90(60)
13	4.58	28.50	25.66	1656	-1(1)	80(50)	0(1)	40(130)
14	3.30	5.41	202.20	1001	12(0)	50(20)	5(4)	40(30)
15	1.23	30.59	55.96	1561	▽	▽	▽	▽
16	2.22	38.66	304.97	734	30(10)	80(20)	20(0)	100(30)

Table 6.4: Observed and simulated data (nitrogen) for two quarters of project operation. Absent data is described by the following symbols: • for PMT saturation, □ no observed Auger data, ▽ no observed TA data, × simulation reconstruction fails to trigger. Total uncertainty is given in parentheses.

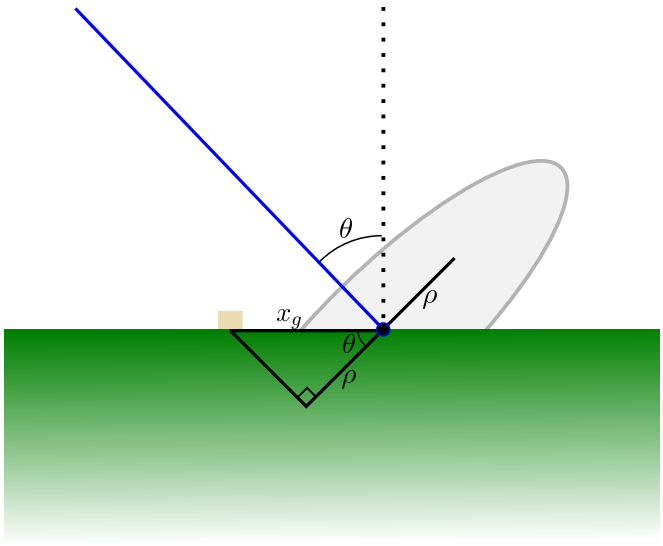


Figure 6.22: Simplified shower geometry to calculate the core distance.

hadronic interaction models, like QGSJETII-04 used in this study, under-predict the number of muons at the ground [212]. In future studies the following prescription will be followed:

- Scale provided TA FD energy by 1/1.104. This brings the energy into the Auger FD energy scale.
- Scale the Auger FD energy by 1.09 to account for the muon deficit in the QGSJET interaction model [212].
- After $\overline{\text{Off}}\underline{\text{line}}$ reconstruction, scale the *muon* signal by $1.59 \times 1.09^{0.9}$ to retrieve the actual VEM signal a WCD would record on the ground.

6.15 Discussion

Since we are dealing with very small statistics for our derived quantities, making finalized, definitive statements is problematic. It must also be stressed the analysis done here is *indirect*. The results should be considered preliminary until a larger data set is assembled. Table 6.3 shows a deficiency for the predicted VEM signal in a majority of events. The relative difference does not cover zero within uncertainty for 9 of 13 events. For MIP signals the relative difference covers zero within uncertainty for 7 of 13 events. There do not appear to be any strong trends between simulated signal discrepancy and shower parameters (e.g. θ, ρ, etc.). Heavier primaries to seem to bring the simulation distribution closer to the observed value, however the improvement is not dramatic. Moreover, at these energies a nitrogen primary is not favored by recent composition studies published by Auger and TA. The heavier primary seems to slightly improve the VEM agreement but at the expense of MIP agreement, suggesting that this shower parameter is likely not the source of the true underlying discrepancy. In future simulations the discrepancy between predicted and observed VEM signals might diminish after correcting for the under-predictions of muons at the ground by the QGSJET high-energy interaction model.

6.16 Empirical models for MIP and VEM based signals

Following the previous analysis sections which show a variety of ways to compare and contrast the datasets, we can use the data to make an empirical relation for

conversion of the two signal types. A variety of models are investigated for the raw data and a quality cut set.

6.16.1 Linear fit

As shown in a previous section from simple arguments, the signals produced by idealized TA and Auger detectors should be roughly linear and are largely zenith angle dependent. Since both signal types have uncertainty instead of using an ordinary least squares (OLS) fitting algorithm, I've opted for an orthogonal distance regression (ODR) fit which can accommodate the x and y error bars using the `ORDPACK` library [191]. The fit to the raw data is shown in Figure 6.23 along with the residuals. Saturated events are excluded, leaving 97 data points in the sample. The fit function is simply

$$y = a_1 x + a_2 \tag{6.12}$$

I also perform a fit for a subsample of the 97 points which pass the following quality cut

$$S(\text{TA}) < 3 \text{ [MIP]} \tag{6.13}$$

$$S(\text{Auger}) < 3.2 \text{ [VEM]} \tag{6.14}$$

Therefore, in a MIP-VEM coordinate pair, if any of the conditions in Equation 6.13 are true, then that pair won't be used in the fit. These threshold values correspond to the minimum hardware levels needed to promote a station level trigger (L1 for TA and T2 for Auger). Although it's still possible to find an integrated signal less than these values due to averaging over PMTs and fluctuations in calibration histograms,

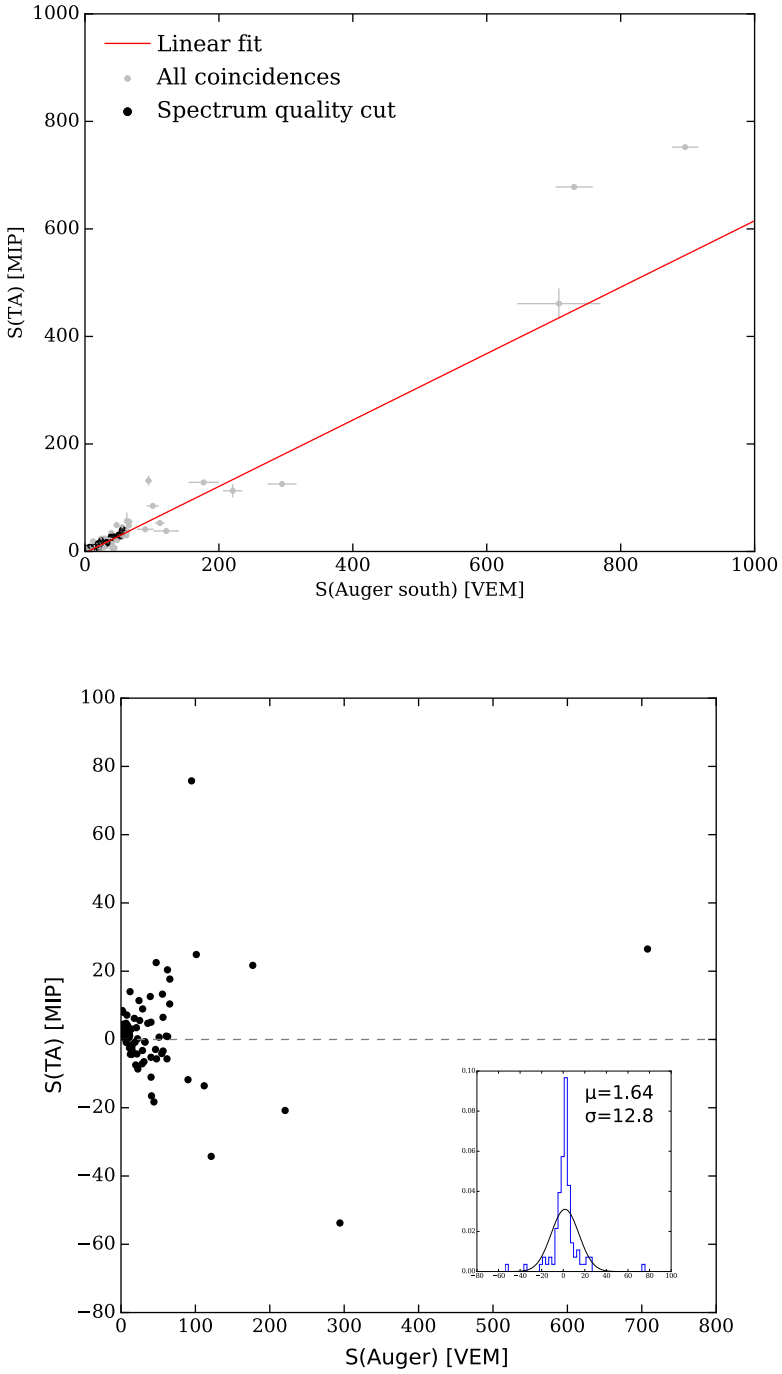


Figure 6.23: Top: Data and linear fit. Bottom: Residuals from fit.

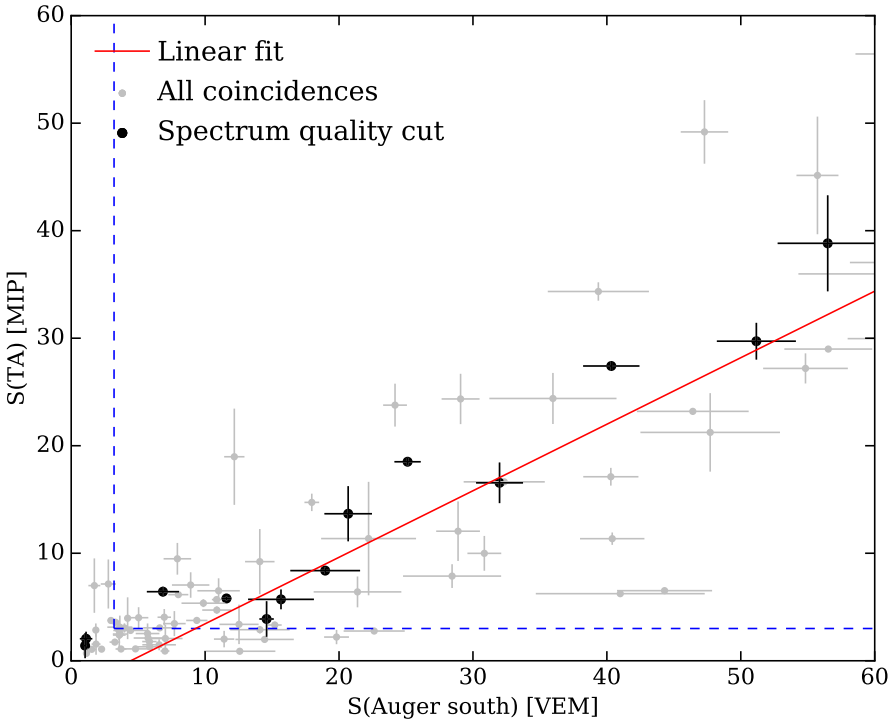


Figure 6.24: Similar to Figure 6.23, but the fit is only applied to points passing the threshold quality cut, the region enclosed by the dashed blue lines. Also, the plot axes have been artificially reduced to show the threshold cuts.

it's a reasonable practical lower limit to set. A plot of the threshold data is shown in Figure 6.24. The linear fit data for both cases is available in Table 6.5.

6.16.2 Piecewise linear fit

There is also motivation for a piecewise linear fit to capture regions which might truly have different slopes. As was shown in the primitive effective area model, the slope of a MIP-VEM curve is expected to change as a function of zenith angle. An additional complication is core distance dependence, which is completely ignored

in the geometrical model. Although a model wasn't presented which takes both effects into account, it seems plausible that they would lead to two, or perhaps more, MIP-VEM regions with different slopes. A piecewise linear model can describe this behavior

$$y = \begin{cases} a_1 x + a_2 - a_1 a_3 & \text{if } x < a_3 \\ a_4 x + a_2 - a_4 a_3 & \text{if } x \geq a_3 \end{cases}$$

A fit to this function of the raw data, along with residuals are shown in Figure 6.25. The fit to threshold cut data is displayed in Figure 6.26.

6.16.3 Power law fit

Although we expect roughly linear behavior the interplay between zenith angle and core distance could lead to some interesting non-linear effects. These perturbations shouldn't result in a highly unusual dependence like $\propto x^4$, but could lead to an exponent in the neighborhood of 1 instead of exactly 1. The plausibility of this model is investigated by fitting

$$y = a_1 x^{a_2} \tag{6.15}$$

to raw and threshold cut data. The raw data fit with residuals is shown in Figure 6.27. See Figure 6.27 for the threshold cut fit.

6.16.4 Broken power law fit

A broken powerlaw can be used to investigate departures from linearity in different regimes. Also, one advantage for this function is that a threshold can automatically be

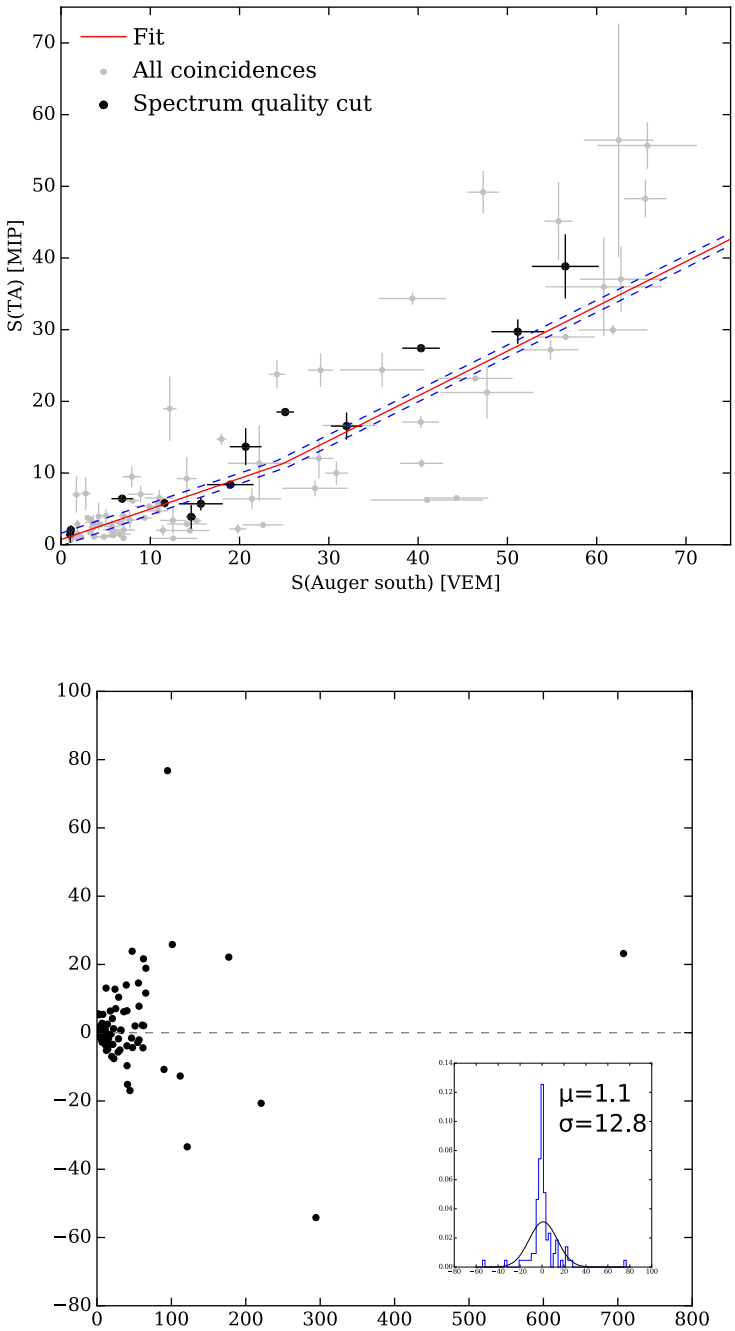


Figure 6.25: Top: Raw data fit with a piecewise line. The plot is limited to max values of 75 MIP and VEM so the breakpoint is visible. Dashed blue lines are the 95% prediction interval. Bottom: Residuals and their distribution properties.

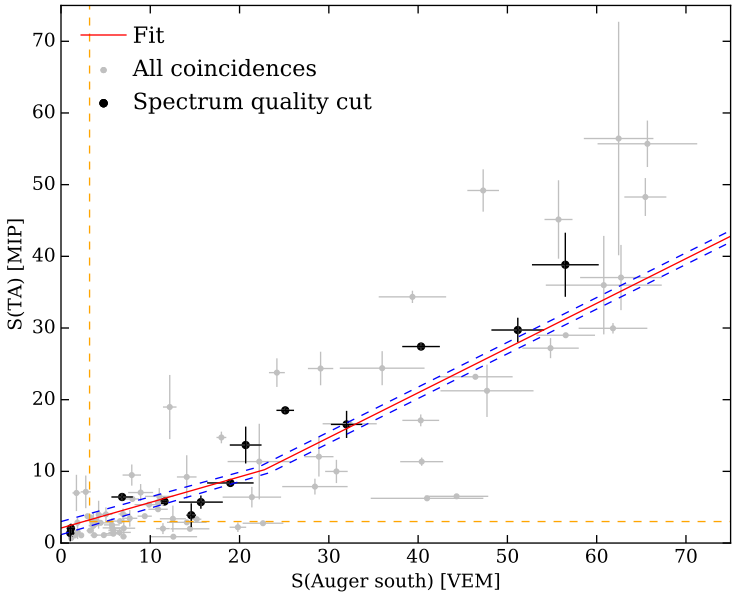


Figure 6.26: Similar to Figure 6.25 but the fit is applied only to data which pass the quality cut. Dashed blue lines are the 95% prediction interval for the fit and the region in the dashed orange box correspond to data above the threshold cut.

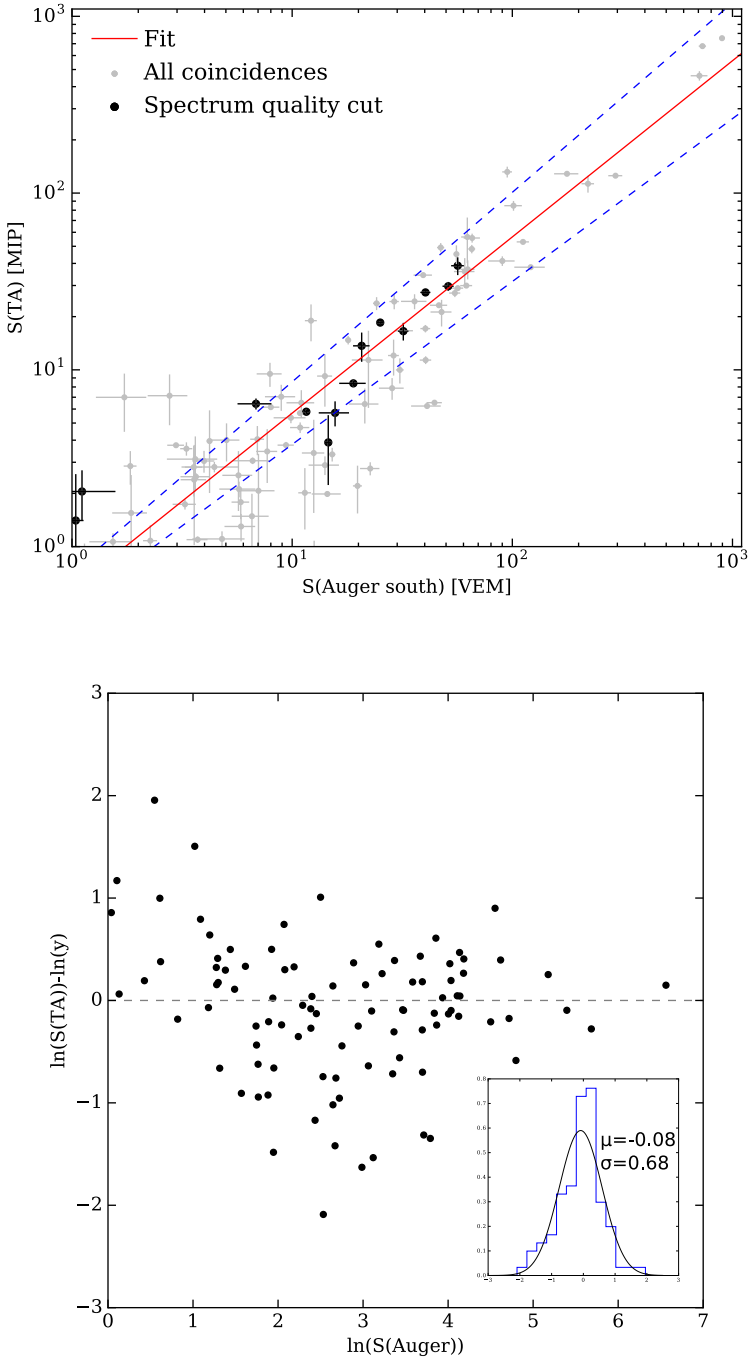


Figure 6.27: Top: Raw data fit with a two parameter powerlaw function. Dashed blue lines are the 95% prediction interval. Bottom: Residuals and their distribution properties.

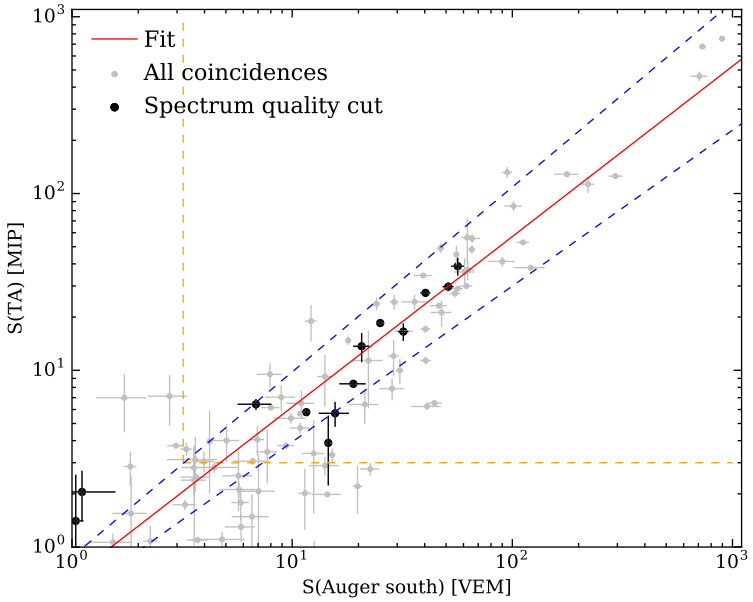


Figure 6.28: Similar to Figure 6.27 but the fit is applied only to data which pass the quality cut. Dashed blue lines are the 95% prediction interval for the fit and the region in the dashed orange box correspond to data above the threshold cut.

determined through the parameter optimization process instead of manually enforcing Equation 6.13. The piecewise linear function is converted below

$$y = \begin{cases} e^{a_2 - a_1 a_3} x^{a_1} & \text{if } x < a_3 \\ \\ e^{a_2 - a_4 a_3} x^{a_4} & \text{if } x \geq a_3 \end{cases}$$

The fit and residuals for raw data can be found in Figure 6.29. It was found that the algorithm fails to converge when using the broken powerlaw and threshold quality cut. This suggests there are no serious deviations from linear behavior for large signals.

6.16.5 Fit parameters and discussion

From the model fitting I find no evidence for significant non-linearities in the data. This is because the power law models returned exponents of 1 within uncertainty, and remained close to 1 even when noisy was introduced. Moreover, when the algorithm was allowed to use separate components to fit the noisy region and larger signal region with a broken power law, the exponent in this region was still close to 1. It was further discovered that the algorithm could not converge for two separate components in the large signal region when a threshold cut was introduced, suggesting that this region is described by a single, linear component. The slopes between the powerlaw fit and linear fits also agree within uncertainty. The intercepts of the linear fits are non-zero, but there is no physical reason to accept this as a real feature, and is likely an artifact of the fit. From this study we've learned that a reasonable empirical model for converting between MIP and VEM based signals is

$$S_{\text{MIP}} = \left(0.55^{+0.12}_{-0.10}\right) S_{\text{VEM}}^{1.00 \pm 0.08} \tag{6.16}$$

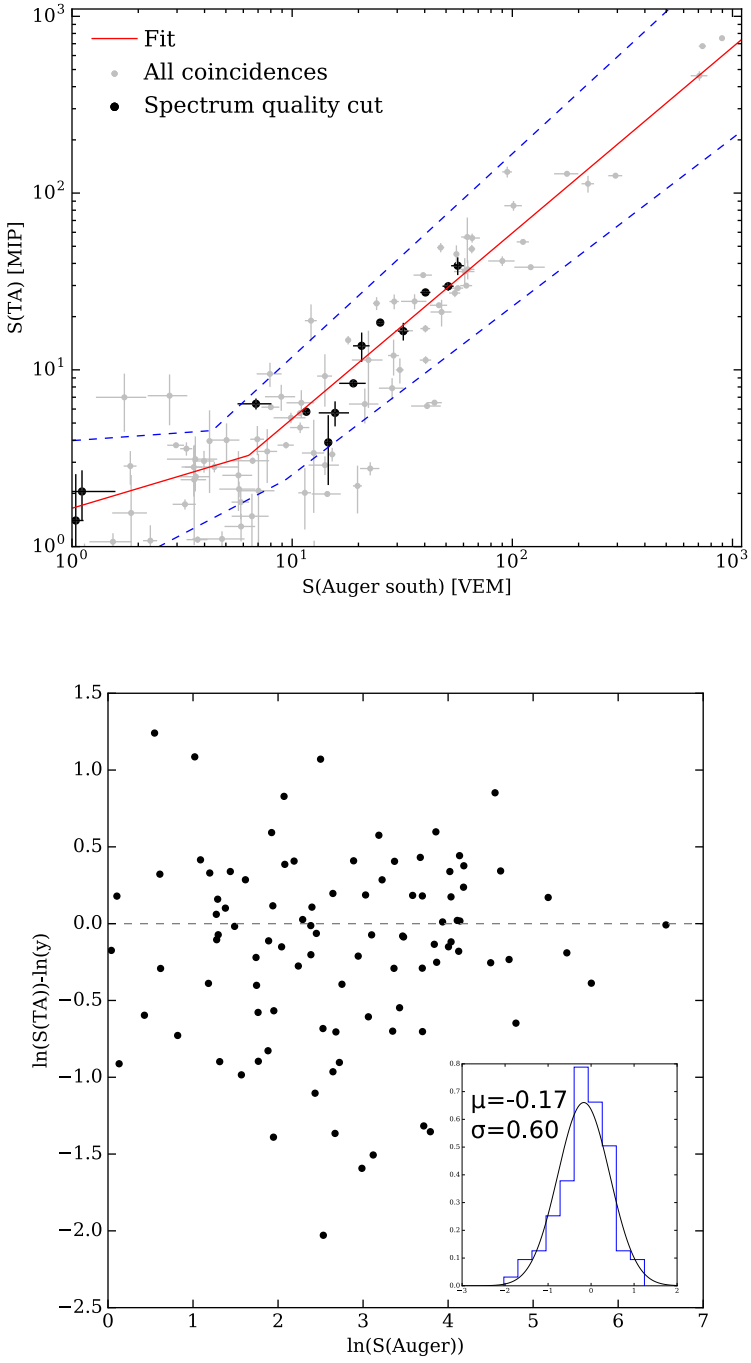


Figure 6.29: Top: Raw data fit with a four parameter broken powerlaw function. Dashed blue lines are the 95% prediction interval. Bottom: Residuals and their distribution properties.

Fit function	a_1	a_2	a_3	a_4	Threshold cut?
linear	0.6176 ± 0.0004	-2.551 ± 0.001			no
linear	0.6187 ± 0.0004	-2.75 ± 0.04			yes
piecewise linear	24.9 ± 0.7	11.3 ± 0.4	0.426 ± 0.008	0.6247 ± 0.0004	no
piecewise linear	22.8 ± 0.7	11.3 ± 0.4	0.426 ± 0.008	0.6247 ± 0.0004	yes
powerlaw	$0.55^{+0.12}_{-0.10}$	1.00 ± 0.08			no
powerlaw	$0.67^{+0.21}_{-0.16}$	0.96 ± 0.08			yes
broken powerlaw	$6.4^{+2.9}_{-2.1}$	$3.3^{+1.2}_{-0.9}$	0.37 ± 0.28	1.05 ± 0.10	no
broken powerlaw				$0.96\pm$	yes

Table 6.5: Fit parameters for a variety of functions with and without threshold cuts applied to the data. Uncertainties are 1σ values.

The expression can be inverted to also convert MIP to VEM

$$S_{\mathrm{VEM}} = \left(1.81^{+0.41}_{-0.33}\right) S_{\mathrm{MIP}}^{1.00\pm0.08} \tag{6.17}$$

where S_{MIP} is an integrated signal in MIP units and S_{VEM} is an integrated signal in VEM units.

6.17 Trace comparison study

From discussions with TA collaborators and colleagues at the recent 2017 ICRC I discovered there is some interest in comparing the waveforms of the Auger South station ("Cosmo") and the co-located TA SD ("DET2421"). A quick visual inspection of the traces might lead one to speculate about particle counting discrepancies, see Figure 6.30. While an exact match is not expected for myriad reasons the differences

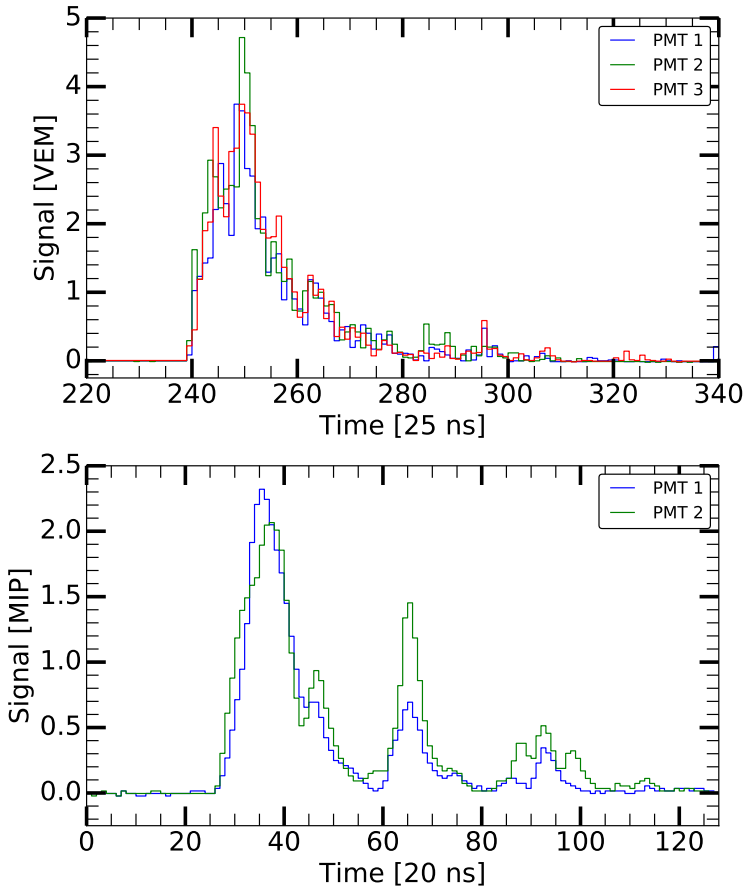


Figure 6.30: Example Auger south and DET2421 waveforms. The x-axis shows elapsed time in terms of FADC bins. At a quick glance it seems as though the TA trace has more structure, but one must be cautious when doing a direct comparison since different sampling rates are used.

in traces could range from detector sensitivities to the circuitry used for digitizing the PMT signals to the PMTs themselves. In this section I attempt to quantify the divergence of traces. The results are interpreted conservatively with an eye toward serious and large discrepancies.

6.17.1 Up-sampling Auger South traces

As mentioned in Chapter 3, Auger and TA have implemented different sampling frequencies in their electronics. In TA the FADCs run at 50 MHz while in Auger these are 40 MHz. Thus for the same time window, a TA station will collect 1.25× more samples than the Auger station. In order to make a fair comparison of traces, it's necessary to either down-sample TA, up-sample Auger, or simultaneously down-sample TA while up-sampling TA (e.g. interpolating both to 45 Hz).

I've made the ad-hoc choice to up-sample the Auger waveform. This is achieved using polyphase filtering. The algorithm up-samples the waveform to 960 bins, applies a zero-phase low-pass finite impulse response (FIR), and down-samples to the original length of 768 bins, giving a trace sampled at 960/768 = 1.25 the original rate of 40 MHz. A Kaiser window is used:

$$w[n] = \begin{cases} \frac{I_0\left(\pi\alpha\sqrt{1-\left(\frac{2n}{N-1}-1\right)^2}\right)}{I_0(\pi\alpha)} & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases} \tag{6.18}$$

where $\alpha = 5$ is the window shape parameter, N is the sequence length, I_0 is the zeroth-order modified Bessel function of the first kind. This technique is used, compared to a traditional fast Fourier transform (FFT), to preserve the temporal locations of the original waveform. The FFT method can introduce lags/delays and other ringing distortion to pulse features, which we wish to avoid. An example up-sampling for an event waveform recorded on Jan. 2017 17:04 UTC is shown in Figure 6.31. The high gain channel of a single PMT is shown, after the baseline has been subtracted.

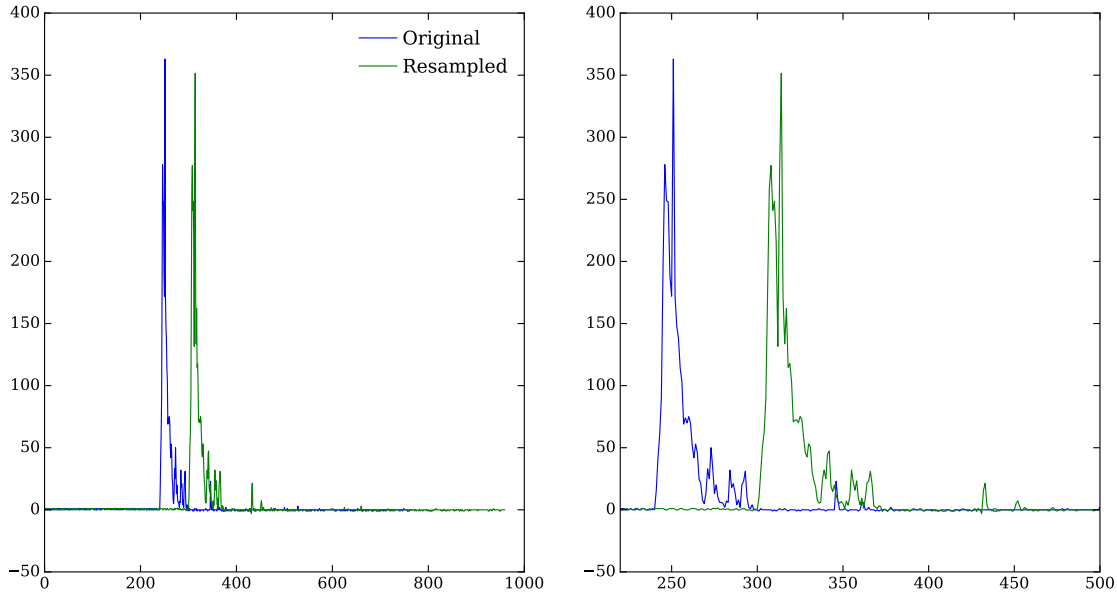


Figure 6.31: Original and re-sampled FADC traces. The y-axis is counts and the x-axis is time. For the blue curve, each time bin is 20 ns. For the green curve, each time bin is 25 ns. The left and right plots are identical, but the right plot shows a region of interest around the pulses.

6.17.2 Aligning the Auger and TA traces

In order to calculate a similarity parameter, the traces must first be aligned. This is accomplished by finding the optimal lag time, τ', from a peak in the cross-correlation function

$$C(\tau) = \sum_n f(n+\tau)\, g^*(n) \tag{6.19}$$

where the asterisk is the complex conjugation operator. Since our functions are real $g^* = g$.

The original TA waveform is sampled over 128 bins. Before the cross-correlation function is calculated, the sequence is zero padded on both ends to bring the total bin count to 960. A plot showing the example PMT traces and their normalized

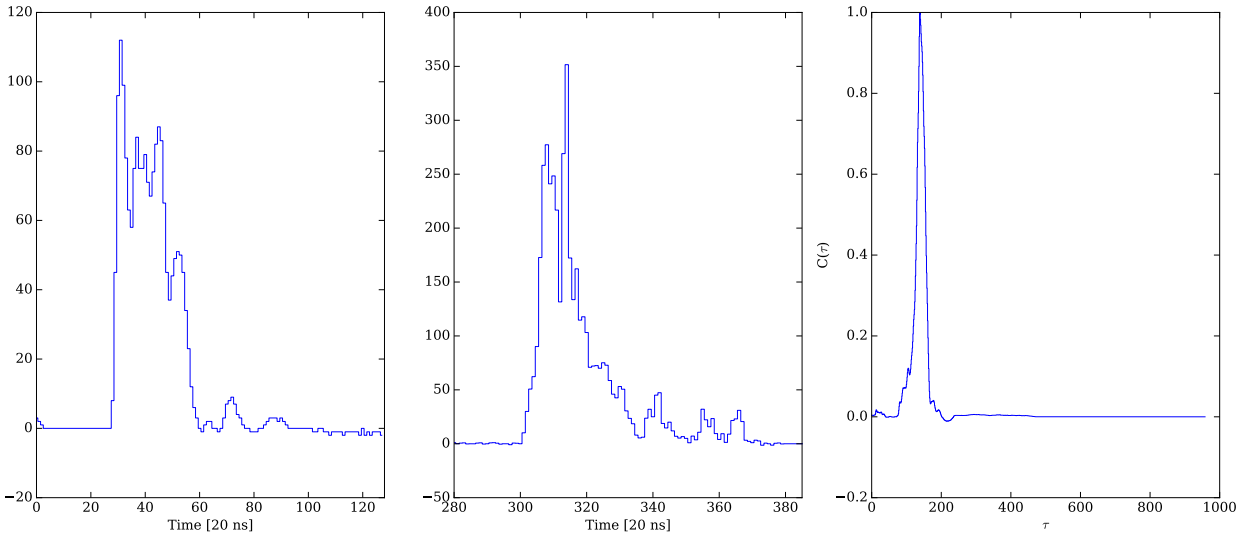


Figure 6.32: Left: TA PMT trace. Middle: Up-sampled Auger South PMT trace. Right: Normalized cross-correlation function for the pair.

correlation function is shown in Figure 6.32.

The peak of the cross-correlation function is the optimal lag time to apply to the leading function in order to align the traces. In this example $\tau' = 139$,so the traces are aligned by shifting the leader back 139 bins. The result of this procedure is shown in Figure 6.33.

Further examples of this procedure in action are shown in Figure 6.34.

6.17.3 Comparing the traces

There are a variety of possible ways to quantify the similarity of functions. If these were probability distributions, one could perform a KS test, Anderson-Darling test, Kullback-Leibler divergence, etc. However, I don't believe it's appropriate to treat the traces as probability distributions since very few samples are made of the

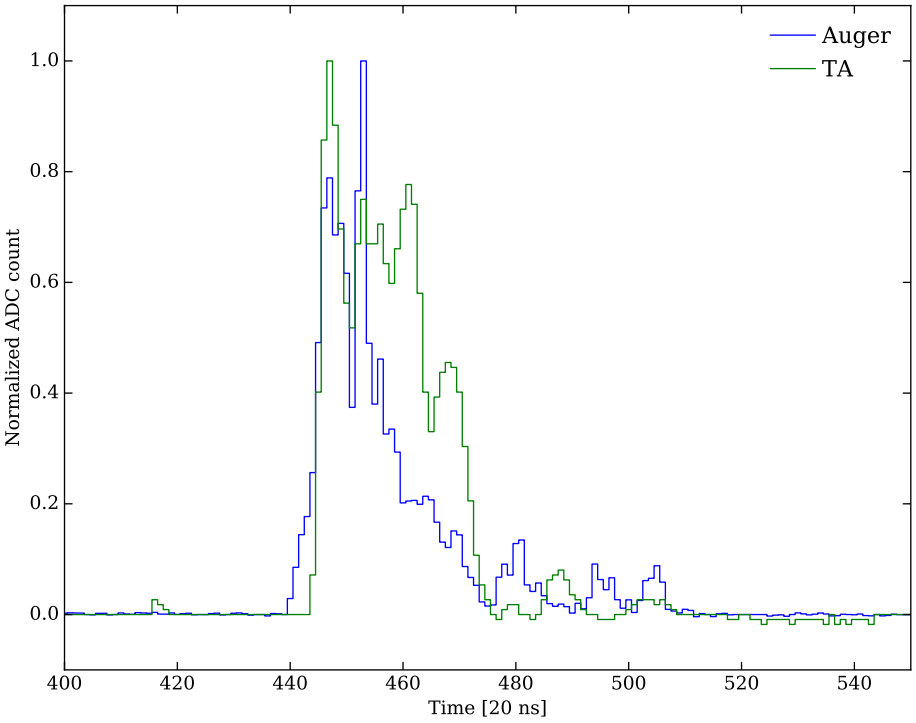


Figure 6.33: Traces after alignment. Each trace has been normalized by its own peak value.

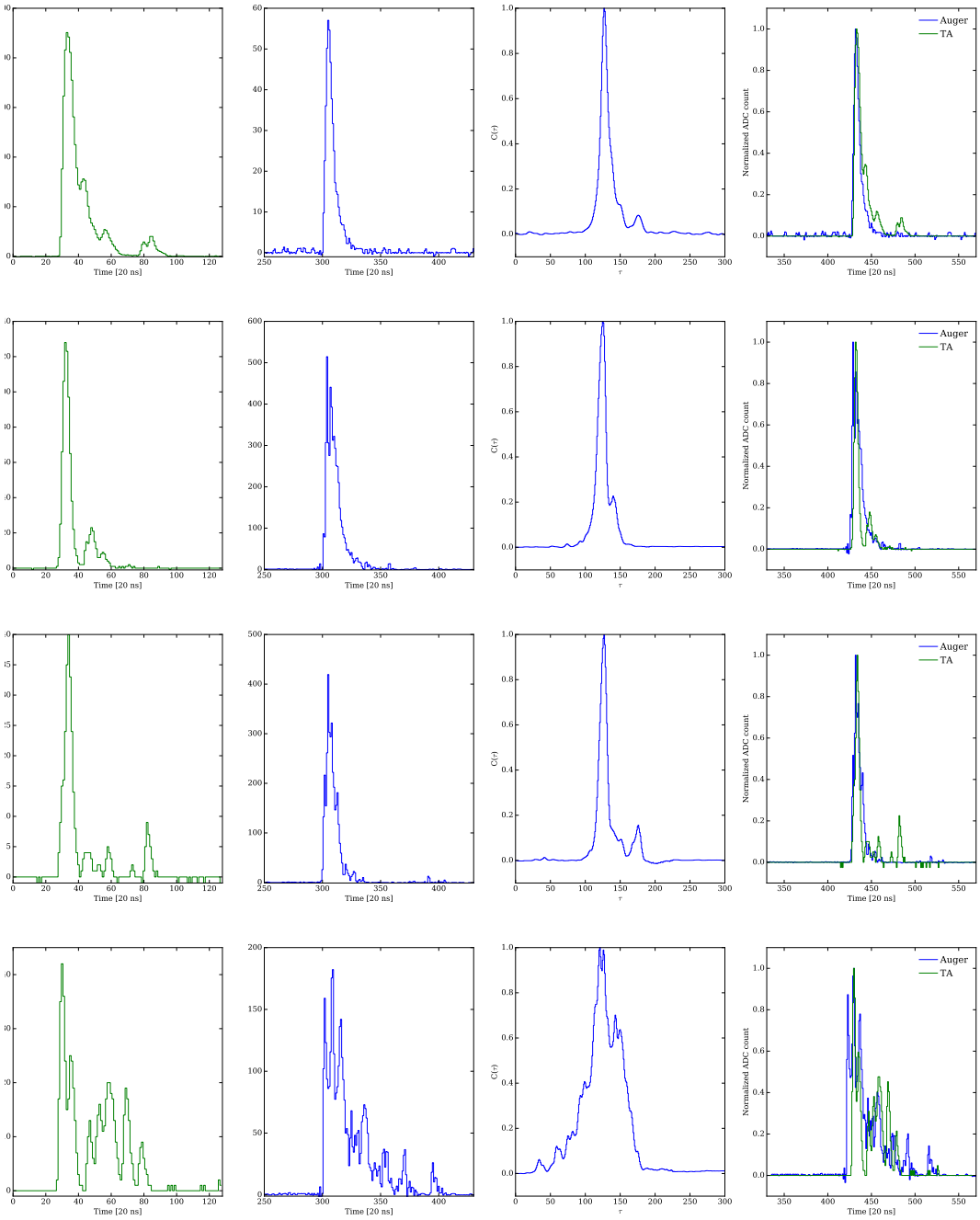


Figure 6.34: Additional examples of trace alignment procedure and the cross-correlation function.

event. Since Auger's PMTs observe the same detection volume one could conceivably build a trace "histogram" from the 3 waveforms, but you are in the small statistics regime. In the case of TA, you only have 1 sample per detection volume.

Instead I believe it's more appropriate to treat the traces as two n-vectors, and ask how similar or correlated these vectors are. In this sense, features of the correlation function from the previous section provides some measure of similarity.

Since there is no universal metric one can use to judge similarity, I consider four measures for diversity:

- cosine distance: a measure of directional similarity $s_1 = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}$ where $|\cdot|$ refers to the Euclidean norm.

- Euclidean distance: a measure of spatial proximity. I use the definition of $s_2 = \frac{1}{1+d}$ where $d = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$

- Generalized Jaccard index: a measure of set overlap $s_3 = J(\boldsymbol{a}, \boldsymbol{b}) = \frac{\sum_i \min(a_i, b_i)}{\sum_i \min(a_i, b_i)}$

- dynamic time warping (DTW), s_4: an algorithm which performs a non-linear "stretching" transformation to a common set of times such that the total Euclidean distance is minimized. The algorithm returns this distance and the set of warping paths which achieves it. The warping path for the example data is shown in Figure 6.35. The algorithm used for this metric can be found here.

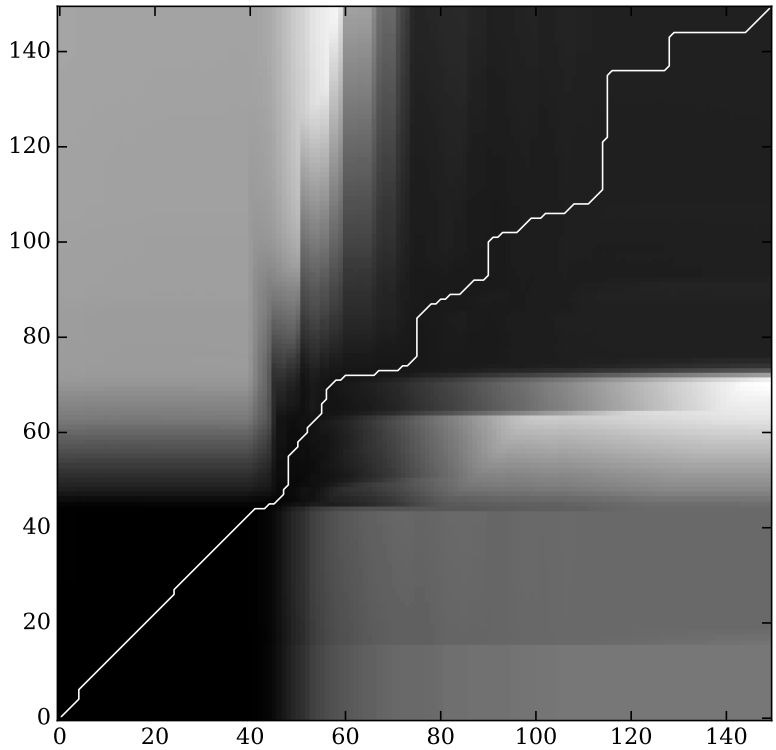


Figure 6.35: Warping path which minimizes the accumulated cost function. I adopt the Euclidean distance as the cost measure.

6.17.4 Results and discussion

For the 97 unsaturated events the traces are in general similar. This can be seen from the consistent values from the four metrics. For s_1 more than half of the events report similarities of > 0.7. Although the s_2 distribution is not as close to 1, this metric could be considered more robust since it's absolute deviation instead of a direction based. The reported values for data in the range of 0.3-0.5 are *much* higher than you would find for random traces (found to be a factor of 10 lower in simulations). The generalized Jaccard covers the range of 0-0.5 suggesting some degree of similarity. Negative values are caused by low signal conditions where FADC fluctuations around the baseline. The DTW results (s_4) also support a majority similar interpretation since a large fraction of the traces have small normalized distances relative to the rest of the population.

A table of the similarities using the four measures along with the optimal correlation lag for the example plots in Figure 6.34 is provided in Table 6.6. For clarity, metrics 1–3 use a scale from 0-1 where 0 is completely dissimilar and 1 is an exact match.

6.18 Estimating the muon component for reconstructed events

One of the principal goals of the Auger upgrade is the ability to discriminate EAS components with the SD. This would permit separating the muonic and electromagnetic components for a large range of shower parameters [184]. This allows for a more precise measurement of the muon density, and the reader will recall from an

Trace ID	s_1	s_2	s_3	s_4
8	0.938	0.527	0.446	4.5×10^{-5}
64	0.811	0.423	0.458	1.1×10^{-4}
76	0.885	0.499	0.4670	9.9×10^{-5}
91	0.773	0.332	0.430	0.0004

Table 6.6: Metrics for some example events.

introductory chapter how this can be used to very accurately determine the $X_{\max}$ parameter. This parameter is critical in shower composition studies, and its improved measurement with the SD has some dramatic implications such as: assignment of primary charge on an event-by-event basis (extremely useful for backtracking studies) and significantly increasing the overall $X_{\max}$ statistics, especially at the highest energies due to the high duty cycle of the SD [184]. While the Auger@TA setup is not an exact replication of the upgraded Auger SSD, the TA design is similar enough that we should be able to look at a preliminary analysis of the methods employed by the Auger upgrade to decompose the shower components.

A technique pioneered by Letessier-Selvon & Billoir et al. uses a matrix inversion formalism to separate the EM and muonic components based on the linear superposition of these signal types into separate layers of a detector [190]. The method was originally derived for a segmented WCD type design and has recently been adapted

to accommodate a top scintillator layer [184]. The formalism is

$$\begin{bmatrix} S_{\mathrm{SSD}} \\ \\ S_{\mathrm{WCD}} \end{bmatrix} = \begin{bmatrix} \lambda A_{\mathrm{SSD}} & A_{\mathrm{SSD}} \\ \\ \beta A_{\mathrm{WCD}} & A_{\mathrm{WCD}} \end{bmatrix} \begin{bmatrix} F_{\mathrm{em}} \\ \\ F_{\mu} \end{bmatrix} \tag{6.20}$$

with the SSD signal in MIP units and WCD in VEM units and where

$$F_{\mathrm{em}} = \left(\frac{S_{\mathrm{SSD}}}{A_{\mathrm{SSD}}} - \frac{S_{\mathrm{WCD}}}{A_{\mathrm{WCD}}} \right) \Big/ \lambda - \beta \tag{6.21}$$

and

$$F_{\mu} = \left(\lambda \frac{S_{\mathrm{WCD}}}{A_{\mathrm{WCD}}} - \beta \frac{S_{\mathrm{SSD}}}{A_{\mathrm{SSD}}} \right) \Big/ \lambda - \beta \tag{6.22}$$

with β being a zenith dependent factor, and λ is related to the average energy deposited into the scintillator by EAS components in terms of a MIP/VEM (also zenith dependent) ratio. Simulation estimates this to be roughly 2.8β [184]. The final expression for the muonic signal contribution, after matrix inversion and applying in the WCD is

$$S^{\mu}_{\mathrm{WCD}} = \delta S_{\mathrm{WCD}} - \gamma S_{\mathrm{SSD}} \tag{6.23}$$

Air shower simulations show δ is mostly stable for a variety of zenith angles, core distances, primary energies and composition with a value of $\delta \approx 1.47$. The γ value is more sensitive to these shower parameters covering a range of 1.1–1.3. For this analysis I adopt $\gamma = 1.2$. An additional correction must be made to the γ since this involves the ratio of detector areas: since the TA SD area is 0.78 that of the SSD, after the appropriate scaling a pre-factor of 1.282 should be applied to γ. For this study I adopt the new value $\gamma' = 1.538$ giving a final expression of

$$S^{\mu}_{\mathrm{WCD}} = 1.47 S_{\mathrm{WCD}} - 1.538 S_{\mathrm{TA}} \tag{6.24}$$

For the 97 non-saturated events I examine the ratios of $f_\mu = \frac{S^\mu_{\rm WCD}}{S_{\rm WCD}}$ to determine the muon fraction of the signal. For 12 (12.4%) ratios $f_\mu < 0$ and 21 (21.6%) ratios were found with $f_\mu > 1$, both of which are non-physical. A histogram showing the remaining 64 ratios is in Figure 6.36. The population is somewhat uniform, indicating the the Auger@TA setup observed a range of different showers. There is an interesting excess around $f_\mu = 0.8$ which could be a large sub-population of large core distance or highly inclined showers. For the 13 global events with reconstructions available 2 ratios are > 1 and 1 is < 0, leaving a total of 10 available, also shown in Figure 6.36. For these events there is about an event split between muonic and EM dominated signals. The ratios are also shown as a function of core distance and zenith angle in Figure 6.37. We see the signals are mostly EM component close to the core and gradually become dominated by muons around $r \approx 1000$ m. It's difficult to interpret the zenith angle results since most of the data is clustered around 38° , but there doesn't seem to be a strong functional dependence.

While the data set is small the results seem to validate the matrix inversion formalism. In particular, the observed muon ratio of around 0.75 at 1 km from the core appear to be in good agreement with extensive simulations studies [213, 214] and the rate of increase vs. core distance is also in rough agreement. While preliminary, these results demonstrate the viability of the proposed Auger upgrade while also providing some insights on expected performance of the detector.

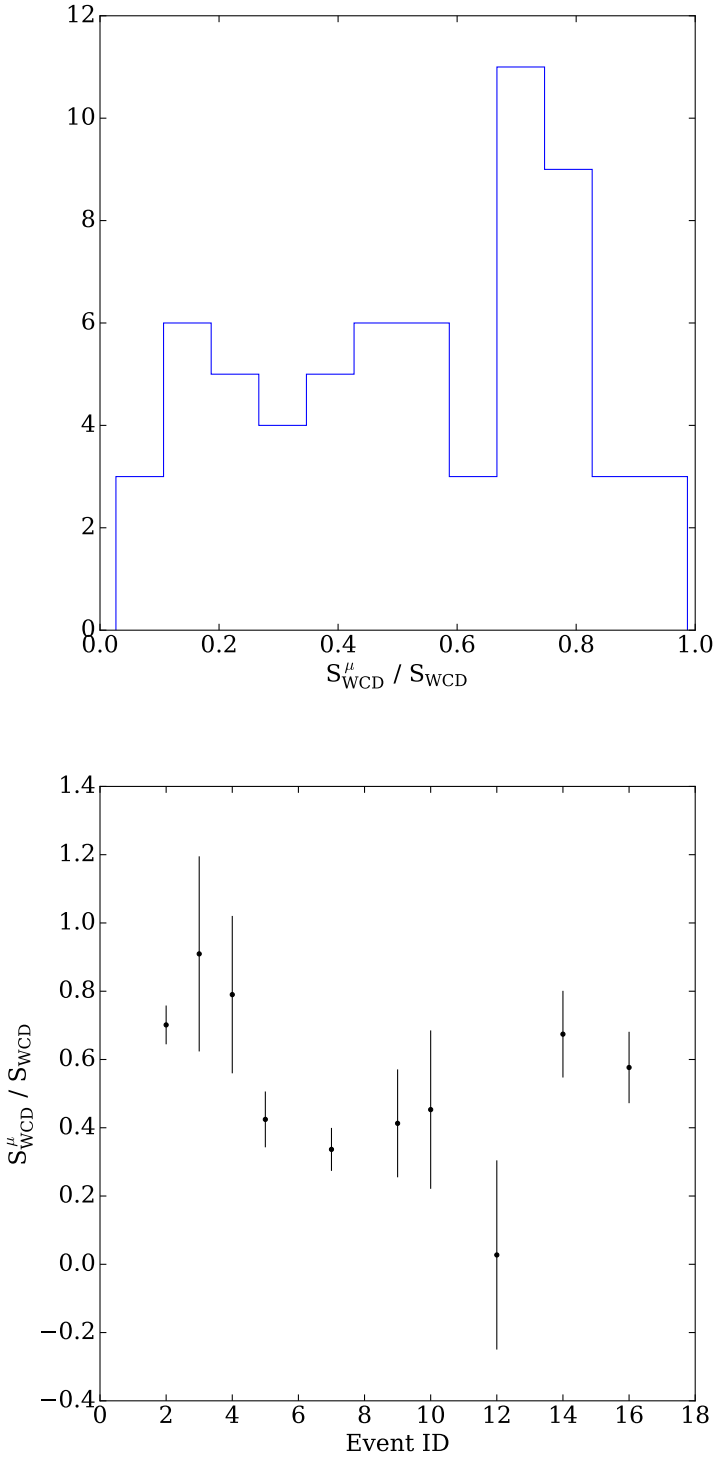


Figure 6.36: Top: Muon fractions for all global trigger events. Bottom: Muon fractions for reconstructed global events. Errorbars are from standard propagation techniques where it's assumed errors between Auger and TA signals are independent.

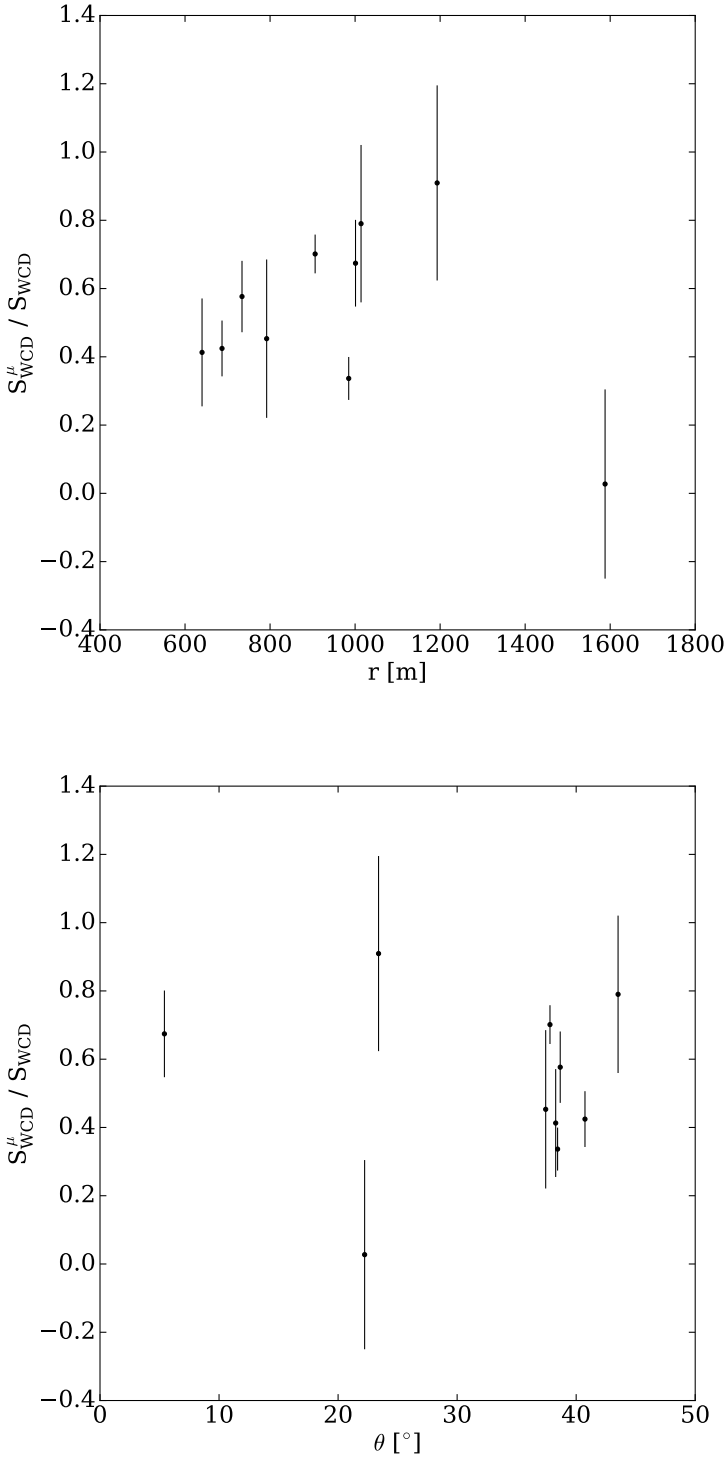


Figure 6.37: Top: Muon fraction vs. core distance. Bottom: Muon fraction vs. zenith angle.

Chapter 7: The Auger@TA Project: upgrades and future plans

7.1 Preamble

This chapter discusses near term improvements that will be made to the setup to improve performance, especially for the local trigger. Details of longer plans and their motivation are also presented.

7.2 New prototype "smart" trigger

The current local trigger design uses a simple threshold comparator circuit which measures the peak amplitude of PMT waveforms. This results in a bias favoring inclined "old" showers dominated by muons, compared to vertical "young" showers where the signal is spread out in time and contains a larger fraction of electromagnetic components. Also, the current setup doesn't provide any information about the waveform morphology initiating the trigger.

To address this limitation we have developed a new prototype local trigger system which uses a Pico Technology 2206B USB oscilloscope connected to a Minnowboard Turbot SBC. Using the Pico software, we can remotely configure programmable triggers to fire on temporal features in addition to threshold levels. This setup offers a

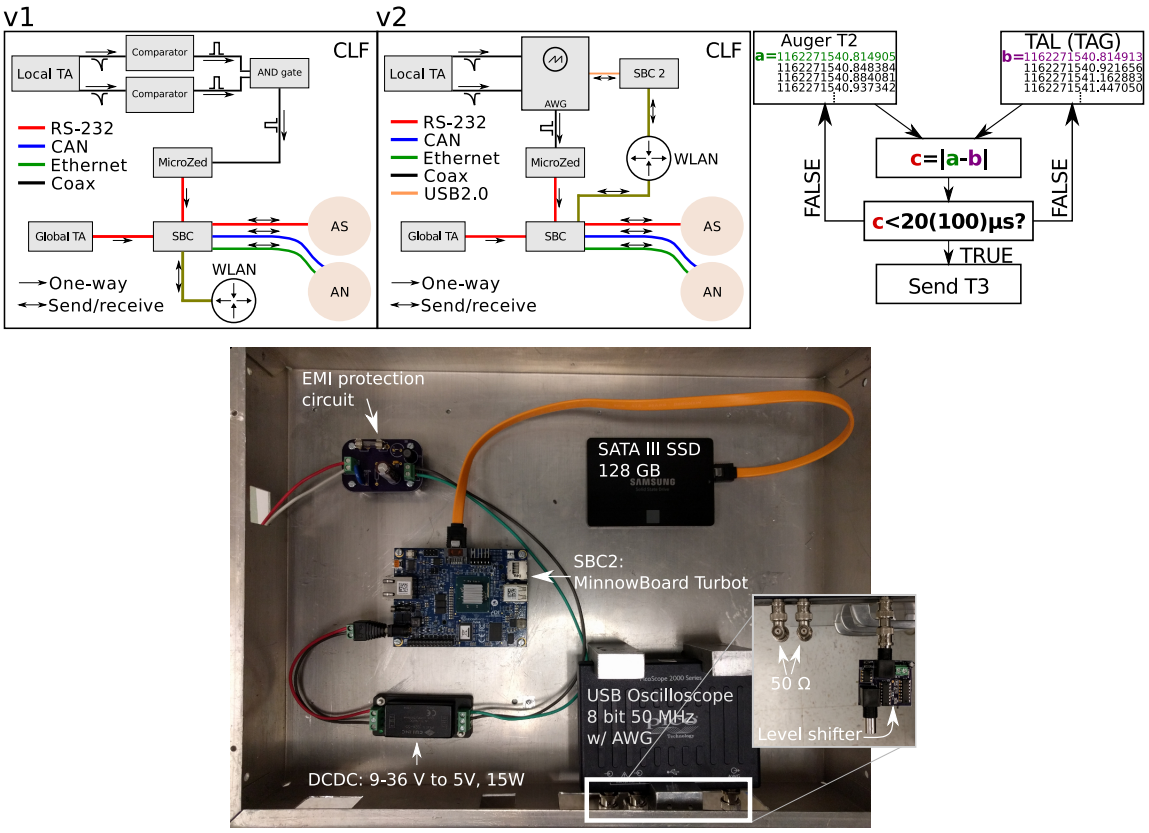


Figure 7.1: Top left: block diagram of current local trigger hardware deployed at the CLF site (see text for details). Top middle: future local trigger setup under development. Top right: trigger-condition flow chart showing the "do while" loop to compare example local time stamps. Items in parentheses refer to the global trigger. Bottom: photo of the future prototype local trigger hardware to be installed in the field.

closer approximation to Auger and TA station level triggers in lieu of TA electronics for the local detector. The new SBC is equipped with a hard drive so calibration data and event waveforms can be stored. The associated time stamp is relayed from the SBC inside the CLF over the local network using the existing time tagging system. The prototype will be a drop-in replacement for the current circuit—a diagram and photo of the design is shown in Figure 7.1.

7.3 Lab testing

A brief field test was performed for around 4 hours at dusk to avoid thermal issues associated with the black light-tight wrap. A 1 ft^2 (929 cm^2) 0.5 inch (1.27 cm) thick PVT scintillator is used, connected to a 2" Photonis PMT. The PMT is run at about 1000 V using an external module developed in our lab. A negative AC coupled cathode is input into the scope using a tee with 50 Ω impedance.

The Picoscope "window dropout" trigger mode was used with thresholds of -2500 mV and -200 mV and time window of 2 μs. This is a flexible type which captures both large amplitude and fast pulses, but will also capture followers in the threshold window. Alternatively, the "runt" trigger can be used with the same threshold levels, but this might result in higher rates since this type doesn't look for the end of the wave train. In principle it's capable of self triggering on young and old showers. Some example events are shown in Figure 7.2. The event rate for this configuration was about 1.15 Hz ($\approx$16400 events in 4 hours). Given the Auger T2-ToT rate is 1.2 Hz, and a majority of real showers are recorded by this trigger, this suggests

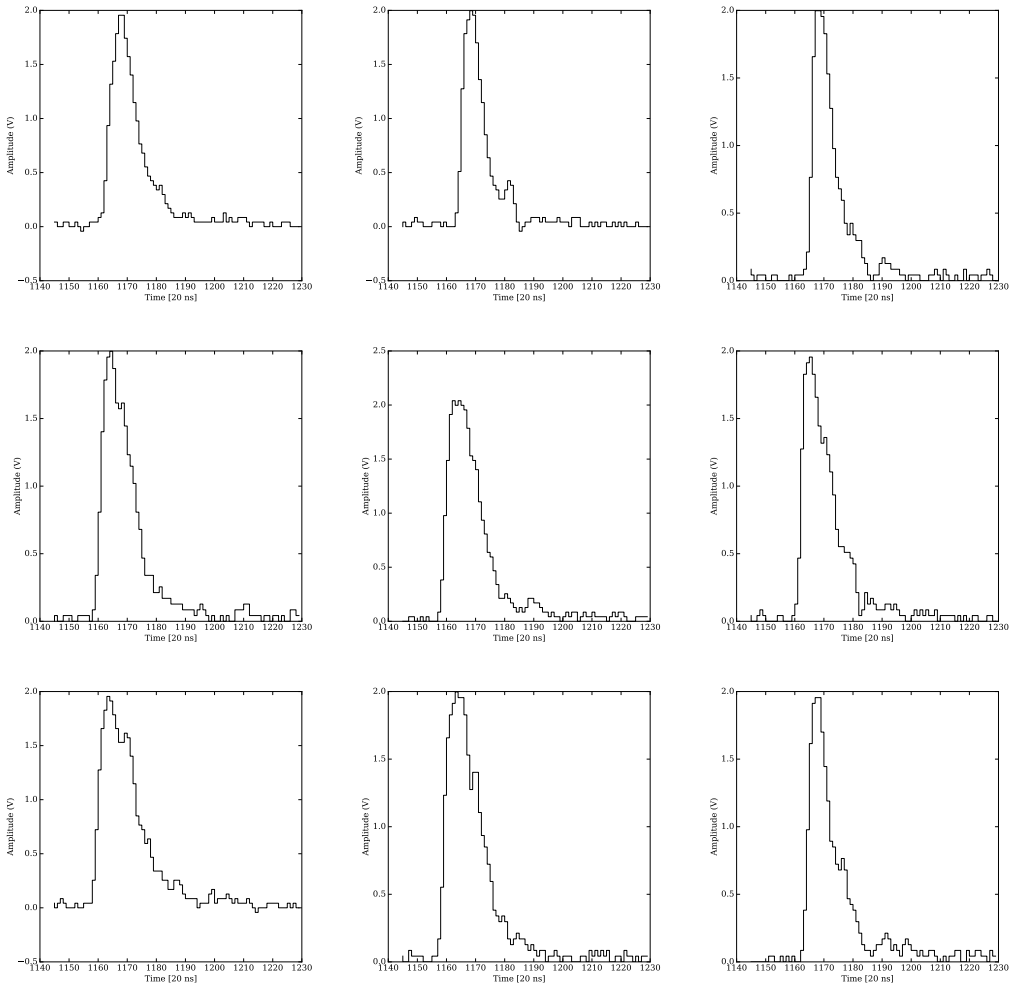


Figure 7.2: Nine example traces using the new trigger system. The advanced trigger leads to a variety of events instead of fast, high pulses.

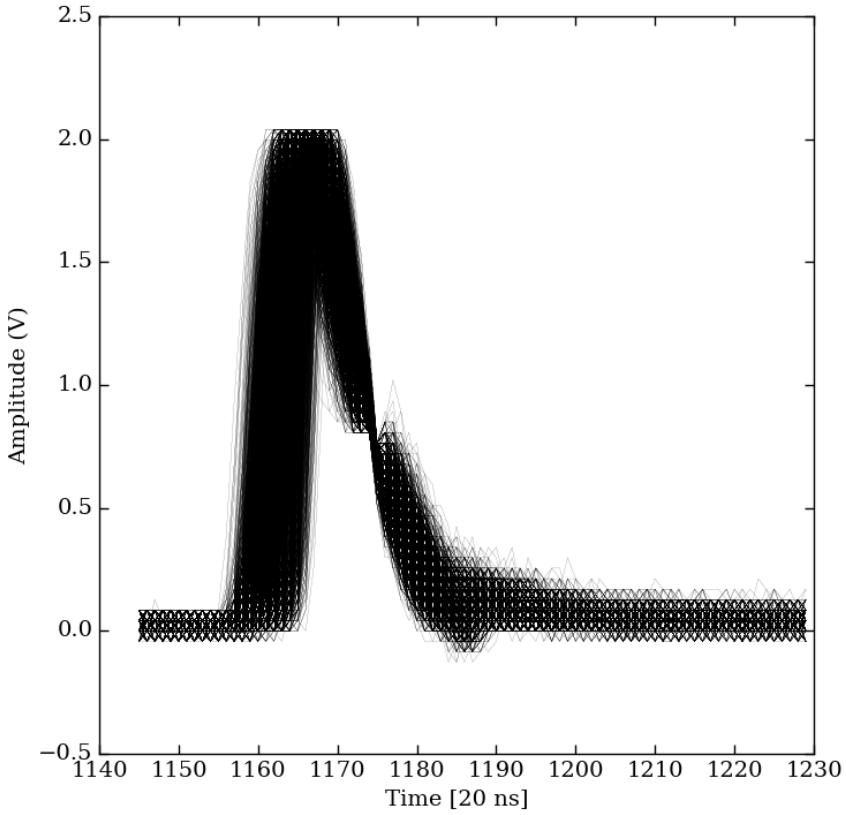


Figure 7.3: A plot of 4000 independent traces from the testing period. Features similar to Figure 7.2 are visible for many different events.

the Picoscope performance is quite good in discriminating shower waveforms from background signals. An ensemble plot showing 4000 traces is shown in Figure 7.3.

7.4 Phase II

The next phase of the project involves embedding 5 AN stations inside the TA array at various detector sites, see Figure 7.4. Showers incident on this array can then be reconstructed using TA scintillator and Auger WCD stations independently.

Additionally, the cross-calibration curve can be used for a "mixed" reconstruction: showers where an LDF can be only be fit using a combination of WCD and TA stations.

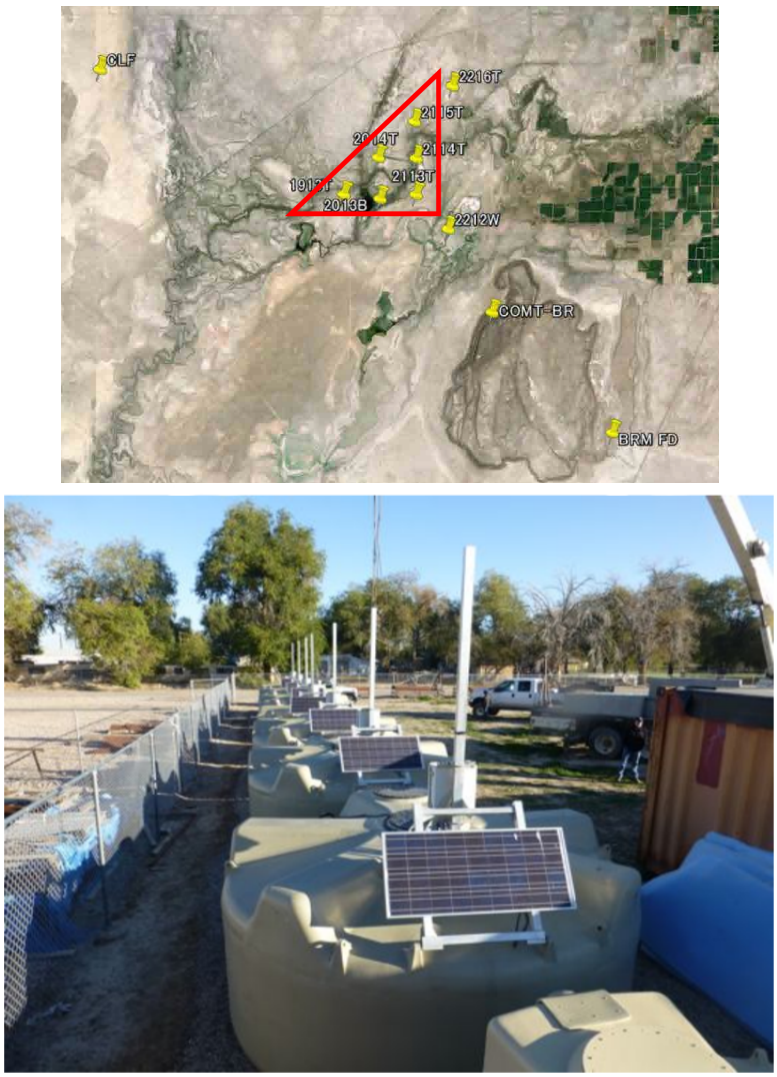


Figure 7.4: Top: Proposed co-location sites for the AN WCDs. Bottom: Photo of the staging area showing the AN stations available for deployment.

Chapter 8: The Auger@TA Project: discussions and conclusions

8.1 Preamble

This chapter summarizes and interprets the main results of the Auger@TA experimental project. I will also discuss the utility of these results for ongoing Auger projects (SSD uprgade) and how one can incorporate our results in a Auger-TA SD data comparison study. I also show how the results can be used for phase II of the project to expand the exposure of the micro-array.

8.2 Interpretation

In the previous chapter the analysis can be broken down into two regimes: (1) raw station level signal response comparison and (2) an indirect comparison of shower parameters by comparing the expected Auger signal to the measured value based on the provided TA experimental air shower parameters. The raw signals were consistent with a primitive geometric model of detectors which assumed an increased efficiency (2-3 above 1:1 ratio) for electromagnetic shower components in a scintillator. The MIP-VEM data do not agree if one assumes an identical efficiency for both detectors in this model. This result indicates both stations correctly record particle densities and

that TA detectors are more sensitive to the EM component, agreeing with predictions of simulations in [184]. Furthermore, it was observed that for every TA physics trigger, an associated Auger T2 is available, indicating that at the very least the Auger trigger algorithms exactly cover those of TA, but also allow for the possibility that TA station level triggers are a subset of Auger.

We also observed that the cross-calibration data are highly correlated over a large dynamic range. A linear relationship is favored, even when testing power-law models, and the best fit slope is found to be ≈ 0.62 MIP/VEM. We recorded 3 saturated global events and found the TA station saturated at roughly 680 MIP while the Auger station saturated at ≈ 3200 VEM. A particle count of 680 MIP when accounting for EM enhancement would be equivalent to around 1360-2040 VEM. The difference in saturation could be related to the PMT collection area, PMT performance (quantum efficiency, face uniformity, etc.) or digitization electronics. This property is a significant piece of SD operation and it determines how close to the shower core a station can survive and report accurate densities for the LDF fit. This result indicates that, on average, more TA stations are excluded from LDF fits due to: (1) the observed reduced dynamic range (2) the TA array's separation of 1250 m.

The raw signal agreement was also tested by comparing the FADC traces of the same shower. In this analysis the Auger traces were first up-sampled to the TA FADC frequency. The traces were then normalized and shifted so they had the same start time. The statistics for the four independent metrics quantifying similarity demonstrate that in a large majority of cases the traces are quite similar. For example, the widely used Euclidean distance metric had a median of 0.38 for 97 events and

88 of the 97 showed $s_3 > 0.3$. This metric is fairly robust, and randomly generated traces give values of ≈ 0.01, further increasing the confidence of this conclusion. Similar does not mean identical, and a visual inspection of all 97 traces would verify this reality. But given the overall morphologies are comparable, the remaining minor trace differences along with the calibration values lead to different overall integrated signals.

To indirectly compare the reconstruction frameworks of the two collaborations a simulation study was done using the 13 global events which had air-shower parameters available. After scaling the primary energy, the air-shower is reproduced in a `CORSIKA` simulation. The detector setup at the TA CLF is formed from the main array in the $\overline{\text{Off}}\underline{\text{line}}$ package and the shower is placed at the same ground coordinates, relative to this doublet, in the simulation. A simulation centroid is formed from 100 `CORSIKA` runs and this is compared to observed signals using the MIP-VEM plane. Three primary particle compositions were studied: proton, helium and nitrogen. For light primaries (proton, helium) the relative error for MIP signals is found to cover 0 (agreement) for all events far from the core, $\rho > 1000$ m, but is discrepant closer to the core in 4 of 6 events. For VEM signals with $\rho > 1000$ m the relative error agrees within uncertainty for 1 of 6 events, and for close signals the relative error agrees for 0 of 6 events. For nitrogen primaries far from the core the VEM relative difference agrees for 3 of 7 events, and MIP relative differences agree for 5 of 7 events. Far from the core 2 of 6 VEM relative differences agree, and 0 of 6 MIP relative differences agree. The results show, albeit for a very small sample, that increasing the primary mass leads to better agreement far from the core at the expense of agreement at

smaller core distances. The discrepancies can't be resolved by tuning this parameter, so the underlying source is likely due to another shower parameter. Using heavier primaries would likely continue this trend, but investigating anything heavier than nitrogen is not merited at such low energies, given the recent composition results published by Auger and TA. A future MC study which finds the optimal shower parameters, given the TA reconstruction as a starting point, that match the observed signals would be very interesting.

In terms of the Auger north and Auger south detector results, statistically the two stations show the same response for simultaneously observed air showers. The correlation improves when considering a sub-sample of showers which satisfy the TA physics trigger. In addition to the integrated signal comparison, it was also shown that the low level T2 trigger algorithms agree over several months of operation. The expected number of 2-fold coincidences for 2 days worth of data at 25 Hz and using a gate of 5 μs is around 2.6×10^4, but the observed number of 2-folds is on the order of 2×10^5.

A preliminary study of the muon fraction of observed MIP and VEM signals using the method of matrix inversion demonstrated that when the setup was close to the shower core the shower components were mostly EM particles, but with increasing distance this transitioned to a muonic dominated signal. This behavior is predicted by simulations. The result serves as a validation and proof of concept for the Auger upgrade design in addition to the analysis methods.

To summarize, the Auger@TA project has provided a variety of insights. We have demonstrated that in terms of particle counting and low level station responses the

SD stations are comparable. In terms of expected signals from air shower parameters there appears to be a discrepancy. We have shown that statistically the response of Auger north and Auger south are compatible and their triggering firmware share the same characteristics. There is a strong case to continue data collection to investigate how these results evolve with time and especially to improve the confidence of the conclusions.

8.3 Caveats

As in any experimental project, there are many possible sources of systematic errors and bias which could influence our results. In this section I take a critical view on the weaknesses in our approach and examine how these issues could negatively impact our conclusions. The aim here is not to undermine the previous section, but to instead inform the reader about the challenges involved in the study.

Perhaps the most obvious criticism is that our setup uses only a single Auger and TA station for comparison, and there could be a fault in one, or both of these stations, leading to garbage data. This is certainly a possibility, but there are several reasons why this is unlikely:

- For available data, the Auger doublet give similar responses to a global trigger shower
- For the entire data operation period the Auger T2 rates and charge histograms are stable
- Other monitoring data do not indicate unusual levels

- TA colleagues monitor DET2421 and have not alerted us to unusual operation of this detector

Nevertheless, to further eliminate doubt it would be interesting to deploy another Auger and TA station pair to verify these results.

There are several points about the simulated response comparison that can be criticized. Perhaps the most glaring issue is that we are using the Auger upgrade module (the SSD) as a proxy for a TA detector, instead of a proper TA station mock up. This is a fair point, but the counter-argument is that it would be a large project in itself to port the TA design, triggers, etc. into the Auger $\overline{\text{Off}}\underline{\text{line}}$ framework especially if the relative differences are very small. There is no advantage in terms of triggering since the SSD is triggered by the WCD and we've already demonstrated that low-level TA triggers are likely contained in the Auger T2 set. There are some differences in the active parts of the detector: the TA scintillators are 2 mm thicker and have about double the number of fibers. It's unlikely this will lead to a significant difference in light yield. The most critical difference is the active area, and this is accounted for by scaling the simulated SSD signal to the TA area. An alternative to using the SSD in $\overline{\text{Off}}\underline{\text{line}}$ is to simply input our `CORSIKA` showers into the TA simulation framework and extract the relevant signals. Another potential issue is the use of QGSJETII-04 while the TA SD Monte-Carlo uses QGSJETII-03. This could lead to potential air-shower parameter prediction discrepancies due to the re-tuning of QGSJETII-04 to LHC data. There might also be a source of error introduced the thinning performed in our `CORSIKA` showers. The shower parameters derived by

TA use a de-thinning method [215] aimed at restoring some information lost by the thinning process. Another potential source of error is the use of customized analysis tools for signal integration of the Auger data instead of the standard Auger tools. While efforts were made to emulate the analysis procedures, it's true that our software is not identical. One example of a weakness is the baseline estimation algorithm. In certain situations the minor differences could lead to an over or underestimation of the signal. However, these would likely be small differences contained within the present signal uncertainties.

To summarize these points, while there are multiple potential sources of error in our setup, given the performance and results so far there is no reason to suspect a critical flaw in any of the instruments. We can verify this by continuing to collect data, or by installing an identical setup elsewhere and comparing the results.

8.4 Empirical MIP-VEM relation for phase II

Phase one of this project, which involves a station level comparison of signals, has been operating for about 8 months. We will continue studying the MIP-VEM correlation which can be used for phase two: embedding WCD stations in the main array of TA. This will enhance the study in a variety of ways: the six available stations will increase the acceptance of our setup and thus the rate of higher energy showers. There are two scenarios for observing EAS in this phase: (1) the shower plane is covers the micro-array, (2) the shower grazes corners or portions of the micro-array, which is more likely. These are depicted in Figure 8.1. The first scenario allows for a direct

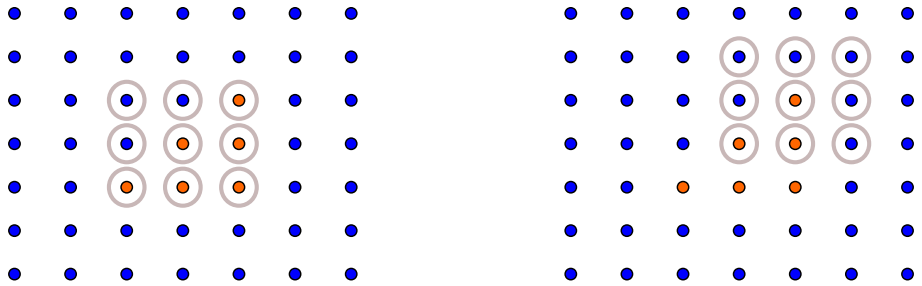


Figure 8.1: Left: An ideal shower covering the micro-array. Participating stations are enclosed by gray circles. Orange grid points are where Auger stations are co-located with TA stations, blue points are TA stations. In this scenario there are sufficient stations to build independent LDFs. Right: A shower where only a portion of the micro-array is covered, making an Auger-only LDF fit difficult.

comparison of station signals and reconstructed shower parameters as determined by TA and Auger. The second scenario allows a comparison of extrapolated Auger shower parameters after applying the MIP↔VEM conversion to adjacent TA signals in order to populate the lateral distribution function.

8.5 Parting thoughts

Phase I of this project, while still ongoing, has provided the first cross-calibration of the world's two largest observatories' surface detector stations. A small collection of tests and analyses were performed to characterize and understand, but there are many more (potentially more interesting) studies that can be done using these results. It's my hope that this data serves as a laboratory for scintillator and WCD based detectors, and informs the larger community about their important differences and similarities.

With this phase we have demonstrated the feasibility of *in-situ* measurements and the power of combining the resources of competing collaborations to better understand their hardware and make progress in our discipline. It's my sincere hope that this phase of the project serves as a signpost for the rich and useful data that can be generated by these collaborations. I will also argue that while the results of phase I might be interesting and in need of additional statistics, the physics reach is very limited due to reliance on TA reconstructions. The much more compelling situation is a *direct* comparison of reconstructed air shower parameters, and this can be achieved by moving on to phase II.

My thanks for reading this far, or if you have skipped to this section first, welcome!

Chapter 9: Miscellany

9.1 Preamble

This chapter discusses work on projects which don't neatly fit into the two categories of arrival direction work or the Auger@TA project. I discuss efforts made on a prototype non-imaging atmospheric Cherenkov detector aimed at further constraining the $X_{\max}$ EAS parameter. I also discuss construction of a muon telescope used for testing the time-tagging board used in the local trigger system.

9.2 Atmospheric Cherenkov detector

The HEA group has made efforts to develop a new air shower measurement technique using Cherenkov light generated by the EAS. The apparatus would be situated directly adjacent to WCD stations. The design uses an optical concentrator and UV optimized PMT directed toward the sky and aims to reconstruct the $X_{\max}$ parameter using Cherenkov light yield. It would likely be tied to Auger T3 triggers.

A prototype instrument was built from a large aluminum cylinder reflector and a 2" PMT. Field test were conducted in the fall of 2013 and winter of 2014 in the Pierre Auger North research and development array (RDA). For these tests the Cherenkov

detector was externally triggered by a 4-fold coincidence unit made from scintillators coupled to smaller PMTs which are connected to discriminators.

9.3 Muon telescope

In this section I present calibration and diagnostic data for the external trigger system to be deployed at the Telescope Array (TA) central laser facility (CLF). As mentioned in previous chapters, an external trigger is required for the Auger SD doublet because self triggering is not possible when the AN station is in local storage mode. A base counting rate and timing resolution is established by analyzing data from a muon telescope. I begin with apparatus hardware, followed by a description of data collected. I conclude with an analysis of the calibration data and offer the parameters of operation for the system.

9.3.1 Apparatus

Atmospheric muons are detected using a photomultiplier (Hamamatsu R580) coupled to a 1 ft^2 scintillator panel. Four of these detectors are arranged in an overlapping vertical stack, separated by about 2 ft. The apparatus is housed in the basement of a 4 story building (A .W . Smith) which provides roughly 150 cm of concrete shielding. Using standard concrete density and energy loss of muons through this material, the telescope observes $E > 590$ MeV muons [216]. Each photomultiplier tube (PMT) is operated at -1500 ± 10 V using power supplies designed by a former graduate student (Ross Burton) in the high energy astrophysics group. An image of the arrangement

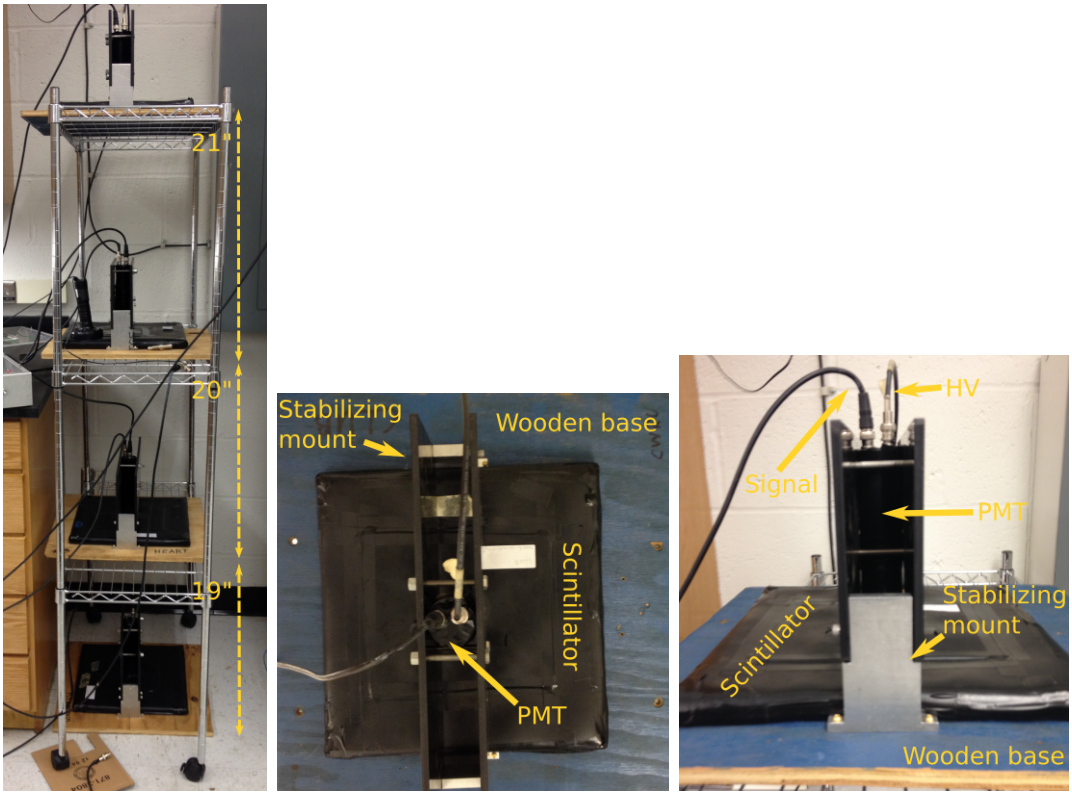


Figure 9.1: Left: muon telescope with separation distances. Middle: Top view of scintillator panel. Right: side view of scintillator panel assembly.

is shown in Figure 9.1

9.3.2 Detectors and electronics

For these tests we're primarily concerned with event rates/muon counting. We don't provide a detailed calibration of the output PMT signal in terms of photo-electrons or associated particle energy. Instead, the electronics and power supplies are tuned based on the rate of coincidence counts. A muon crossing generates a brief, nanosecond pulse of fluorescent light in the scintillator panel. The photons are collected at the PMT photo-cathode coupled to the panel. A cascade of photoelec-

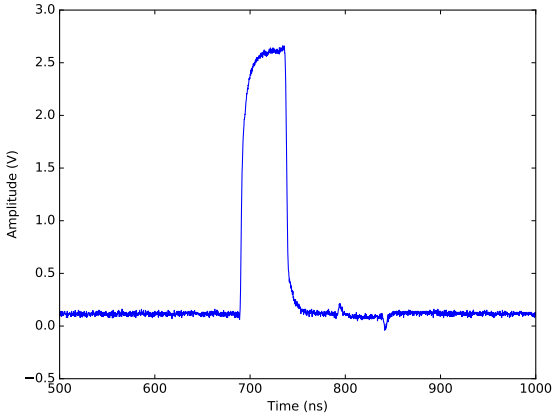


Figure 9.2: Characteristic pulse from a discriminator recorded on a Tektronix TDS 3054B digital oscilloscope at 50 Ω impedance.

trons are produced and accelerated across the large potential difference. An amplified analog signal is then piped from the photo-anode to a discriminator.

The discriminator has been tuned to emit a 5 V logic pulse (see Figure 9.2) when the PMT signal crosses a certain threshold. To focus on high energy muons, this threshold, along with HV power into the PMT, has been tuned to reduce the counting rate to agree with expectations of atmospheric muons. To verify the instrument is recording downward traveling atmospheric muons we insist on 4-fold coincidence between the discriminators; this is implemented using a dual 4 input positive AND gate (Texas Instruments SN74HC21N). When this condition is met the gate will output a 5 V pulse observed which is then recorded by the oscilloscope. A block diagram of our apparatus is provided in Figure 9.3. The circuit for coincidencing is also show in this figure.

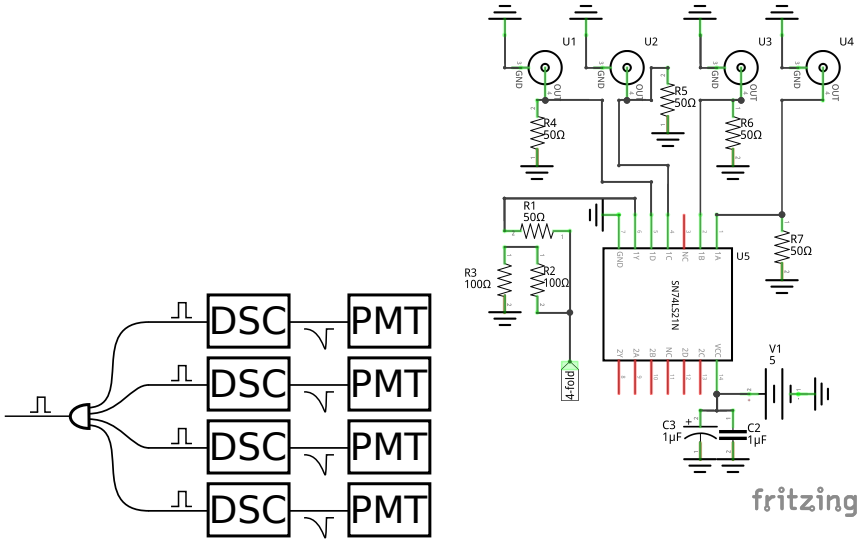


Figure 9.3: A simplified sketch (left) of our setup. Each block represents an assembly built of multiple sub-components. PMT and DSC are photomultipliers and discriminators, respectively. If the 4 fold condition is satisfied the logic gate (left) generates a 5 V pulse. Circuit diagram for coincidencing (right).

9.3.3 Data

In this section I describe two modes of data collection for our setup. In the first mode we setup our coincidencing circuit on a breadboard and used the oscilloscope to record output amplitudes. Timing information was recorded using a computer. We also calculate the timing resolution of the telescope using this setup. In the second test, we used the complete time tagging apparatus to record precise time stamps. For both tests we compute the event rate.

For the breadboard test we assembled the circuit shown in Figure 9.3 on a RAL-BOARD, but without the voltage divider. The 4-fold output was piped to the scope using banana cables, and the scope was triggered in high impedance (1 MΩ) mode. This resulted in moderate ripple, likely due to impedance mismatch between the logic gate and scope.

The scope was then connected to our LAN using an Ethernet cable. Remote data acquisition software was used to download and archive events which triggered the scope. The program saved scope traces (amplitude vs time) and also time stamped events using the `date` program, in Unix time. The apparatus collected data for 25.5 hours, resulting in about 20,000 events—data can be accessed here. This gives an average event rate of $\langle \dot{n} \rangle = 0.233$ Hz. For this sample we find an amplitude value of 4.778 ± 0.002 V which is within 5% of gate V_{cc}.

Since this is a counting experiment, the appropriate statistical model to apply to the data is the Poisson distribution. This can be seen from the following argument. The probability of adjacent events having a time separation of $t + \Delta t$ can be written

as $p\,(t)\,dt$. The density function $p\,(t)$ is constructed to have zero events from 0 to t, followed by 1 event at t in the interval dt. The total probability for this scenario is the product ot those individual probabilities

$$p\,(t)\,dt = e^{-\lambda t}\lambda\ dt \tag{9.1}$$

It is also important to consider accidental coincidences between the 4 scintillators. The counting rate for individual discriminators is on the order of $\lambda_D \approx 1$ kHz with a resolution of roughly $\delta t \approx 1\ \mu$s. The accidental coincidence probability is then $\propto (\delta t \lambda_D)^4$ where we've assumed identical counting rates for all four discriminators. This results in a marginal probability which can be safely neglected.

We infer the underlying event rate by constructing a histogram of time differences, and finding the maximum likelihood estimator (MLE), $\hat{\lambda} = \hat{\dot{n}}$, for the true rate, $\lambda = \dot{n}$. The data for this run is shown in Figure 9.4. We find a rate of $\hat{\lambda} = 0.238$ Hz.

9.3.3.1 Timing resolution

I also used this configuration to determine the timing resolution of the apparatus. For this test we placed a small circular scintillator (the "hockey puck") coupled to a Burle 83112-502 PMT on the second level. For this setup we disconnected the square panel that sits at ground level. The -1500 ± 10 V power supply and discriminator normally used for that panel was instead used for the hockey puck. A photo of the hockey puck is shown in Figure For this test the top three square panels were configured in 3-fold mode by holding the unused pin at a constant 5 V. Data was collected on the scope using two channels: CH1 for the discriminator connected to

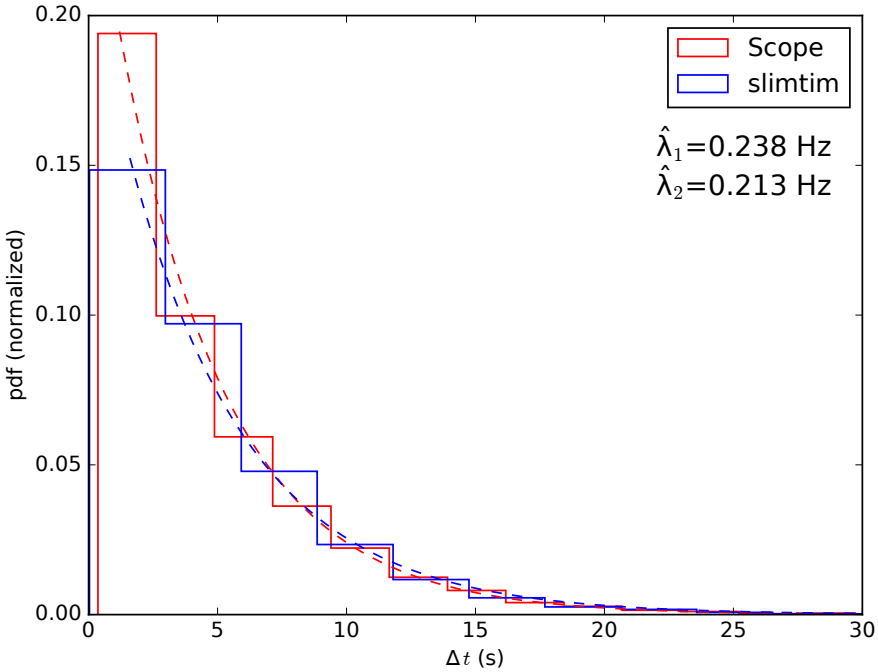


Figure 9.4: Histogram (solid lines) and MLE fit (dashed lines) of counting data for two setups. Data for the RALBOARD + scope is the red curve, while data collected using the 4XT apparatus is shown in the blue curve. The "slimtim" legend refers to the MicroZed time tagging board.

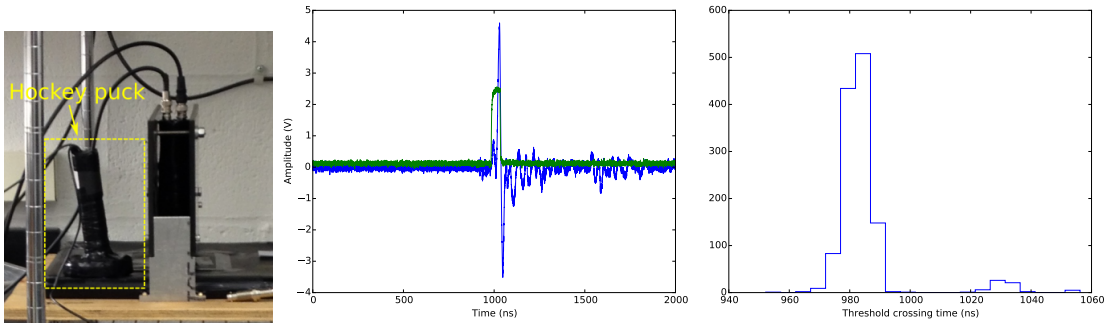


Figure 9.5: Left: small circular PMT used for timing resolution study: the hockey puck. Center: scope trace of CH1 (green) and CH2 (blue). Since the scope trigger is using high impedance (1 MΩ) the amplitude of CH1 is reduced from the expected 5 V. Right: Histogram of threshold crossing times.

the hockey puck and CH2 for the output of the 3-fold coincidence circuit. We used a threshold trigger region of 1.6–1.7 V on CH1 and record the associated crossing time of the scope trace in that region. An example event trace, and histogram of crossing times can be found in Figure 9.5. Data was collected for 16 hours resulting in 1289 events. There are fewer events compared to the 4-fold setup due to the hockey puck's significantly reduced acceptance. We find the timing resolution of the setup, which is defined to be 2σ of the histogram's primary population, to be $\delta t = 23.3$ ns.

9.3.4 "slimtim" setup

This test is similar to the breadboard setup; however, instead of implementing the circuit on a RALBOARD we have soldered all the components onto a custom PCB design. Additionally, time tagging is done using the "slimtim" module (short for slim timing instrumentation machine), shown in Figure 9.6. This module uses a MicroZed development board and specialized firmware to time stamp signals sent to a user GPIO pin. Timing is synchronized to atomic clocks on board GPS satellites using a M12+ receiver in a SynPaq III sensor board. Timing resolution for the device is limited by the dead time between events of about 1 μs. Dead time is a result of FPGA clock cycles needed to run the compiled `C++` program that communicates with the receiver and time stamps signals. This resolution is precise enough for external triggering.

Slimtim outputs ASCII text over a serial connection. For this test data was recorded onto a Raspberry Pi (running `wheezy` Raspbian OS) computer using a serial

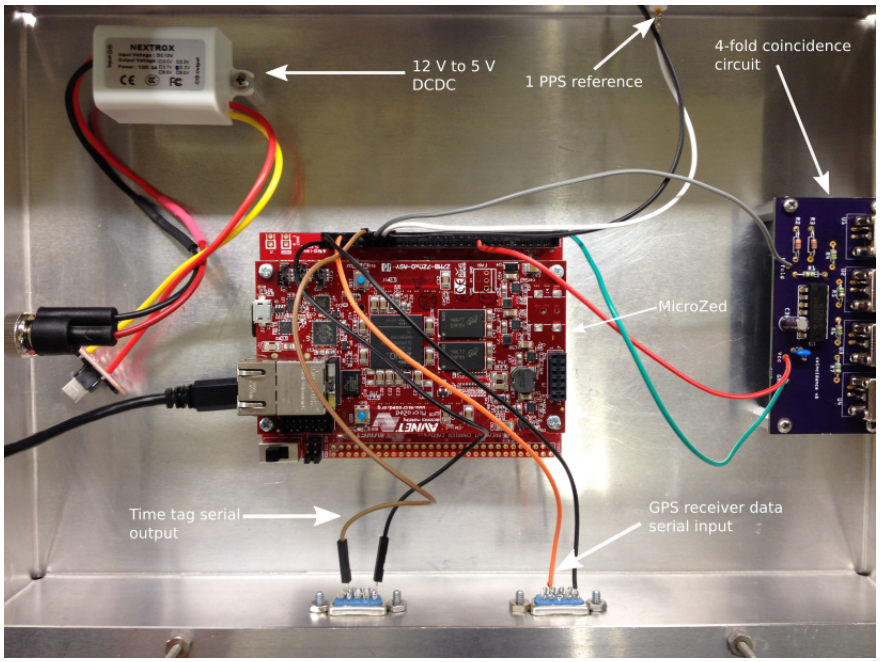


Figure 9.6: Photo of slimtim module with labeled parts. Discriminators are connected to the 4-fold coincidence circuit (right) using BNC cables.

to USB adapter. A program was developed to parse slimtim's output into GPS time stamps with microsecond precision. The code can be accessed here.

Data was collected for about 19 hours and can be accessed here. A plot of the results is shown in Figure 9.4. We find the MLE rate from slimtim to be $\hat{\lambda} = 0.213$ which is about 10% less than the RALBOARD setup. As seen in Figure 9.4, the discrepancy is the result of a higher value in the pdf for the smallest time difference bin. The data collection for the scope+RALBOARD trial which records all events, even if separated by microseconds. For example, there are situations where the logic circuit will send a bundle of signals, most likely related to the same event. The data acquisition program for that setup will nonetheless count each individual signal, resulting in artificial inflation of small time differences. The slimtim DAQ program, on

the other hand, will record at most two events with a time difference of microseconds.

9.3.5 Validity of data

In this section we compare the counting rates of our setup to accepted values in the literature. The integral intensity of vertical muons with $p_\mu > 1$ GeV/c at sea level is roughly $I_\mu = 70$ m^{-2} s^{-1} sr^{-1} [217, 218]. Our vertical muon telescope is a stack of four 12"×12" panels. The expected flux of the apparatus can be calculated using

$$J = \int I(\theta) \cos\theta \; d\Omega \tag{9.2}$$

where $I(\theta)$ is the zenith dependent muon intensity which we take as $I(\theta) \approx I_\mu \cos^2\theta$ and $d\Omega$ is the solid angle element. Using symmetry we can rewrite the integral as

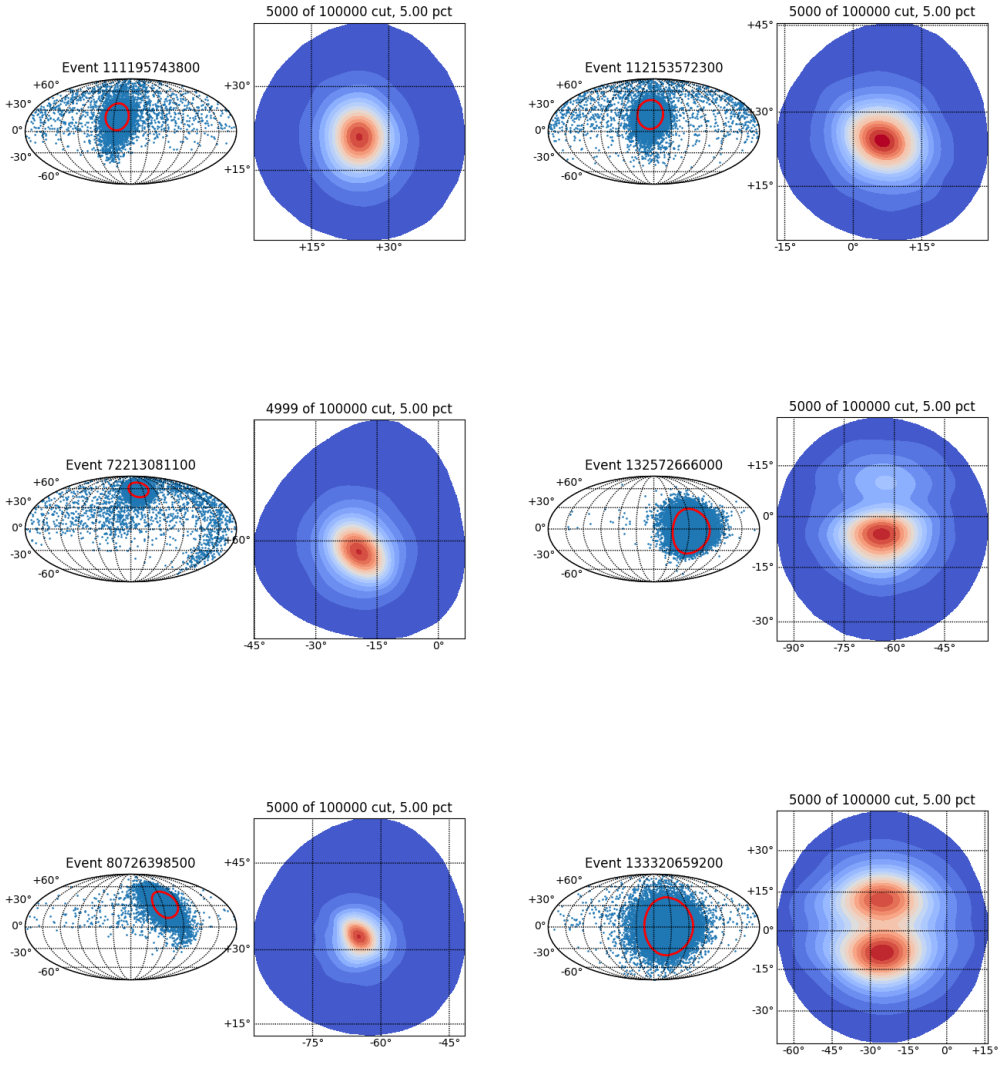
$$J = 4I_\mu \int_{0\leq x,y\leq a/2} \cos^3\theta \sin\theta d\theta d\phi \tag{9.3}$$

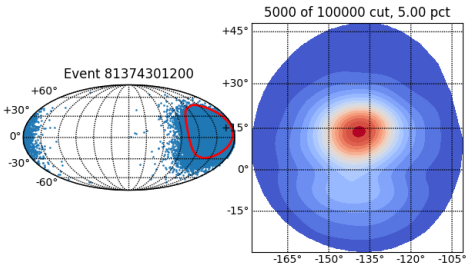
where we exploit azimuthal symmetry by only considering one quadrant of the square detector. Following the derivation presented in [219], this integral can be simplified as

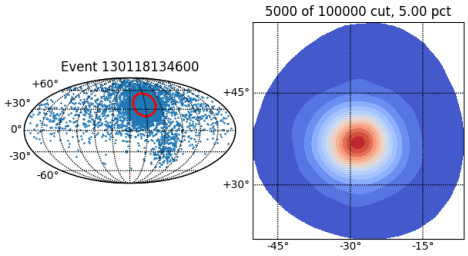
$$J = 4I_\mu \left(\frac{\pi}{8} - \frac{1}{2} \int_0^{\pi/4} \frac{d\phi}{\left(1 + \frac{a^2}{4d^2} \sec^2\phi\right)} \right) \tag{9.4}$$

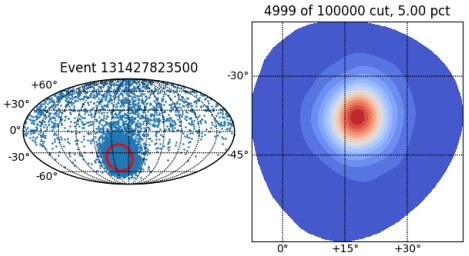
Here d and a refers to the height and square base of the telescope, respectively. In our setup $d = 60$" and $a = 12$". Evaluating the integral gives $J = 0.0392 I_\mu$ sr $= 2.744$ m^{-2} s^{-1} where we've used the accepted value of vertical muon intensity. From our setup, using the slimtim result reported in the previous section, we find $\hat{J} = 2.559$ m^{-2} s^{-1}, which agrees with the accepted value within 6.75%.

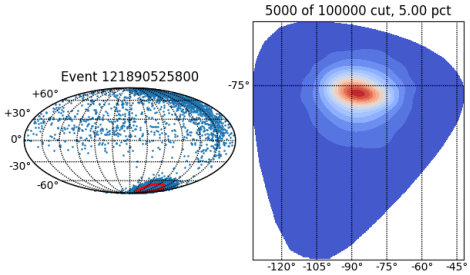
Chapter A: Appendix A

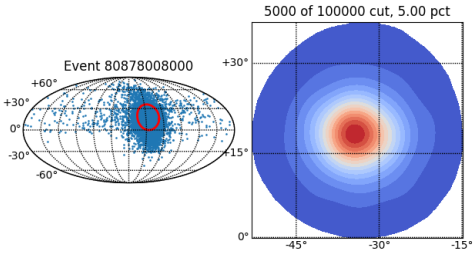


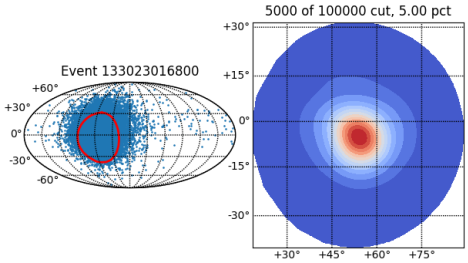

5000 of 100000 cut, 5.00 pct
Event 81374301200

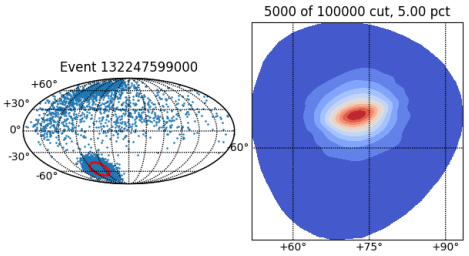

5000 of 100000 cut, 5.00 pct
Event 130118134600

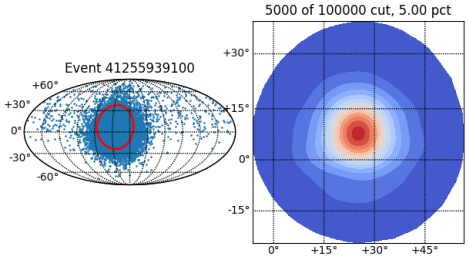

4999 of 100000 cut, 5.00 pct
Event 131427823500

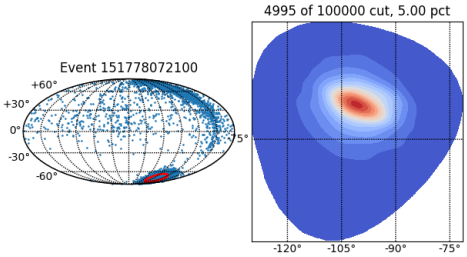

5000 of 100000 cut, 5.00 pct
Event 121890525800

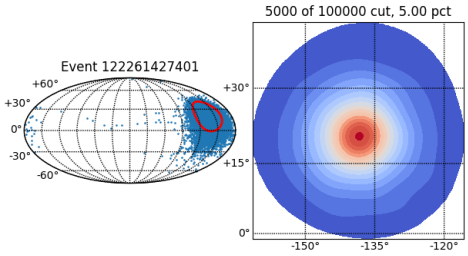

5000 of 100000 cut, 5.00 pct
Event 80878008000

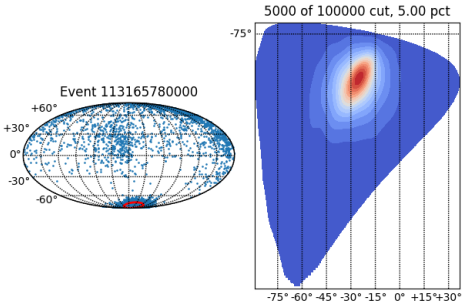

5000 of 100000 cut, 5.00 pct
Event 133023016800

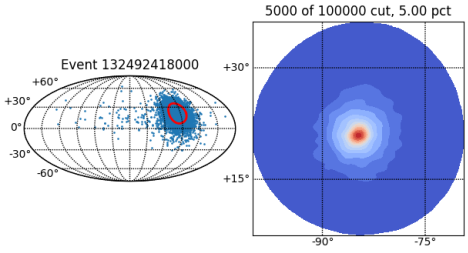

5000 of 100000 cut, 5.00 pct
Event 132247599000

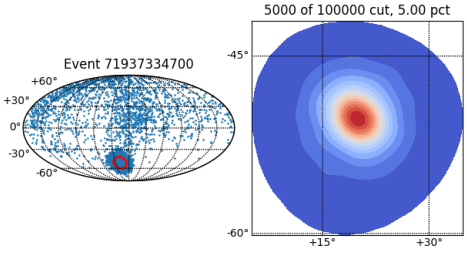

5000 of 100000 cut, 5.00 pct
Event 41255939100

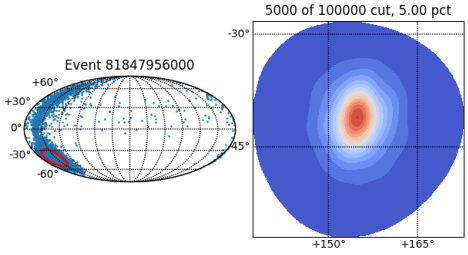

4995 of 100000 cut, 5.00 pct
Event 151778072100
+60°
+30°
0°
-30°
-60°
-75°
-120°
-105°
-90°
-75°

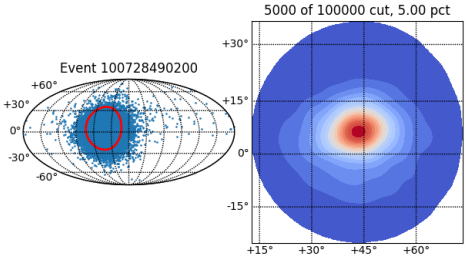

5000 of 100000 cut, 5.00 pct
Event 122261427401
+60°
+30°
0°
-30°
-60°
+30°
+15°
0°
-150°
-135°
-120°

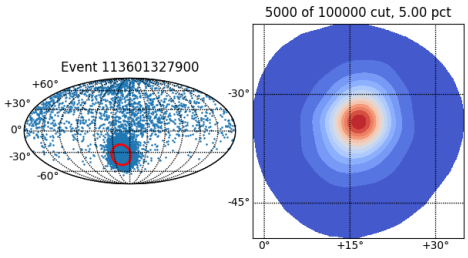

5000 of 100000 cut, 5.00 pct
Event 113165780000
+60°
+30°
0°
-30°
-60°
-75°
-75° -60° -45° -30° -15° 0° +15° +30°

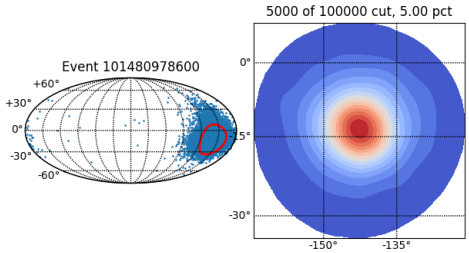

5000 of 100000 cut, 5.00 pct
Event 132492418000
+60°
+30°
0°
-30°
-60°
+30°
+15°
-90°
-75°

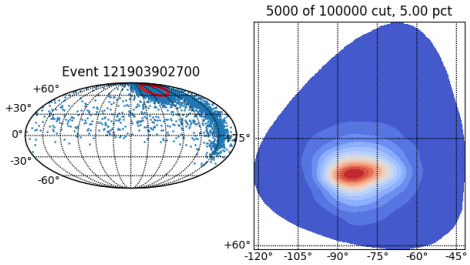

5000 of 100000 cut, 5.00 pct
Event 71937334700
+60°
+30°
0°
-30°
-60°
-45°
-60°
+15°
+30°

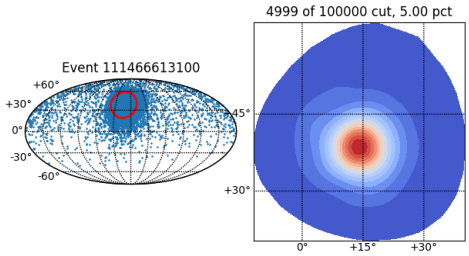

5000 of 100000 cut, 5.00 pct
Event 81847956000
+60°
+30°
0°
-30°
-60°
-30°
-45°
+150°
+165°

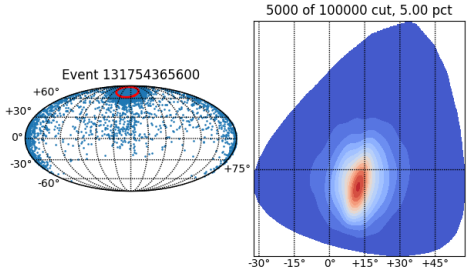

5000 of 100000 cut, 5.00 pct
Event 100728490200
+60°
+30°
0°
-30°
-60°
+30°
+15°
0°
-15°
+15°
+30°
+45°
+60°

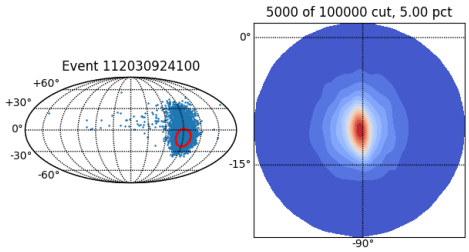

5000 of 100000 cut, 5.00 pct
Event 113601327900
+60°
+30°
0°
-30°
-60°
-30°
-45°
0°
+15°
+30°

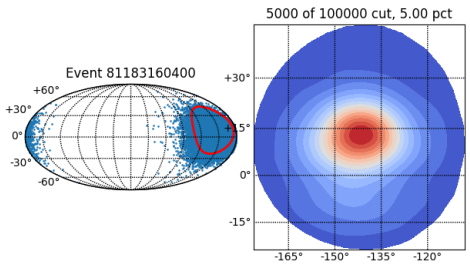

5000 of 100000 cut, 5.00 pct
Event 101480978600
+60°
+30°
0°
-30°
-60°
0°
-15°
-30°
-150°
-135°

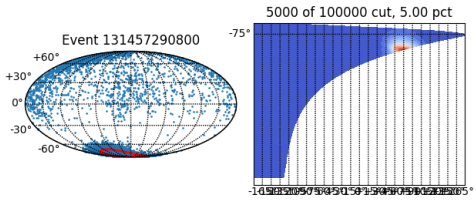

5000 of 100000 cut, 5.00 pct
Event 121903902700
+60°
+30°
0°
-30°
-60°
+75°
+60°
-120°
-105°
-90°
-75°
-60°
-45°

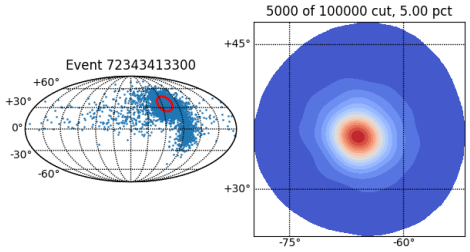

4999 of 100000 cut, 5.00 pct
Event 111466613100
+60°
+30°
0°
-30°
-60°
-45°
+30°
0°
+15°
+30°

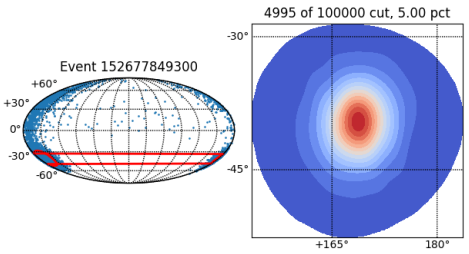

5000 of 100000 cut, 5.00 pct
Event 131754365600
+60°
+30°
0°
-30°
-60°
+75°
-30°
-15°
0°
+15°
+30°
+45°

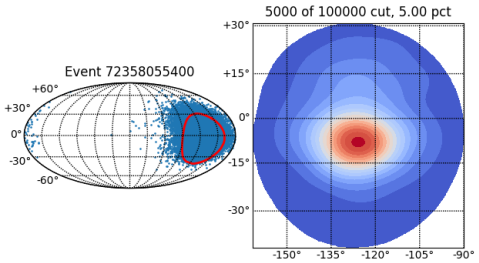

5000 of 100000 cut, 5.00 pct
Event 112030924100
+60°
+30°
0°
-30°
-60°
0°
-15°
-90°

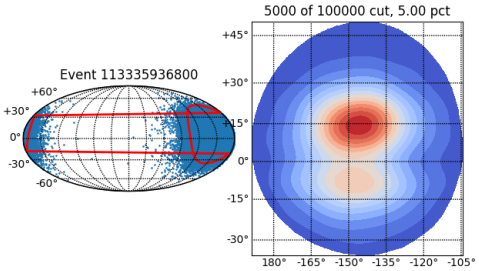

5000 of 100000 cut, 5.00 pct
Event 81183160400
+60°
+30°
0°
-30°
-60°
+30°
+15°
0°
-15°
-165°
-150°
-135°
-120°

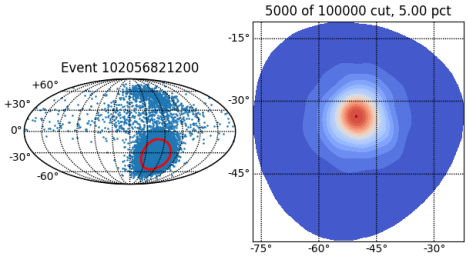

5000 of 100000 cut, 5.00 pct
Event 131457290800
+60°
+30°
0°
-30°
-60°
-75°

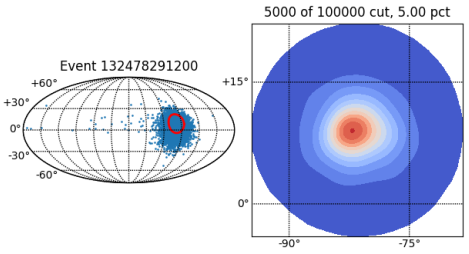

5000 of 100000 cut, 5.00 pct
Event 72343413300
+60°
+30°
0°
-30°
-60°
+45°
+30°
-75°
-60°

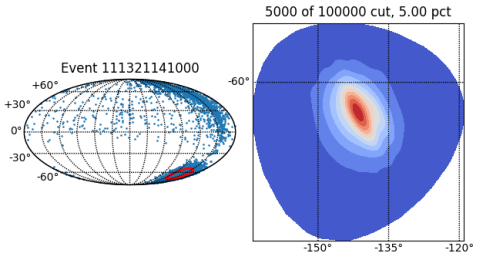

4995 of 100000 cut, 5.00 pct
Event 152677849300

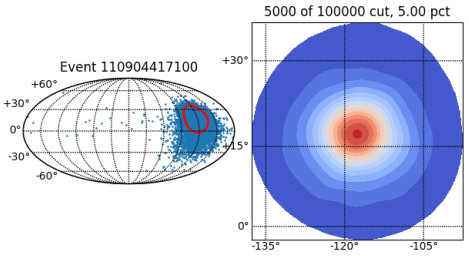

5000 of 100000 cut, 5.00 pct
Event 72358055400

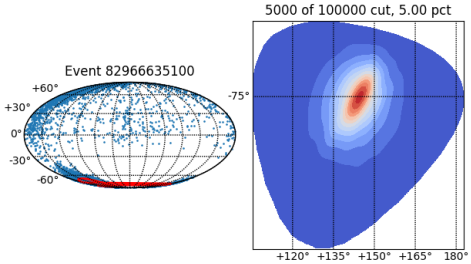

5000 of 100000 cut, 5.00 pct
Event 113335936800

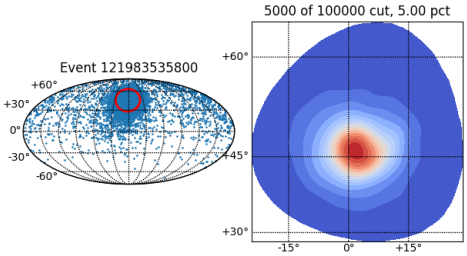

5000 of 100000 cut, 5.00 pct
Event 102056821200

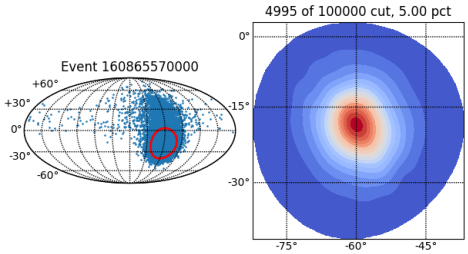

5000 of 100000 cut, 5.00 pct
Event 132478291200

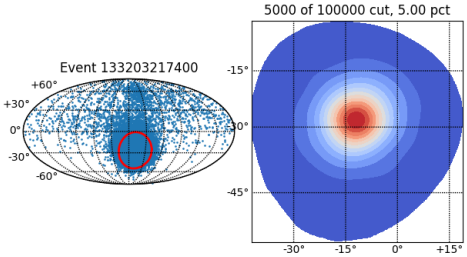

5000 of 100000 cut, 5.00 pct
Event 111321141000

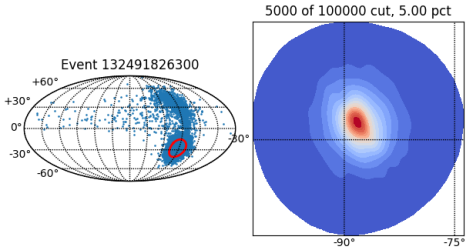

5000 of 100000 cut, 5.00 pct
Event 110904417100

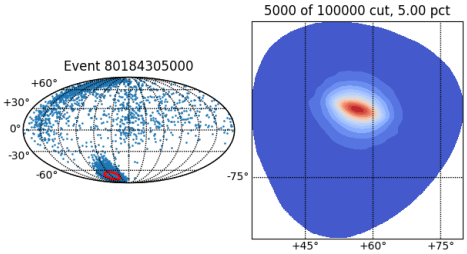

5000 of 100000 cut, 5.00 pct
Event 82966635100

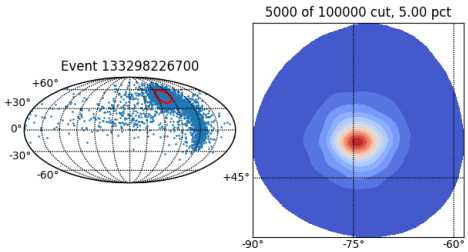

5000 of 100000 cut, 5.00 pct
Event 121983535800

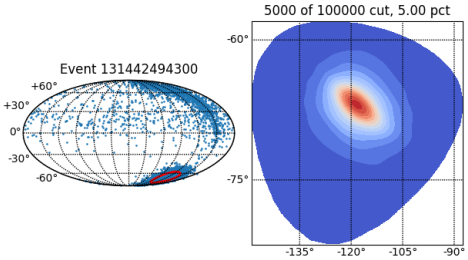

4995 of 100000 cut, 5.00 pct
Event 160865570000

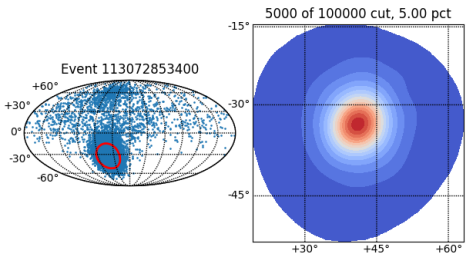

5000 of 100000 cut, 5.00 pct
Event 133203217400

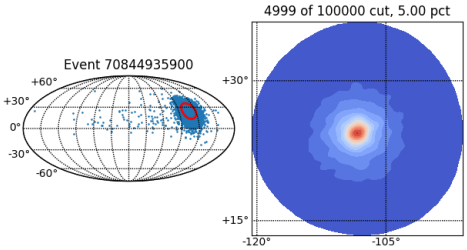

5000 of 100000 cut, 5.00 pct
Event 132491826300

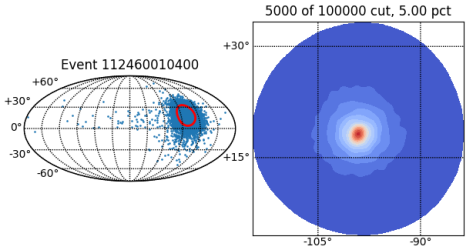

5000 of 100000 cut, 5.00 pct
Event 80184305000

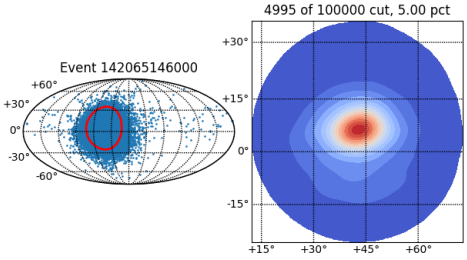

5000 of 100000 cut, 5.00 pct
Event 133298226700

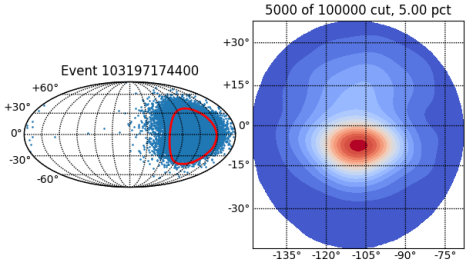

5000 of 100000 cut, 5.00 pct
Event 131442494300

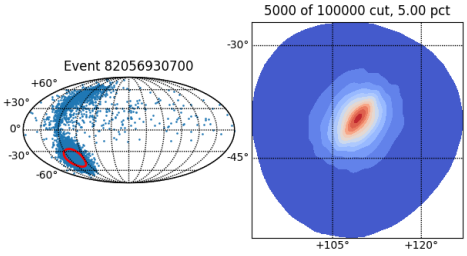

5000 of 100000 cut, 5.00 pct
Event 113072853400

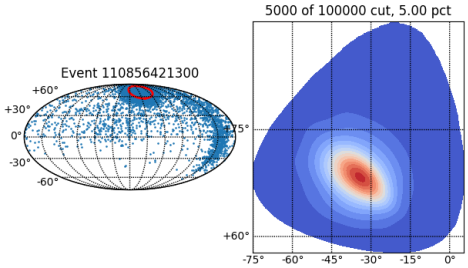

4999 of 100000 cut, 5.00 pct
Event 70844935900

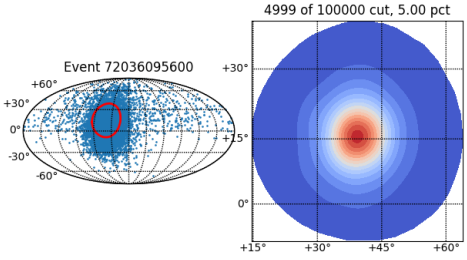

5000 of 100000 cut, 5.00 pct
Event 112460010400

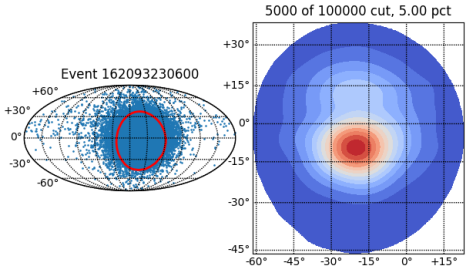

4995 of 100000 cut, 5.00 pct
Event 142065146000

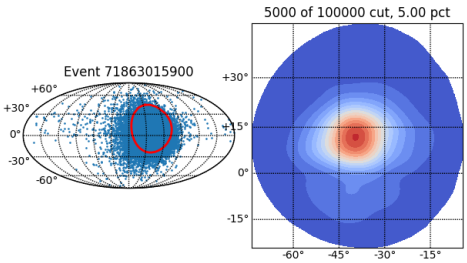

5000 of 100000 cut, 5.00 pct
Event 103197174400

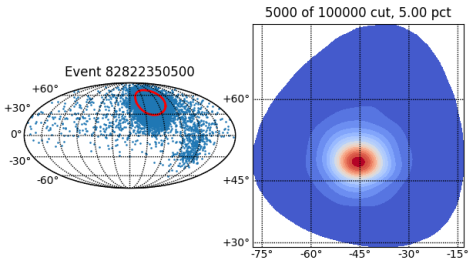

5000 of 100000 cut, 5.00 pct
Event 82056930700

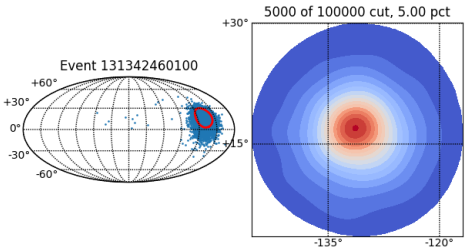

5000 of 100000 cut, 5.00 pct
Event 110856421300

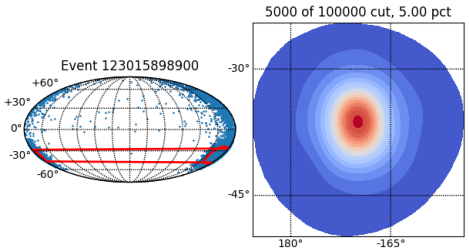

4999 of 100000 cut, 5.00 pct
Event 72036095600

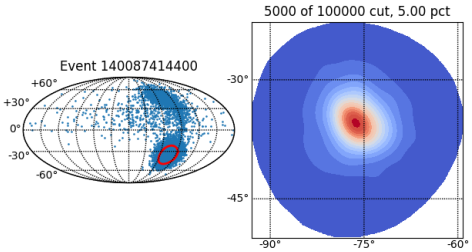

5000 of 100000 cut, 5.00 pct
Event 162093230600

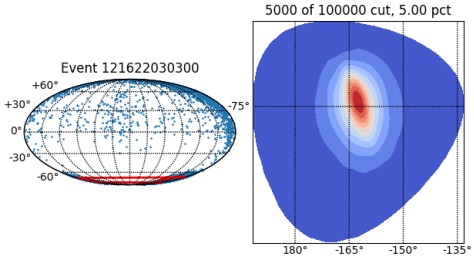

5000 of 100000 cut, 5.00 pct
Event 71863015900
+60°
+30°
0°
-30°
-60°
+30°
+15°
0°
-15°
-60°
-45°
-30°
-15°

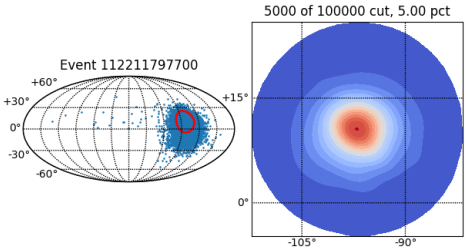

5000 of 100000 cut, 5.00 pct
Event 82822350500
+60°
+30°
0°
-30°
-60°
+60°
+45°
+30°
-75°
-60°
-45°
-30°
-15°

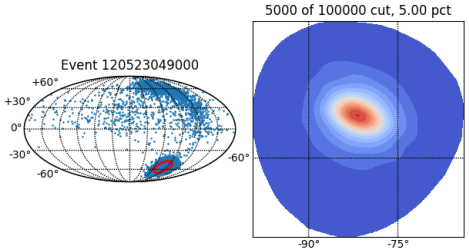

5000 of 100000 cut, 5.00 pct
Event 131342460100
+60°
+30°
0°
-30°
-60°
+30°
+15°
-135°
-120°

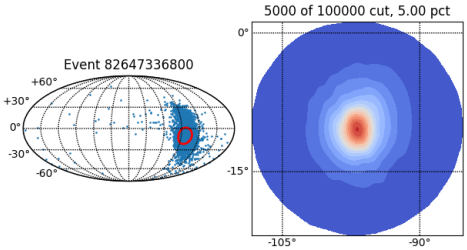

5000 of 100000 cut, 5.00 pct
Event 123015898900
+60°
+30°
0°
-30°
-60°
-30°
-45°
180°
-165°

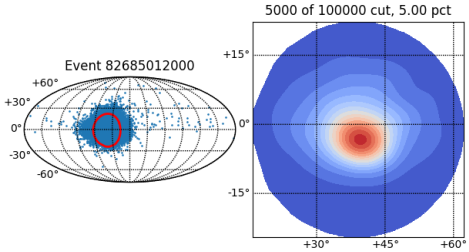

5000 of 100000 cut, 5.00 pct
Event 140087414400
+60°
+30°
0°
-30°
-60°
-30°
-45°
-90°
-75°
-60°

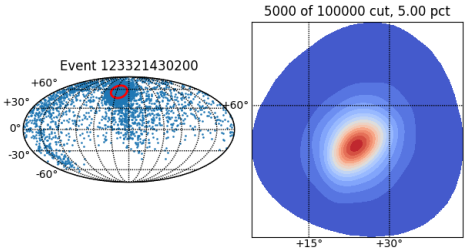

5000 of 100000 cut, 5.00 pct
Event 121622030300
+60°
+30°
0°
-30°
-60°
-75°
180°
-165°
-150°
-135°

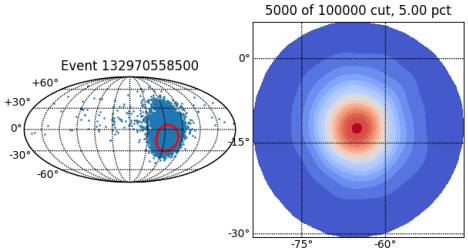

5000 of 100000 cut, 5.00 pct
Event 112211797700
+60°
+30°
0°
-30°
-60°
+15°
0°
-105°
-90°

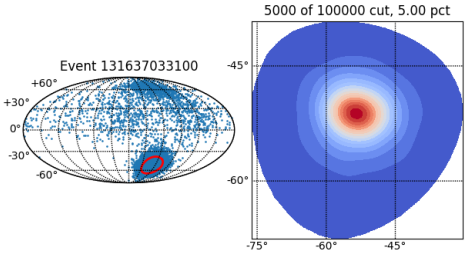

5000 of 100000 cut, 5.00 pct
Event 120523049000
+60°
+30°
0°
-30°
-60°
-60°
-90°
-75°

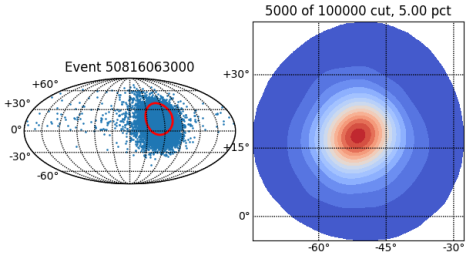

5000 of 100000 cut, 5.00 pct
Event 82647336800
+60°
+30°
0°
-30°
-60°
0°
-15°
-105°
-90°

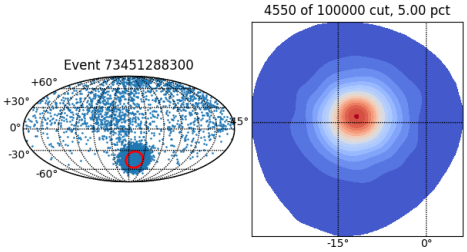

5000 of 100000 cut, 5.00 pct
Event 82685012000
+60°
+30°
0°
-30°
-60°
+15°
0°
-15°
+30°
+45°
+60°

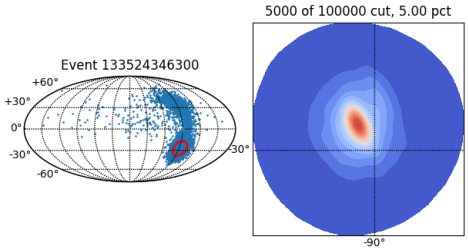

5000 of 100000 cut, 5.00 pct
Event 123321430200
+60°
+30°
0°
-30°
-60°
+60°
+15°
+30°

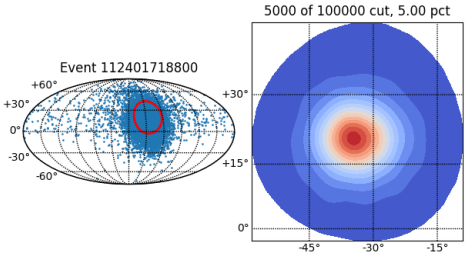

5000 of 100000 cut, 5.00 pct
Event 132970558500
+60°
+30°
0°
-30°
-60°
0°
-15°
-30°
-75°
-60°

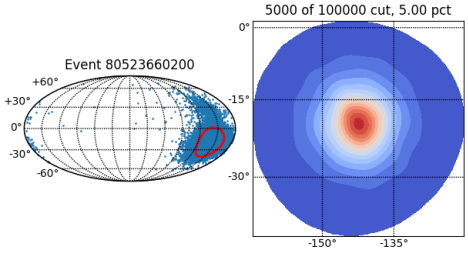

5000 of 100000 cut, 5.00 pct
Event 131637033100
+60°
+30°
0°
-30°
-60°
-45°
-60°
-75°
-60°
-45°

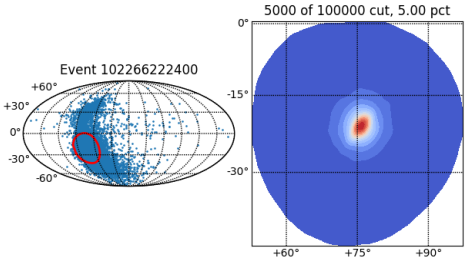

5000 of 100000 cut, 5.00 pct
Event 50816063000
+60°
+30°
0°
-30°
-60°
+30°
+15°
0°
-60°
-45°
-30°

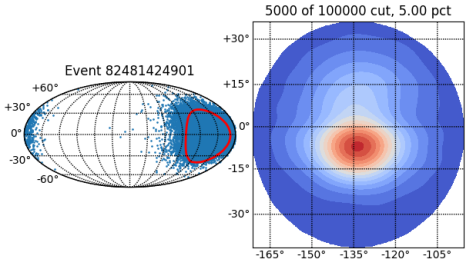

4550 of 100000 cut, 5.00 pct
Event 73451288300
+60°
+30°
0°
-30°
-60°
-15°
0°

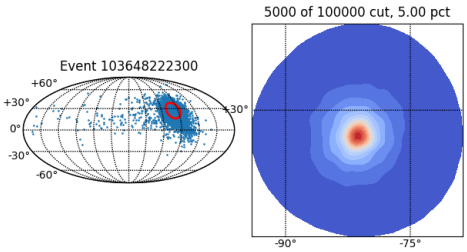

5000 of 100000 cut, 5.00 pct
Event 133524346300
+60°
+30°
0°
-30°
-60°
-90°

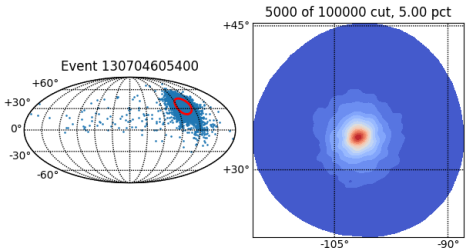

5000 of 100000 cut, 5.00 pct
Event 112401718800

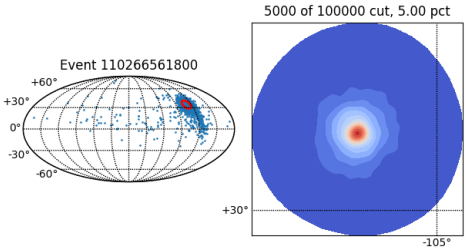

5000 of 100000 cut, 5.00 pct
Event 80523660200

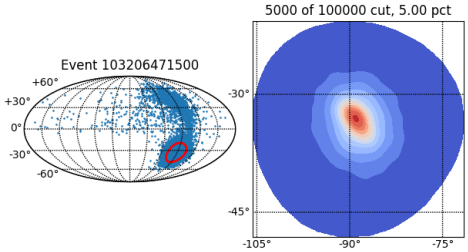

5000 of 100000 cut, 5.00 pct
Event 102266222400

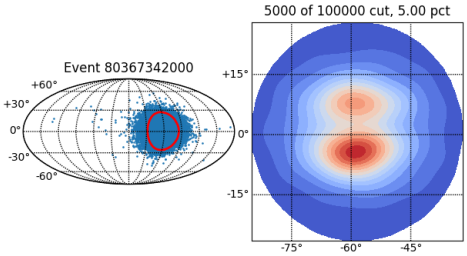

5000 of 100000 cut, 5.00 pct
Event 82481424901

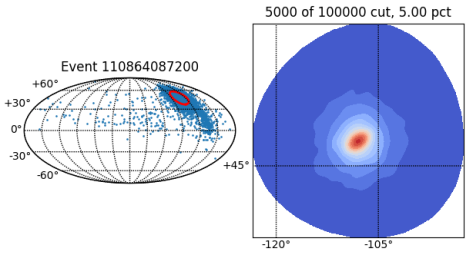

5000 of 100000 cut, 5.00 pct
Event 103648222300

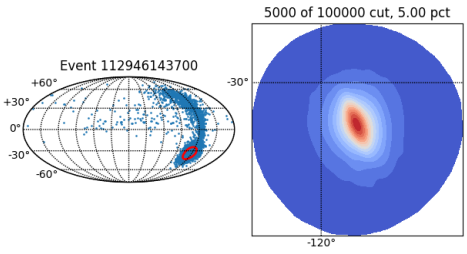

5000 of 100000 cut, 5.00 pct
Event 130704605400

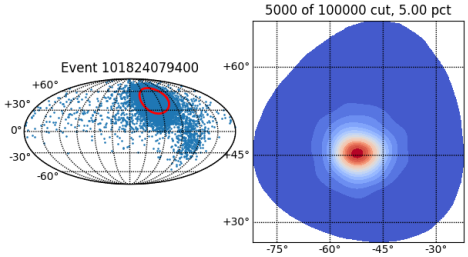

5000 of 100000 cut, 5.00 pct
Event 110266561800

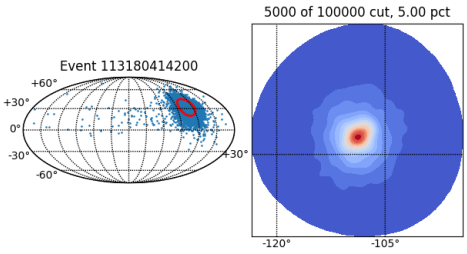

5000 of 100000 cut, 5.00 pct
Event 103206471500

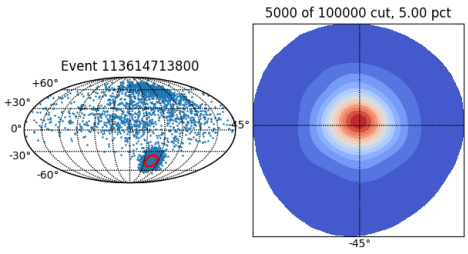

5000 of 100000 cut, 5.00 pct
Event 80367342000
+60°
+30°
0°
-30°
-60°
+15°
0°
-15°
-75°
-60°
-45°

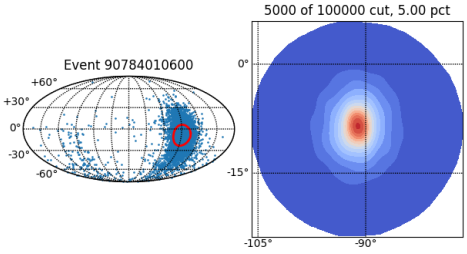

5000 of 100000 cut, 5.00 pct
Event 110864087200
+60°
+30°
0°
-30°
-60°
+45°
-120°
-105°

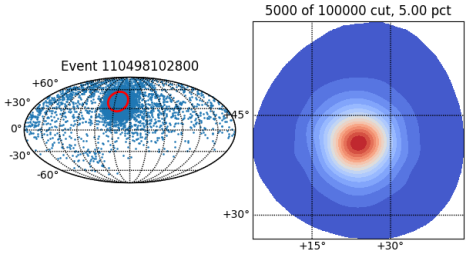

5000 of 100000 cut, 5.00 pct
Event 112946143700
+60°
+30°
0°
-30°
-60°
-30°
-120°

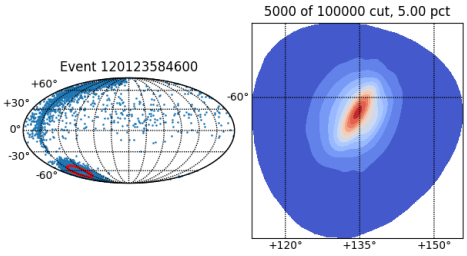

5000 of 100000 cut, 5.00 pct
Event 101824079400
+60°
+30°
0°
-30°
-60°
+60°
+45°
+30°
-75°
-60°
-45°
-30°

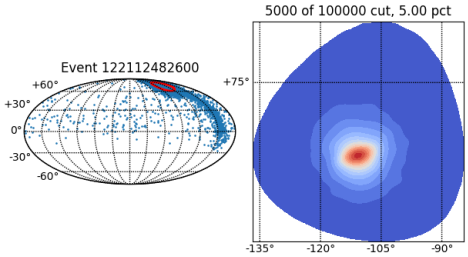

5000 of 100000 cut, 5.00 pct
Event 113180414200
+60°
+30°
0°
-30°
-60°
+30°
-120°
-105°

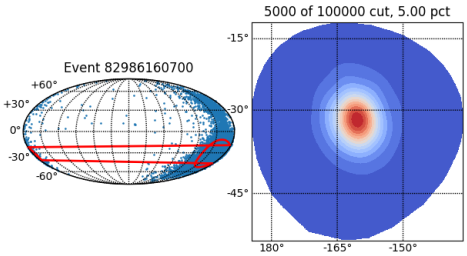

5000 of 100000 cut, 5.00 pct
Event 113614713800
+60°
+30°
0°
-30°
-60°
-45°

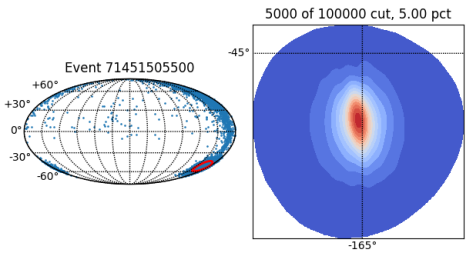

5000 of 100000 cut, 5.00 pct
Event 90784010600
+60°
+30°
0°
-30°
-60°
0°
-15°
-105°
-90°

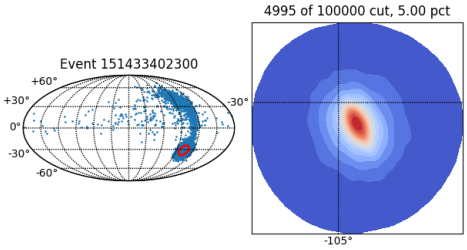

5000 of 100000 cut, 5.00 pct
Event 110498102800
+60°
+30°
0°
-30°
-60°
+45°
+30°
+15°
+30°

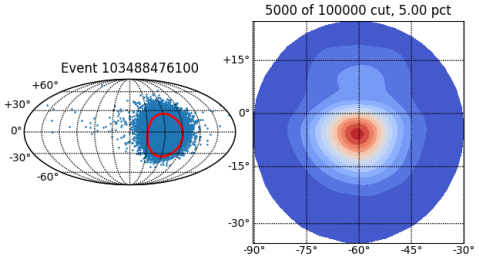

5000 of 100000 cut, 5.00 pct
Event 120123584600
+60°
+30°
0°
-30°
-60°
-60°
+120°
+135°
+150°

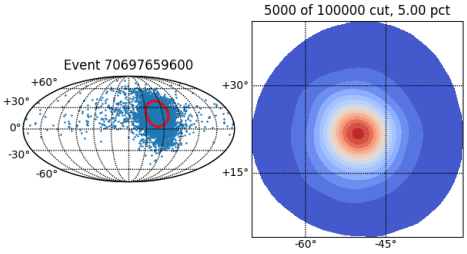

5000 of 100000 cut, 5.00 pct
Event 122112482600
+60°
+30°
0°
-30°
-60°
+75°
-135°
-120°
-105°
-90°

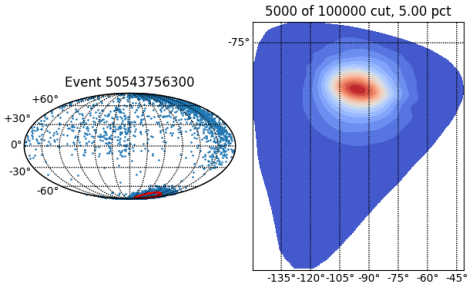

5000 of 100000 cut, 5.00 pct
Event 82986160700
+60°
+30°
0°
-30°
-60°
-15°
-30°
-45°
180°
-165°
-150°

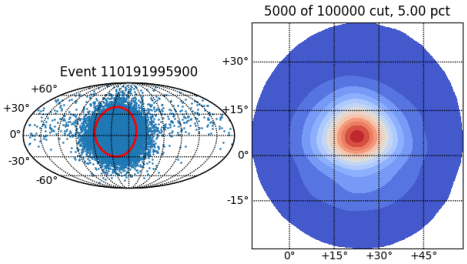

5000 of 100000 cut, 5.00 pct
Event 71451505500
+60°
+30°
0°
-30°
-60°
-45°
-165°

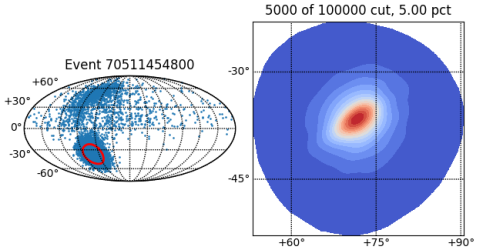

4995 of 100000 cut, 5.00 pct
Event 151433402300
+60°
+30°
0°
-30°
-60°
-30°
-105°

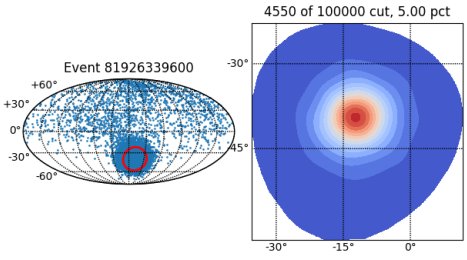

5000 of 100000 cut, 5.00 pct
Event 103488476100
+60°
+30°
0°
-30°
-60°
+15°
0°
-15°
-30°
-90°
-75°
-60°
-45°
-30°

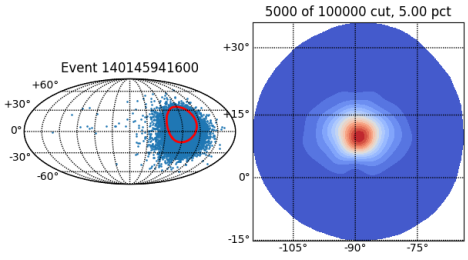

5000 of 100000 cut, 5.00 pct
Event 70697659600
+60°
+30°
0°
-30°
-60°
+30°
+15°
-60°
-45°

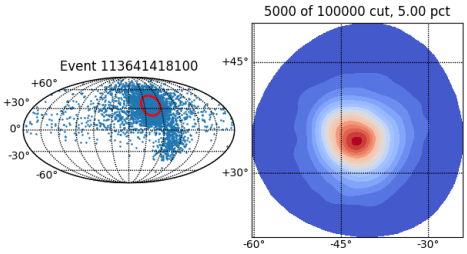

5000 of 100000 cut, 5.00 pct
Event 50543756300
+60°
+30°
0°
-30°
-60°
-75°
-135°
-120°
-105°
-90°
-75°
-60°
-45°

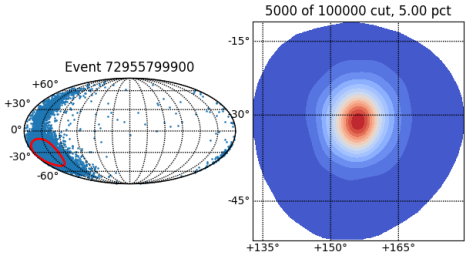

5000 of 100000 cut, 5.00 pct
Event 110191995900

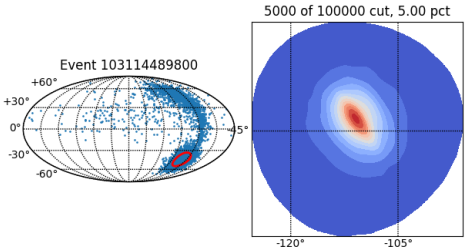

5000 of 100000 cut, 5.00 pct
Event 70511454800

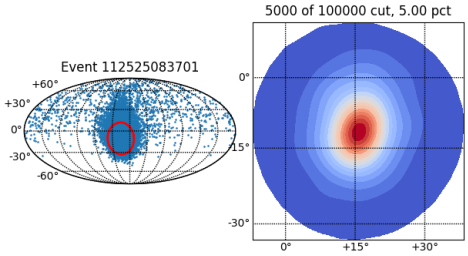

4550 of 100000 cut, 5.00 pct
Event 81926339600

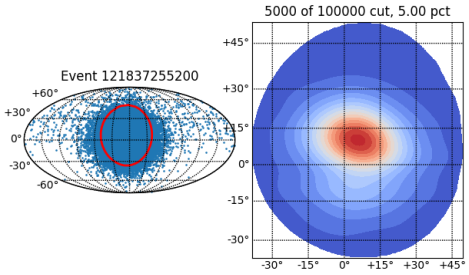

5000 of 100000 cut, 5.00 pct
Event 140145941600


5000 of 100000 cut, 5.00 pct
Event 113641418100

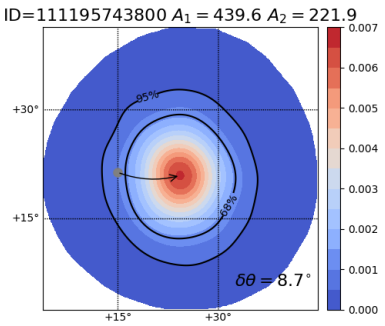

5000 of 100000 cut, 5.00 pct
Event 72955799900

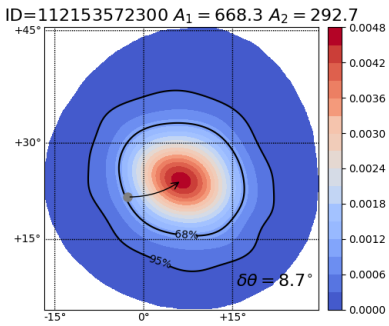

5000 of 100000 cut, 5.00 pct
Event 103114489800

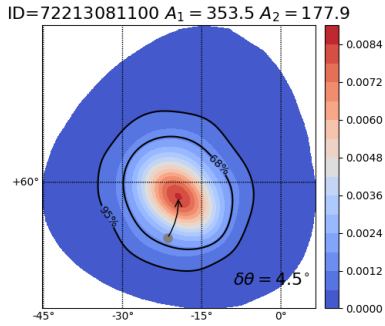

5000 of 100000 cut, 5.00 pct
Event 112525083701

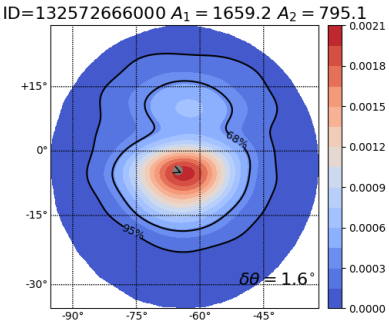

Event 121837255200
+60°
+30°
0°
-30°
-60°
5000 of 100000 cut, 5.00 pct
+45°
+30°
+15°
0°
-15°
-30°
-30°
-15°
0°
+15°
+30°
+45°

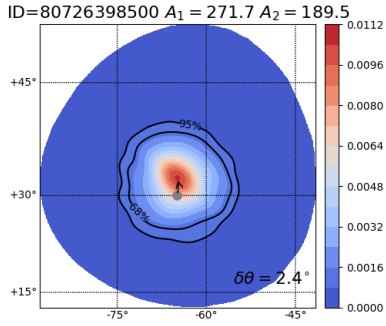
Chapter B: Appendix B

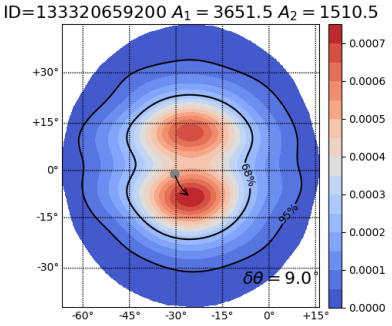


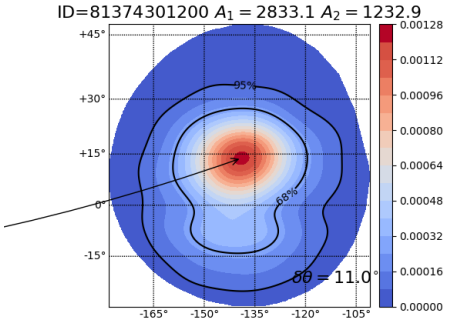


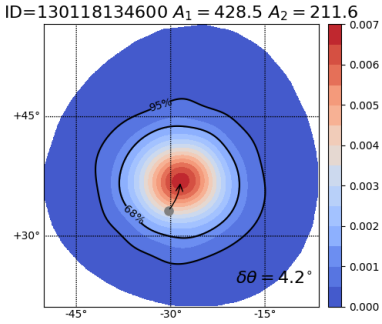


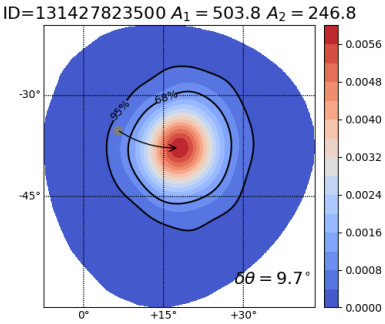


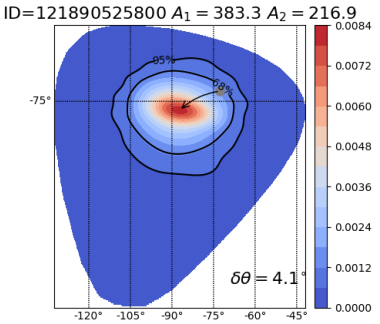


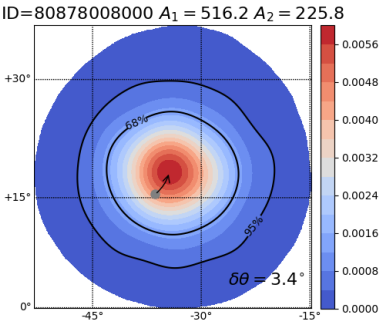


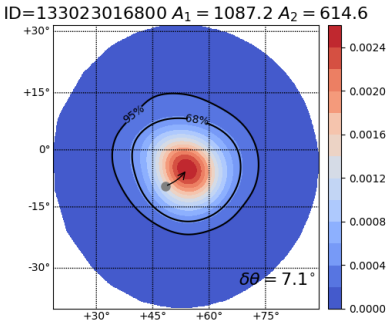

ID=81374301200 $A_1 = 2833.1$ $A_2 = 1232.9$
95%
68%
$\delta\theta = 11.0^\circ$

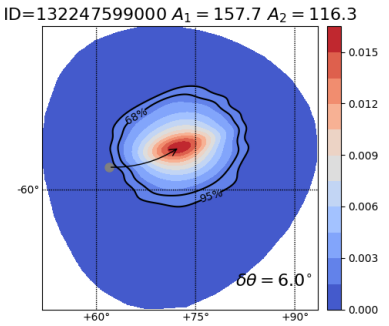

ID=130118134600 $A_1 = 428.5$ $A_2 = 211.6$
95%
68%
$\delta\theta = 4.2^\circ$

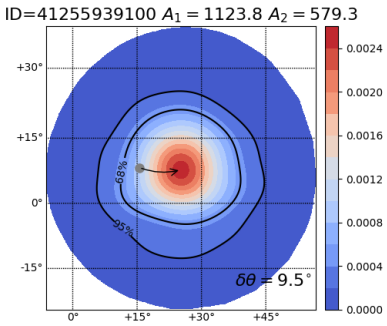

ID=131427823500 $A_1 = 503.8$ $A_2 = 246.8$
95%
68%
$\delta\theta = 9.7^\circ$

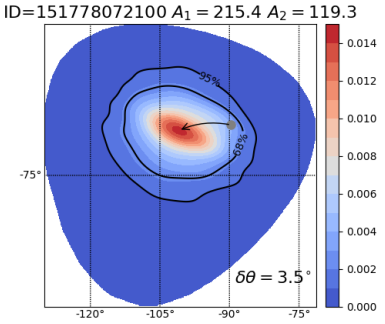

ID=121890525800 $A_1 = 383.3$ $A_2 = 216.9$
95%
68%
$\delta\theta = 4.1^\circ$

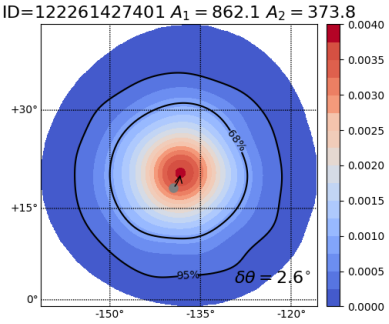

ID=80878008000 $A_1 = 516.2$ $A_2 = 225.8$
68%
95%
$\delta\theta = 3.4^\circ$

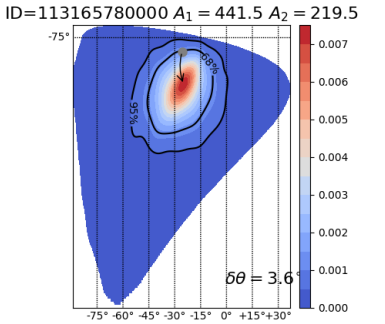

ID=133023016800 $A_1 = 1087.2$ $A_2 = 614.6$
95%
68%
$\delta\theta = 7.1^\circ$

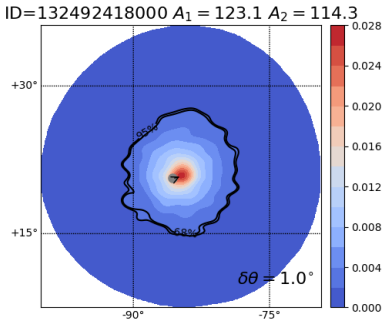

ID=132247599000 $A_1 = 157.7$ $A_2 = 116.3$
68%
95%
$\delta\theta = 6.0^\circ$

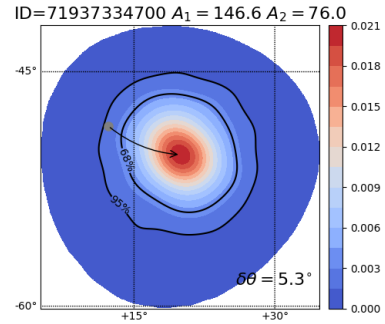

ID=41255939100 $A_1 = 1123.8$ $A_2 = 579.3$
68%
95%
$\delta\theta = 9.5^\circ$

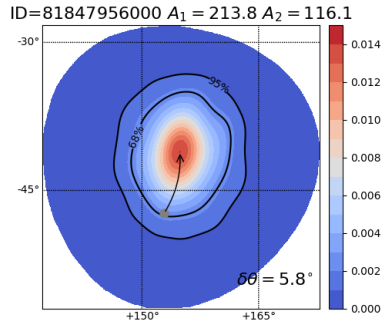

ID=151778072100 $A_1 = 215.4$ $A_2 = 119.3$
95%
68%
$\delta\theta = 3.5^\circ$
-120°
-105°
-90°
-75°
-75°
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000

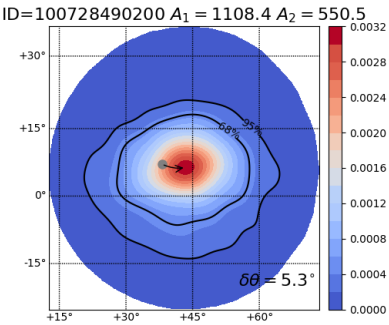

ID=122261427401 $A_1 = 862.1$ $A_2 = 373.8$
68%
95%
$\delta\theta = 2.6^\circ$
+30°
+15°
0°
-150°
-135°
-120°
0.0040
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000

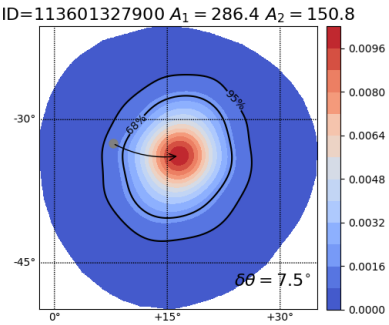

ID=113165780000 $A_1 = 441.5$ $A_2 = 219.5$
68%
95%
$\delta\theta = 3.6^\circ$
-75°
-75° -60° -45° -30° -15° 0° +15° +30°
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000

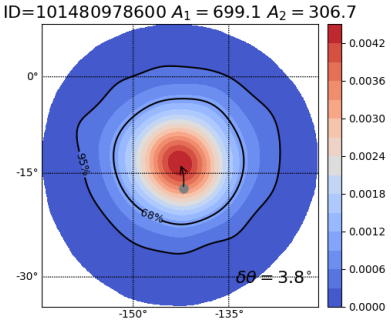

ID=132492418000 $A_1 = 123.1$ $A_2 = 114.3$
95%
68%
$\delta\theta = 1.0^\circ$
+30°
+15°
-90°
-75°
0.028
0.024
0.020
0.016
0.012
0.008
0.004
0.000

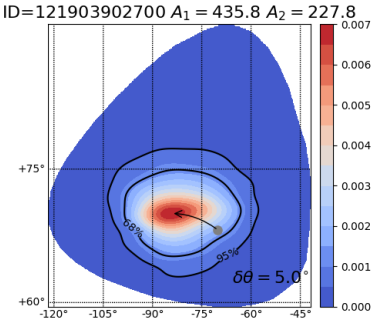

ID=71937334700 $A_1 = 146.6$ $A_2 = 76.0$
68%
95%
$\delta\theta = 5.3^\circ$
-45°
-60°
+15°
+30°
0.021
0.018
0.015
0.012
0.009
0.006
0.003
0.000

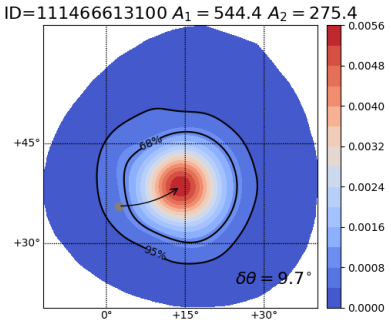

ID=81847956000 $A_1 = 213.8$ $A_2 = 116.1$
95%
68%
$\delta\theta = 5.8^\circ$
-30°
-45°
+150°
+165°
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000

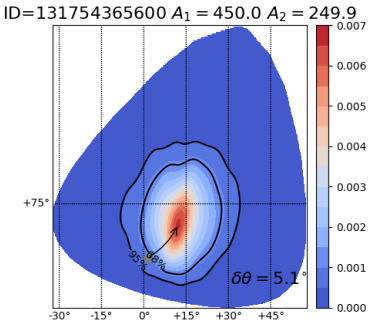

ID=100728490200 $A_1 = 1108.4$ $A_2 = 550.5$
68%
95%
$\delta\theta = 5.3^\circ$
+30°
+15°
0°
-15°
+15°
+30°
+45°
+60°
0.0032
0.0028
0.0024
0.0020
0.0016
0.0012
0.0008
0.0004
0.0000

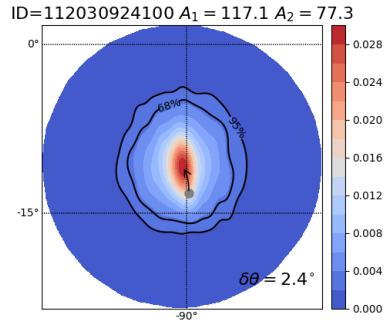

ID=113601327900 $A_1 = 286.4$ $A_2 = 150.8$
95%
68%
$\delta\theta = 7.5^\circ$
-30°
-45°
0°
+15°
+30°
0.0096
0.0080
0.0064
0.0048
0.0032
0.0016
0.0000

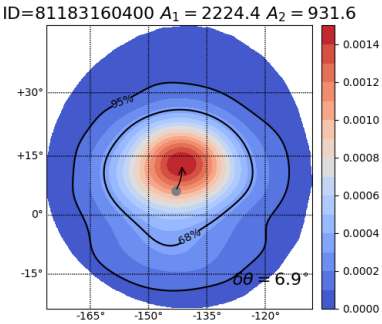

ID=101480978600 $A_1 = 699.1$ $A_2 = 306.7$
$\delta\theta = 3.8^\circ$
95%
68%

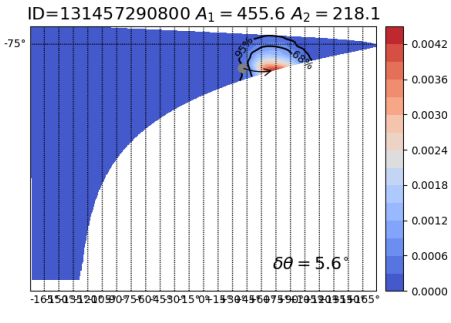

ID=121903902700 $A_1 = 435.8$ $A_2 = 227.8$
$\delta\theta = 5.0^\circ$
68%
95%

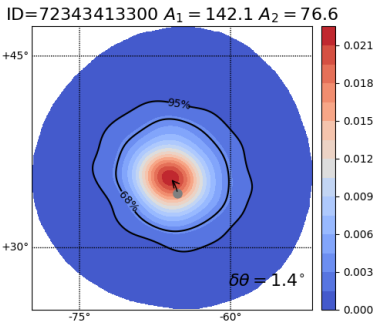

ID=111466613100 $A_1 = 544.4$ $A_2 = 275.4$
$\delta\theta = 9.7^\circ$
68%
95%

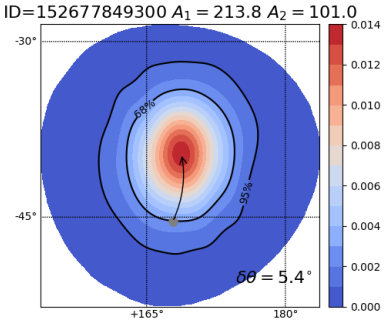

ID=131754365600 $A_1 = 450.0$ $A_2 = 249.9$
$\delta\theta = 5.1^\circ$
95%
68%

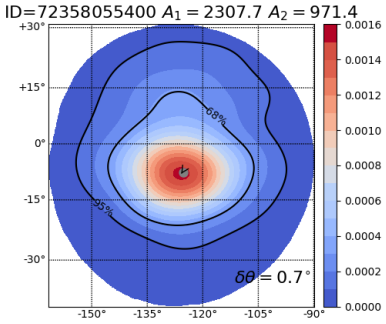

ID=112030924100 $A_1 = 117.1$ $A_2 = 77.3$
$\delta\theta = 2.4^\circ$
68%
95%

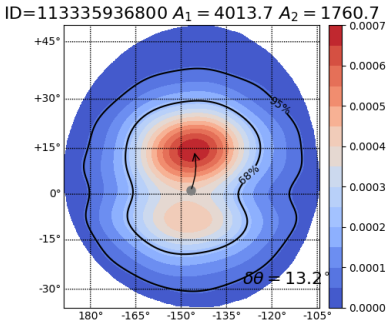

ID=81183160400 $A_1 = 2224.4$ $A_2 = 931.6$
$\delta\theta = 6.9^\circ$
95%
68%

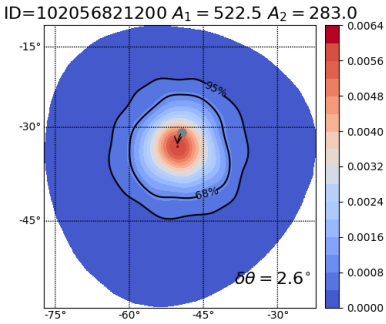

ID=131457290800 $A_1 = 455.6$ $A_2 = 218.1$
$\delta\theta = 5.6^\circ$
95%
68%

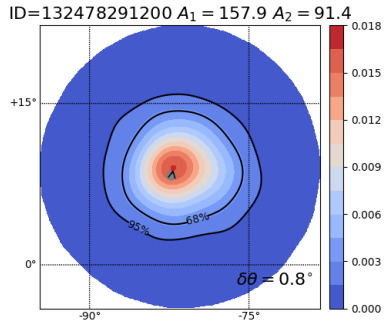

ID=72343413300 $A_1 = 142.1$ $A_2 = 76.6$
$\delta\theta = 1.4^\circ$
95%
68%

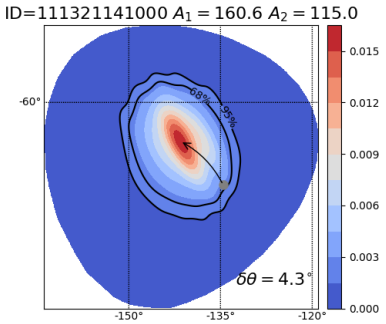

ID=152677849300 $A_1 = 213.8$ $A_2 = 101.0$
$\delta\theta = 5.4^\circ$

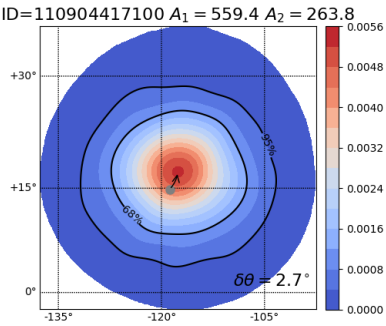

ID=72358055400 $A_1 = 2307.7$ $A_2 = 971.4$
$\delta\theta = 0.7^\circ$

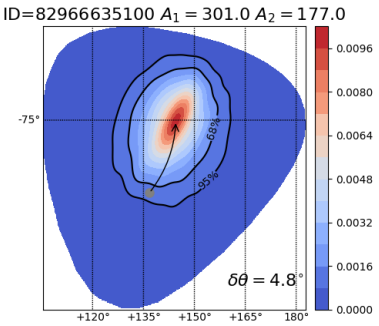

ID=113335936800 $A_1 = 4013.7$ $A_2 = 1760.7$
$\delta\theta = 13.2^\circ$

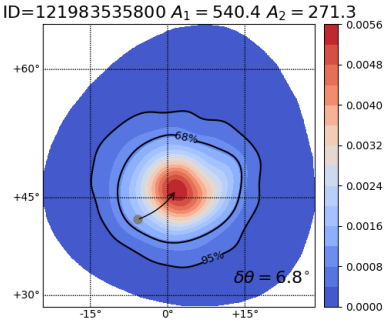

ID=102056821200 $A_1 = 522.5$ $A_2 = 283.0$
$\delta\theta = 2.6^\circ$

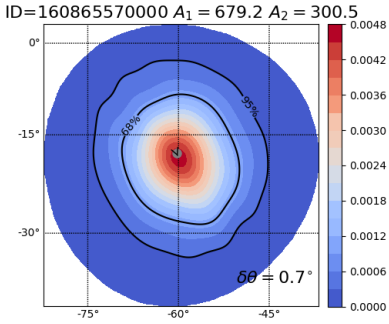

ID=132478291200 $A_1 = 157.9$ $A_2 = 91.4$
$\delta\theta = 0.8^\circ$

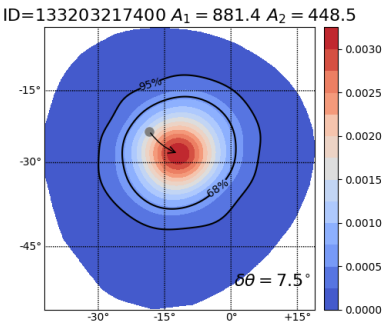

ID=111321141000 $A_1 = 160.6$ $A_2 = 115.0$
$\delta\theta = 4.3^\circ$

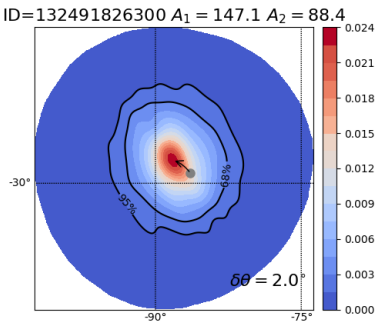

ID=110904417100 $A_1 = 559.4$ $A_2 = 263.8$
$\delta\theta = 2.7^\circ$

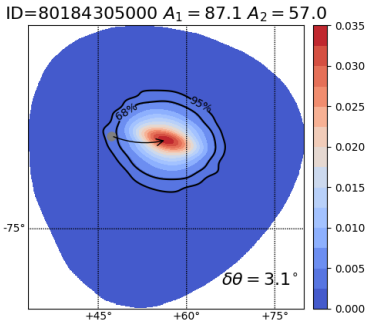

ID=82966635100 $A_1 = 301.0$ $A_2 = 177.0$
$\delta\theta = 4.8^\circ$

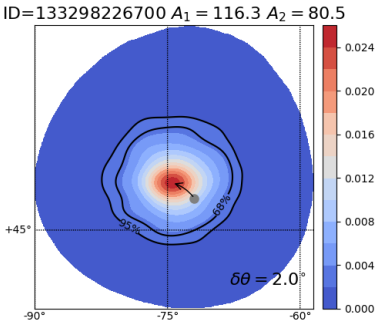

ID=121983535800 $A_1 = 540.4$ $A_2 = 271.3$
+60°
+45°
+30°
-15°
0°
+15°
68%
95%
$\delta\theta = 6.8^\circ$
0.0056
0.0048
0.0040
0.0032
0.0024
0.0016
0.0008
0.0000

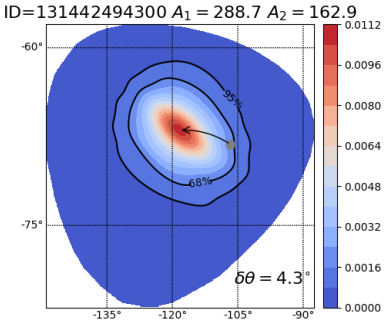

ID=160865570000 $A_1 = 679.2$ $A_2 = 300.5$
0°
-15°
-30°
-75°
-60°
-45°
68%
95%
$\delta\theta = 0.7^\circ$
0.0048
0.0042
0.0036
0.0030
0.0024
0.0018
0.0012
0.0006
0.0000

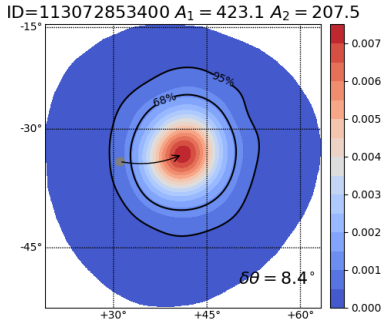

ID=133203217400 $A_1 = 881.4$ $A_2 = 448.5$
-15°
-30°
-45°
-30°
-15°
0°
+15°
95%
68%
$\delta\theta = 7.5^\circ$
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000

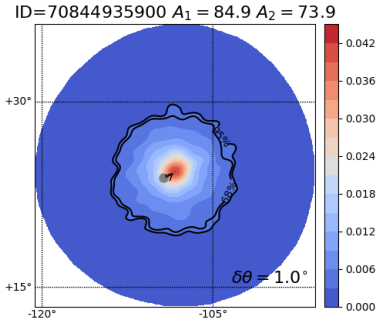

ID=132491826300 $A_1 = 147.1$ $A_2 = 88.4$
-30°
-90°
-75°
68%
95%
$\delta\theta = 2.0^\circ$
0.024
0.021
0.018
0.015
0.012
0.009
0.006
0.003
0.000

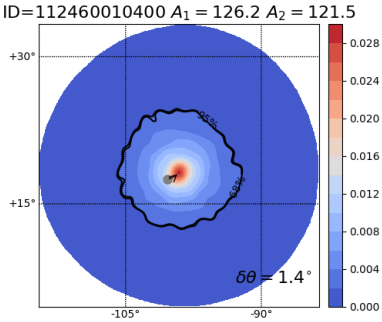

ID=80184305000 $A_1 = 87.1$ $A_2 = 57.0$
-75°
+45°
+60°
+75°
68%
95%
$\delta\theta = 3.1^\circ$
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

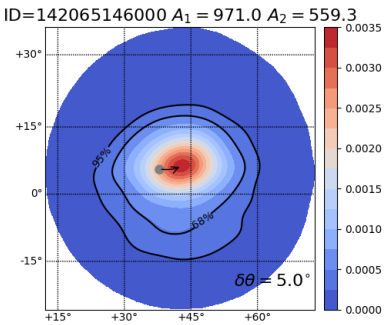

ID=133298226700 $A_1 = 116.3$ $A_2 = 80.5$
+45°
-90°
-75°
-60°
68%
95%
$\delta\theta = 2.0^\circ$
0.024
0.020
0.016
0.012
0.008
0.004
0.000

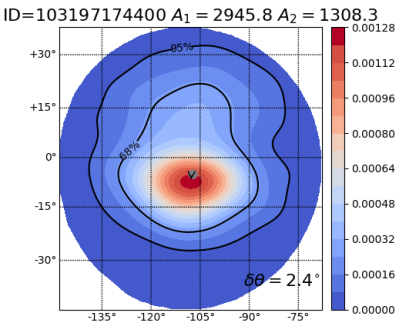

ID=131442494300 $A_1 = 288.7$ $A_2 = 162.9$
-60°
-75°
-135°
-120°
-105°
-90°
95%
68%
$\delta\theta = 4.3^\circ$
0.0112
0.0096
0.0080
0.0064
0.0048
0.0032
0.0016
0.0000

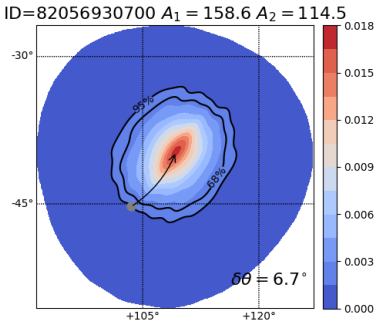

ID=113072853400 $A_1 = 423.1$ $A_2 = 207.5$
-15°
-30°
-45°
+30°
+45°
+60°
68%
95%
$\delta\theta = 8.4^\circ$
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000

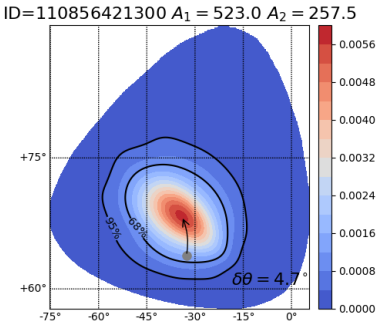

ID=70844935900 $A_1 = 84.9$ $A_2 = 73.9$
$\delta\theta = 1.0^\circ$

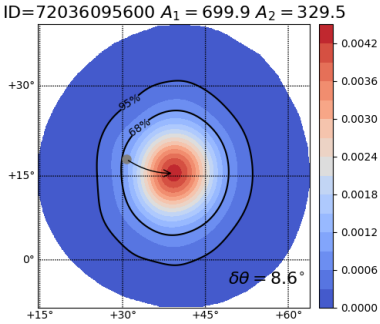

ID=112460010400 $A_1 = 126.2$ $A_2 = 121.5$
$\delta\theta = 1.4^\circ$

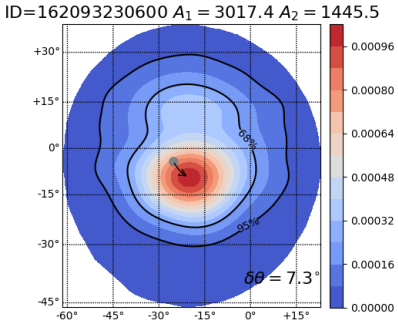

ID=142065146000 $A_1 = 971.0$ $A_2 = 559.3$
$\delta\theta = 5.0^\circ$


ID=103197174400 $A_1 = 2945.8$ $A_2 = 1308.3$
$\delta\theta = 2.4^\circ$

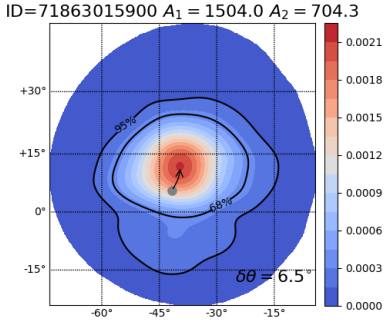

ID=82056930700 $A_1 = 158.6$ $A_2 = 114.5$
$\delta\theta = 6.7^\circ$


ID=110856421300 $A_1 = 523.0$ $A_2 = 257.5$
$\delta\theta = 4.7^\circ$

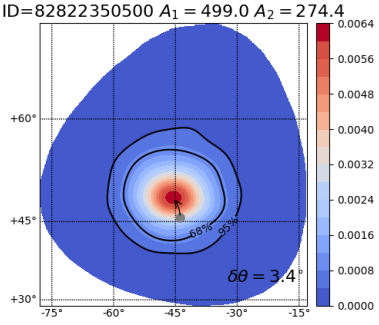

ID=72036095600 $A_1 = 699.9$ $A_2 = 329.5$
$\delta\theta = 8.6^\circ$


ID=162093230600 $A_1 = 3017.4$ $A_2 = 1445.5$
$\delta\theta = 7.3^\circ$

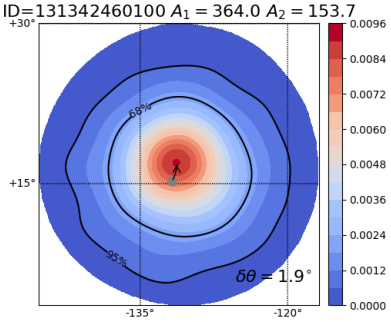
ID=71863015900 $A_1 = 1504.0$ $A_2 = 704.3$



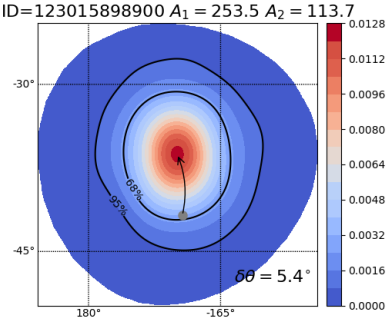
ID=82822350500 $A_1 = 499.0$ $A_2 = 274.4$



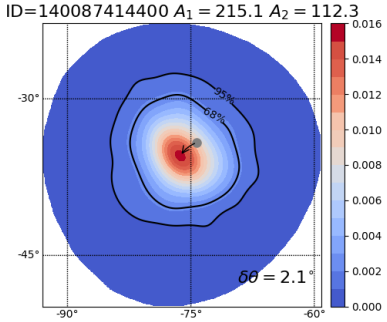
ID=131342460100 $A_1 = 364.0$ $A_2 = 153.7$



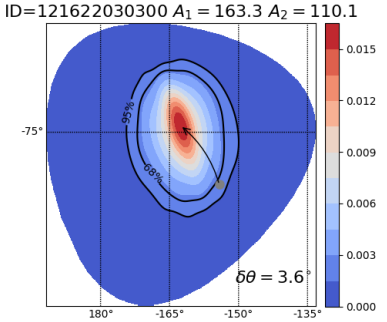
ID=123015898900 $A_1 = 253.5$ $A_2 = 113.7$



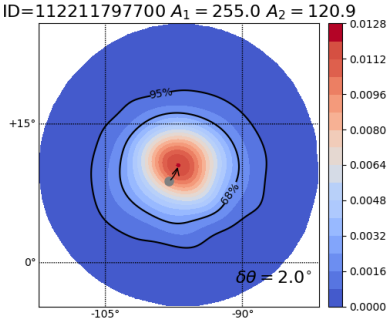
ID=140087414400 $A_1 = 215.1$ $A_2 = 112.3$



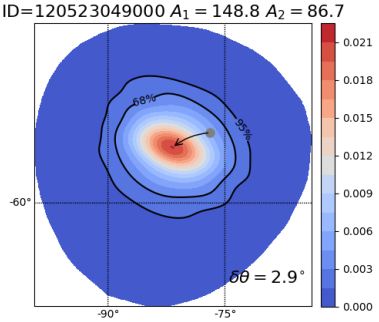
ID=121622030300 $A_1 = 163.3$ $A_2 = 110.1$

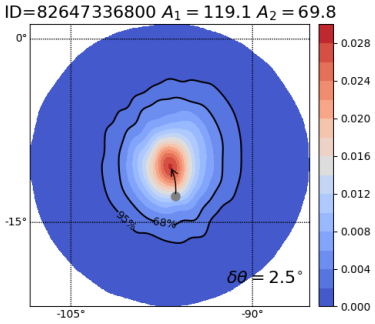


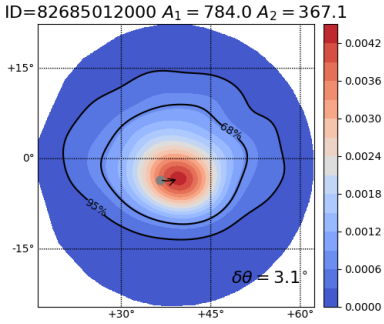
ID=112211797700 $A_1 = 255.0$ $A_2 = 120.9$

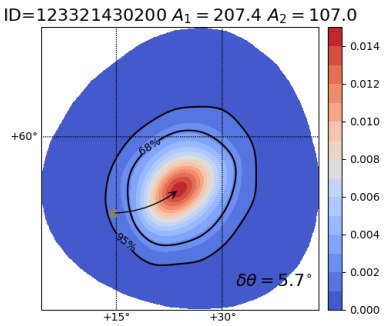


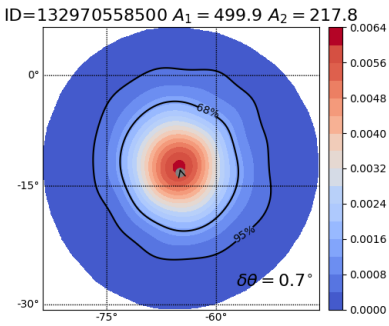
ID=120523049000 $A_1 = 148.8$ $A_2 = 86.7$

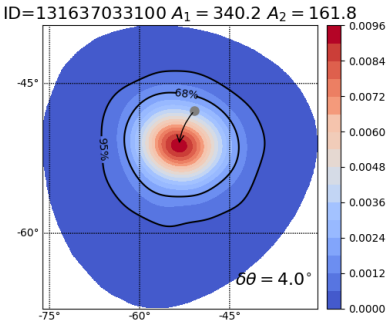


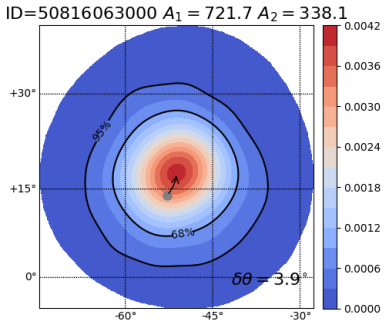

ID=82647336800 $A_1 = 119.1$ $A_2 = 69.8$
0°
-15°
-105°
-90°
95%
68%
$\delta\theta = 2.5^\circ$
0.028
0.024
0.020
0.016
0.012
0.008
0.004
0.000

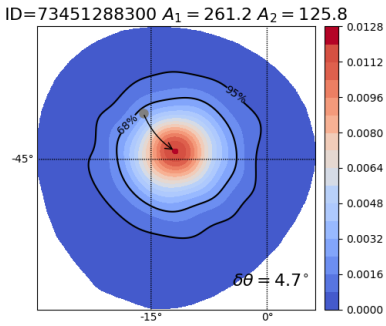

ID=82685012000 $A_1 = 784.0$ $A_2 = 367.1$
+15°
0°
-15°
+30°
+45°
+60°
68%
95%
$\delta\theta = 3.1^\circ$
0.0042
0.0036
0.0030
0.0024
0.0018
0.0012
0.0006
0.0000

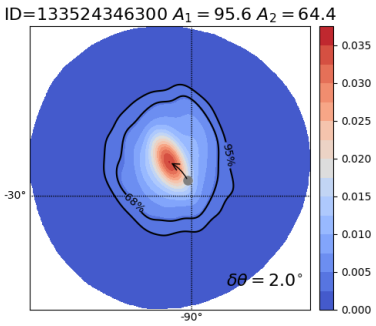

ID=123321430200 $A_1 = 207.4$ $A_2 = 107.0$
+60°
+15°
+30°
68%
95%
$\delta\theta = 5.7^\circ$
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000

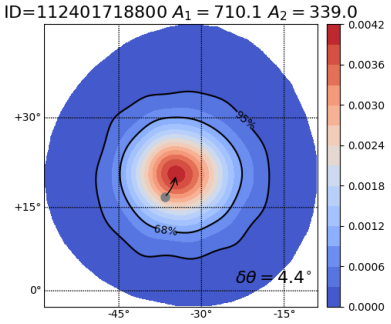

ID=132970558500 $A_1 = 499.9$ $A_2 = 217.8$
0°
-15°
-30°
-75°
-60°
68%
95%
$\delta\theta = 0.7^\circ$
0.0064
0.0056
0.0048
0.0040
0.0032
0.0024
0.0016
0.0008
0.0000

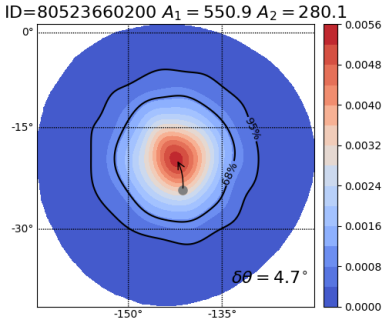

ID=131637033100 $A_1 = 340.2$ $A_2 = 161.8$
-45°
-60°
-75°
-60°
-45°
68%
95%
$\delta\theta = 4.0^\circ$
0.0096
0.0084
0.0072
0.0060
0.0048
0.0036
0.0024
0.0012
0.0000

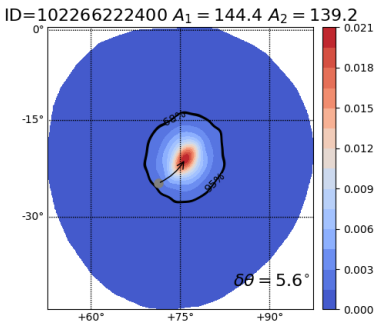

ID=50816063000 $A_1 = 721.7$ $A_2 = 338.1$
+30°
+15°
0°
-60°
-45°
-30°
95%
68%
$\delta\theta = 3.9^\circ$
0.0042
0.0036
0.0030
0.0024
0.0018
0.0012
0.0006
0.0000

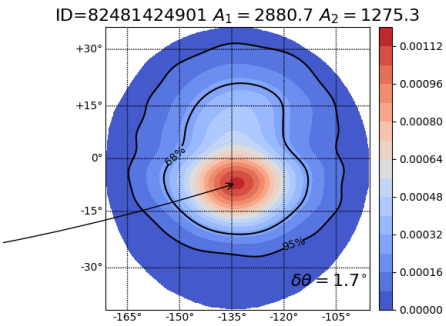

ID=73451288300 $A_1 = 261.2$ $A_2 = 125.8$
-45°
-15°
0°
95%
68%
$\delta\theta = 4.7^\circ$
0.0128
0.0112
0.0096
0.0080
0.0064
0.0048
0.0032
0.0016
0.0000

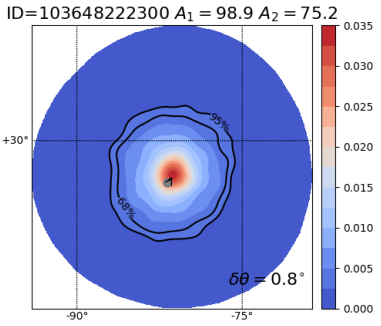

ID=133524346300 $A_1 = 95.6$ $A_2 = 64.4$
-30°
-90°
95%
68%
$\delta\theta = 2.0^\circ$
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

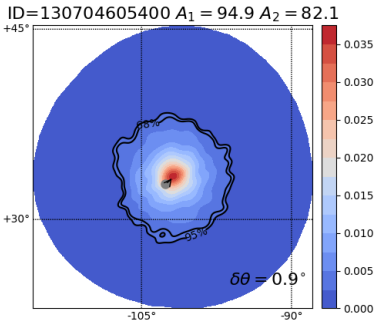

ID=112401718800 $A_1 = 710.1$ $A_2 = 339.0$
+30°
+15°
0°
-45°
-30°
-15°
95%
68%
$\delta\theta = 4.4^\circ$
0.0042
0.0036
0.0030
0.0024
0.0018
0.0012
0.0006
0.0000

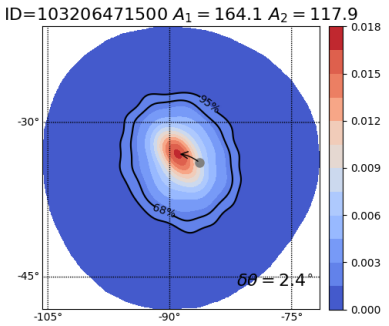

ID=80523660200 $A_1 = 550.9$ $A_2 = 280.1$
0°
-15°
-30°
-150°
-135°
95%
68%
$\delta\theta = 4.7^\circ$
0.0056
0.0048
0.0040
0.0032
0.0024
0.0016
0.0008
0.0000

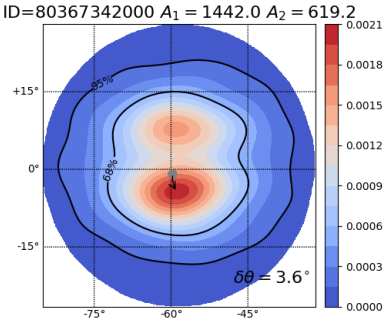

ID=102266222400 $A_1 = 144.4$ $A_2 = 139.2$
0°
-15°
-30°
+60°
+75°
+90°
68%
95%
$\delta\theta = 5.6^\circ$
0.021
0.018
0.015
0.012
0.009
0.006
0.003
0.000

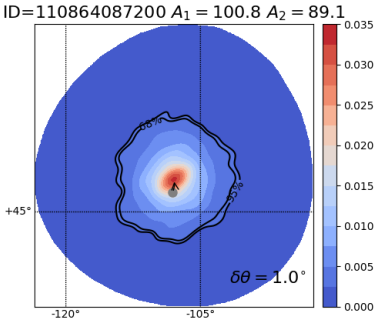

ID=82481424901 $A_1 = 2880.7$ $A_2 = 1275.3$
+30°
+15°
0°
-15°
-30°
-165°
-150°
-135°
-120°
-105°
68%
95%
$\delta\theta = 1.7^\circ$
0.00112
0.00096
0.00080
0.00064
0.00048
0.00032
0.00016
0.00000

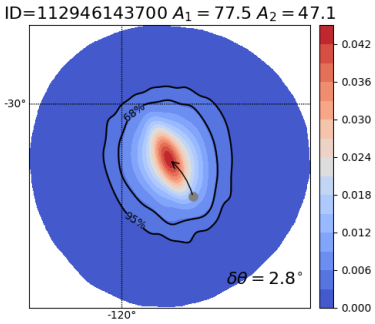

ID=103648222300 $A_1 = 98.9$ $A_2 = 75.2$
+30°
-90°
-75°
95%
68%
$\delta\theta = 0.8^\circ$
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

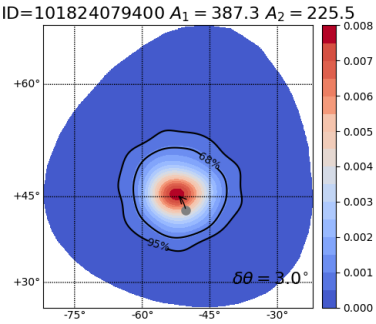

ID=130704605400 $A_1 = 94.9$ $A_2 = 82.1$
+45°
+30°
-105°
-90°
68%
95%
$\delta\theta = 0.9^\circ$
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

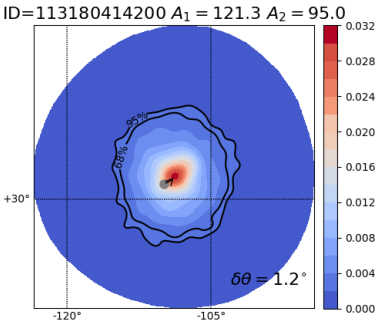

ID=103206471500 $A_1 = 164.1$ $A_2 = 117.9$
-30°
-45°
-105°
-90°
-75°
95%
68%
$\delta\theta = 2.4^\circ$
0.018
0.015
0.012
0.009
0.006
0.003
0.000

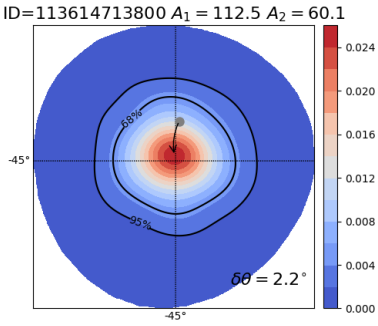

ID=80367342000 $A_1 = 1442.0$ $A_2 = 619.2$
+15°
0°
-15°
-75°
-60°
-45°
68%
95%
$\delta\theta = 3.6^\circ$
0.0021
0.0018
0.0015
0.0012
0.0009
0.0006
0.0003
0.0000

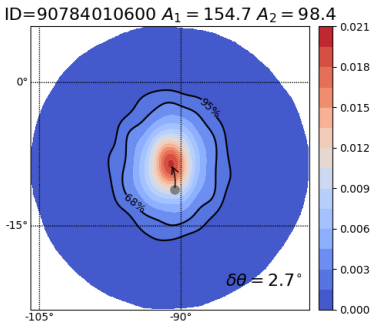

ID=110864087200 $A_1 = 100.8$ $A_2 = 89.1$
+45°
-120°
-105°
68%
95%
$\delta\theta = 1.0^\circ$
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

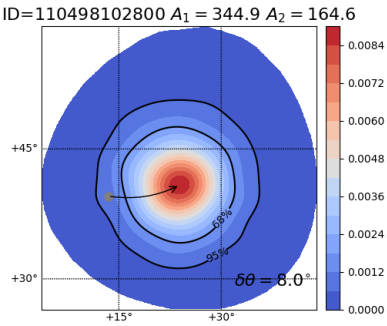

ID=112946143700 $A_1 = 77.5$ $A_2 = 47.1$
-30°
-120°
68%
95%
$\delta\theta = 2.8^\circ$
0.042
0.036
0.030
0.024
0.018
0.012
0.006
0.000

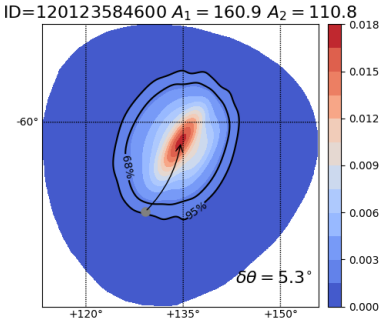

ID=101824079400 $A_1 = 387.3$ $A_2 = 225.5$
+60°
+45°
+30°
-75°
-60°
-45°
-30°
68%
95%
$\delta\theta = 3.0^\circ$
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000

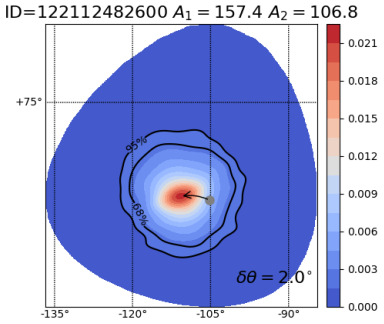

ID=113180414200 $A_1 = 121.3$ $A_2 = 95.0$
+30°
-120°
-105°
68%
95%
$\delta\theta = 1.2^\circ$
0.032
0.028
0.024
0.020
0.016
0.012
0.008
0.004
0.000

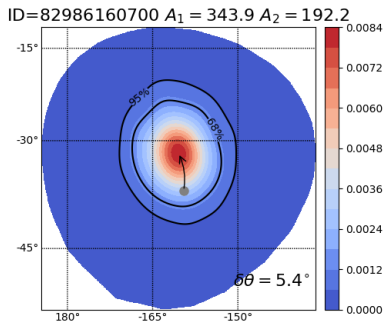

ID=113614713800 $A_1 = 112.5$ $A_2 = 60.1$
-45°
-45°
68%
95%
$\delta\theta = 2.2^\circ$
0.024
0.020
0.016
0.012
0.008
0.004
0.000

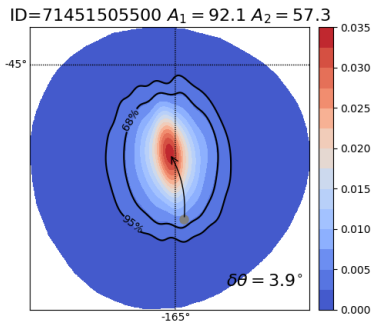

ID=90784010600 $A_1 = 154.7$ $A_2 = 98.4$
0°
-15°
-105°
-90°
68%
95%
$\delta\theta = 2.7^\circ$
0.021
0.018
0.015
0.012
0.009
0.006
0.003
0.000

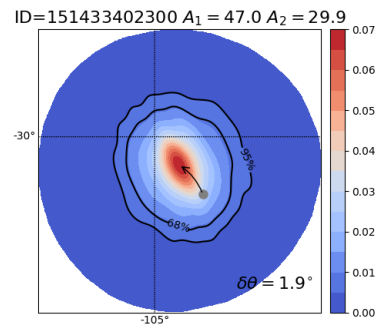

ID=110498102800 $A_1 = 344.9$ $A_2 = 164.6$
+45°
+30°
+15°
+30°
68%
95%
$\delta\theta = 8.0^\circ$
0.0084
0.0072
0.0060
0.0048
0.0036
0.0024
0.0012
0.0000

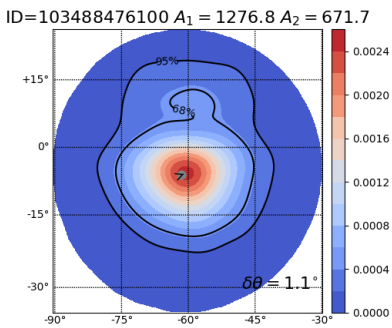

ID=120123584600 $A_1 = 160.9$ $A_2 = 110.8$
$\delta\theta = 5.3^\circ$
68%
95%
-60°
+120°
+135°
+150°
0.018
0.015
0.012
0.009
0.006
0.003
0.000

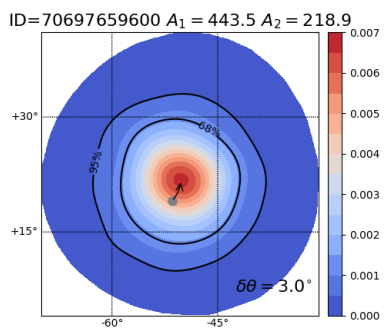

ID=122112482600 $A_1 = 157.4$ $A_2 = 106.8$
$\delta\theta = 2.0^\circ$
68%
95%
+75°
-135°
-120°
-105°
-90°
0.021
0.018
0.015
0.012
0.009
0.006
0.003
0.000

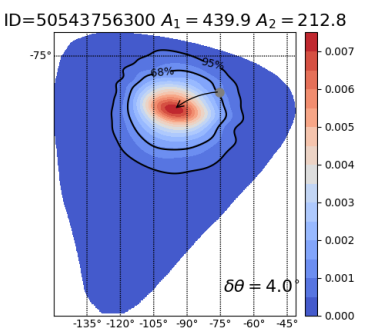

ID=82986160700 $A_1 = 343.9$ $A_2 = 192.2$
$\delta\theta = 5.4^\circ$
68%
95%
-15°
-30°
-45°
180°
-165°
-150°
0.0084
0.0072
0.0060
0.0048
0.0036
0.0024
0.0012
0.0000

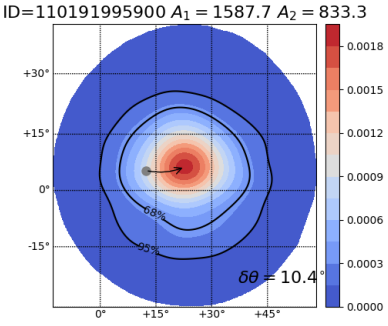

ID=71451505500 $A_1 = 92.1$ $A_2 = 57.3$
$\delta\theta = 3.9^\circ$
68%
95%
-45°
-165°
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

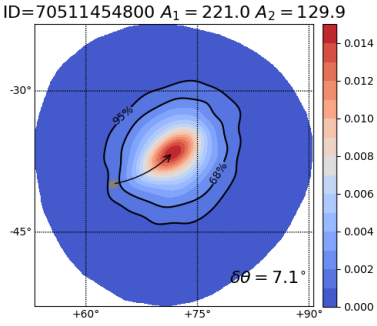

ID=151433402300 $A_1 = 47.0$ $A_2 = 29.9$
$\delta\theta = 1.9^\circ$
68%
95%
-30°
-105°
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00

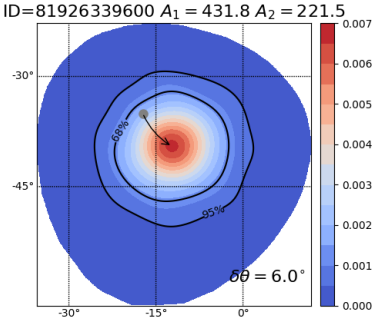

ID=103488476100 $A_1 = 1276.8$ $A_2 = 671.7$
$\delta\theta = 1.1^\circ$
68%
95%
+15°
0°
-15°
-30°
-90°
-75°
-60°
-45°
-30°
0.0024
0.0020
0.0016
0.0012
0.0008
0.0004
0.0000

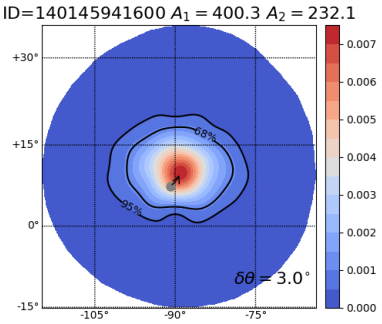

ID=70697659600 $A_1 = 443.5$ $A_2 = 218.9$
$\delta\theta = 3.0^\circ$
68%
95%
+30°
+15°
-60°
-45°
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000

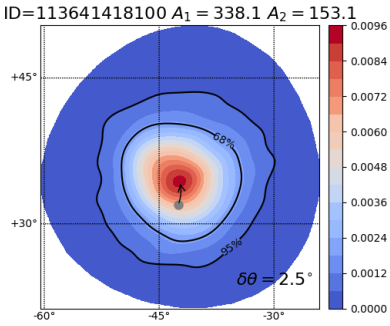

ID=50543756300 $A_1 = 439.9$ $A_2 = 212.8$
$\delta\theta = 4.0^\circ$
68%
95%
-75°
-135°
-120°
-105°
-90°
-75°
-60°
-45°
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000

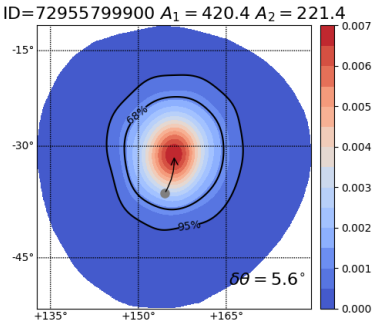

ID=110191995900 $A_1 = 1587.7$ $A_2 = 833.3$
$\delta\theta = 10.4^\circ$
68%
95%

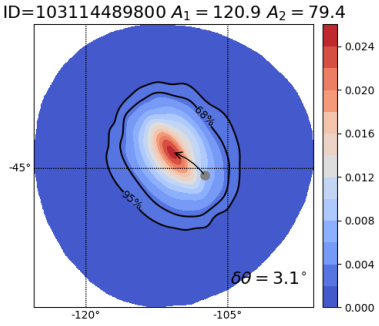

ID=70511454800 $A_1 = 221.0$ $A_2 = 129.9$
$\delta\theta = 7.1^\circ$
68%
95%

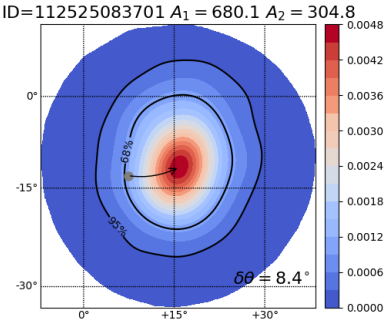

ID=81926339600 $A_1 = 431.8$ $A_2 = 221.5$
$\delta\theta = 6.0^\circ$
68%
95%

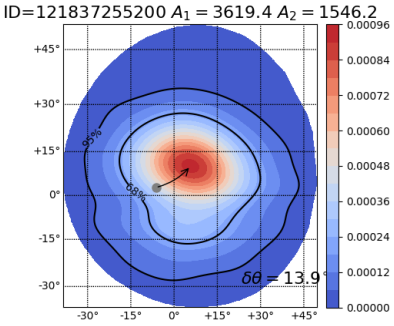

ID=140145941600 $A_1 = 400.3$ $A_2 = 232.1$
$\delta\theta = 3.0^\circ$
68%
95%


ID=113641418100 $A_1 = 338.1$ $A_2 = 153.1$
$\delta\theta = 2.5^\circ$
68%
95%

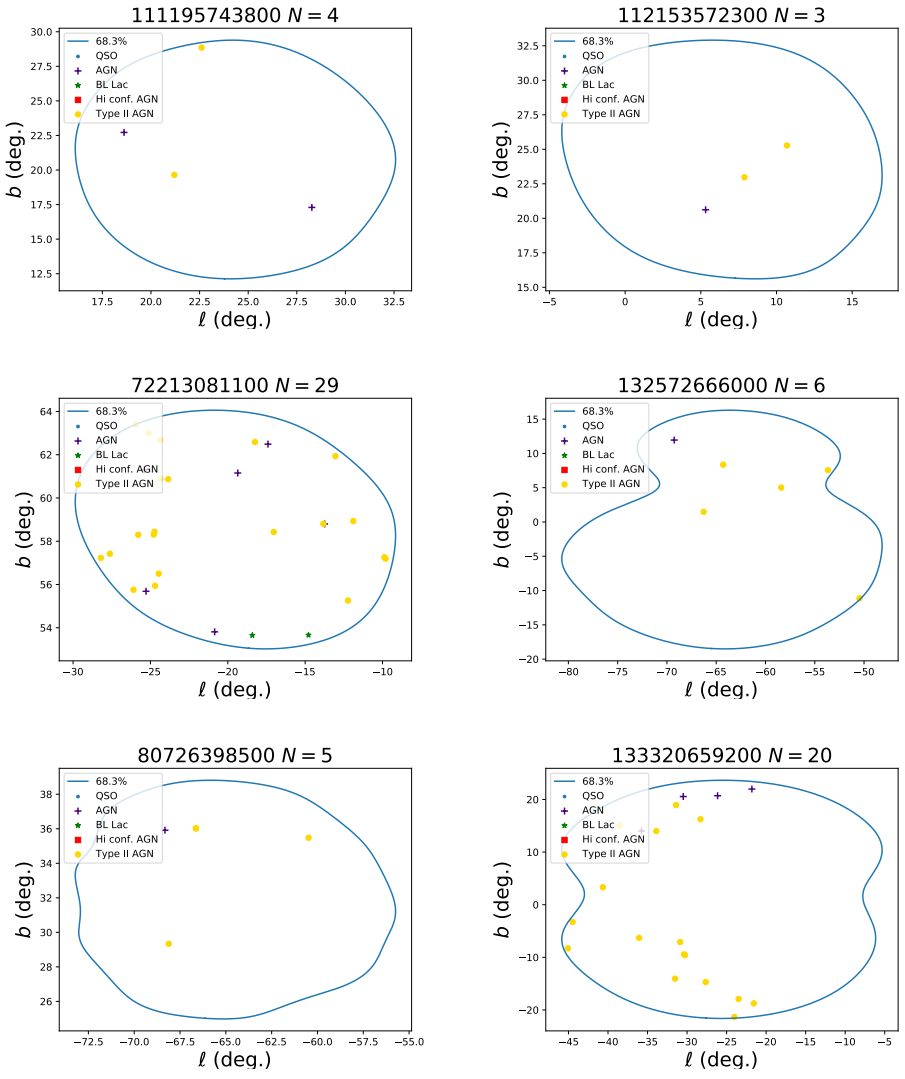

ID=72955799900 $A_1 = 420.4$ $A_2 = 221.4$
$\delta\theta = 5.6^\circ$
68%
95%

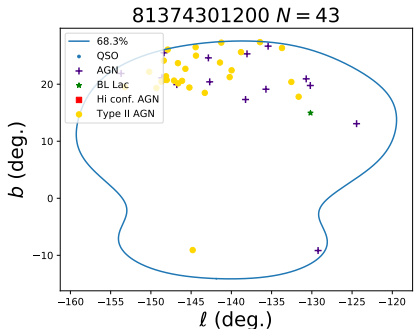

ID=103114489800 $A_1 = 120.9$ $A_2 = 79.4$
$\delta\theta = 3.1^\circ$
68%
95%

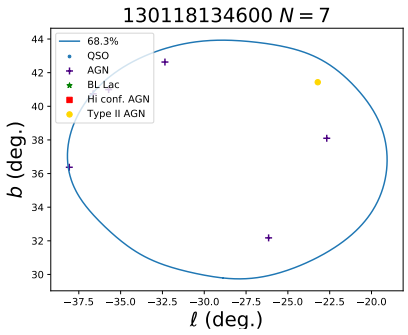

ID=112525083701 $A_1 = 680.1$ $A_2 = 304.8$
$\delta\theta = 8.4^\circ$
68%
95%

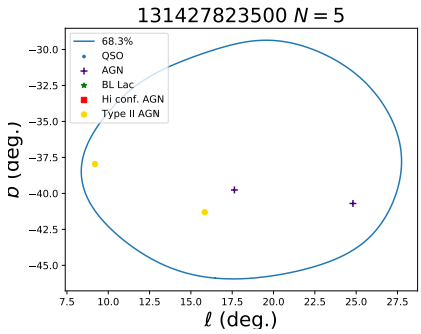

ID=121837255200 $A_1 = 3619.4$ $A_2 = 1546.2$
+45°
+30°
+15°
0°
-15°
-30°
95%
68%
$\delta\theta = 13.9°$
-30°
-15°
0°
+15°
+30°
+45°
0.00096
0.00084
0.00072
0.00060
0.00048
0.00036
0.00024
0.00012
0.00000

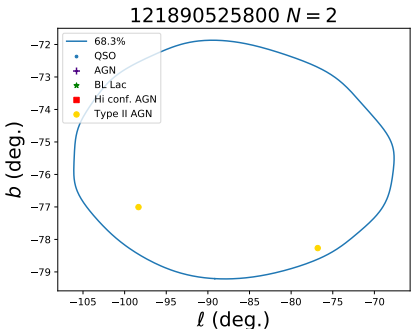
Chapter C: Appendix C

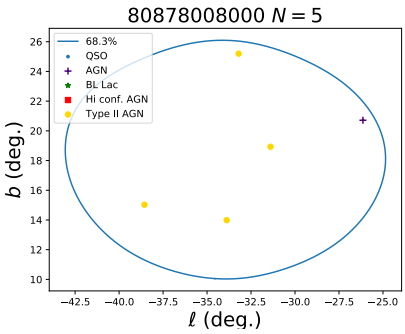


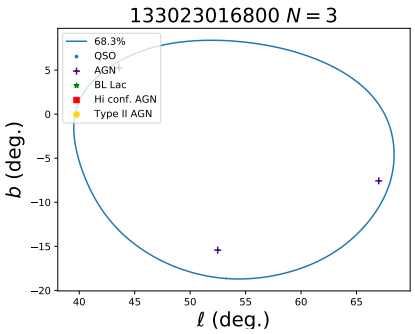

81374301200 $N = 43$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

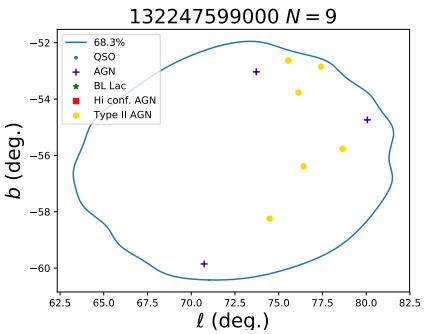

130118134600 $N = 7$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

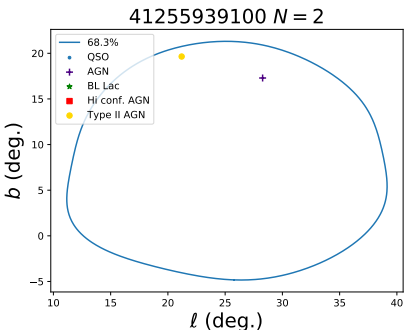

131427823500 $N = 5$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

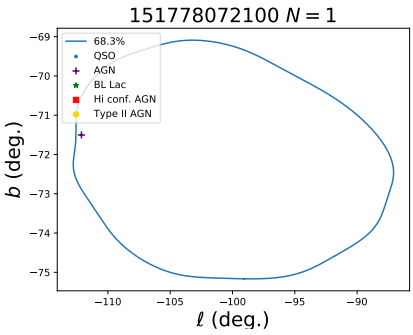

121890525800 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

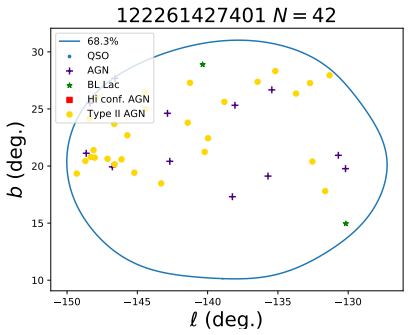

80878008000 $N = 5$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

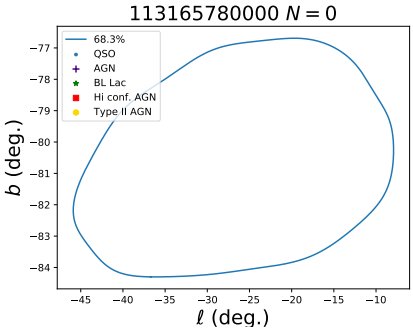

133023016800 $N = 3$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

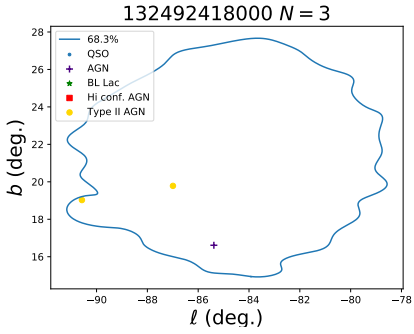

132247599000 $N = 9$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

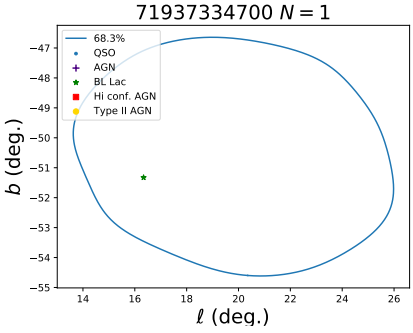

41255939100 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

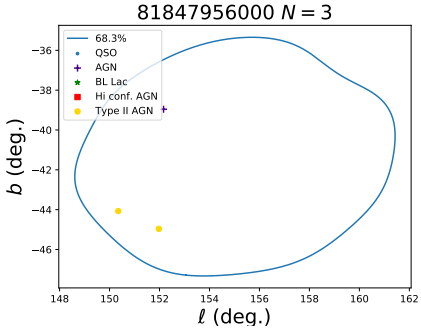

151778072100 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

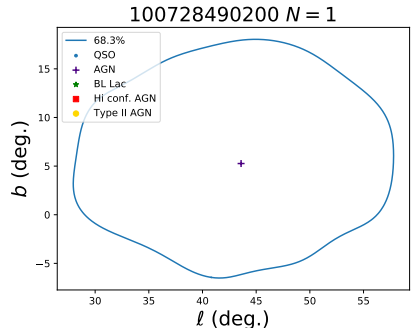

122261427401 $N = 42$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

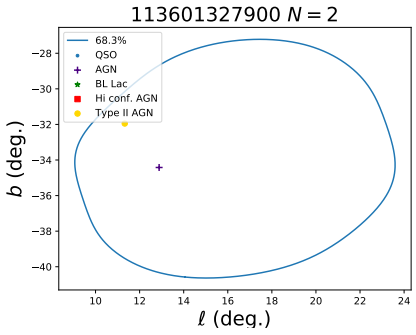

113165780000 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

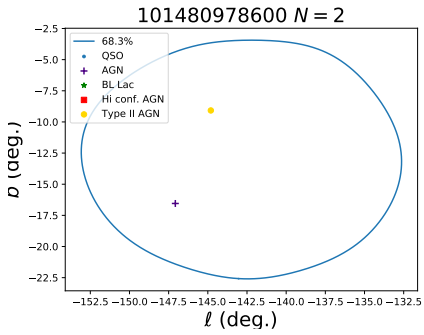

132492418000 $N = 3$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

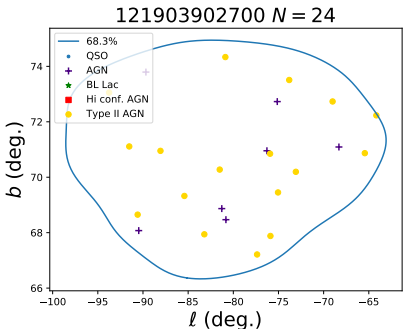

71937334700 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

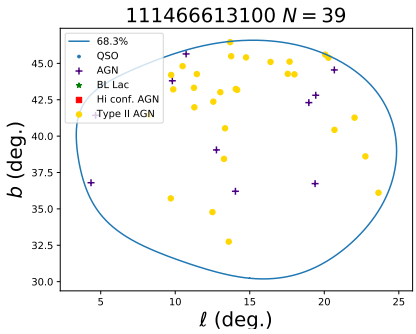

81847956000 $N = 3$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

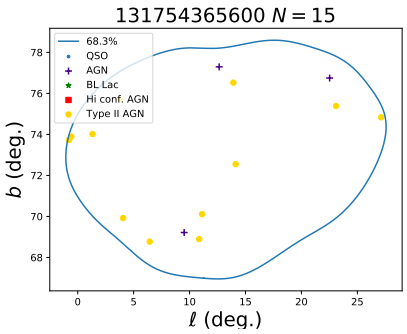

100728490200 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

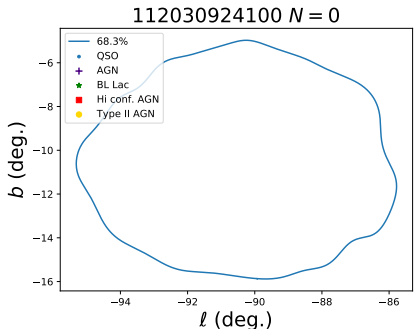

113601327900 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

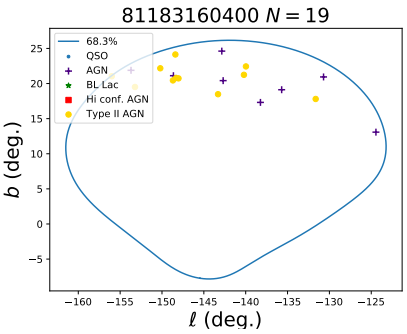

101480978600 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

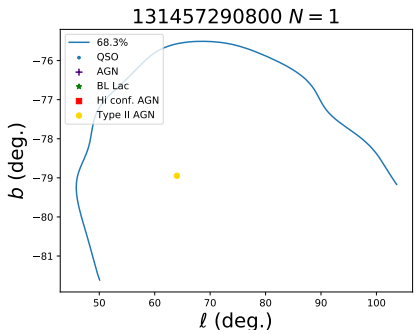

121903902700 $N = 24$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

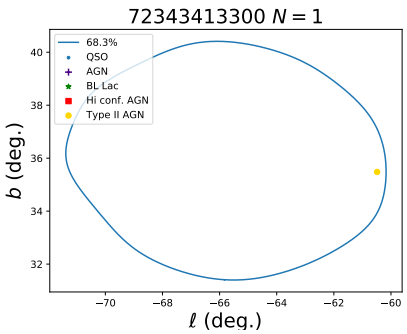

111466613100 $N = 39$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

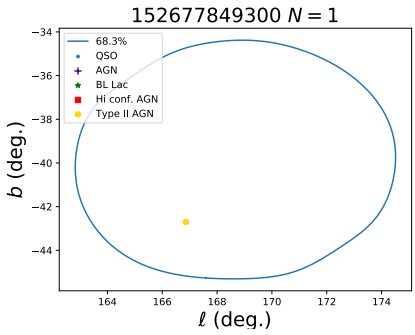

131754365600 $N = 15$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

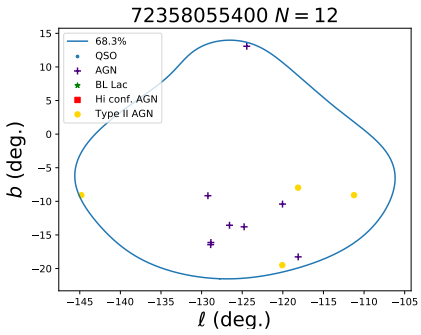

112030924100 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

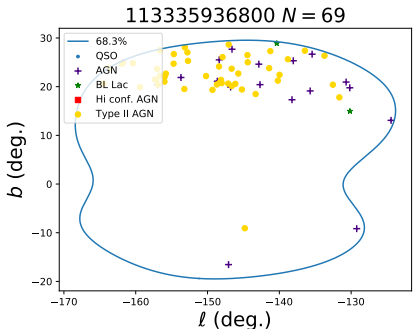

81183160400 $N = 19$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

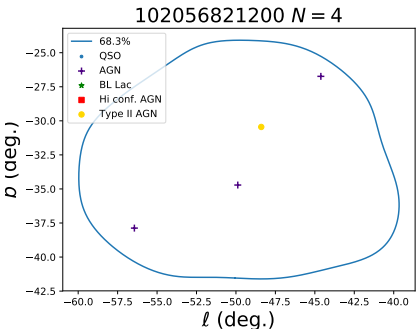

131457290800 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

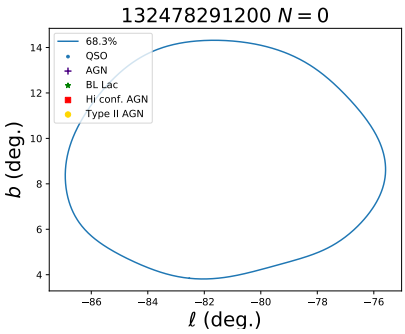

72343413300 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

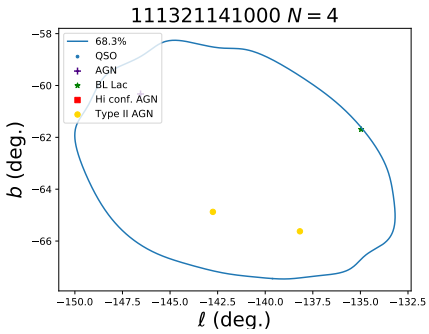

152677849300 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

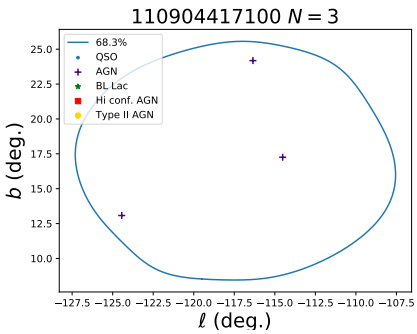

72358055400 $N = 12$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

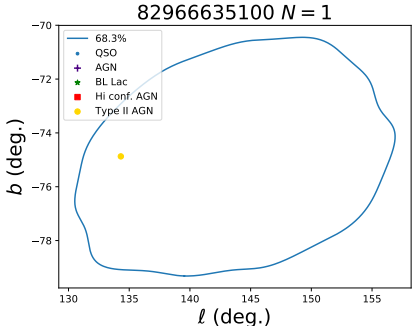

113335936800 $N = 69$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

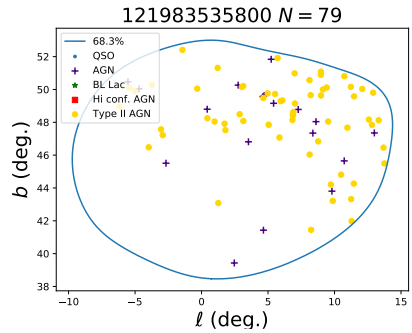

102056821200 $N = 4$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

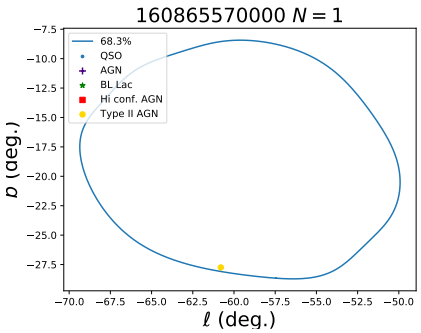

132478291200 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

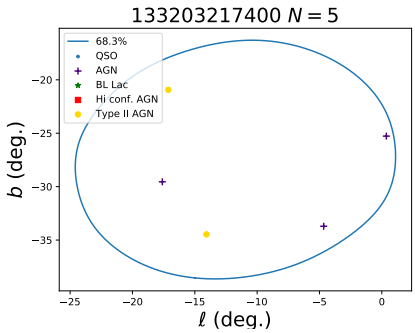

111321141000 $N = 4$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

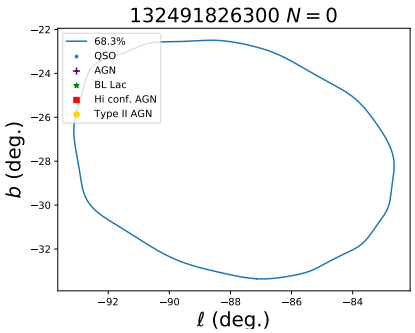

110904417100 $N = 3$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

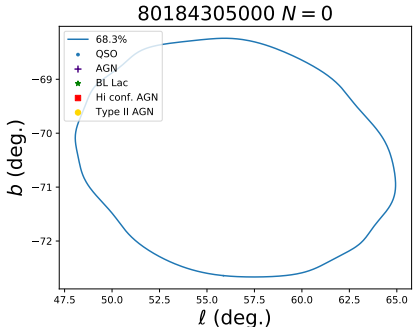

82966635100 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

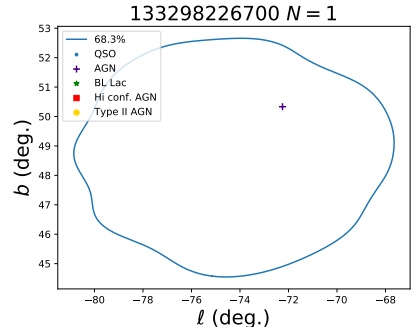

121983535800 $N = 79$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

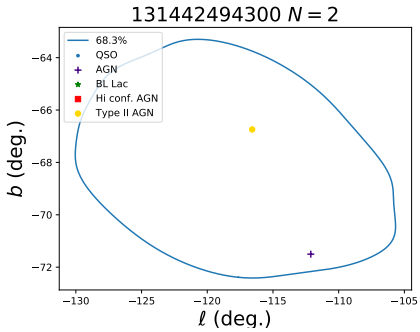

160865570000 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

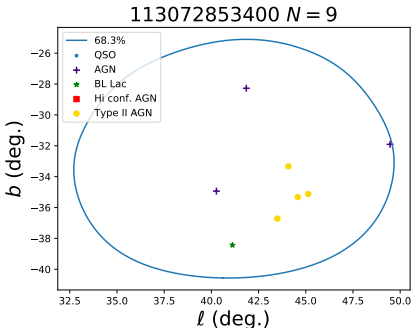

133203217400 $N = 5$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

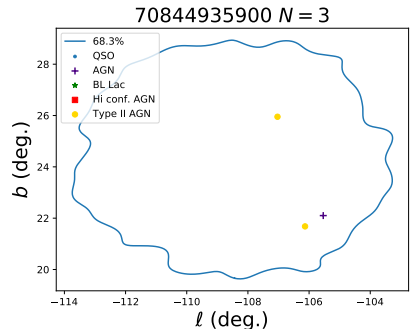

132491826300 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

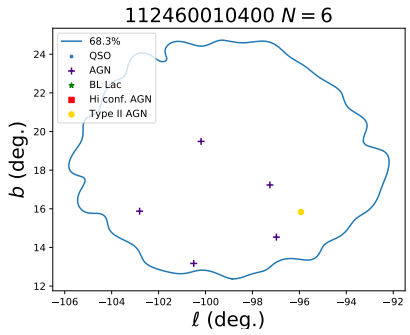

80184305000 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

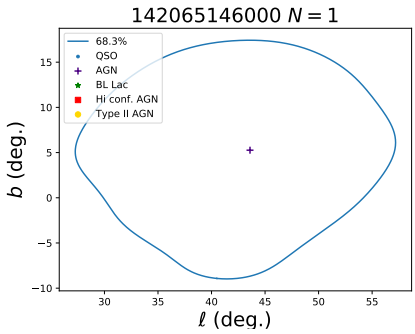

133298226700 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

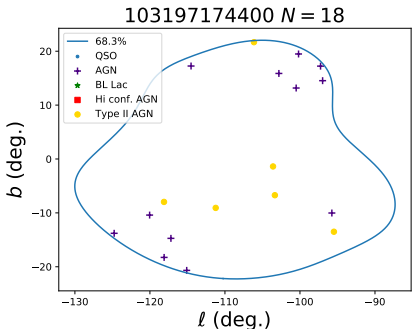

131442494300 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

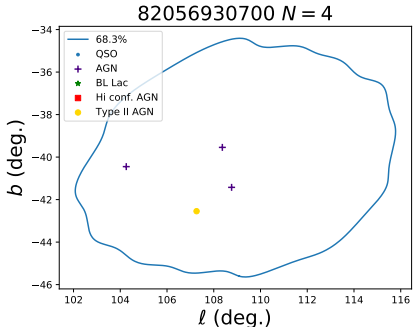

113072853400 $N = 9$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

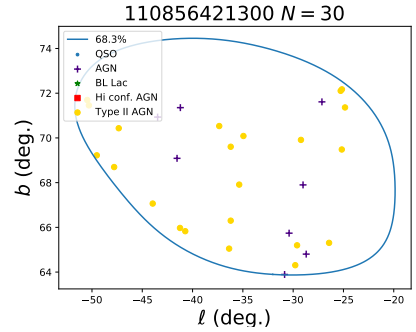

70844935900 $N = 3$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

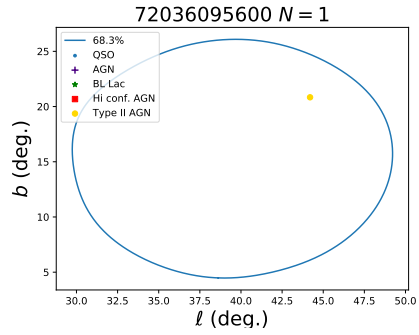

112460010400 $N = 6$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

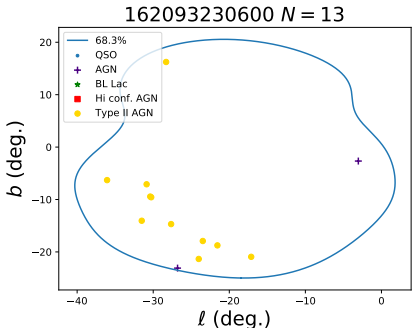

142065146000 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

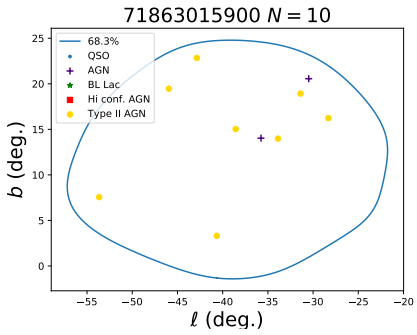

103197174400 $N = 18$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

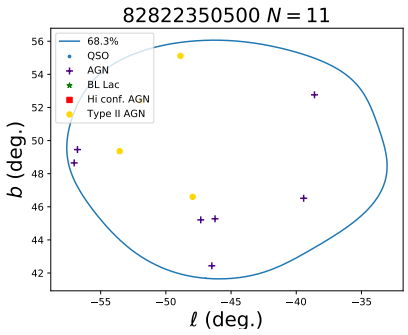

82056930700 $N = 4$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

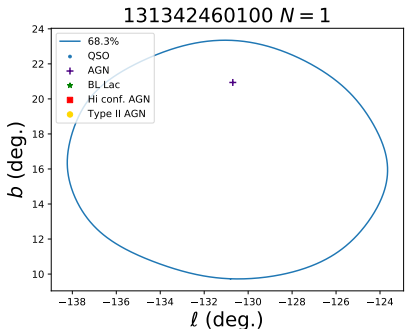

110856421300 $N = 30$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

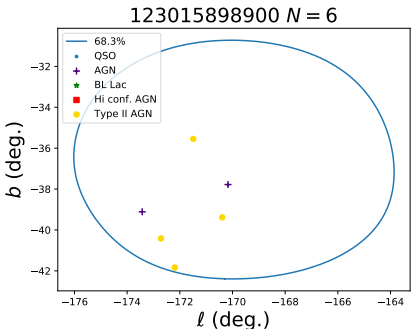

72036095600 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

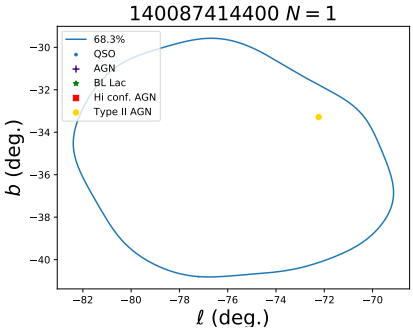

162093230600 $N = 13$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

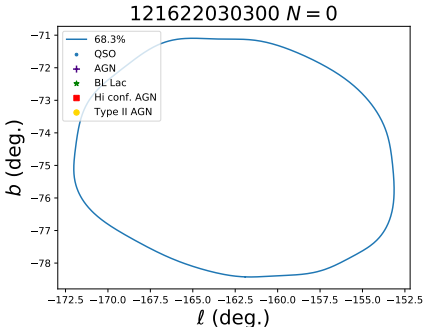

71863015900 $N = 10$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

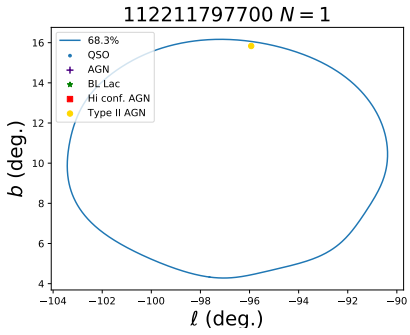

82822350500 $N = 11$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

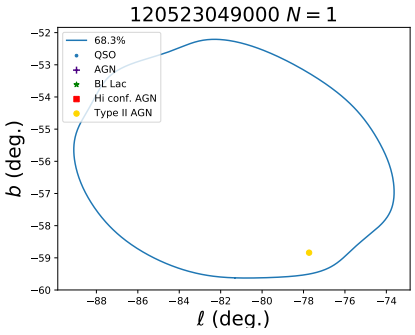

131342460100 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

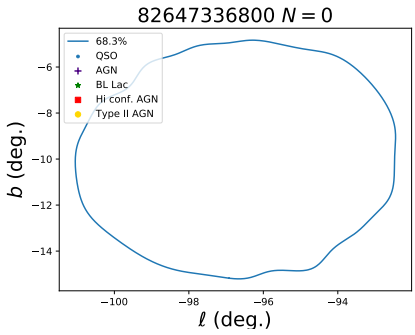

123015898900 $N = 6$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

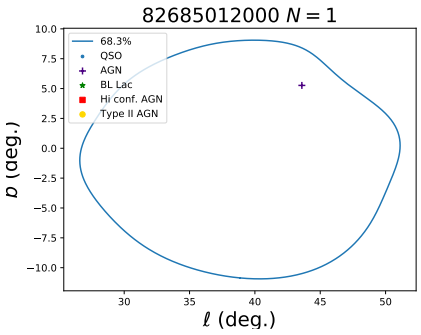

140087414400 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

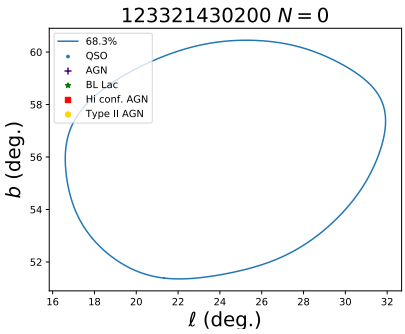

121622030300 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

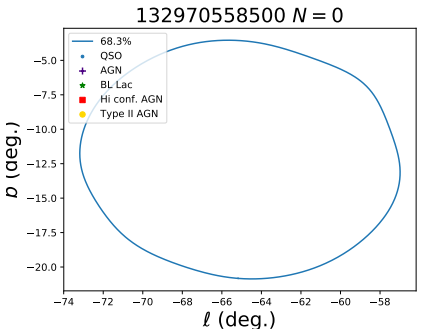

112211797700 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

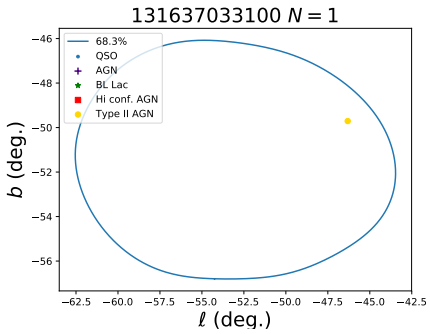

120523049000 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

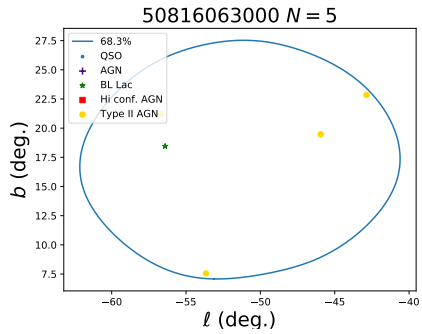

82647336800 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

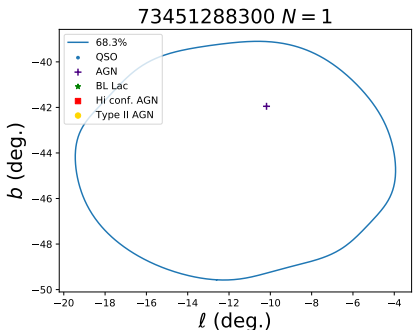

82685012000 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

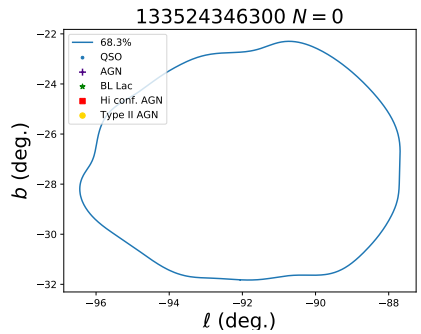

123321430200 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

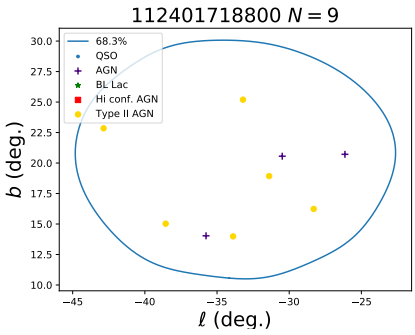

132970558500 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

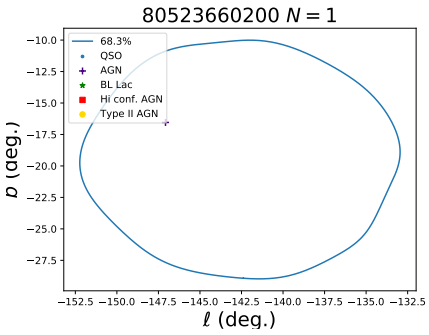

131637033100 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

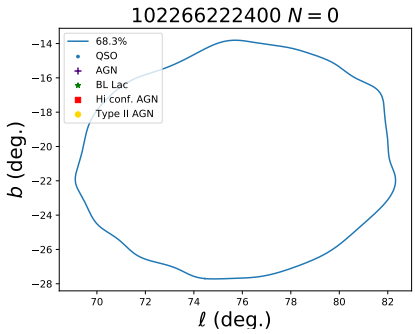

50816063000 $N = 5$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

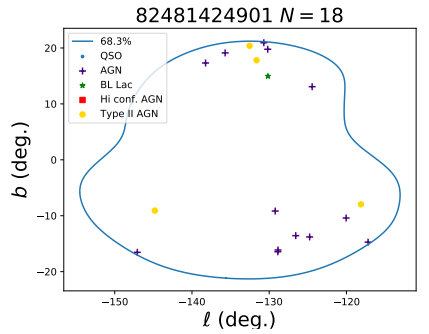

73451288300 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

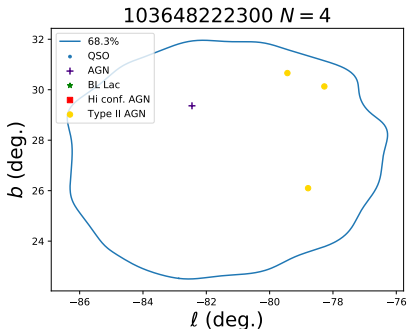

133524346300 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

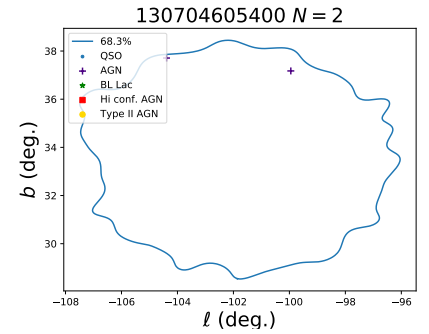

112401718800 $N = 9$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

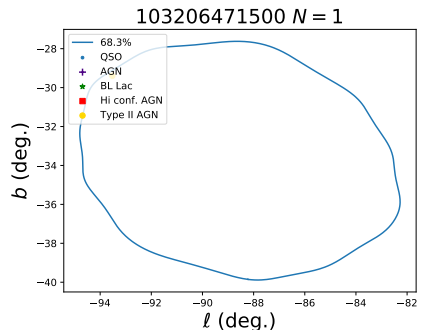

80523660200 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

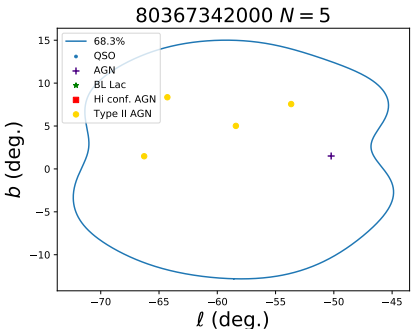

102266222400 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

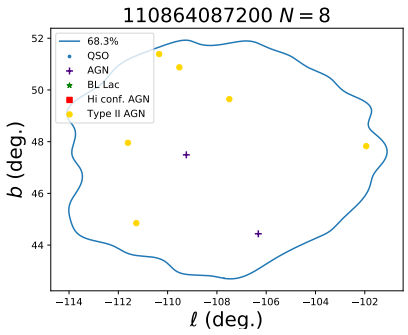

82481424901 $N = 18$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

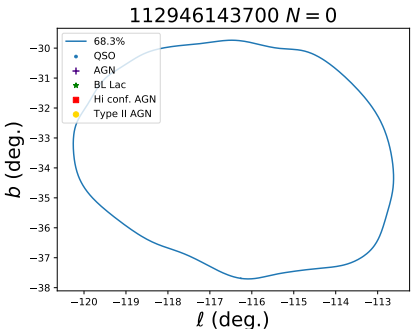

103648222300 $N = 4$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

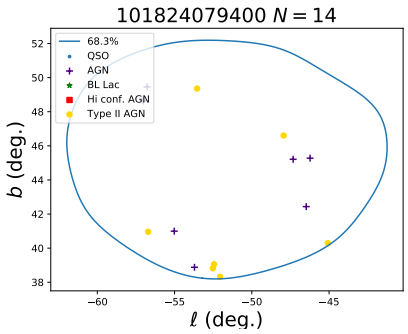

130704605400 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

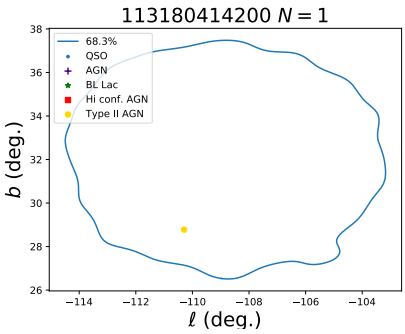

103206471500 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

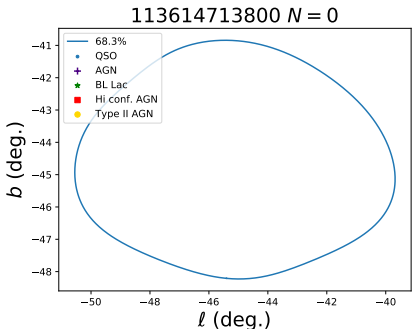

80367342000 $N = 5$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

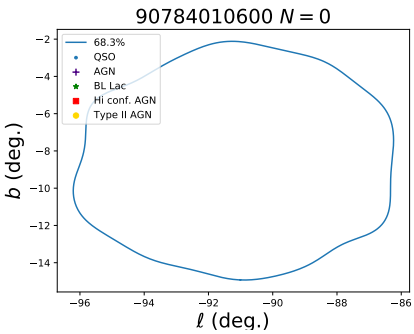

110864087200 $N = 8$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

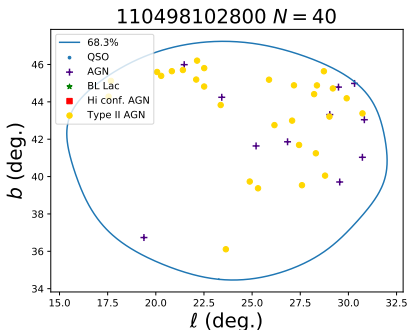

112946143700 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

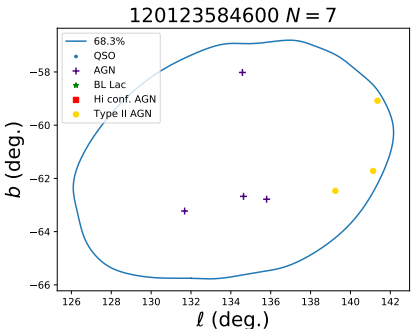

101824079400 $N = 14$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

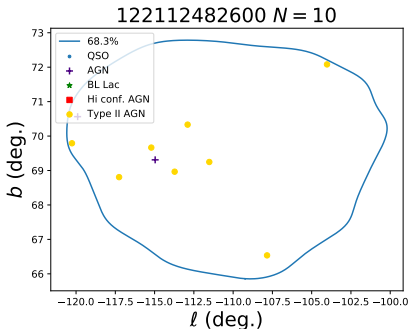

113180414200 $N = 1$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

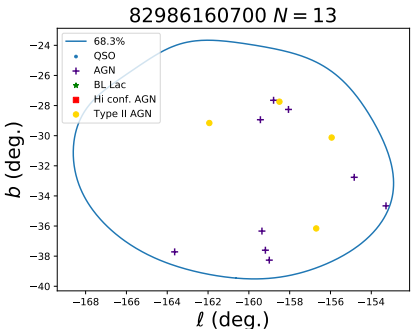

113614713800 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

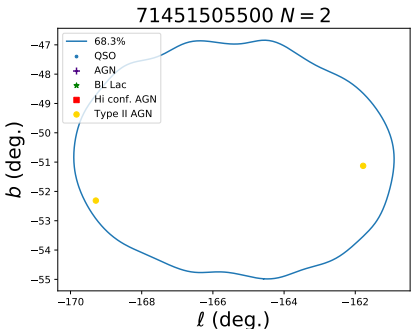

90784010600 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

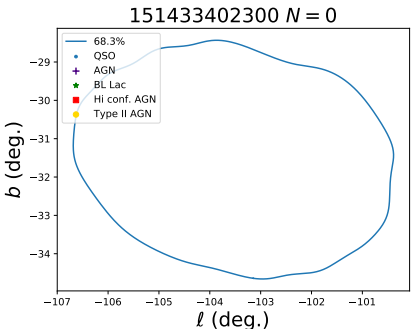

110498102800 $N = 40$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

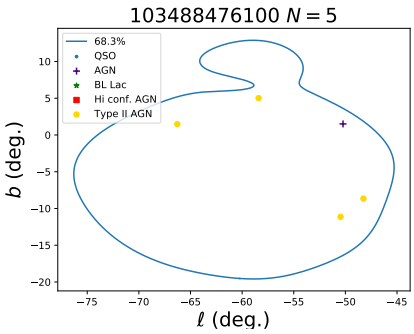

120123584600 $N = 7$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

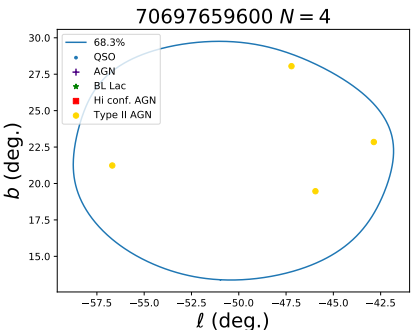

122112482600 $N = 10$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

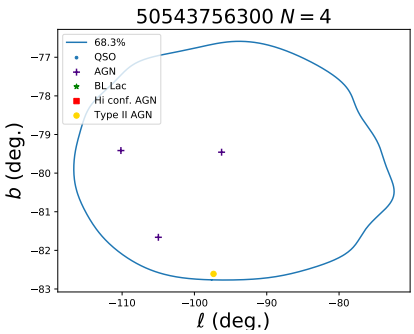

82986160700 $N = 13$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

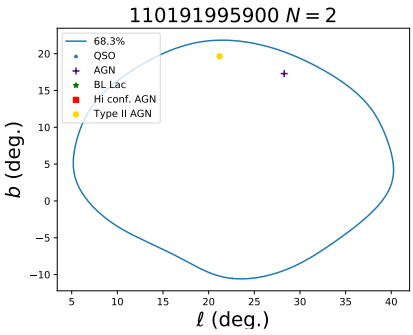

71451505500 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

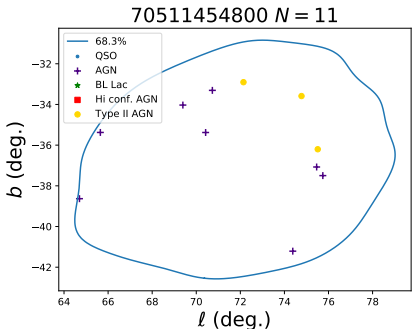

151433402300 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

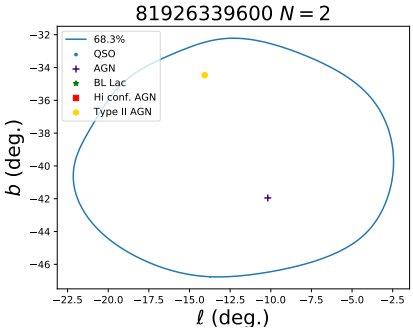

103488476100 $N = 5$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

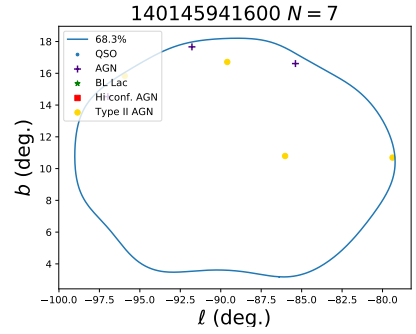

70697659600 $N = 4$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

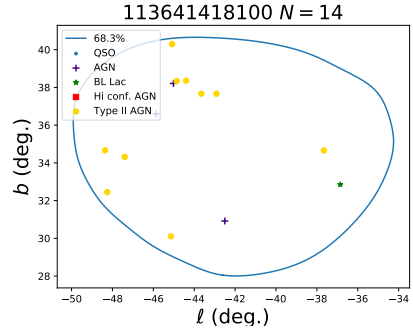

50543756300 $N = 4$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

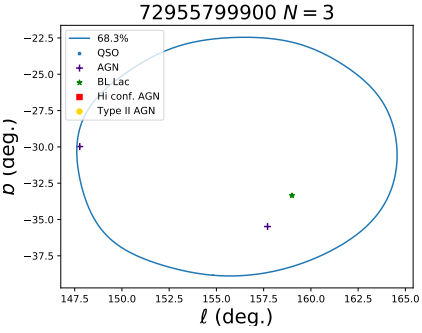

110191995900 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

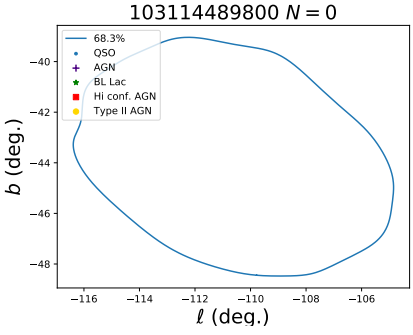

70511454800 $N = 11$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

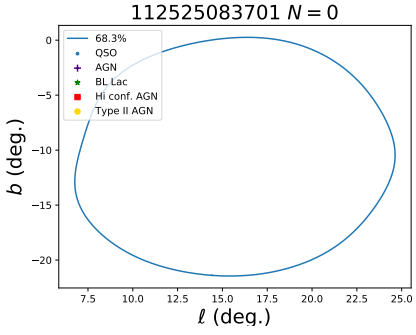

81926339600 $N = 2$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

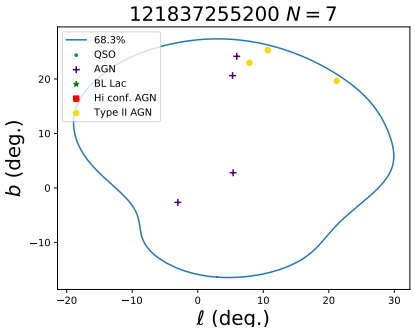

140145941600 $N = 7$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)


113641418100 $N = 14$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)

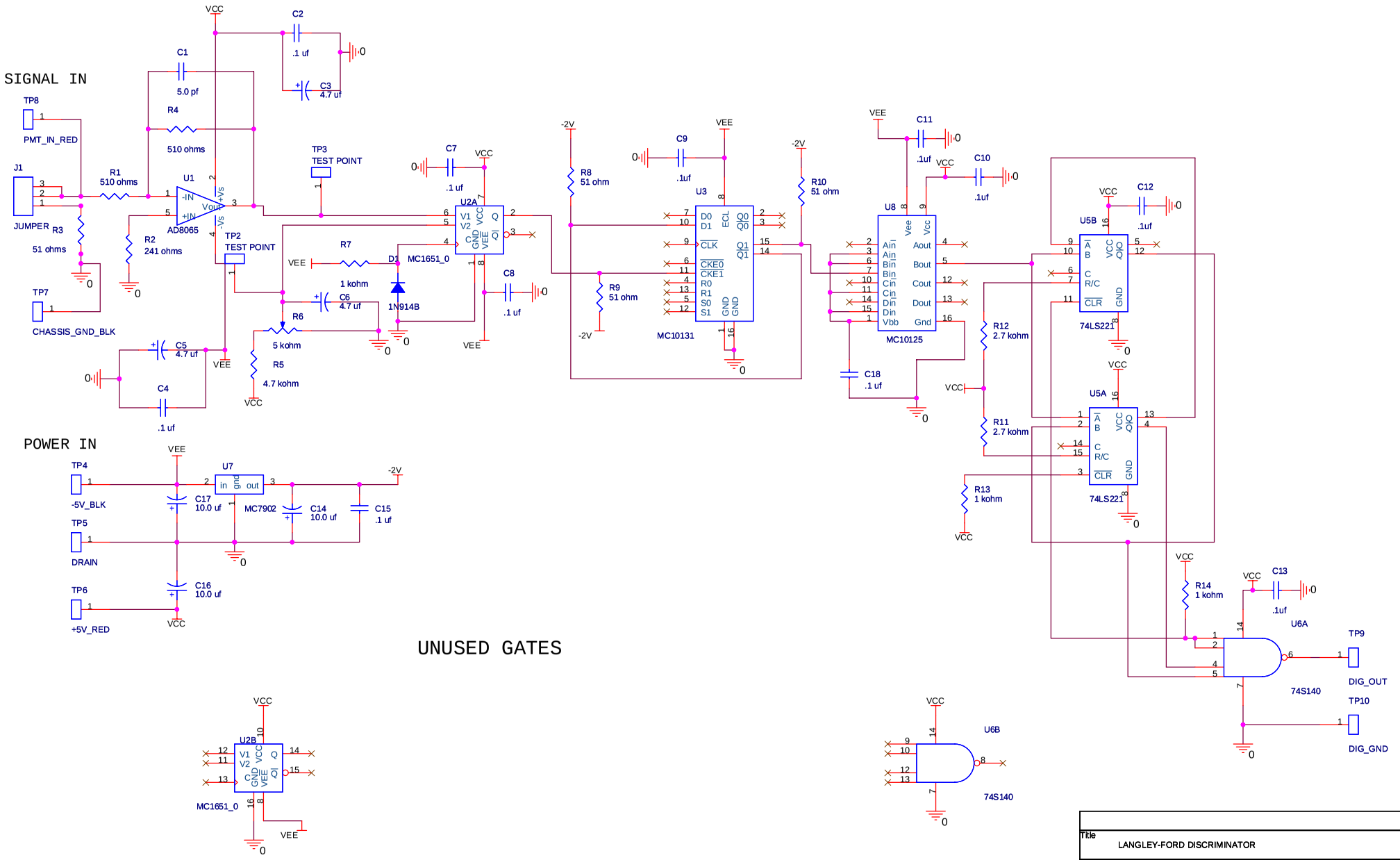

72955799900 $N = 3$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)


103114489800 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)


112525083701 $N = 0$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)


121837255200 $N = 7$
68.3%
QSO
AGN
BL Lac
Hi conf. AGN
Type II AGN
b (deg.)
ℓ (deg.)
20
10
0
−10
−20
−10
0
10
20
30

Chapter D: Appendix D


SIGNAL IN
POWER IN
UNUSED GATES
VCC
VEE
C1
5.0 pf
C2
.1 uf
C3
4.7 uf
R4
510 ohms
TP8
PMT_IN_RED
J1
JUMPER
R1
510 ohms
R3
51 ohms
R2
241 ohms
U1
AD8065
TP3
TEST POINT
TP2
TEST POINT
TP7
CHASSIS_GND_BLK
C5
4.7 uf
C4
.1 uf
R7
1 kohm
C6
4.7 uf
R6
5 kohm
R5
4.7 kohm
D1
1N914B
C7
.1 uf
U2A
MC1651_0
C8
.1 uf
R8
51 ohm
R9
51 ohm
C9
.1uf
U3
MC10131
R10
51 ohm
C11
.1uf
C10
.1uf
U8
MC10125
C18
.1 uf
R12
2.7 kohm
R11
2.7 kohm
R13
1 kohm
C12
.1uf
U5B
74LS221
U5A
74LS221
R14
1 kohm
C13
.1uf
U6A
74S140
TP9
DIG_OUT
TP10
DIG_GND
TP4
-5V_BLK
TP5
DRAIN
TP6
+5V_RED
U7
MC7902
C17
10.0 uf
C14
10.0 uf
C15
.1 uf
C16
10.0 uf
-2V
U2B
MC1651_0
U6B
74S140
Title
LANGLEY-FORD DISCRIMINATOR
Size
B
Document Number
<Doc>
Rev
1
Date:
Tuesday, September 05, 2006
Sheet
1
of
1

Chapter E: Appendix E

E.1 Local TA parser

```
import math
import time
import subprocess as sp

def utcFromString(x):
  split_x = x.split(',') # Break up position/date message
  split_x = split_x[0:6] # Only need d,m,y,h,min,s. Ignore northing, easting, alt
  int_x = map(int,split_x)
  right_order = [2,1,0,3,4,5]
  int_x = [int_x[i] for i in right_order]
  return int_x

secsInWeek = 604800
secsInDay = 86400
gpsEpoch = (1980, 1, 6, 0, 0, 0)  # (year, month, day, hh, mm, ss)

def gpsFromUTC(year, month, day, hour, minute, sec, leapSecs=18):
  """converts UTC to GPS second
```

```
19
20      Original function can be found at: http://software.ligo.org/docs/glue/frames.html
21
22      GPS time is basically measured in (atomic) seconds since
23      January 6, 1980, 00:00:00.0  (the GPS Epoch)
24
25      The GPS week starts on Saturday midnight (Sunday morning), and runs
26      for 604800 seconds.
27
28      Currently, GPS time is 17 seconds ahead of UTC
29      While GPS SVs transmit this difference and the date when another leap
30      second takes effect, the use of leap seconds cannot be predicted.  This
31      routine is precise until the next leap second is introduced and has to be
32      updated after that.
33
34      SOW = Seconds of Week
35      SOD = Seconds of Day
36
37      Note:  Python represents time in integer seconds, fractions are lost!!!
38      """
39      secFract = sec % 1
40      epochTuple = gpsEpoch + (-1, -1, 0)
41      t0 = time.mktime(epochTuple)
42      t = time.mktime((year, month, day, hour, minute, sec, -1, -1, 0))
43      # Note: time.mktime strictly works in localtime and to yield UTC, it should be
44      #       corrected with time.timezone
45      #       However, since we use the difference, this correction is unnecessary.
```

```
46    # Warning:  trouble if daylight savings flag is set to -1 or 1 !!!
47    t = t + leapSecs
48    tdiff = t - t0
49    gpsSOW = (tdiff % secsInWeek)  + secFract
50    gpsWeek = int(math.floor(tdiff/secsInWeek))
51    gpsDay = int(math.floor(gpsSOW/secsInDay))
52    gpsSOD = (gpsSOW % secsInDay)
53    gps_tuple = (gpsWeek, gpsSOW, gpsDay, gpsSOD)
54    return int(gps_tuple[0] * secsInWeek + gps_tuple[1])
55
56  N = 749995494 # Precision oscillator frequency
57
58  def computeMicro(n1,n2):
59    if n2>n1:
60      micsec = float(n2-n1) / N
61      micsec_str = "%.6f" %micsec
62      return int(micsec_str.split('.')[1])
63    else:
64      micsec = float(2**32-n1+n2) / N
65      micsec_str = "%.6f" %micsec
66      return int(micsec_str.split('.')[1])
67
68  dev_id = '/dev/ttyUSB1'
69  p = sp.Popen(['stdbuf','-oL','cat',dev_id],stdout=sp.PIPE)
70  # Main loop
71  int_list = [0] * 150 # Micros buffer. Should only need ~70, but
72                       # allocate more just in case
```

```
73  ii = 0
74  gps_n = 0
75  gps_sec = 0
76  while True:
77    msg = p.stdout.readline()
78    if "GPS" in msg:
79      if ii > 0:
80        with open("TA_LOCAL.txt",'a') as F:
81          for jj in range(ii):
82          #print "%i.%06d" %(gps_sec,int_list[jj])
83            F.write("%i.%06d\n" %(gps_sec,int_list[jj]))
84        ii = 0
85      gps_n = int(msg.split(' ')[-1])
86      continue
87    elif ",2017," in msg:
88      utc_list = utcFromString(msg)
89      gps_sec = gpsFromUTC(*utc_list)
90      continue
91    elif "TEST" in msg:
92      microint = int(msg.split(' ')[-1])
93      microsec = computeMicro(gps_n,microint)
94      int_list[ii] = microsec
95      ii += 1
96      continue
97    else:
98      continue
```

Chapter F: Appendix F

F.1 AS unhex and sort

```
# unhex_and_sort_jul2017_v1.py
# Copyright (c) Case Western Reserve University 2017
# This software is distributed under Apache License 2.0
# Consult the file LICENSE.txt
# Author: Sean Quinn spq@case.edu
# Mar 23 2017

import binascii
import bz2
import struct
import os
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
import numpy as np
import subprocess as sp
import time
import sys
```

```
19 import scipy.signal as spysig
20
21 """Small Python3 script that parses the master T3.out list. The file contains
22 all the relevant T3 data: Event ID, GPS time, Trigger type, and FADC traces for
   ↪  the
23 3 PMTs. Ricardo Sato has written a program "x2" which recovers the trace data
24 only from a decompressed T3 message. This script performs the following function
25 1.) Isolates individual events in the master list
26 2.) Converts ASCII hex data to raw binary
27 2.) Decompresses (bz2 format) T3 data message (still raw binary)
28 3.) Creates a folder for individual events with following name format
29     GPSTIME_MICROSECOND
30     Example: 1117640005_616863
31 4.) Places output of x2 program, which is an ASCII text file containing the
32     PMT traces (presumably mV levels for dynode/anode?)
33 """
34
35
36 def rle(inarray):
37         """ run length encoding. Partial credit to R rle function.
38             Multi datatype arrays catered for including non Numpy
39             returns: tuple (runlengths, startpositions, values) """
40         ia = np.array(inarray)                  # force numpy
41         n = len(ia)
42         if n == 0:
43             return (None, None, None)
44         else:
```

```
45              y = np.array(ia[1:] != ia[:-1])     # pairwise unequal (string safe)
46              i = np.append(np.where(y), n - 1)   # must include last element posi
47              z = np.diff(np.append(-1, i))       # run lengths
48              p = np.cumsum(np.append(0, z))[:-1] # positions
49              return(z, p, ia[i])
50
51  def save_calib(bytes):
52    f = open('calib_info.txt','w')
53    calsize = struct.unpack('>I',bytes[8212:8216])[0]
54    if calsize == 0:
55      calsize = struct.unpack('>I',bytes[8216:8220])[0]
56      if calsize == 84 or calsize == 104:
57        f.write('Version {}\n'.format(struct.unpack('>H',bytes[8220:8222])[0]))
58        f.write('TubeMask {}\n'.format(struct.unpack('>H',bytes[8222:8224])[0]))
59        f.write('StartSecond {}\n'.format(struct.unpack('>I',bytes[8224:8228])[0]))
60        f.write('EndSecond {}\n'.format(struct.unpack('>I',bytes[8228:8232])[0]))
61        f.write('NbT1 {}\n'.format(struct.unpack('>H',bytes[8232:8234])[0]))
62        f.write('NbT2 {}\n'.format(struct.unpack('>H',bytes[8234:8236])[0]))
63        evol = [0,0,0] # Last 8 minutes of calibration evolution
64        for i in range(3):
65          evol[i] = struct.unpack('>H',bytes[8236+2*i:8236+2*(i+1)])[0]
66        f.write('Evolution {0} {1} {2}\n'.format(*evol))
67        # Finish at 8242
68        dynode_base = [0,0,0]
69        for i in range(3):
70          dynode_base[i] = struct.unpack('>H',bytes[8242+2*i:8242+2*(i+1)])[0]*0.01
71        f.write('Dynode Base {0:.3f} {1:.3f} {2:.3f}\n'.format(*dynode_base))
```

```
72        # Finish at 8248
73        anode_base = [0,0,0]
74        for i in range(3):
75          anode_base[i] = struct.unpack('>H',bytes[8248+2*i:8248+2*(i+1)])[0]*0.01
76        f.write('Anode Base {0} {1} {2}\n'.format(*anode_base))
77        #Finish at 8254
78        dynode_base_var = [0,0,0]
79        for i in range(3):
80          dynode_base_var[i] =
    ↪ struct.unpack('>H',bytes[8254+2*i:8254+2*(i+1)])[0]*0.01
81        f.write('Dynode Base Var {0} {1} {2}\n'.format(*dynode_base_var))
82        # Finish at 8260
83        anode_base_var = [0,0,0]
84        for i in range(3):
85          anode_base_var[i] =
    ↪ struct.unpack('>H',bytes[8260+2*i:8260+2*(i+1)])[0]*0.01
86        f.write('Anode Base Var {0} {1} {2}\n'.format(*anode_base_var))
87        #Finish at 8266
88        block = ['VemPeak','Rate','NbTDA','DA','SigmaDA','VemCharge']
89        vem_peak = [0,0,0]
90        for i in range(3):
91          vem_peak[i] = struct.unpack('>H',bytes[8266+2*i:8266+2*(i+1)])[0]*0.1
92        f.write('VemPeak ' + '{0} {1} {2}\n'.format(*vem_peak))
93        # Finish at 8272
94        rate70Hz = [0,0,0]
95        for i in range(3):
96          rate70Hz[i] = struct.unpack('>H',bytes[8272+2*i:8272+2*(i+1)])[0]*0.01
```

```
97          f.write('70 Hz Rate ' + '{0} {1} {2}\n'.format(*rate70Hz))
98          # Finish at 8278
99          trigger_DA = [0,0,0]
100         for i in range(3):
101           trigger_DA[i] = struct.unpack('>H',bytes[8278+2*i:8278+2*(i+1)])[0]
102         f.write('Trigger D/A ' + '{0} {1} {2}\n'.format(*trigger_DA))
103         # Finish at 8284
104         DA = [0,0,0]
105         for i in range(3):
106           DA[i] = struct.unpack('>H',bytes[8284+2*i:8284+2*(i+1)])[0]*0.01
107         f.write('D/A ' + '{0} {1} {2}\n'.format(*DA))
108         # Finish at 8290
109         DA_var = [0,0,0]
110         for i in range(3):
111           DA_var[i] = struct.unpack('>H',bytes[8290+2*i:8290+2*(i+1)])[0]*0.01
112         f.write('D/A var ' + '{0} {1} {2}\n'.format(*DA_var))
113         # Finish at 8296
114         Area = [0,0,0]
115         for i in range(3):
116           Area[i] = struct.unpack('>H',bytes[8296+2*i:8296+2*(i+1)])[0]*0.1
117         f.write('VemCharge ' + '{0} {1} {2}\n'.format(*Area))
118         # Finish at 8302
119         totRate = struct.unpack('>H',bytes[8302:8304])[0]*0.01
120         f.write('TotRate ' + '{0}\n'.format(totRate))
121         #Finish at 8304
122         f.write('NbTOT {}\n'.format(struct.unpack('>H',bytes[8302:8304])[0]))
123         if calsize == 104:
```

```
124           block = ['DADt','SigmaDADt','DAChi2']
125           for i in range(3):
126             vals = [0,0,0]
127             for j in range(3):
128               vals[j]=struct.unpack('>H',bytes[8266+2*(3*i+j):8266+2*(3*i+j+1)])[0]
    ↪  /100.
129           f.write(block[i]+' '+'{0} {1} {2}\n'.format(*vals))
130       else:
131         f.write("BAD COMPRESS\n")
132         f.close()
133     else:
134       f.write("BAD COMPRESS\n")
135       f.close()
136     f.close()
137     return calsize
138
139 def save_mon(bytes,si):
140     #si is fondly known as start index
141     si += 8220
142     mon_size = struct.unpack('>I',bytes[si:si+4])[0]
143     si += 4
144     if mon_size == 6080:
145       f = open('mon_hist_offset.txt','w')
146       offsets = np.zeros(10,dtype=int)
147       for i in range(10):
148         val = struct.unpack('>H',bytes[si+2*i:si+2*(i+1)])[0]
149         offsets[i] = val
```

```
150         f.write('{}\n'.format(val))
151       si += 2 * 10
152       f.close()
153   # -------------BASELINE HISTOGRAMS-------------
154       pmt_base = np.zeros((20,6),dtype=int)
155       pmt_base[:,0] = np.arange(offsets[0],offsets[0]+20)
156       pmt_base[:,2] = np.arange(offsets[1],offsets[1]+20)
157       pmt_base[:,4] = np.arange(offsets[2],offsets[2]+20)
158       tmp_labels=['','PMT 1','','PMT 2','','PMT 3']
159       for j in [1,3,5]:
160         for i in range(20):
161           pmt_base[i,j] = struct.unpack('>H',bytes[si+2*i:si+2*(i+1)])[0]
162         si += 2*20
163         plt.step(pmt_base[:,j-1],pmt_base[:,j],where='pre',label=tmp_labels[j])
164       plt.xlabel('FADC channels')
165       plt.ylabel('Counts')
166       plt.title('Baseline histograms')
167       plt.legend()
168       plt.savefig('mon_hist_base.png')
169       plt.close('all')
170       np.savetxt('mon_hist_base.txt',pmt_base,fmt="%i")
171   # -------------PULSE HEIGHT HISTOGRAMS-------------
172       mon_peak = np.zeros((150,6),dtype=int)
173       mon_peak[:,0] = np.arange(offsets[3],offsets[3]+150)
174       mon_peak[:,2] = np.arange(offsets[4],offsets[4]+150)
175       mon_peak[:,4] = np.arange(offsets[5],offsets[5]+150)
176       tmp_labels=['','PMT 1','','PMT 2','','PMT 3']
```

```
177     for j in [1,3,5]:
178       for i in range(150):
179         mon_peak[i,j] = struct.unpack('>H',bytes[si+2*i:si+2*(i+1)])[0]
180       si += 2*150
181       lab = tmp_labels[j]
182       plt.step(mon_peak[:,j-1],mon_peak[:,j],where='pre',label=lab)
183     plt.xlabel('FADC channels')
184     plt.ylabel('Counts')
185     plt.title('Pulse height histograms')
186     plt.legend()
187     plt.savefig('mon_hist_pulse_height.png')
188     plt.close('all')
189     np.savetxt('mon_hist_pulse_height.txt',mon_peak,fmt="%i")
190 # -------------CHARGE HISTOGRAMS-------------
191     mon_charge = np.zeros((600,8),dtype=int)
192     mon_charge[:,0] = np.arange(offsets[6],offsets[6]+600)
193     mon_charge[:,2] = np.arange(offsets[7],offsets[7]+600)
194     mon_charge[:,4] = np.arange(offsets[8],offsets[8]+600)
195     mon_charge[:,6] = np.arange(offsets[9],offsets[9]+600)
196     tmp_labels=['','PMT 1','','PMT 2','','PMT 3','','PMT SUM']
197     for j in [1,3,5,7]:
198       for i in range(600):
199         mon_charge[i,j] = struct.unpack('>H',bytes[si+2*i:si+2*(i+1)])[0]
200       si += 2*600
201       lab = tmp_labels[j]
202       if j != 7:
203         plt.step(mon_charge[:,j-1],mon_charge[:,j],where='pre',label=lab)
```

```
204     plt.xlabel('FADC channels')
205     plt.ylabel('Counts')
206     plt.title('Charge histograms')
207     plt.legend()
208     plt.savefig('mon_hist_charge.png')
209     plt.close('all')
210     np.savetxt('mon_hist_charge.txt',mon_charge,fmt="%i")
211     plt.step(mon_charge[:,6],mon_charge[:,7],where='pre',label=tmp_labels[-1])
212     plt.xlabel('FADC channels')
213     plt.ylabel('Counts')
214     plt.title('Sum charge histogram')
215     plt.legend()
216     plt.savefig('mon_hist_charge_sum.png')
217     plt.close('all')
218   # -------------SHAPE HISTOGRAMS-------------
219     mon_shape = np.zeros((20,4),dtype=int)
220     mon_shape[:,0] = np.arange(0,500,25)
221     for j in range(1,4):
222       for i in range(20):
223         mon_shape[i,j] = struct.unpack('>I',bytes[si+4*i:si+4*(i+1)])[0]
224       si += 4*20
225       plt.step(mon_shape[:,0],mon_shape[:,j],where='pre',label='PMT %i' %j)
226     plt.xlabel('FADC bins [25 ns]')
227     plt.ylabel('Counts')
228     plt.title('PMT Shape')
229     plt.legend()
230     plt.savefig('mon_hist_pmt_shape.png')
```

```
231       plt.close('all')
232       np.savetxt('mon_hist_pmt_shape.txt',mon_shape,fmt="%i")
233     elif mon_size == 0:
234       return 0
235     else:
236       f = open('BAD_MON_COMPRESS','w')
237       f.close()
238       return mon_size
239     return si
240
241   def save_gps(bytes,si):
242     gpssize = struct.unpack('>I',bytes[si:si+4])[0]
243     si += 4
244     block = ['Current100','Next100','Current40','Next40','PreviousST',
245           'CurrentST','NextST']
246     f = open('gps_info.txt','w')
247     for i in range(7):
248       val = struct.unpack('>I',bytes[si+4*i:si+4*(i+1)])[0]
249       f.write(block[i]+' '+'{}\n'.format(val))
250     si += 4*7
251     if gpssize == 30:
252       val = struct.unpack('>H',bytes[si:si+2])[0]
253       f.write('Offset {}'.format(val))
254     f.close()
255
256   def find_baseline(x,y):
257     """Determine baseline from ADC trace.
```

```
258 Looks at a subsample (0-125) of pretrigger ADC counts.
259 The bin count with the largest weight is taken as the baseline.
260 Algorithm is based on GAP2016_044.
261
262 Parameters
263 ----------
264 x,y : array_like
265   Dynode and anode channel arrays. Converts to int.
266
267 Returns
268 -------
269 a,b : array_like
270   The baseline values for dynode and anode, respectively
271
272 c,d,e,f : scalars
273   Start bins for dynode and anode, stop bins for dynode and anode
274
275 """
276   dyn = x.astype(int)
277   ano = y.astype(int)
278   sigma = 2
279   #Find high gain baseline pieces
280   dyn_b = np.zeros(768)
281   #Determine most likely baseline
282   binval = np.arange(dyn.min(),dyn.min()+5)
283   counts = np.array([len(dyn[dyn==i]) for i in binval])
284   likely_base = binval[counts.argmax()]
```

```
285     for i in range(768):
286       if abs(dyn[i] - likely_base) < 2:
287         dyn_b[i] = 1
288     num_vals,start,val = rle(dyn_b)
289     base_i = np.where(val==1)[0]
290     num_vals,start=num_vals[base_i],start[base_i]
291     n_pieces = len(num_vals)
292     for i in range(n_pieces):
293       delta = num_vals[i]
294       if delta > 10:
295         base_mean = dyn[start[i]:start[i]+num_vals[i]].mean()
296         dyn_b[start[i]:start[i]+num_vals[i]] = base_mean
297       else:
298         dyn_b[start[i]:start[i]+num_vals[i]] = 0
299     #Interpolate between pieces
300     zeros = np.where(dyn_b == 0.)[0]
301     logical = np.zeros(768,dtype=bool)
302     logical[zeros] = True
303     tz = lambda z: z.nonzero()[0]
304     #Interp might fail in some situations
305     try:
306       dyn_b[logical] = np.interp(tz(logical),tz(~logical),dyn_b[~logical])
307     except:
308       if len(zeros) > 0:
309         dyn_b[logical] = dyn_b[760]
310     #Signal start search
311     dyn2 = dyn-dyn_b
```

```
312     dyn_start = 150 #Default in case problems
313     for i in range(100,768-1):
314       w0 = dyn2[i]
315       w1 = dyn2[i+1]
316       if w0 > 10 and w1 > 10:
317         dyn_start = i - 2
318         break
319     #Signal stop search
320     dyn_finish = 350 #Default in case of problems
321     #Don't care about spurious muons near end either
322     for i in range(767,dyn_start,-1):
323       w0 = dyn2[i]
324       if w0 > 4 and i < 400:
325         dyn_finish = i + 10
326         break
327     ano_b = np.zeros(768)
328     #Determine most likely baseline
329     binval = np.arange(ano.min(),ano.min()+5)
330     counts = np.array([len(ano[ano==i]) for i in binval])
331     likely_base = binval[counts.argmax()]
332     for i in range(768):
333       if abs(ano[i] - likely_base) < 2:
334         ano_b[i] = 1
335     num_vals,start,val = rle(ano_b)
336     base_i = np.where(val==1)[0]
337     num_vals,start=num_vals[base_i],start[base_i]
338     n_pieces = len(num_vals)
```

```
339   for i in range(n_pieces):
340     delta = num_vals[i]
341     if delta > 10:
342       base_mean = ano[start[i]:start[i]+num_vals[i]].mean()
343       ano_b[start[i]:start[i]+num_vals[i]] = base_mean
344     else:
345       ano_b[start[i]:start[i]+num_vals[i]] = 0
346   #Interpolate between pieces
347   zeros = np.where(ano_b == 0.)[0]
348   logical = np.zeros(768,dtype=bool)
349   logical[zeros] = True
350   tz = lambda z: z.nonzero()[0]
351   #Interp might fail in some situations
352   try:
353     ano_b[logical] = np.interp(tz(logical),tz(~logical),ano_b[~logical])
354   except:
355     if len(zeros) > 0:
356       ano_b[logical] = ano_b[760]
357   #Signal start search
358   ano2 = ano-ano_b
359   ano_start = 150 #Default in case problems
360   for i in range(100,768-1):
361     w0 = ano2[i]
362     w1 = ano2[i+1]
363     if w0 > 10 and w1 > 10:
364       ano_start = i - 2
365       break
```

```
366     #Signal stop search
367     ano_finish = 350 #Default in case of problems
368     #Don't care about spurious muons near end either
369     for i in range(767,ano_start,-1):
370       w0 = ano2[i]
371       w1 = ano2[i-1]
372       if w0 > 7 and w1 > 7 and i < 400:
373         ano_finish = i + 12
374         break
375     if len(np.where(dyn > 1020)[0]) < 2:
376       ano_start = dyn_start
377       ano_finish = dyn_finish
378     return dyn_b,ano_b,dyn_start,ano_start,dyn_finish,ano_finish
379
380   def find_vem(p):
381     """Determine peak of the charge histogram for PMT `p`.
382   Loads mon_hist_charge.txt file for input PMT. The histogram
383   is smoothed using a 45th order 3rd degree polynomial Savitzky-Golay filter.
384   The second peak of the smoothed signal is selected.
385
386   Parameters
387   ----------
388   p : int
389     PMT number. I.e. 0,1 or 2 for PMT#1,#2,#3
390
391   Returns
392   -------
```

```
393  y : int
394    Estimated location of charge histogram peak
395  """
396    xax,yax =
     ↪  np.loadtxt("mon_hist_charge.txt",usecols=(p*2,2*p+1),dtype=int,unpack=True)
397    xax = xax - xax[0]
398    Y = spysig.savgol_filter(yax,45,3)
399    ped_peak = Y[:60].argmax()
400    q_peak = Y[60:].argmax() + 60
401    return q_peak,ped_peak
402
403  def plot_vem():
404    fig = plt.figure(figsize=(16,9))
405    for p in range(3):
406      xax,yax =
     ↪  np.loadtxt("mon_hist_charge.txt",usecols=(p*2,2*p+1),dtype=int,unpack=True)
407      xax = xax - xax[0]
408      Y = spysig.savgol_filter(yax,45,3)
409      ped_peak = Y[:60].argmax()
410      q_peak = Y[60:].argmax() + 60
411      ax = fig.add_subplot(1,3,p+1)
412      plt.step(xax,yax)
413      plt.plot(xax,Y)
414      plt.xlabel('FADC channels')
415      plt.title('PMT %i charge histogram' %(p+1))
416      plt.ylim(ymin=0)
417      ymax = plt.ylim()[1]
```

```
418       plt.vlines(ped_peak,0,ymax)
419       plt.vlines(q_peak,0,ymax)
420       s='Pedestal peak = %i\nCharge peak=%i' %(ped_peak,q_peak)
421       plt.text(0.65,0.75,s,fontsize=10,transform=ax.transAxes)
422     plt.tight_layout()
423     plt.savefig('hist_charge_fit.png')
424     plt.close('all')
425
426   def make_plots(evt_num,gps):
427     # Get *estimated* offsets from calib data
428     a_base = [0,0,0]
429     d_base = [0,0,0]
430     v_peaks = [0,0,0]
431     da = [0,0,0]
432     v_charge = [0,0,0]
433     with open('calib_info.txt','r') as F:
434       for line in F:
435         if "Dynode Base" in line and "Var" not in line:
436           ss = line.split(' ')
437           for j in range(3):
438             d_base[j] = float(ss[j+2])
439         elif "Anode Base" in line and "Var" not in line:
440           ss = line.split(' ')
441           for j in range(3):
442             a_base[j] = float(ss[j+2])
443         elif "VemPeak" in line:
444           ss = line.split(' ')
```

```
445           for j in range(3):
446             v_peaks[j] = float(ss[j+1])
447         elif "D/A" in line and "var" not in line and "Trigger" not in line:
448           ss = line.split(' ')
449           for j in range(3):
450             da[j] = float(ss[j+1])
451         elif "VemCharge" in line:
452           ss = line.split(' ')
453           for j in range(3):
454             v_charge[j] = float(ss[j+1])
455         else:
456           continue
457     fadc_hist = np.loadtxt('FADC_trace',dtype=int)
458     xaxis = np.arange(0,768)
459     plt.figure(figsize=(19,8))
460     for i in range(3):
461       plt.subplot(1,3,i+1)
462       plt.plot(xaxis,fadc_hist[:,i+1],
463                drawstyle='steps-pre')
464       plt.title('PMT {}'.format(i+1))
465       plt.xlabel(r'Time [25 ns]')
466       plt.ylabel('ADC counts')
467     plt.tight_layout()
468     plt.savefig('dynode_adc.png')
469     plt.close()
470     plt.figure(figsize=(19,8))
471     for i in range(3):
```

```
472       plt.subplot(1,3,i+1)
473       plt.plot(xaxis,fadc_hist[:,i+4],
474             drawstyle='steps-pre')
475       plt.title('PMT {}'.format(i+1))
476       plt.xlabel(r'Time [25 ns]')
477       plt.ylabel('ADC counts')
478     plt.tight_layout()
479     plt.savefig('anode_adc.png')
480     plt.close()
481     # Make Signal plot
482     sga = [0.]*3
483     sgd = [0.]*3
484     f,axs = plt.subplots(nrows=2,ncols=3,sharex='col',sharey='row',figsize=(22,14))
485     axs[0][0].set_ylabel('ANODE Signal [VEM peak]')
486     axs[1][0].set_ylabel('DYNODE Signal [VEM peak]')
487     ano_sat = [0]*3
488     dyn_sat = [0]*3
489     signal_only_trace = []
490     for i in range(3):
491       y=np.empty(0) #Anode y-axis
492       y2 = np.empty(0) #Dynode y-axis
493       #Get ADC traces for anode
494       y = fadc_hist[:,i+4]
495       max_ano_adc = np.where(y>1020)[0]
496       #Get ADC traces for dynode
497       y2 = fadc_hist[:,i+1]
498       max_dyn_adc = np.where(y2>1020)[0]
```

```
499     #Determine baseline
500     dyn_b,ano_b,d_start,a_start,d_end,a_end = find_baseline(y2,y)
501     qvem_peak, pdl_peak = find_vem(i)
502     #Calculate signals
503     y_sig = y - ano_b
504     sga[i] = y_sig[a_start:a_end].sum() / (qvem_peak / 32)
505     y2_sig = y2 - dyn_b
506     sgd[i] = y2_sig[d_start:d_end].sum() / qvem_peak
507     #Put ADC traces into normalized units
508     y_peak_in_vem = np.max(y_sig / (qvem_peak / 32))
509     y2_peak_in_vem = np.max(y2_sig / qvem_peak)
510     y2_max_val = y2_sig.max()
511     xnew = xaxis[d_start-10:d_end+10]
512     plot_y = y_sig[d_start-10:d_end+10] * y_peak_in_vem / (qvem_peak / 32)
513     plot_y2 = y2_sig[d_start-10:d_end+10] * y2_peak_in_vem / qvem_peak
514     #Add isolated signal window to list
515     #Format: [ANODE1,DYNODE1,ANODE2,DYNODE2,ANODE3,DYNODE3]
516     signal_only_trace.append(y_sig[a_start:a_end]  / (qvem_peak / 32))
517     signal_only_trace.append(y2_sig[d_start:d_end] / qvem_peak)
518     #Plot anode
519     axs[0][i].step(xnew,plot_y)
520     axs[0][i].vlines(a_start,0,1.2*plot_y.max(),linestyle='dashed',color='green')
521     axs[0][i].vlines(a_end,0,1.2*plot_y.max(),linestyle='dashed',color='green')
522     axs[1][i].step(xnew,plot_y2)
523     axs[1][i].vlines(d_start,0,1.2*plot_y2.max(),linestyle='dashed',color='green')
524     axs[1][i].vlines(d_end,0,1.2*plot_y2.max(),linestyle='dashed',color='green')
525     axs[1][i].set_xlabel(r'Time [25 ns]')
```

```
526     boxstr = 'S=%.1f VEM\nD/A=%.1f\nVEM Charge=%.1f\nVEM Peak=%.1f'
    ↪  %(sga[i],da[i],v_charge[i],v_peaks[i])
527     boxstr2 = 'S=%.1f VEM\nD/A=%.1f\nVEM Charge=%.1f\nVEM Peak=%.1f'
    ↪  %(sgd[i],da[i],v_charge[i],v_peaks[i])
528     axs[0][i].text(0.7,0.75,boxstr,fontsize=10,transform=axs[0][i].transAxes)
529     axs[1][i].text(0.7,0.75,boxstr2,fontsize=10,transform=axs[1][i].transAxes)
530     axs[0][i].set_title('PMT %i' %(i+1))
531     axs[0][i].set_xlim(xmin=xnew.min(),xmax=xnew.max())
532     axs[1][i].set_xlim(xmin=xnew.min(),xmax=xnew.max())
533     if len(max_ano_adc) > 2:
534       axs[0][i].text(0.1,0.75,'SATURATED',color='red',fontsize=10,transform=axs[0
    ↪  ][i].transAxes)
535       ano_sat[i] = 1
536     if len(max_dyn_adc) > 2:
537       axs[1][i].text(0.1,0.75,'SATURATED',color='red',fontsize=10,transform=axs[1
    ↪  ][i].transAxes)
538       dyn_sat[i] = 1
539   plt.tight_layout()
540   plt.savefig('%i_signal.png' %evt_num)
541   plt.close('all')
542   #Put isolated signal waveforms into numpy array
543   #Format:
544   #A1_AOP D1_AOP A2_AOP D2_AOP A3_AOP D3_AOP
545   #A1_RISET D1_RISET A2_RISET D2_RISET A3_RISET D3_RISET
546   aop_rt_array = np.zeros((2,6))
547   for i in range(6):
548     #np_signal_only_trace[:,i] = signal_only_trace[i]
```

```
549       np.savetxt("signal_%i.txt" %i, signal_only_trace[i],fmt="%.2f")
550       #Calculate area over peak and rise time
551       max_time_bin = signal_only_trace[i].argmax()
552       max_time = max_time_bin * 25
553       aop_peak = signal_only_trace[i][max_time_bin] / max_time
554       curr_signal = signal_only_trace[i].sum()
555       aop = curr_signal / aop_peak
556       signal_cdf = signal_only_trace[i].cumsum()
557       #Find closest point that gives 10%
558       signal_10pct = np.abs(signal_cdf - 0.1 * curr_signal)
559       i10 = signal_10pct.argmin()
560       if signal_cdf[i10] > 0.1 * curr_signal:
561         xp = np.array([i10-1,i10])*25
562         fp = np.array([signal_cdf[i10-1],signal_cdf[i10]])
563         t10 = int(np.interp(curr_signal * 0.1,fp,xp))
564       else:
565         xp = np.array([i10,i10+1])*25
566         fp = np.array([signal_cdf[i10],signal_cdf[i10+1]])
567         t10 = int(np.interp(curr_signal * 0.1,fp,xp))
568       #Find closest point that gives 50%
569       signal_50pct = np.abs(signal_cdf - 0.5 * curr_signal)
570       i50 = signal_50pct.argmin()
571       if signal_cdf[i50] > 0.5 * curr_signal:
572         xp = np.array([i50-1,i50])*25
573         fp = np.array([signal_cdf[i50-1],signal_cdf[i50]])
574         t50 = int(np.interp(curr_signal * 0.5,fp,xp))
575       else:
```

```
576         xp = np.array([i50,i50+1])*25
577         fp = np.array([signal_cdf[i50],signal_cdf[i50+1]])
578         t50 = int(np.interp(curr_signal * 0.5,fp,xp))
579       rise_time = t50-t10
580       aop_rt_array[0,i] = aop
581       aop_rt_array[1,i] = rise_time
582     np.savetxt("aop_risetime.txt",aop_rt_array,fmt="%.3f")
583     #Make calibration plots
584     plot_vem()
585     return "%i %.3f %.3f %.3f %.3f %.3f %.3f %i %i %i %i %i %i"
      ↪  %(gps,sga[0],sga[1],sga[2],sgd[0],sgd[1],sgd[2],ano_sat[0],ano_sat[1],ano_sa⌋
      ↪  t[2],dyn_sat[0],dyn_sat[1],dyn_sat[2])
586
587 #Read through T3 file collecting events into a large list
588 #Since writing the original script I've found a more elegant approach
589 #thanks to inspectorG4dget at stackoverflow
590 #http://stackoverflow.com/questions/18865058/
591
592 fdate = sp.check_output(['date','--date','-1 day','+%Y%m%d'])
593 fdate = fdate.strip()
594
595 t3list = []
596 num_evts = 0
597 print("Reading T3 event file ...")
598
599 yr = int(fdate[:4])
600 mo = int(fdate[4:6])
```

```
601  dy = int(fdate[6:])
602
603  os.chdir('/home/augta/web_monitor/tmp')
604  fname = "%i_%02d_%02d" %(yr,mo,dy)
605  sp.call(['cp','/home/augta/data/south/t3/%s.T3.gz' %fname,'.'])
606
607  sp.call(['gunzip',"%s.T3.gz" %fname])
608
609  filename = "%s.T3" %fname
610
611  with open(filename,'r') as t3file:
612    copy = False
613    for line in t3file:
614      if line.strip() == "Event ...":
615        copy = True
616        data_str = ''
617        num_evts += 1
618      elif line.strip() == "----------":
619        copy = False
620        #Avoid clipped data streams. Compressed event should be more than 10kB
621        if len(data_str)>9000:
622          t3list.append(data_str)
623      elif copy:
624        data_str += line.strip().replace(' ','')
625
626  time.sleep(0.5)
627  #print("[ OK ]")
```

```
628 #print("Found {} events".format(num_evts))
629
630 evt_count = 1
631 signal_data = []
632 for t3 in t3list:
633   print("Event %i of %i" %(evt_count,len(t3list)))
634   evt_id = int(t3[:4],16)
635   error_code = int(t3[4:6],16) #Value of 1 indicates no error for T3
636   packed=binascii.unhexlify(bytes(t3[8:],'ascii'))
637   dec_t3 = bz2.decompress(packed)
638   # Now that we have uncompressed message let's get some information
639   # The PowerPC hardware uses big endian format
640   gps_YMDHMnS = struct.unpack('>I', dec_t3[:4])[0] #First 4 bytes are GPS sec
641   gps_TICK = struct.unpack('>I', dec_t3[4:8])[0] #Next 4 are GPS clock cycles
642   try:
643     os.mkdir("{0}_{1}".format(evt_id,gps_YMDHMnS))
644   except OSError as e:
645     #Catch folder already exists error
646     if e.errno==17:
647       print("This event has already been unpacked and saved, skipping ...")
648       continue
649   os.chdir('{0}_{1}'.format(evt_id,gps_YMDHMnS))
650   f=open('T3_{}.bin'.format(evt_id), 'bw')
651   f.write(dec_t3)
652   f.close()
653   sp.call(["../../x2", "T3_{}.bin".format(evt_id)])
654   monstart = save_calib(dec_t3)
```

```
655     gpsstart = save_mon(dec_t3,monstart)
656     save_gps(dec_t3,gpsstart)
657     signal_data.append(make_plots(evt_count,gps_YMDHMnS))
658     evt_count += 1
659     os.chdir('..')
660
661   sp.call(['rm',filename])
662   dirlist = os.listdir('.')
663   dirlist.sort()
664   sp.call(['mkdir','/var/www/html/monitor/data/global_south/%s' %fname])
665   sp.call(['mkdir','/var/www/html/monitor/data/local_south/%s' %fname])
666
667   sp.call(['cp',"/home/augta/data/coincidence/%s.CTAG.gz" %fname,
668     '/var/www/html/monitor/data/global_south/%s' %fname])
669   sp.call(['cp',"/home/augta/data/coincidence/%s.CTAL.gz" %fname,
670     '/var/www/html/monitor/data/local_south/%s' %fname])
671
672   sp.call(['cp',"/home/augta/data/coincidence/%s.CTAG.gz" %fname,'.'])
673
674   global_gps = np.loadtxt('%s.CTAG.gz'%fname,usecols=(6,),dtype='S500',
675     comments=None)
676
677   sp.call(['rm',"%s.CTAG.gz" %fname])
678
679   num_glob = global_gps.size
680
681   dirlist_gpsonly = []
```

```
682 for j in dirlist:
683   dirlist_gpsonly.append(j.split('_')[1])
684
685 # Make list of GPS seconds from signal data list
686 signal_data_gpsonly = []
687 for j in signal_data:
688   signal_data_gpsonly.append(j.split(' ')[0])
689
690 # Locate and move all global events
691 #Also, write event data to global text file
692 if num_glob > 0:
693   if num_glob == 1:
694     new_glob = np.zeros(1,dtype='S500')
695     new_glob[0] = global_gps
696     global_gps = new_glob
697
698   for g in global_gps:
699     gps_sec = g.decode('ascii').split('.')[0]
700     try:#Event might not have associated T3 (sad, but happens rarely it seems)
701       fold_ind = dirlist_gpsonly.index(gps_sec)
702     except:#Skip to next event
703       continue
704     d = dirlist[fold_ind]
705     sp.call(['mv',d+'/','/var/www/html/monitor/data/global_south/%s/' %fname])
706     data_ind = signal_data_gpsonly.index(gps_sec)
707     s = signal_data[data_ind]
708     with open('/home/augta/web_monitor/south_global_signal.txt','a') as f:
```

```
709         f.write(s+'\n')
710
711 # Get new dirlist which should, in principle, contain only local events
712 dirlist = os.listdir('.')
713 dirlist.sort()
714 dirlist_gpsonly = []
715 for j in dirlist:
716   dirlist_gpsonly.append(j.split('_')[1])
717
718 N = len(dirlist)
719 for i in range(N):
720   d = dirlist[i]
721   sp.call(['mv',d+'/','/var/www/html/monitor/data/local_south/%s/' %fname])
722   gps_sec = dirlist_gpsonly[i]
723   data_ind = signal_data_gpsonly.index(gps_sec)
724   s = signal_data[data_ind]
725   with open('/home/augta/web_monitor/south_local_signal.txt','a') as f:
726     f.write(s+'\n')
```

F.2 North daily events

```
1 # north_daily_events_final_2017_clarke_v5.py
2 # Copyright (c) Case Western Reserve University 2017
3 # This software is distributed under Apache License 2.0
4 # Consult the file LICENSE.txt
5 # Author: Sean Quinn spq@case.edu
```

```
6  # July 2017
7
8  import time
9  import math
10 import datetime
11 import matplotlib
12 matplotlib.use('Agg')
13 import matplotlib.pyplot as plt
14 from matplotlib.dates import DateFormatter, date2num
15 import numpy as np
16 import subprocess as sp
17 import time
18 import sys
19 import os
20 from itertools import groupby
21 import scipy.stats as scistats
22
23 #*****NOTICE*****: LEAPSEC MUST BE CHANGED TO 18 ON JAN 1 2017
24 def gpsFromUTC(year, month, day, hour, minute, sec, leapSecs=18):
25   """converts UTC to GPS second
26
27   Original function can be found at:
   ↪  http://software.ligo.org/docs/glue/frames.html
28
29   GPS time is basically measured in (atomic) seconds since
30   January 6, 1980, 00:00:00.0  (the GPS Epoch)
31
```

```
    The GPS week starts on Saturday midnight (Sunday morning), and runs
    for 604800 seconds.

    Currently, GPS time is 17 seconds ahead of UTC
    While GPS SVs transmit this difference and the date when another leap
    second takes effect, the use of leap seconds cannot be predicted.  This
    routine is precise until the next leap second is introduced and has to be
    updated after that.

    SOW = Seconds of Week
    SOD = Seconds of Day

    Note:  Python represents time in integer seconds, fractions are lost!!!
    """
    secFract = sec % 1
    epochTuple = gpsEpoch + (-1, -1, 0)
    t0 = time.mktime(epochTuple)
    t = time.mktime((year, month, day, hour, minute, sec, -1, -1, 0))
    # Note: time.mktime strictly works in localtime and to yield UTC, it should be
    #       corrected with time.timezone
    #       However, since we use the difference, this correction is unnecessary.
    # Warning:  trouble if daylight savings flag is set to -1 or 1 !!!
    t = t + leapSecs
    tdiff = t - t0
    gpsSOW = (tdiff % secsInWeek)  + secFract
    gpsWeek = int(math.floor(tdiff/secsInWeek))
    gpsDay = int(math.floor(gpsSOW/secsInDay))
```

```
59    gpsSOD = (gpsSOW % secsInDay)
60    gps_tuple = (gpsWeek, gpsSOW, gpsDay, gpsSOD)
61    return int(gps_tuple[0] * secsInWeek + gps_tuple[1])
62
63
64
65  def get_calibs(thedate):
66    flist = os.listdir('/clarke_local/imogen/Muons/%s' %thedate)
67    mu_only = []
68    for i in flist:
69      if '.dat' in i:
70        mu_only.append(i)
71    # Build histograms for each file.
72    nloc_file = '/clarke_local/imogen/data/local_north/%s/calib.txt' %thedate
73    nglob_file = '/clarke_local/imogen/data/global_north/%s/calib.txt' %thedate
74    nloc_dir = '/clarke_local/imogen/data/local_north/%s' %thedate
75    nglob_dir = '/clarke_local/imogen/data/global_north/%s' %thedate
76    # Remove calibration files if they exist
77    sp.call(['rm',nloc_file])
78    sp.call(['rm',nglob_file])
79    mu_only.sort()
80    for mu_file in mu_only:
81      mu_full_path = '/clarke_local/imogen/Muons/%s/%s' %(thedate,mu_file)
82      out_file = '/clarke_local/imogen/web_monitor/tmp/muon_tmp.txt'
83      sp.call(['./anamu','-i','%s' %mu_full_path,'-o','%s' %out_file])
84      # Get the starting time of muon run
85      temp_p = sp.Popen(['head','-n','1',out_file],stdout=sp.PIPE)
```

```
86      start_str = temp_p.stdout.readline()
87      start_str = start_str.split(' ')[-1]
88      start_str.replace('\n','')
89      start_str = start_str.split('.')[0]
90      # Get the finishing time of muon run
91      temp_p = sp.Popen(['tail','-n','1',out_file],stdout=sp.PIPE)
92      end_str = temp_p.stdout.readline()
93      end_str = end_str.split(' ')[-1]
94      end_str.replace('\n','')
95      end_str = end_str.split('.')[0]
96      # Load the A30 muons!
97      try:
98        mu = np.loadtxt(out_file,usecols=(1,))
99      except:
100       continue
101     # Get number of buffers
102     n = len(mu)
103     # This should be a multiple of 63!
104     if n % 63 != 0:
105       # Problem! make this an integer multiple of 63
106       n = n - n % 63
107     # Convert buffer number to number of muons
108     N = n / 63
109     # Instantiate charge and peak vectors
110     charge = []
111     peaks = []
112     # Build histogram!
```

```
113     for i in range(N):
114       tmp_mu = mu[i*63:(i+1)*63] - 511.5
115       charge.append(tmp_mu.sum())
116       peaks.append(tmp_mu.max())
117     # Save calibration data to local and global folders
118     charge = np.array(charge)
119     peaks = np.array(peaks)
120     charge = charge[charge>0]
121     charge = charge[charge<1200]
122     num_samps = len(charge)
123     k = scistats.gaussian_kde(charge,bw_method='silverman')
124     pos = np.arange(1200)
125     # KDE vector
126     z = k(pos)
127     # Argmax of KDE vector used as VEM PEAK
128     VEM = z[150:].argmax() + 150 #Ignore first 150 bins since they are pedestal
129     # Get bin size from histogram
130     y,x = np.histogram(charge,bins='fd') # Freedman Diaconis Estimator
131     bin_size = x[1] - x[0]
132     py,px = np.histogram(peaks,bins='fd')
133     PEAK = px[py.argmax()]
134     with open(nloc_file,'a') as f:
135       # start stop num_samps CHARGE_PEAK UNC_CHARGE_PEAK PEAK
136       f.write('%s %s %i %.2f %.2f %.2f\n'
    ↪  %(start_str,end_str,num_samps,VEM,bin_size,PEAK))
137     plt.figure(figsize=(16,9))
138     plt.subplot(121)
```

```
139       plt.hist(peaks,bins='fd',histtype='step')
140       plt.xlabel('A30  ADC Peak [ADC counts]')
141       plt.subplot(122)
142       #plt.hist(charge,bins=600,histtype='step',normed=True,label='600 bins')
143       plt.hist(charge,bins='fd',histtype='step',normed=True,label='FD')
144       plt.plot(pos,z,label='FD smoothed',lw=4)
145       plt.vlines(VEM,plt.ylim()[0],plt.ylim()[1])
146       plt.xlabel('VEM Charge [Integrated ADC counts]')
147       plt.ylabel('Normalized samples')
148       plt.title(r'1 A30 VEM = %i $\pm$ %i int. ADC counts' %(VEM,bin_size))
149       plt.legend()
150       plt.savefig(nloc_dir + '/%s_%s_calib.png' %(start_str,end_str))
151       plt.tight_layout()
152       plt.close('all')
153       # Copy this plot to global folder too
154       sp.call(['cp',nloc_dir + '/%s_%s_calib.png' %(start_str,end_str),nglob_dir])
155     # Finally, copy the calibration data to the global dir
156     sp.call(['cp',nloc_file,nglob_dir])
157 fdate = sp.check_output(['date','--date=-2 days','+%Y%m%d'])
158 fdate = fdate.replace('\n','')
159 yr = int(fdate[:4])
160 mo = int(fdate[4:6])
161 dy = int(fdate[6:])
162 secsInWeek = 604800
163 secsInDay = 86400
164 gpsEpoch = (1980, 1, 6, 0, 0, 0)  # (year, month, day, hh, mm, ss)
165 fname = "%i%02d%02d" %(yr,mo,dy)
```

```
166 os.chdir('/clarke_local/imogen')
167 # Collect events
168 evt_list = os.listdir('/clarke_local/imogen/Events/%s/' %fname)
169 evt_list = [k for k in evt_list if '.evt' in k]
170 glob_evt_list = [k for k in evt_list if 'global' in k]
171 glob_evt_list.sort()
172 local_evt_list = [k for k in evt_list if 'local' in k]
173 local_evt_list.sort()
174 # Directory stuff
175 nloc_dir = '/clarke_local/imogen/data/local_north/%s/' %fname
176 nglob_dir = '/clarke_local/imogen/data/global_north/%s/' %fname
177 sp.call(['mkdir',nloc_dir])
178 sp.call(['mkdir',nglob_dir])
179 sp.call(['rm',nloc_dir+'signal.txt'])
180 sp.call(['rm',nglob_dir+'signal.txt'])
181 # Get calibration information
182 get_calibs(fdate)
183 # Get list of muon run start times
184 muon_file = '/clarke_local/imogen/data/local_north/%s/calib.txt' %fdate
185 muon_runs = np.loadtxt(muon_file,usecols=(0,1))
186 vems = np.loadtxt(muon_file,usecols=(3,))
187 unc_vems = np.loadtxt(muon_file,usecols=(4,))
188 # Get event information
189 bl_evt = 514
190 if len(local_evt_list) > 0:
191   for m in local_evt_list:
192     curr_file = '/clarke_local/imogen/Events/%s/%s' %(fname,m)
```

```
193     fout = nloc_dir + m[:-4] + '.txt'
194     with open(fout,'w') as f:
195       sp.call(['./testevt',curr_file],stdout=f)
196     with open(fout,'r') as f:
197       gps_line = f.readline() # Get GPS information
198       if len(gps_line) == 0:
199         continue
200       gps_ts = gps_line.split(' ')[-2].replace(',','') # Select time stamp
201       utc_line = f.readline()
202       utc_date = utc_line.split(' ')[-2]
203       utc_time = utc_line.split(' ')[-1].split('.')[0]
204     utc = utc_date+'_'+utc_time
205     # Bracket event time between muon run start and finish
206     evt_gps = int(gps_ts.split('.')[0])
207     VEM_CHARGE = 0
208     UNC_VEM_CHARGE = 0
209     for j in range(len(muon_runs)):
210       if evt_gps > muon_runs[j,0] and evt_gps < muon_runs[j,1]:
211         VEM_CHARGE = vems[j]
212         UNC_VEM_CHARGE = unc_vems[j]
213     if VEM_CHARGE == 0:
214       print "Unable to find associated calibration value. Event information:"
215       print "%s" %gps_line
216       continue
217     a30_a01_a1 = np.loadtxt(fout,usecols=(1,2,4),skiprows=3)
218     # Saturation check
219     sat = [0]*3
```

```
220     for i in range(3):
221       max_adc = np.where(a30_a01_a1[:,i]>1020)[0]
222       if len(max_adc) > 2:
223         sat[i] = 1
224     a30_a01_a1[:,0] = a30_a01_a1[:,0] - 512
225     a30_a01_a1[:,1] = a30_a01_a1[:,1] - 515.2
226     a30_a01_a1[:,2] = a30_a01_a1[:,2] - 514
227     scale_factors = [1.,300.,30.]
228     signals = np.zeros(3)
229     low_signals = np.zeros(3)
230     high_signals = np.zeros(3)
231     f,axs = plt.subplots(nrows=1,ncols=3,figsize=(20,12))
232     ylabs = ['A x 30 Signal [VEM peak]','A x 0.1 Signal [VEM peak]','Anode Signal
    ↪  [VEM peak]']
233     x = np.arange(1024)
234     for i in range(3):
235       ymax = a30_a01_a1[:,i].argmax()
236       signals[i] = a30_a01_a1[ymax-30:ymax+60,i].sum() / (VEM_CHARGE /
    ↪  scale_factors[i])
237       low_signals[i] = a30_a01_a1[ymax-30:ymax+60,i].sum() / ((VEM_CHARGE -
    ↪  UNC_VEM_CHARGE) / scale_factors[i])
238       high_signals[i] = a30_a01_a1[ymax-30:ymax+60,i].sum() / ((VEM_CHARGE +
    ↪  UNC_VEM_CHARGE) / scale_factors[i])
239       trace_normed = a30_a01_a1[:,i] / (VEM_CHARGE / scale_factors[i])
240       trace_normed = trace_normed.max() * trace_normed
241       axs[i].set_ylabel(ylabs[i])
242       axs[i].set_xlabel('Time [10 ns]')
```

```
243         axs[i].step(x,trace_normed)
244         axs[i].set_xlim(ymax-30,ymax+60)
245         axs[i].set_title('Signal: %.3f VEM' %signals[i])
246         if sat[i] == 1:
247           axs[i].text(0.075,0.75,'SATURATED',color='red',fontsize=11,transform=axs[
    ↪  i].transAxes)
248       plt.tight_layout()
249       plt.savefig(nloc_dir + '%s_trace.png' %m[:-4])
250       plt.close('all')
251       locindex = m.split('_')[0]
252       out_str='%s %s %.3f %.3f %.3f %.3f %.3f %i %i %i\n'
    ↪  %(locindex,gps_ts,signals[-1],low_signals[-1],high_signals[-1],VEM_CHARGE,UN
    ↪  C_VEM_CHARGE,sat[0],sat[1],sat[2])
253       with open(nloc_dir + 'signal.txt','a') as f:
254         f.write(out_str)
255       with open('north_local_signal.txt','a') as f:
256         f.write(out_str)
257   # Same stuff, but for global events
258   if len(glob_evt_list) > 0:
259     for m in glob_evt_list:
260       curr_file = '/clarke_local/imogen/Events/%s/%s' %(fname,m)
261       fout = nglob_dir + m[:-4] + '.txt'
262       with open(fout,'w') as f:
263         sp.call(['./testevt',curr_file],stdout=f)
264       with open(fout,'r') as f:
265         gps_line = f.readline() # Get GPS information
266         if len(gps_line) == 0:
```

```
267           continue
268         gps_ts = gps_line.split(' ')[-2].replace(',','') # Select time stamp
269         utc_line = f.readline()
270         utc_date = utc_line.split(' ')[-2]
271         utc_time = utc_line.split(' ')[-1].split('.')[0]
272       utc = utc_date+'_'+utc_time
273       # Bracket event time between muon run start and finish
274       evt_gps = int(gps_ts.split('.')[0])
275       VEM_CHARGE = 0
276       UNC_VEM_CHARGE = 0
277       for j in range(len(muon_runs)):
278         if evt_gps > muon_runs[j,0] and evt_gps < muon_runs[j,1]:
279           VEM_CHARGE = vems[j]
280           UNC_VEM_CHARGE = unc_vems[j]
281       if VEM_CHARGE == 0:
282         print "Unable to find associated calibration value. Event information:"
283         print "%s" %gps_line
284         continue
285       a30_a01_a1 = np.loadtxt(fout,usecols=(1,2,4),skiprows=3)
286       # Saturation check
287       sat = [0]*3
288       for i in range(3):
289         max_adc = np.where(a30_a01_a1[:,i]>1020)[0]
290         if len(max_adc) > 2:
291           sat[i] = 1
292       a30_a01_a1[:,0] = a30_a01_a1[:,0] - 512
293       a30_a01_a1[:,1] = a30_a01_a1[:,1] - 515.2
```

```
294     a30_a01_a1[:,2] = a30_a01_a1[:,2] - 514
295     scale_factors = [1.,300.,30.]
296     signals = np.zeros(3)
297     low_signals = np.zeros(3)
298     high_signals = np.zeros(3)
299     f,axs = plt.subplots(nrows=1,ncols=3,figsize=(20,12))
300     ylabs = ['A x 30 Signal [VEM peak]','A x 0.1 Signal [VEM peak]','Anode Signal
    ↪  [VEM peak]']
301     x = np.arange(1024)
302     for i in range(3):
303       ymax = a30_a01_a1[:,i].argmax()
304       signals[i] = a30_a01_a1[ymax-30:ymax+60,i].sum() / (VEM_CHARGE /
    ↪  scale_factors[i])
305       low_signals[i] = a30_a01_a1[ymax-30:ymax+60,i].sum() / ((VEM_CHARGE -
    ↪  UNC_VEM_CHARGE) / scale_factors[i])
306       high_signals[i] = a30_a01_a1[ymax-30:ymax+60,i].sum() / ((VEM_CHARGE +
    ↪  UNC_VEM_CHARGE) / scale_factors[i])
307       trace_normed = a30_a01_a1[:,i] / (VEM_CHARGE / scale_factors[i])
308       trace_normed = trace_normed.max() * trace_normed
309       axs[i].set_ylabel(ylabs[i])
310       axs[i].set_xlabel('Time [10 ns]')
311       axs[i].step(x,trace_normed)
312       axs[i].set_xlim(ymax-30,ymax+60)
313       axs[i].set_title('Signal: %.3f VEM' %signals[i])
314       if sat[i] == 1:
315         axs[i].text(0.075,0.75,'SATURATED',color='red',fontsize=11,transform=axs[
    ↪  i].transAxes)
```

```
316     plt.savefig(nglob_dir + '%s_trace.png' %m[:-4])
317     plt.close('all')
318     globindex = m.split('_')[0]
319     out_str='%s %s %.3f %.3f %.3f %.3f %.3f %i %i %i\n'
    ↪  %(globindex,gps_ts,signals[-1],low_signals[-1],high_signals[-1],VEM_CHARGE,U
    ↪  NC_VEM_CHARGE,sat[0],sat[1],sat[2])
320     with open(nglob_dir + 'signal.txt','a') as f:
321       f.write(out_str)
322     with open('north_global_signal.txt','a') as f:
323       f.write(out_str)
```

Chapter G: Appendix G

G.1 Auger@TA MOU

Annex #3. Sharing of data to allow study of station-to-station comparisons between the TA scintillator SD and Auger water Cherenkov SD stations Auger@TA setup in Delta, UT

The scope of this annex is to identify the data to be shared between the Auger and TA collaborations as part of the Auger@TA project along the line of the umbrella MOU for data sharing already agreed upon by both collaborations.

The Auger@TA joint effort:

The Auger@TA project currently comprises four dedicated SD stations, 2 TA stations and 2 Auger stations (one Auger station and one Auger North prototype station) installed near the TA CLF location. The first part of this project, for which data sharing is requested, aims at collecting coincidence signals between WCD and scintillator detectors at the station level and, with the help of simulations, understand the correlations between VEM-based and MIP-based signals.

G.2 Data samples

Data are collected using two separate triggers:

- Local trigger (under development): a local trigger formed by one "local" TA station and the two Auger stations allows for the collection of coincident events. Given the MIP and VEM calibrations of the TA and Auger stations respectively, one can make limited station level comparisons since no reconstruction is possible. All the data (waveforms and timing information) from those stations are to be shared[1].

- Global trigger (in operation): when a global TA trigger is formed, the data from the non-participating "global" TA station (DET2421) at the CLF is read out and a trigger is provided to the Auger stations, which can then be read out as well. For coincident events, the following data are to be shared:

  - Waveforms and timing information from the two Auger SD stations
  - Waveforms and timing information from the "global" TA station (DET2421)
  - Timing and integrated MIP signal from the four closest neighbors (DET1216, DET1217, DET1316, DET1317) and only the timing from all the other participating stations that are identified by TA to reconstruct a coarse shower geometry.
  - Later, and according to the TA reconstruction release schedule, the full

[1] At present no local TA waveforms are saved

reconstruction parameters including zenith angle, azimuth angle, core location and energy

If some associated events are found to be special by TA, the data sharing of those events can be postponed with notification.

Chapter H: Appendix H

H.1 Build master

```
# build_master_v4.py
# Copyright (c) Case Western Reserve University 2017
# This software is distributed under Apache License 2.0
# Author: Sean Quinn spq@case.edu
# Feb 15 2017
import numpy as np
import matplotlib.pyplot as plt

#Edit variable below to point to file with Auger south global data
auger_global_file = "south_global_signal_extra.txt"
auger_data = np.loadtxt(auger_global_file,dtype='S300')

#http://stackoverflow.com/questions/1066758/find-length-of-sequences-of-identical
    -values-in-a-numpy-array
def rle(inarray):
        """ run length encoding. Partial credit to R rle function.
            Multi datatype arrays catered for including non Numpy
            returns: tuple (runlengths, startpositions, values) """
```

```
        ia = np.array(inarray)                  # force numpy
        n = len(ia)
        if n == 0:
            return (None, None, None)
        else:
            y = np.array(ia[1:] != ia[:-1])     # pairwise unequal (string safe)
            i = np.append(np.where(y), n - 1)   # must include last element posi
            z = np.diff(np.append(-1, i))       # run lengths
            p = np.cumsum(np.append(0, z))[:-1] # positions
            return(z, p, ia[i])

def get_trace(x):
  #Separate layers
  p1 = x[:16]
  p2 = x[16:]
  xvals = np.arange(128)
  #Put traces into arrays
  p1_trace = np.empty(0)
  p2_trace = np.empty(0)
  for i in p1:
    tmp = np.fromstring(i.split(':')[1],sep=' '*4)
    p1_trace = np.append(p1_trace,tmp)
  for i in p2:
    tmp = np.fromstring(i.split(':')[1],sep=' '*4)
    p2_trace = np.append(p2_trace,tmp)
  #Establish base lines
  #Expected baseline for both channels between 3-9
```

```
45    N = len(p1_trace)
46    s1 = False
47    s2 = False
48    if p1_trace.max() >= 4095:
49      s1 = True
50    if p2_trace.max() >= 4095:
51      s2 = True
52    p1_b_logic = np.zeros(N,dtype=int)
53    p2_b_logic = np.zeros(N,dtype=int)
54    p1_b = np.zeros(N,dtype=float)
55    p2_b = np.zeros(N,dtype=float)
56    p1_cnt_bins = np.histogram(p1_trace,bins=np.arange(40))[0]
57    p2_cnt_bins = np.histogram(p2_trace,bins=np.arange(40))[0]
58    p1_mean_base = np.average(p1_cnt_bins.argsort()[-2:][::-1],weights=p1_cnt_bins[
      ↪  p1_cnt_bins.argsort()[-2:][::-1]])
59    p2_mean_base = np.average(p2_cnt_bins.argsort()[-2:][::-1],weights=p2_cnt_bins[
      ↪  p2_cnt_bins.argsort()[-2:][::-1]])
60    p1_trace -= p1_mean_base
61    p2_trace -= p2_mean_base
62    return p1_trace,p2_trace,s1,s2
63
64  with open("MOU2_20170511_DET2421_globals_calib_2017.txt",'r') as f:
65    det2421 = f.read()
66
67  det2421_list = det2421.split('\n')
68
69  xbin = np.arange(128)
```

```
with open("auger_ta_joint_master_table_new.txt",'w') as F:
  evtid = 0
  for j in auger_data:
    as_tsp = j[1].split('.')[0]
    # Convert to TA time stamp convention
    yr,mo,d,tm = as_tsp.split('-')
    as_tsp_new = yr+"/"+mo+"/"+d+" "+tm
    # See if neighboring detectors share this timestamp
    if as_tsp_new in det2421:
      # Fetch TA timestamp
      loc = det2421.index(as_tsp_new)
      raw_ta = det2421[loc:loc+29]
      # Convert to ISO 8601 format
      ymd,tm,nano = raw_ta.split(' ')[:3]
      micro = "%.06f" %(float(nano)*1e-9)
      micro = micro.split('.')[1]
      ta_tsp = ymd.replace('/','-')+"T"+tm+".%s" %micro
      # Integrate FADC for both PMTs
      for k in range(len(det2421_list)):
        if as_tsp_new in det2421_list[k]:
          trace = det2421_list[k+1:k+1+32]
          evt_info = det2421_list[k]
          try:
            cal_str = evt_info[evt_info.index('mip0:'):]
          except:
            print "EXCEPTION: NO CALIBRATION DATA AVAILABLE FOR THIS EVENT.
    ↪  SKIPPING."
```

```
96              print "EVENT INFO: %s" %evt_info
97              ta_tsp = "inf"
98              break
99          cal0 = float(cal_str.split(' ')[1])
100         cal1 = float(cal_str.split(' ')[3])
101         pmt0,pmt1,sat0,sat1 = get_trace(trace)
102         # Saturation check
103         ta_sat = False
104         if sat0 == True or sat1 == True:
105           ta_sat = True
106         # Calculate signals
107         s0 = pmt0.sum() / cal0
108         s1 = pmt1.sum() / cal1
109         y0 = pmt0 / cal0
110         y1 = pmt1 / cal1
111         f,ax = plt.subplots(nrows=1,ncols=1)
112         ax.step(xbin,y0,'r-',where='mid',label='PMT0')
113         ax.hlines(0,0,128,color='black')
114 #           ax.vlines(v1,0,plt.ylim()[1]/1.5,color='red',linestyle='dashed')
115         ax.step(xbin,y1,'b--',where='mid',label='PMT1')
116 #           ax.vlines(v2,0,plt.ylim()[1]/1.5,color='blue',linestyle='dashed')
117         ax.set_xlabel('Time [20 ns]')
118         ax.set_ylabel('Signal [MIP]')
119         ta_time_str = det2421_list[k].split(' ')[1]+" "+det2421_list[k].split('
    ↪   ')[5]
120         ax.set_title(ta_time_str)
121         ax.set_ylim(ymin=-0.1)
```

```
122             ax.set_xlim(xmax=128)
123             ax.text(0.1,0.82,'PMT0\nS=%.2f MIP\nMIP Charge=%.2f'
    ↪  %(s0,cal0),fontsize=10,transform=ax.transAxes)
124             ax.text(0.45,0.82,'PMT1\nS=%.2f MIP\nMIP Charge=%.2f'
    ↪  %(s1,cal1),fontsize=10,transform=ax.transAxes)
125             ax.legend()
126             plt.savefig('%i_DET2421.png' %evtid)
127             plt.close('all')
128             break
129       det2421_out = "%i " %evtid + ta_tsp + " " + "%.2f " %s0 + "%.2f " %s1
130     else:
131       det2421_out = "%i " %evtid + "X X X "
132     ta_str = det2421_out
133     # Check for problems like missing calibration values
134     if "X "*3 in ta_str or "inf" in ta_str:
135       continue
136     #Convert Auger UTC string to ISO 8601
137     as_tsp_iso = j[1][:10]+"T"+j[1][11:]
138     #Check for high gain saturation
139     sat_flag = j[9:].astype('int')
140     #First 3 are low gain, last 3 are high gain
141     #If high gain flagged as saturated, use low gain
142     if 1 in sat_flag[3:]:
143       auger_str = as_tsp_iso+" %s"*4 %tuple(j[2:6])
144       as_pmt = j[3:6].astype('float')
145     else:
146       auger_str = as_tsp_iso+" %s" %j[2] + " %s"*3 %tuple(j[6:9])
```

```
147         as_pmt = j[6:9].astype('float')
148       auger_sat_str = " 0"
149       if 1 in sat_flag[:3]:
150         auger_sat_str = " 1"
151       ta_sat_str = " 0"
152       if ta_sat == True:
153         ta_sat_str = " 1"
154       pmt_mean = np.mean(as_pmt)
155       pmt_err = np.std(as_pmt)/np.sqrt(3)
156       pmt_err_str = " %s" %float("%.1g" %pmt_err)
157       # Figure out number of decimal places
158       prec = r" %."+str(len(pmt_err_str.split('.')[1]))+"f"
159       pmt_mean_str = prec %pmt_mean
160       auger_str += pmt_mean_str + pmt_err_str
161       final_str = ta_str + auger_str + ta_sat_str + auger_sat_str
162       evtid += 1
163       #Print GPS second for PMT trace retrieval
164       print "'%s'," %j[0].split('.')[0]
165       F.write(final_str + "\n")
```

Chapter I: Appendix I

```
RUNNR     101

NSHOW     25

PRMPAR    14

ESLOPE    0

ERANGE    2.603500e+09   2.603500e+09

SEED      7523175  0  0

SEED      5977659  0  0

SEED      9995598  0  0

SEED      3500117  0  0

DIRECT    /home/seanserver/CORSIKA/out/QGSJETII-04/proton/10/

THIN      1.000000E-07 12500. 58215

THINH     1.000E+00 1.000E+02

THETAP    37.44   37.44

PHIP      168.46   168.46

ATMOD     17

OBSLEV    139100.0

MAGNET    21.8787  45.8686
```

```
ECUTS     5.000E-02 5.000E-02 2.000E-04 2.000E-04

MUADDI    T

MUMULT    T

ELMFLG    F   T

STEPFC    1.0

RADNKG    5.0E+05

LONGI     T  5.  T  T

ECTMAP    2.5E+5

MAXPRT    1000

PAROUT    T T

DATBAS    T

EXIT
```

Chapter J: Appendix J

J.1 EventGenerator.xml.in

```
<?xml version="1.0" encoding="iso-8859-1"?>

<!DOCTYPE EventGenerator [

<!ENTITY UTMData "<zone> 19 </zone>

<band> H </band>">

]>


<!-- Configuration of Module EventGenerator -->

<!-- In the EventGenerator tag we define the path and name of the schema file

(EventGenerator.xsd) used to validate this xml file. When make is invoked,

\@SCHEMALOCATION\@ gets replaced by the path to the directory in the install area

where configuration files are copied. -->


<EventGenerator

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xsi:noNamespaceSchemaLocation='@SCHEMALOCATION@/EventGenerator.xsd'>


<mode> SD </mode>


<eventIdentifier>
```

```
<libraryIdentifier> Test </libraryIdentifier>

<format> Library_%1%:Run_%2%:Shower_%3%:Use_%4% </format>

<!-- digits to reserve in SD id for shower and use -->

<sdIdFormat> 2 2 </sdIdFormat>

</eventIdentifier>


<!-- Specify how to randomize the core position -->

<coreRandomization>

<listOfCorePositions>

<core>

<northing unit="meter"> 6092553 </northing>

<easting unit="meter"> 480227 </easting>

<altitude unit="meter"> 1391 </altitude>

&UTMData;

</core>

<core>

<northing unit="meter"> 6092553 </northing>

<easting unit="meter"> 480227 </easting>

<altitude unit="meter"> 1391 </altitude>

&UTMData;

</core>

.

.

.

</listOfCorePositions>

</coreRandomization>
```

```
<!-- Time stamp of the core impact on ground -->

<eventTime> 2008-01-01T04:33:12.5 </eventTime>


<!-- Use core position and/or time from radio event data -->

<useRadioCorePosition>0</useRadioCorePosition>

<useRadioEventTime>0</useRadioEventTime>


</EventGenerator>
```

J.2 bootstrap.xml.in

```
<?xml version="1.0" encoding="iso-8859-1"?>

<!DOCTYPE bootstrap [
<!ENTITY standardSdIdealDetConfig SYSTEM '@CONFIGDIR@/standardSdIdealDetConfig.xml'>
<!ENTITY testSdSimModuleConfig SYSTEM '@CONFIGDIR@/testSdSimModuleConfig.xml'>
<!ENTITY standardSdRecModuleConfig SYSTEM '@CONFIGDIR@/standardSdRecModuleConfig.xml'>
]>

<bootstrap
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation='@SCHEMALOCATION@/bootstrap.xsd'
xmlns:xlink="http://www.auger.org/schema/types">

&standardSdIdealDetConfig;
&testSdSimModuleConfig;
&standardSdRecModuleConfig;

<centralConfig>
<configLink
id         = "ModuleSequence"
type       = "XML"
xlink:href = "./ModuleSequence_Geant410.xml"/>
<configLink
id         = "EventFileReader"
type       = "XML"
xlink:href = "./EventFileReader.xml"/>
<configLink
id         = "EventGenerator"
```

```
type       = "XML"
xlink:href = "./EventGenerator.xml"/>
<configLink
id         = "G4StationSimulator"
type       = "XML"
xlink:href = "./G4StationSimulator.xml"/>
<configLink
id         = "SManagerRegister"
type       = "XML"
xlink:href = "./standardSdIdealSManagerRegisterConfig.xml"/>
<configLink
id         = "SStationListXMLManager"
type       = "XML"
xlink:href = "@CONFIGDIR@/SIdealUpgradedStationList.xml"/>
<configLink
id         = "DataWriter"
type       = "XML"
xlink:href = "./DataWriter.xml"/>
<configLink
id         = "CachedShowerRegenerator"
type       = "XML"
xlink:href = "@CONFIGDIR@/CachedShowerRegenerator.xml"/>
</centralConfig>

<parameterOverrides>

<configLink id="CachedShowerRegenerator">
<CachedShowerRegenerator>
<EnergyCuts>
<ElectronEnergyCut unit="MeV"> 0.2 </ElectronEnergyCut>
<PhotonEnergyCut   unit="MeV"> 0.2  </PhotonEnergyCut>
</EnergyCuts>
</CachedShowerRegenerator>
</configLink>

<configLink id="SStationListXMLManager">
<stationList>
<station id="5040">
<altitude unit="meter"> 1391 </altitude>
<northing unit="meter"> 6091805 </northing>
<easting  unit="meter"> 480594 </easting>
</station>
<station id="5048">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5084">
```

```
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5083">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5085">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5047">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5050">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5049">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5012">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5013">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="5014">
<altitude unit="meter"> 1391 </altitude>
<easting  unit="meter"> 482000 </easting>
</station>
<station id="5015">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="4979">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="4980">
<altitude unit="meter"> 1391 </altitude>
</station>
<station id="4981">
<altitude unit="meter"> 1391 </altitude>
</station>
</stationList>
</configLink>

</parameterOverrides>

</bootstrap>
```

Bibliography

[1] C. de Coulomb. *Mdm. de l'Acad. des Sciences*, page 612, 1875.

[2] J W Cronin. *Astroparticle Physics*, 53:6–18, 2014.

[3] A d Angelis. *Astroparticle Physics*, 53:19–26, 2014.

[4] M Faraday. *Experimental Researches in Electricity Volume 2*. University of London, 1844.

[5] W Crookes. *Proceedings of the Royal Society*, 38:347, 1879.

[6] Q Xu and L M Brown. *American Journal of Physics*, 55:23–33, 1987.

[7] H Becquerel. *Comptes Rendus de l'Acad. des Sci.*, 122:420, 1896.

[8] P Curie, Mme P Curie, and G Bémont. *Comptes Rendus de l'Acad. des Sci.*, 127:1215, 1898.

[9] T. Wulf. *Phys. Zeit.*, 10:152–157, 1909.

[10] T. Wulf. *Phys. Zeit.*, 11:811–813, 1910.

[11] D. Pacini. *Ann. Uff. Centr. Meteor. XXXII parte 1*, 1910.

[12] D. Pacini. *Le Radium VIII*, 307, 1911. URL: http://arxiv.org/abs/1101.3015.

[13] A. Gockel. *Phys. Zeit.*, 11:280–282, 1910.

[14] V. Hess. *Phys. Zeit.*, 13:1084, 1912.

[15] A. M. Hillas. *Cosmic Rays*. Pergamon Press, London, 1972.

[16] W. Kolhorster. *Verh. Dtsch. Phys. Ges.*, 16:719, 1914.

[17] R.A. Millikan and R.M. Otis. *Phys. Rev. Lett.*, 23:778, 1924.

[18] H. Geiger and E. Marsden. *Proc. of the Royal Soc. Series A*, 82:495–500, 1909.

[19] B. Rossi. *Ziet. fur Phys.*, 82:151, 1933.

[20] The ATLAS Collaboration. *Eur. Phys. J. C*, 74:3034, 2014.

[21] P.M.S Blackett and G. Occhialini. *Proc. Roy. Soc.*, 139:699, 1933.

[22] C.D. Anderson. *Phys. Rev. Lett.*, 43:491, 1933.

[23] Studies in cosmic ray muons. http://physics.okstate.edu/rpl/muons.htm. Accessed: 2016-10-26.

[24] E. Regener and G. Pfotzer. *Nature*, 136:718, 1935.

[25] P. Auger et al. *Comptes Rendus*, 208:1641, 1939.

[26] R.W. Williams. *Phys. Rev. Lett.*, 74:1689, 1948.

[27] A.A. Watson. *Astroparticle Phys.*, 53:107–114, 2014.

[28] P. Bassi et al. *Phys. Rev. Lett.*, 92:441, 1953.

[29] G. W. Clark et al. *Phys. Rev. Lett.*, 122:637–654, 1960.

[30] K. Kamata and J. Nishimura. *Prog. Theor. Phys. Supplement*, 6:93–155, 1958. doi:10.1143/PTPS.6.93.

[31] K. Greisen. *Progress in Cosmic Ray Physics*, volume 3. North-Holland Publishing Company, Amsterdam, 1956.

[32] M. Nagano et al. *J. Phys.*, G10:1295, 1984. doi:10.1143/JPSJ.53.1667.

[33] J. Linsley et al. *Phys. Rev. Lett.*, 6:485–487, 1961. doi:10.1103/PhysRevLett.6.485.

[34] J. Linsley. *Phys. Rev. Lett.*, 10:146–148, 1963. doi:10.1103/PhysRevLett.10.146.

[35] M. A. Lawrence et al. *J. Phys. G*, 17:733–757, 1991. doi:10.1088/0954-3899/17/5/019.

[36] M. M. Winn et al. *J. Phys. G: Nucl. Phys.*, 12:653, 1986. doi:10.1088/0305-4616/12/7/015.

[37] M. M. Winn et al. *J. Phys. G: Nucl. Phys.*, 12:675, 1986. doi:10.1088/0305-4616/12/7/016.

[38] K. Suga. *5th Inter-American Seminar on Cosmic Rays (Bolivia)*, 1962.

[39] A. Chudakov. *5th Inter-American Seminar on Cosmic Rays (Bolivia)*, 1962.

[40] T. Hara et al. *Atca Phys. Acad. Sci. Hung.*, 29:369–376, 1970. doi:10.1088/0954-3899/17/5/019.

[41] B. Dawson. *arXiv:1112.5686v1*, 2011. arXiv:1112.5686v1.

[42] H. E. Bergeson et al. *Phys. Rev. Lett.*, 39:847–849, 1977. doi:10.1103/PhysRevLett.39.847.

[43] N. Chiba et al. *Nucl. Inst. Meth. in Phys. Res. A*, 311:338–349, 1992. doi:10.1016/0168-9002(92)90882-5.

[44] K. Shinozaki et al. *Nucl. Phys. B Proc. Supp.*, 151:3–10, 2006. doi:10.1016/j.nuclphysbps.2005.07.002.

[45] M. Takeda et al. *Astro. Part. Phys.*, 19:447–462, 2003. doi:10.1016/S0927-6505(02)00243-8.

[46] A. A. Ivanov, S. P. Knurenko, and I. Ye. Sleptsov. *Nucl. Phys. B (Proc. Suppl.)*, 122:226, 2003. doi:10.1016/S0920-5632(03)80385-7.

[47] A. A. Ivanov et al. *EPJ Web of Conferences*, 53:04003, 2013. doi:10.1051/epjconf/20135304003.

[48] R.M. Baltrusaitis et al. *Nucl. Inst. Meth. in Phys. Res. A*, 240:410–428, 1985.

[49] D.J. Bird et al. *Astrophysical Journal*, 441:144, 1995. doi:10.1086/175344.

[50] M. Nagano and A. A. Watson. *Rev. Mod. Phys.*, 72:689, 2000. doi:10.1103/RevModPhys.72.689.

[51] R. Abbasi et al. *Astropart. Phys.*, 30:175, 2008. doi:10.1016/j.astropartphys.2008.08.004.

[52] J. Belz. *Nucl. Phys. B (Proc. Suppl.)*, 190:5, 2009. doi:10.1016/j.nuclphysbps.2009.03.061.

[53] M. Kachelriess V. Berezinsky and A. Vilenkin. *Phys. Rev. Lett.*, 79:22, 1997. doi:10.1103/PhysRevLett.79.4302.

[54] Th. M. Nieuwenhuizen. *EPL*, 86:59001, 2009. doi:10.1209/0295-5075/86/59001.

[55] T. J. Weiler. *arXiv*, 1999. arXiv:hep-ph/9910316.

[56] P. Bhattacharjee and G. Sigl. *Phys. Rep.*, 327:109–247, 2000. doi:10.1016/S0370-1573(99)00101-5.

[57] M. Ackermann. *ApJ*, 787:15, 2014. doi:10.1088/0004-637X/787/1/15.

[58] E. Fermi. *Phys. Rev.*, 75:1169, 1949. doi:10.1103/PhysRev.75.1169.

[59] L. Davis Jr. *Phys. Rev.*, 101:351, 1956. doi:10.1103/PhysRev.75.1169.

[60] T. Gaisser. *Cosmic Rays and Particle Physics*.

[61] V. N. Zirakashvili and H. J. Volk. *Adv. in Space Res.*, 37:1923, 2006. doi:10.1016/j.asr.2005.06.013.

[62] M. T. Dova. *arXiv*, 2016. arXiv:1604.07584.

[63] A. Reisenegger. *arXiv:1305.2542*, 2013.

[64] B. C. Lacki. *Astronomy & Astrophysics*, 444:L39, 2014. doi:10.1093/mnrasl/slu107.

[65] A. Wandel, B. M. Peterson, and M. A. Malkan. *The Astrophysical Journal*, 526(2):579, 1999. URL http://stacks.iop.org/0004-637X/526/i=2/a=579.

[66] F. Halzen and E. Zas. *The Astrophysical Journal*, 488(2):669, 1997. URL http://stacks.iop.org/0004-637X/488/i=2/a=669.

[67] K. Greisen. *Phys. Rev. Lett.*, 16:748–750, 1966. 10.1103/PhysRevLett.16.748.

[68] V.A. Kuzmin G.T. Zatsepin. *JETP. Lett.*, 4:78, 1966.

[69] IceCube collaboration. *arXiv:1701.03731*, 2017. arXiv:1701.03731v1.

[70] P. P. Kronberg. *Astronomische Nachrichten*, 327:517–522, 2006. doi: 10.1002/asna.200610576. URL http://dx.doi.org/10.1002/asna.200610576.

[71] M. S. Pshirkov, P. G. Tinyakov, and F. R. Urban. *Phys. Rev. Lett.*, 116: 191302, 2016. doi: 10.1103/PhysRevLett.116.191302. URL https://link.aps.org/doi/10.1103/PhysRevLett.116.191302.

[72] T. E. Clarke, P. P. Kronberg, and Hans Bhringer. *The Astrophysical Journal Letters*, 547:L111, 2001. URL http://stacks.iop.org/1538-4357/547/i=2/a=L111.

[73] T. R. Jaffe et al. *Monthly Notices of the Royal Astronomical Society*, 401: 1013–1028, 2010. doi: 10.1111/j.1365-2966.2009.15745.x.

[74] M. Krause. *IAU Symposium*, 140:187–196, 1990. URL http://adsabs.harvard.edu/abs/1990IAUS..140..187K.

[75] A. Boulares and D. P. Cox. *ApJ*, 365:544, 1990.

[76] D. H. F. M. Schnitzeler. *MNRAS*, 409:L99, 2010.

[77] W. Heitler. *Quantum Theory of Radiation*. Oxford University Press, 1944.

[78] J. W. Elbert. *Proc. DUMAND Summer Workshop*, 2:101, 1978.

[79] H. Bilokon et al. *Proc. 21st Int. Cosmic Ray Conf.*, 9:366, 1990.

[80] Pierre Auger Collaboration. *Phys. Rev. D*, 90, 2014. doi: 10.1103/PhysRevD.90.122005.

[81] T. K. Gaisser and A. M. Hillas. *Proc. 15th Int. Cosmic Ray Conf.*, 8:353, 1977.

[82] Pierre Auger Collaboration. *NIMA*, 798:172, 2015. doi: 10.1016/j.nima.2015.06.058.

[83] X. Bertou et al. *NIMA*, 568:839–846, 2006. doi: https://doi.org/10.1016/j.nima.2006.07.066.

[84] K. Martens. The telescope array and its low energy extension. *Nuclear Physics B - Proceedings Supplements*, 165:33 – 36, 2007. doi: http://dx.doi.org/10.1016/j.nuclphysbps.2006.11.006. URL http://www.sciencedirect.com/science/article/pii/S0920563206008395.

[85] Telescope Array Collaboration. *NIMA*, 689:87 – 97, 2012. doi: http://dx.doi.org/10.1016/j.nima.2012.05.079. URL http://www.sciencedirect.com/science/article/pii/S0168900212005931.

[86] D. Nitz. *IEEE Transactions on Nuclear Science*, 51(3):413–419, 2004.

[87] Pierre Auger Collaboration. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 523:50 – 95, 2004. doi: http://dx.doi.org/10.1016/j.nima.2003.12.012. URL http://www.sciencedirect.com/science/article/pii/S0168900203033497.

[88] M. Allen. Ultra high energy cosmic ray energy spectrum and composition using hybrid analysis with telescope array. *University of Utah*, 2012. URL http://www.telescopearray.org/media/theses/mallen_thesis.pdf.

[89] Dmitri Ivanov. *Energy Spectrum Measured by the Telescope Array Surface Detector*. PhD thesis, Rutgers University, 2012.

[90] The Pierre Auger collaboration. *Journal of Instrumentation*, 6:P01003, 2011. URL http://stacks.iop.org/1748-0221/6/i=01/a=P01003.

[91] Pierre Auger Collaboration. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 613:29 – 39, 2010. doi: http://dx.doi.org/10.1016/j.nima.2009.11.018. URL http://www.sciencedirect.com/science/article/pii/S0168900209021688.

[92] Valerio Verzi for the Pierre Auger Collaboration. *Measurement of the Energy Spectrum of Cosmic Rays with the Pierre Auger Observatory*. doi: 10.7566/JPSCP.9.010001. URL http://journals.jps.jp/doi/abs/10.7566/JPSCP.9.010001.

[93] Mariangela Settimo. *The European Physical Journal Plus*, 127:87, 2012. doi: 10.1140/epjp/i2012-12087-9. URL https://doi.org/10.1140/epjp/i2012-12087-9.

[94] Giorgio Matthiae. *New Journal of Physics*, 12:075009, 2010. URL http://stacks.iop.org/1367-2630/12/i=7/a=075009.

[95] D. Veberic for the Pierre Auger Collaboration. *Proceedings of the 33rd ICRC*, page 23, 2013. URL https://arxiv.org/abs/1307.5059.

[96] J. Hersil, I. Escobar, D. Scott, G. Clark, and S. Olbert. *Phys. Rev. Lett.*, 6: 22–23, 1961. doi: 10.1103/PhysRevLett.6.22. URL https://link.aps.org/doi/10.1103/PhysRevLett.6.22.

[97] S Yoshida et al. *Journal of Physics G: Nuclear and Particle Physics*, 20:651, 1994. URL http://stacks.iop.org/0954-3899/20/i=4/a=011.

[98] D. Newton, J. Knapp, and A.A. Watson. *Astroparticle Physics*, 26:414 – 419, 2007. doi: http://dx.doi.org/10.1016/j.astropartphys.2006.08.003. URL http://www.sciencedirect.com/science/article/pii/S0927650506001216.

[99] Pierre Sokolski. *SPIE Newsroom*, 2015. doi: http://dx.doi.org/10.1117/2.1201508.006098. URL http://spie.org/newsroom/6098-searching-for-a-cosmic-ray-hotspot-with-the-telescope-array?SSO=1.

[100] Telescope Array Collaboration. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 676:54 – 65, 2012. doi: http://dx.doi.org/10.1016/j.nima.2012.02.044. URL http://www.sciencedirect.com/science/article/pii/S0168900212002422.

[101] Pierre Auger Collaboration. *Astroparticle Physics*, 35:591 – 607, 2012. doi: http://dx.doi.org/10.1016/j.astropartphys.2011.12.002. URL http://www.sciencedirect.com/science/article/pii/S0927650511002271.

[102] Telescope Array Collaboration. *The Astrophysical Journal Letters*, 768:L1, 2013. URL http://stacks.iop.org/2041-8205/768/i=1/a=L1.

[103] The Pierre Auger Observatory: Contributions to the 35th International Cosmic Ray Conference (ICRC 2017). 2017. URL https://inspirehep.net/record/1617990/files/arXiv:1708.06592.pdf.

[104] *The Pierre Auger Observatory: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*, 2015. URL https://inspirehep.net/record/1393211/files/arXiv:1509.03732.pdf.

[105] Pierre Auger Collaboration. *The Astrophysical Journal*, 804:15, 2015. URL http://stacks.iop.org/0004-637X/804/i=1/a=15.

[106] *TA Spectrum Summary: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*, 2015. URL https://pos.sissa.it/236/349/pdf.

[107] *Energy Spectrum and Mass Composition of Ultra-High Energy Cosmic Rays Measured with the Telescope Array Fluorescence Detector: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*, 2015. URL https://pos.sissa.it/236/320/pdf.

[108] *Energy Spectrum and Mass Composition of Ultra-High Energy Cosmic Rays Measured by the hybrid technique in Telescope Array: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*, 2015. URL https://pos.sissa.it/236/362/pdf.

[109] The Telescope Array Collaboration. *The Astrophysical Journal Letters*, 790: L21, 2014. URL http://stacks.iop.org/2041-8205/790/i=2/a=L21.

[110] *Telescope Array anisotropy summary: Contributions to the 35th International Cosmic Ray Conference (ICRC 2017)*, 2017. URL https://pos.sissa.it/301/548/pdf.

[111] *Evidence of Intermediate-Scale Energy Spectrum Anisotropy in the Northern Hemisphere from Telescope Array: Contributions to the 35th International Cosmic Ray Conference (ICRC 2017)*, 2017. URL https://pos.sissa.it/301/513/pdf.

[112] C. Patrignani et al. *Chin. Phys.*, C40:100001, 2016. doi: 10.1088/1674-1137/40/10/100001. URL http://dx.doi.org/10.1088/1674-1137/40/10/100001.

[113] *I.C. Maris for the Pierre Auger and Telescope Array Collaborations, High Energy Spectrum Working Group Report, presentation at the UHECR symposium (2014), Springdale, USA*, 2014. URL uhecr2014.telescopearray.org/maris/TAAuger_Springdale.pdf.

[114] *V. Verzi for the Pierre Auger and Telescope Array Collaborations, Energy Spectrum Working Group Report, presentation at the UHECR symposium (2016), Kyoto, Japan*, 2016. URL https://indico.cern.ch/event/504078/contributions/2290831/attachments/1352191/2041629/SpectrunWG-KyotoOct2016.pdf.

[115] *Report of theWorking Group on the Composition of Ultra-High Energy Cosmic Rays: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*, 2015. URL https://pos.sissa.it/236/307/pdf.

[116] *Testing the agreement between the Xmax distributions measured by the Pierre Auger and Telescope Array Observatories: Contributions to the 35th International Cosmic Ray Conference (ICRC 2017)*, 2017. URL https://pos.sissa.it/301/522/pdf.

[117] *Large-Scale Distribution of Arrival Directions of Cosmic Rays Detected at the Pierre Auger Observatory and the Telescope Array above* 10^{19} *eV: Contributions to the 35th International Cosmic Ray Conference (ICRC 2015)*, 2015. URL https://pos.sissa.it/236/395/pdf.

[118] *Classical Mechanics Third Edition*. Addison Wesley, 2002.

[119] V. López-Barquero, S. Xu, P. Desiati, A. Lazarian, N. V. Pogorelov, and H. Yan. *The Astrophysical Journal*, 842:54, 2017. URL http://stacks.iop.org/0004-637X/842/i=1/a=54.

[120] Ronnie Jansson and Glennys R. Farrar. *The Astrophysical Journal*, 757:14, 2012. URL http://stacks.iop.org/0004-637X/757/i=1/a=14.

[121] Sun, X. H., Reich, W., Waelkens, A., and Ensslin, T. A. *A&A*, 477:573–592, 2008. doi: 10.1051/0004-6361:20078671. URL https://doi.org/10.1051/0004-6361:20078671.

[122] M. S. Pshirkov, P. G. Tinyakov, P. P. Kronberg, and K. J. Newton-McGee. *The Astrophysical Journal*, 738:192, 2011. URL http://stacks.iop.org/0004-637X/738/i=2/a=192.

[123] Fauvet, L., Macías-Pérez, J. F., et al. *A&A*, 540:A122, 2012. doi: 10.1051/0004-6361/201016349. URL https://doi.org/10.1051/0004-6361/201016349.

[124] T. R. Jaffe, K. M. Ferrière, A. J. Banday, et al. *Monthly Notices of the Royal Astronomical Society*, 431:683–694, 2013. doi: 10.1093/mnras/stt200. URL http://dx.doi.org/10.1093/mnras/stt200.

[125] J. C. Brown et al. *The Astrophysical Journal*, 663:258, 2007. URL http://stacks.iop.org/0004-637X/663/i=1/a=258.

[126] R. Beck. *Ap&SS*, 320:77, 2009.

[127] M. Krause. *Revista Mexicana de Astronomia y Astrofisica Conference Series*, 36:25, 2009.

[128] Ronnie Jansson and Glennys R. Farrar. *The Astrophysical Journal Letters*, 761:L11, 2012. URL http://stacks.iop.org/2041-8205/761/i=1/a=L11.

[129] D. Larson et al. *The Astrophysical Journal Supplement Series*, 192(2):16, 2011. URL http://stacks.iop.org/0067-0049/192/i=2/a=16.

[130] A. W. Strong et al. *arxiv*, (0907.0559), 2009.

[131] Waelkens, A. et al. *A&A*, 495:697–706, 2009. doi: 10.1051/0004-6361:200810564. URL https://doi.org/10.1051/0004-6361:200810564.

[132] Todor Stanev. *The Astrophysical Journal*, 479:290, 1997. URL http://stacks.iop.org/0004-637X/479/i=1/a=290.

[133] A. Gelman et al. *Bayesian Data Analysis*. CRC Press, 2014.

[134] J. Geweke. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. *Bayesian Statistics 4*, 1992.

[135] Stephen P. Brooks and Andrew Gelman. *Journal of Computational and Graphical Statistics*, 7:434–455, 1998. doi: 10.1080/10618600.1998.10474787. URL http://www.tandfonline.com/doi/abs/10.1080/10618600.1998.10474787.

[136] M.S. Sutherland, B.M. Baughman, and J.J. Beatty. *Astroparticle Physics*, 34: 198 – 204, 2010. URL http://dx.doi.org/10.1016/j.astropartphys.2010.07.002.

[137] A. Saltelli, K. Chan, and E. M. Scott. *Sensitivity Analysis*. John Wiley & Sons, Ltd., 2000.

[138] P. Mahalanobis. On the generalized distance in statistics. *Proc. Nat. Inst. Sci. India (Calcutta)*, 2:49 – 55, 1936.

[139] D. W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, 1992.

[140] A. Neronov and I. Vovk. Evidence for strong extragalactic magnetic fields from fermi observations of tev blazars. *Science*, 328:73, 2010. doi: 10.1126/science.1184192.

[141] A. Neronov and I. Vovk. Extragalactic magnetic fields constraints from simultaneous gev-tev observations of blazars. *Astronomy & Astrophysics*, 529: A144, 2011. doi: 10.1051/0004-6361/201116441.

[142] E. W. Flesch. The half million quasars (HMQ) catalogue. *Publications of the Astronomical Society of Australia*, 32:e010, 2015. doi: 10.1017/pasa.2015.10.

[143] S. Quinn and C. Covault. A monte carlo study of arrival direction uncertainty for the highest energy herald events using the jf12 model (v1). *Auger GAP Notes*, page 012, 2016. URL https://www.auger.org/index.php/document-centre/finish/139-gap-notes-2016/3994-gap2016-012.

[144] A. Kolmogorov. The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers. *Akademiia Nauk SSSR Doklady*, 30: 301–305, 1941. URL http://adsabs.harvard.edu/abs/1941DoSSR..30..301K.

[145] S. R. Spangler and C. R. Gwinn. Evidence for an inner scale to the density turbulence in the interstellar medium. *Astrophysical Journal Letters*, 1990.

[146] Daniel R. Stinebring et al. Five years of pulsar flux density monitoring: Refractive scintillation and the interstellar medium. *The Astrophysical Journal*, 539:300, 2000. URL http://stacks.iop.org/0004-637X/539/i=1/a=300.

[147] W. Wang et al. Contribution to diffuse gamma-rays in the galactic centre region from unresolved millisecond pulsars. *Monthly Notices of the Royal Astronomical Society*, 358:263–269, 2005. doi: 10.1111/j.1365-2966.2005.08816.x. URL http://dx.doi.org/10.1111/j.1365-2966.2005.08816.x.

[148] X. P. You et al. An improved solar wind electron density model for pulsar timing. *The Astrophysical Journal*, 671:907, 2007. URL http://stacks.iop.org/0004-637X/671/i=1/a=907.

[149] W. H. Matthaeus and M. L. Goldstein. Measurement of the rugged invariants of magnetohydrodynamic turbulence in the solar wind. *J. Geophys. Res*, 87: 6011, 1982.

[150] D. A. Roberts and M. L. Goldstein. Turbulence and waves in the solar wind. *Rev. Geophys.*, 29:932, 1991.

[151] Robert H. Kraichnan. Inertialrange spectrum of hydromagnetic turbulence. *The Physics of Fluids*, 8:1385–1387, 1965. doi: 10.1063/1.1761412. URL http://aip.scitation.org/doi/abs/10.1063/1.1761412.

[152] S. GALTIER et al. A weak turbulence theory for incompressible magnetohydrodynamics. *Journal of Plasma Physics*, 63(5):447488, 2000.

[153] P. Goldreich and S. Sridhar. Toward a theory of interstellar turbulence. 2: Strong alfvenic turbulence. *Astrophysical Journal*, 438:763–775, 1995. doi: 10.1086/175121. URL http://adsabs.harvard.edu/abs/1995ApJ...438..763G.

[154] J. R. Jokipii and I. Lerche. Faraday Rotation, Dispersion in Pulsar Signals, and the Turbulent Structure of the Galaxy. *Astrophysical Journal*, 157:1137, 1969. doi: 10.1086/150142.

[155] M. Haverkorn et al. The outer scale of turbulence in the magnetoionized galactic interstellar medium. *The Astrophysical Journal*, 680:362, 2008. URL http://stacks.iop.org/0004-637X/680/i=1/a=362.

[156] A. Chepurnov et al. Velocity spectrum for h i at high latitudes. *The Astrophysical Journal*, 714:1398, 2010. URL http://stacks.iop.org/0004-637X/714/i=2/a=1398.

[157] Hiroshi Ohno and Shinpei Shibata. The random magnetic field in the galaxy. *Monthly Notices of the Royal Astronomical Society*, 262:953–962, 1993. doi: 10.1093/mnras/262.4.953. URL http://dx.doi.org/10.1093/mnras/262.4.953.

[158] A. W. Clegg et al. *The Astrophysical Journal*, 714:1398, 2010.

[159] Iacobelli, M. et al. Studying galactic interstellar turbulence through fluctuations in synchrotron emission - first lofar galactic foreground detection. *A&A*, 558:A72, 2013. doi: 10.1051/0004-6361/201322013. URL https://doi.org/10.1051/0004-6361/201322013.

[160] Beck, R. et al. *Plasma Turbulence and Energetic Particles in Astrophysics: Proceedings of the International Conference, Cracow*.

[161] A. Fletcher et al. *Monthly Notices of the Royal Astronomical Society*, 412: 2396–2416, 2011. doi: 10.1111/j.1365-2966.2010.18065.x. URL http://dx.doi.org/10.1111/j.1365-2966.2010.18065.x.

[162] M. A. Malkov et al. *The Astrophysical Journal*, 721:750, 2010. URL http://stacks.iop.org/0004-637X/721/i=1/a=750.

[163] Rafael Alves Batista et al. *Journal of Cosmology and Astroparticle Physics*, 2016:038, 2016. URL http://stacks.iop.org/1475-7516/2016/i=05/a=038.

[164] S. Gillessen et al. *The Astrophysical Journal*, 837:30, 2017. URL http://stacks.iop.org/0004-637X/837/i=1/a=30.

[165] K. M. Górski et al. *The Astrophysical Journal*, 622:759, 2005. URL http://stacks.iop.org/0004-637X/622/i=2/a=759.

[166] Rainer Beck. *AIP Conference Proceedings*, 1085:83–96, 2008. doi: 10.1063/1.3076806. URL http://aip.scitation.org/doi/abs/10.1063/1.3076806.

[167] Eli Waxman and Jordi Miralda-Escudé. *The Astrophysical Journal Letters*, 472:L89, 1996. URL http://stacks.iop.org/1538-4357/472/i=2/a=L89.

[168] M De Domenico and A Insolia. *Journal of Physics G: Nuclear and Particle Physics*, 40:015201, 2013. URL http://stacks.iop.org/0954-3899/40/i=1/a=015201.

[169] N. I. Fisher. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47:342–348, 1985. URL http://www.jstor.org/stable/2345577.

[170] Travis A. OBrien, Karthik Kashinath, Nicholas R. Cavanaugh, William D. Collins, and John P. O'Brien. *Computational Statistics & Data Analysis*, 101: 148 – 160, 2016. URL http://www.sciencedirect.com/science/article/pii/S0167947316300408.

[171] Travis A. OBrien, William D. Collins, Sara A. Rauscher, and Todd D. Ringler. *Computational Statistics & Data Analysis*, 79:222 – 234, 2014. URL http://www.sciencedirect.com/science/article/pii/S016794731400173X.

[172] Pablo Lemos, Anthony Challinor, and George Efstathiou. *Journal of Cosmology and Astroparticle Physics*, 2017:014, 2017. URL http://stacks.iop.org/1475-7516/2017/i=05/a=014.

[173] P. W. Hodge. *Publications of the Astronomical Society of the Pacific*, 95:721, 1983. URL http://stacks.iop.org/1538-3873/95/i=572/a=721.

[174] Correlation of the highest-energy cosmic rays with nearby extragalactic objects. *Science*, 318(5852):938–943, 2007. URL http://science.sciencemag.org/content/318/5852/938.

[175] Planck Collaboration. Planck 2015 results - xiii. cosmological parameters. *A&A*, 594:A13, 2016. URL https://doi.org/10.1051/0004-6361/201525830.

[176] R. D. Blandford and R. L. Znajek. Electromagnetic extraction of energy from kerr black holes. *Monthly Notices of the Royal Astronomical Society*, 179: 433–456, 1977. URL http://dx.doi.org/10.1093/mnras/179.3.433.

[177] J.-P. Lasota et al. Extracting black-hole rotational energy: The generalized penrose process. *Phys. Rev. D*, 89:024041, 2014. URL https://link.aps.org/doi/10.1103/PhysRevD.89.024041.

[178] M. Ostrowski. Acceleration of ultra-high energy cosmic ray particles in relativistic jets in extragalactic radio sources. *Astronomy and Astrophysics*, 335:134–144, 1998. URL http://adsabs.harvard.edu/abs/1998A%26A...335..134O.

[179] T.-P. Li and Y.-Q. Ma. Analysis methods for results in gamma-ray astronomy. *Astrophysical Journal*, 272:317–324, 1983. URL http://adsabs.harvard.edu/abs/1983ApJ...272..317L.

[180] Planck Collaboration et al. Planck intermediate results - xlii. large-scale galactic magnetic fields. *A&A*, 596:A103, 2016. doi: 10.1051/0004-6361/201528033. URL https://doi.org/10.1051/0004-6361/201528033.

[181] Elena Orlando and Andrew Strong. Galactic synchrotron emission with cosmic ray propagation models. *Monthly Notices of the Royal Astronomical*

Society, 436:2127–2142, 2013. doi: 10.1093/mnras/stt1718. URL http://dx.doi.org/10.1093/mnras/stt1718.

[182] Marijke Haverkorn and Steven R. Spangler. *Space Science Reviews*, 178: 483–511, 2013. URL https://doi.org/10.1007/s11214-013-0014-6.

[183] *Physics of the Dark Universe*, 4:23 – 30, 2014. doi: https://doi.org/10.1016/j.dark.2014.04.003. URL http://www.sciencedirect.com/science/article/pii/S2212686414000120.

[184] Alexander Aab et al. The Pierre Auger Observatory Upgrade - Preliminary Design Report. *arXiv*, (1604.03637), 2016. URL https://arxiv.org/abs/1604.03637.

[185] Sarazin, Fred et al. *EPJ Web of Conferences*, 53:08017, 2013. URL https://doi.org/10.1051/epjconf/20135308017.

[186] Patrick Allison and Sean Quinn. Quick question about signal calculation from adc trace. Private communication, Feb. 16 2017.

[187] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL http://www.scipy.org/. [Online; accessed Jan. 2017].

[188] I. Lhenry-Yvon and Q. Luce. *Auger GAP note*, page 044, 2016. URL https://www.auger.org/index.php/document-centre/viewdownload/139-gap-notes-2016/4256-gap2016-044.

[189] B.W. Silverman. *Density Estimation for Statistics and Data Analysis, Vol. 26, Monographs on Statistics and Applied Probability*. Chapman and Hall, London, 1986.

[190] Antoine Letessier-Selvon et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 767:41 – 49, 2014. URL http://www.sciencedirect.com/science/article/pii/S0168900214009589.

[191] P. T. Boggs and J. E. Rogers. *Proceedings of the AMS-IMS-SIAM joint summer research conference*, 112:186.

[192] S. Quinn and T. Fujii. Auger@ta station level data sharing. Private communication, May 18 2017.

[193] S. Quinn and T. Fujii. Request for april, may, june 2017 mou shared materials. Private communication, Sep. 4 2017.

[194] S. Quinn for the Pierre Auger and Telescope Array collaborations. *Proc. of the 35th Int. Cosmic Ray Conf., Busan, Korea*, page 395, 2017. URL https://pos.sissa.it/301/395/.

[195] Herbert A. Sturges. *Journal of the American Statistical Association*, 21:65–66, 1926. URL http://dx.doi.org/10.1080/01621459.1926.10502161.

[196] David P. Doane. *The American Statistician*, 30:181–183, 1976. URL http://amstat.tandfonline.com/doi/abs/10.1080/00031305.1976.10479172.

[197] David Freedman and Persi Diaconis. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57:453–476, 1981. URL https://doi.org/10.1007/BF01025868.

[198] K. H. Knuth. *ArXiv Physics e-prints*, 2006. URL https://arxiv.org/abs/physics/0605197.

[199] Jeffrey D. Scargle et al. *The Astrophysical Journal*, 764:167, 2013. URL http://stacks.iop.org/0004-637X/764/i=2/a=167.

[200] Pierre Auger Collaboration. *Phys. Rev. D*, 93:072006, 2016. URL https://link.aps.org/doi/10.1103/PhysRevD.93.072006.

[201] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. N. Capdevielle. 1998. URL https://web.ikp.kit.edu/corsika/physics_description/corsika_phys.pdf.

[202] S. Argirò et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 580: 1485–1496, 2007. doi: http://dx.doi.org/10.1016/j.nima.2007.07.010. URL http://www.sciencedirect.com/science/article/pii/S0168900207014106.

[203] T.T. Bohlen et al. *Nuclear Data Sheets*, 120:211–214, 2014. URL http://www.fluka.org/fluka.php.

[204] A. Ferrari et al. *CERN-2005-10 (2005), INFN-TC-05-11, SLAC-R-773*, 2005. URL http://www.fluka.org/fluka.php.

[205] S. Ostapchenko. *Phys. Rev. D*, 83:014018, 2011. doi: 10.1103/PhysRevD.83.014018. URL https://link.aps.org/doi/10.1103/PhysRevD.83.014018.

[206] E. Thébault et al. *Earth, Planets and Space*, 67:79, 2015. URL https://www.ngdc.noaa.gov/IAGA/vmod/igrf.html.

[207] S. Quinn, T. Fujii, and F. Sarazin. More comments. Private communication, June 4 2017.

[208] D. Schmidt for the Pierre Auger Collaboration. *Contributions to the 35th International Cosmic Ray Conference*, PoS(ICRC2017)353, 2017. URL https://pos.sissa.it/301/353/pdf.

[209] D. Schmidt. *Private communication*, May 29 2017.

[210] D. Schmidt. *Private communication*, Aug 3 2017.

[211] D. Ivanov for the Telescope Array and Pierre Auger Collaborations. *Contributions to the 35th International Cosmic Ray Conference*, PoS(ICRC2017)498, 2017. URL https://pos.sissa.it/301/498/pdf.

[212] Pierre Auger Collaboration. *Phys. Rev. Lett.*, 117:192001, 2016. URL https://link.aps.org/doi/10.1103/PhysRevLett.117.192001.

[213] Maximo Ave et al. *Proceedings of the 32nd International Cosmic Ray Conference, ICRC 2011*, 2, 2011.

[214] Markus Roth, Maximo Ave, and Alexander Schulz. *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015): The Hague, The Netherlands, July 30-August 6, 2015*, PoS(ICRC2015)378, 2015. URL http://pos.sissa.it/archive/conferences/236/378/ICRC2015_378.pdf.

[215] B.T. Stokes et al. Dethinning extensive air shower simulations. *Astroparticle Physics*, 35:759 – 766, 2012. URL http://www.sciencedirect.com/science/article/pii/S0927650512000576.

[216] DONALD E. GROOM, NIKOLAI V. MOKHOV, and SERGEI I. STRIGANOV. Muon stopping power and range tables 10 mev–100 tev. *Atomic Data and Nuclear Data Tables*, 78(2):183 – 356, 2001. URL http://www.sciencedirect.com/science/article/pii/S0092640X01908617.

[217] P. K. F. Grieder. *Cosmic Rays at Earth*. Elsevier Science, 2001.

[218] M. P. De Pascale et al. *J. Geophys. Res.*, 98:3501, 1993.

[219] R.J. Mathar. Solid angle of a rectangular plate. *Technical note*. URL http://www.mpia.de/~mathar/public/mathar20051002.pdf.