ELECTRONICS AND TIMING FOR THE AUGERPRIME UPGRADE AND CORRELATION OF STARBURST GALAXIES WITH ARRIVAL DIRECTIONS OF ULTRA HIGH ENERGY COSMIC RAYS

ROBERT HALLIDAY

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Dissertation Adviser: Prof. Corbin E. Covault

Department of Physics

CASE WESTERN RESERVE UNIVERSITY

May 2019

CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

We hereby approve the dissertation of

Robert Halliday

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

Committee Chair

Corbin E. Covault

Committee Member

Benjamin Monreal

Committee Member

John Ruhl

Committee Member

David Kazdan

Date of Defense

March 19, 2019

[*]We also certify that written approval has been obtained for any proprietary material contained therein.

Contents

List of Figures

List of Tables

To my family, past, present and future.

Acknowledgements

While only my name will appear as the author of this document, many people have supported and guided me on my way through the writing of this dissertation. There is no good way to name every person who helped me get here but to say the very least, I would like to thank my advisor, my research colleagues, the support staff in the department, my partner, my family, my friends and my teachers. I have also had the privilege of working with a number of talented CWRU undergraduates, all of whom I thank for their contributions.

Amongst those who have guided and assisted the work in this document are Corbin Covault, Sean Quinn, Robert Sobin, Tate Deskins, David Nitz, Ricardo Sato, Patrick Allison, Xavier Bertou and Sam Ehrenstein, all of whom deserve my deepest thanks for going well out of their way to help me. Finally, I thank the NIH and the CWRU Department of Biomedical Engineering for their instruction and financial support.

Electronics and Timing for the AugerPrime Upgrade and Correlation of Starburst Galaxies with Arrival Directions of Ultra High Energy Cosmic Rays

ROBERT HALLIDAY

Abstract

In this dissertation, we will describe work completed towards the Pierre Auger Observatory's AugerPrime Upgrade as well as auxiliary timing work, hardware design and finally a test of correlations of Starburst Galaxies with the arrival directions of Ultra High Energy Cosmic Rays (UHECRs). In the first three chapters, we review the history, observables and detection techniques of UHECR physics, both past and present. We then look at the future upgrade of Auger and give an in depth description of the firmware, software and hardware that make up the Upgraded Unified Board (UUB), which is to be at the heart of AugerPrime. A discussion of the scientific mechanisms and merits of event-by-event composition measurements is presented, and the necessity of a new board to support this is exposed. We then move into the precision timing implementation in AugerPrime, discussing GPS receiver selection and time-tagging system performance. We find that the timing resolution of the UUB is $\sigma_{det} = 8.44 \pm .15$ ns, and confirm it using two methods.

Subsequent to this, we discuss auxiliary timing projects which support Auger as well as the Cherenkov Telescope Array. Results are shown for an experiment to determine spatial correlations of GPS timing errors, and hardware for timing at CTA and in the Auger@TA cross calibration is described. In the final chapter of this work, we move on to examining the recent Starburst correlation result of the Auger Collaboration, and cross check this by

invoking a magnetic field model and back-tracing the arrival directions of UHECRs seen by Auger. We test to see how likely it is that the observed UHECR sky is more correlated with the observed Starburst Galaxy (SBG) sky than an isotropically chosen set of random sources. The test shows a deviation from isotropy at the 1.6σ level. Finally, we describe future directions for SBG correlation tests.

Chapter 1

Introduction

In the development of modern physics throughout the last century the science of cosmic rays, particles originating outside of Earth's atmosphere, has been a driving force. Since their first observation by Coulomb and subsequent exploration by Domenico Pacini and separately Victor Hess, the study of particles accelerated by extraterrestrial mechanisms has been a major source of physics knowledge, leading to the founding of particle physics, astrophysics, and more recently particle astronomy [1]. Cosmic rays are known to originate from a variety of sources, in particular low and medium energy cosmic rays are believed to originate from supernovae and their remnants, while the highest energy cosmic rays are still of unknown origin [2].

The study of Ultra High Energy Cosmic Rays (UHECRs) originated with experiments by Pierre Auger and his colleagues to detect extensive air showers, and was solidified by the work of John Linsley and collaborators in the experiments at Volcano Ranch [3]. The developments in the detection mechanisms and source identification that will be discussed in this document, follow along the line of experimentation propagated by Linsley, using large arrays of detectors spread over vast distances to detect extensive air showers. Throughout the latter half of the 20th century, various techniques were used to detect UHECRs includ-

ing films and electronic arrays of scintillators and water tanks. These efforts culminated in the early 2000's in the construction of the Pierre Auger Observatory.

Initially built as an array of 1660 Water Cherenkov detectors and 24 Nitrogen Fluorescence telescopes, covering 3600 km^2, approximately the size of the American state of Rhode Island, the Pierre Auger Observatory has been the flagship detector for particles above an energy of 10^{18} eV. As time progressed, the observatory evolved, integrating many efforts to improve its ability to characterize and detect extensive air showers [4]. After about a decade of operation, the possibility of upgrading the observatory's capabilities across all detectors was explored, and today the collaboration is on the threshold of executing an array-wide improvement in the electronics and detection abilities of the apparatus [5]. This upgrade and an effort to correlate the arrival directions of UHECRs with starburst galaxies will be the main focus of this work.

1.1 A brief history of Cosmic Rays and Particle Physics

In the post World War One era of physics, many efforts were made to understand and utilize the advances of Quantum Mechanics and General and Special Relativity. In 1926, Erwin Schrödinger posited his eponymous equation, describing the dynamics of wavefunctions, the fundamental description of quantum particles. Shortly thereafter, in 1928, Paul A. M. Dirac proposed his equation which united the principles of Quantum Mechanics and Special Relativity [6]. With this advance, cosmic rays took center stage in the development of modern physics.

Integral to the Dirac equation, is the prediction of opposite-charge equal-mass particles for each known elementary particle. At the time, knowledge o the full complement of elementary particles was extremely limited, but at the very least this led to the prediction

of the positron. In 1933, Carl Anderson published his paper confirming the existence of a "positive electron". This was quickly followed, just 3 years later, by the discovery of the muon [7, 8]. In the years after this, cloud chambers with scintillator triggered cameras led to the discovery of many new particles, most of which were types of mesons. All of these discoveries were the direct result of the study of cosmic rays.

Meanwhile, Pierre Auger and his collaborators were working to uncover particles of extremely high initial energies causing showers of lower energy particles spread over great distances [9]. He and his team did this by setting up Geiger-Muller and later scintillation counters at increasing separations from each other and watching the rate at which coincidences were detected as a function of their separation. They tried this at multiple sites while monitoring external conditions such as barometric pressure. With this experiment, the existence of particles estimated to have energies upwards of 10^{16} eV was confirmed. While many of the advances in particle physics at this time were driven by lower energy cosmic rays, this represents the first venture towards detecting Ultra High Energy Cosmic Rays.

Post World War II

After the second world war, the technological and science funding situations conspired to provide a previously-unseen level of funding for experimentation (particularly into the mid and late '50s) [10]. The advent of the photomultiplier tube (PMT) and the application of the first computers gave cosmic ray and particle physicists new tools [11]. While versions of the photomultiplier tube existed in the 1930's, the products available in the late 40's and 50's had been refined from their predecessors. The photomultiplier tube is a device consisting of a photocathode– a thin piece of metal off of which free electrons are generated from photons through the photoelectric effect– multiple metal leaves and an anode collect

the freed charge. A large voltage is held between the anode and the cathode such that when an electron comes off of the cathode, it is pulled towards the anode. It excites many more electrons off of the metal leaves and these electrons are similarly accelerated towards the anode by the high voltage. Through this process, a photomultiplier tube takes a photon and produces a measurable current on the order of nano or micro-amperes.
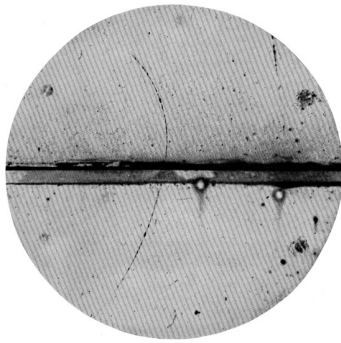
Figure 1.1: The photo published by Carl Anderson in [7] in 1932, which led to the confirmation of the positron's existence. The cosmic rays entering his cloud chamber were put under a magnetic field, hence the bending trajectory, and pictured in the center is a lead plate, meant to stop electrons.

As physics advanced, the twin fields of particle physics and cosmic ray physics began to diverge. While particle accelerators had existed since the 20's (or earlier depending on how rigorously 'particle acceleration' is defined), they were now becoming much more useful as fundamental probes of physics in ways that cosmic rays simply could not match up to [12]. Fundamentally, the difference between studying particle physics using a collider versus using cosmic rays is in the predictability of the point and time of interaction. Cosmic rays, even today, provide access to much higher energy interactions, but they occur more-or-less randomly, while interactions in a collider occur in a predetermined location and time (some like to say "in a jar"). Due to this predictability, particle physicists need to only build one detection apparatus and guarantee the interesting portions of the collision will occur inside this apparatus. Meanwhile, cosmic ray physics at this time, especially at the highest energies, began to be engineered in a distributed way, opting for dispersed detectors in the vein of Pierre Auger's experiments in the '30s. Perhaps the most notable experiment[1] pursuing this direction was the Volcano Ranch Ex-

---

[1]There were many noteworthy experiments that shaped the field, but we will concentrate on an outline of experiments to show the evolution of the field of UHECR physics, leading to the Auger Observatory.

periment, led by John Linsley [3]. His work at Volcano Ranch began with the setup of an array of scintillation counters, commonly referred to as scintillators (although this is a slight misnomer, see section 3.2.2). These scintillators were placed at uniform distances apart from each other in a triangular grid. The grid spacing changed as the experiment progressed, but to give context, in 1962 it was 884m [13]. Due to the lower speed of the electronics of his day, it was still debated whether the arrival times of the signals from each scintillation detector could be used to triangulate the arrival directions of the showers, thus increasing the grid spacing gave longer differences in arrival times and therefore made shower arrival direction reconstruction via timing a more viable approach. That said, for many of the great discoveries of Volcano Ranch, the coincidence window for detection was set such that only showers of $<10°$ zenith would be recorded.

It was in this configuration that Linsley's group detected the first 10^{20} eV comic ray. Along with a handful of other data points in this regime, Linsley was able to see the flattening of the spectrum [13], as will be discussed in the coming sections of this chapter. The effective radius of curvature of cosmic ray showers in various experiments was also noted at this time, and it was correctly assumed by Linsley and others in the field that this curvature relates to the point of first interaction in the atmosphere, and also to the position of the maximum of particle production in the shower.

Figure 1.2: A diagram of Volcano Ranch in 1962 ([13]): the dark circle represents a new detector being added for in depth shower timing studies.

Moving forward in the evolution of cosmic ray physics, we now begin to see the development of the Air Fluorescence camera/technique. At Volcano Ranch by Linsley's group, as well as the Sydney Air Shower Array detectors called

Fluorescence Telescopes (see subsection 3.2.3) were implemented. The design comes from research by the groups of Kenneth Greissen and Bruno Rossi and although it was initially unsuccessful, it ultimately resulted in the construction of the Fly's Eye detector. [14]. The fluorescence telescope is one of the inventions that led to a much deeper understanding of cosmic ray showers as one could now measure how the shower develops as it goes through the atmosphere. This allowed confirmation and further study of important quantities such as X_{max} and the elongation rate.

Before discussing the parallel development of emulsion film detectors, we must discuss the advancement of the water Cherenkov technique by the Haverah park experiment of the University of Leeds in England. This apparatus became one of the major High Energy Cosmic Ray detectors around the same time as the deployment of Volcano Ranch (although it only truly came into full operation about 5 years after both experiments were commissioned). The array initially consisted of 4 water Cherenkov detectors in a semi-triangular grid (as triangular as it can be with only 4 detectors) and was triggered in much the same way as the Volcano Ranch Experiment [15]. Data from the Haverah Park experiment remained relevant well into the 1990's ([16]), however the lasting legacy of Haverah Park is the advancement of the water Cherenkov detection method, which previously had problems with fungus and water purity [17].

In a competing lineage to electronic detectors, emulsion film detectors were developed and deployed in the late 1930's by Marietta Blau, who deployed them high into the Alps [18]. These films acted much like a film picture from a camera, except that they are exposed by either high energy ionizing radiation or X-rays (as in the analogous medical application). These films allowed cosmic ray researchers from the late thirties well into the eighties to record tracks left by showers over long exposure times [19]. While these measurements are not inherently calorimetric, the precision given by being able to visually see the tracks

6

of individual ionizing particles allows for particle counting and shower geometry methods to attempt to determine the primary energy. Film based experiments allowed researchers to directly visualize the detailed nature and progression of cosmic ray showers, allowing them to confirm and give insight into the results of other experiments which were better at collecting larger numbers of showers at higher energies.

Dawn of the Modern Era

Looking forward from the early days of Volcano Ranch and Haverah Park, the late 1960's through the 80's opened new doors for cosmic rays physics, especially in terms of accessing new messenger particles in the form of neutrinos and very high energy photons. In this vein, new detection techniques were provided by both the Whipple Telescope, the first Imaging Atmospheric Cherenkov Telescope (IACT) and the KamiokaNDE experiment, a proton decay experiment that accidentally became the first effective neutrino telescope. The Whipple telescope is based on a Davies-Cotton optics design from the mid 1950's which uses segmented mirrors to create a large aperture that is in turn used to collect the Cherenkov light directly from air showers caused by GeV to TeV gamma rays. This proved the viability of what are now frequently referred to as gamma ray telescopes and almost all such designs are derived from the Whipple design. These include VERITAS (at the same site as the Whipple Telescope), MAGIC and HESS, as well as the upcoming CTA project [14].

On the neutrino front, the KamiokaNDE I and II experiments ran through the 80's, beginning in 1983, and while they intended to detect proton decay, their detectors were sensitive enough that they were able to detect atmospheric, solar and astrophysical neutrinos. The apparatus is a 3000 metric ton pool of water, which is ideal for a proton decay experiment as it contains a high concentration of hydrogen atoms and facilitates Cherenkov

7

light production [20]. However, the detector must be able to detect and discern signals from neutrinos and cosmic rays for background detection, and was therefore able to take data on them. Famously, this experiment caught neutrinos from Supernova 1987A, which were likely the first ever astrophysical neutrinos measured and identified.

In the transition from the late 1980's to the early 1990's, the field of particle astrophysics, especially UHECR astrophysics, began to resemble its current form. Through the interest of Nobel laureate James W. Cronin, and other particle physicists, as well as X-ray astronomers, the goal of figuring out the acceleration mechanisms of UHECRs came to the front of the field. Accordingly, a number of sensitive and precisely executed experiments were built; for the sake of brevity, I will highlight two: the Akeno Giant Air Shower Array (1990-2004) and CASA-MIA (1990-1997). The Akeno Giant Air Shower Array (AGASA) was a 100 km^2 array in Japan at the Akeno Observatory consisting of surface and buried scintillation detectors and was built with the explicit intention of observing cosmic rays of 10^{17} eV or higher. Some infrastructure and detectors were already installed as early as 1984 from other ventures [21]. AGASA paved the way for the Auger Observatory, but was the direct predecessor, along with Fly's Eye, to the Telescope Array [14].

CASA-MIA, or the Chicago Air Shower Array-MIchigan Anti-coincidence muon array, set out with a slightly different objective. Instead of looking for the showers caused by hadronic primaries, the objective of CASA-MIA was to explore the arrival directions of Very High Energy (VHE) gamma rays. In the optical to low gamma ray energy regimes, only direct detection of photons is possible, usually by balloon- or space-borne detectors, but gamma rays above ∼10 GeV cause air showers much like UHECRs except scaled down and of much lower muon content and higher X_{max}. To detect these, about one thousand surface stations consisting of 4 scintillation counters (CASA) each, topped with an ancient lead (i.e. low radioactivity) sheet to encourage pair production in shower photons, and

about one thousand more buried scintillation counters (MIA) were employed in a .23 km^2 array [22].

Looking in the energy range of 10^{14} eV - 10^{16} eV, CASA-MIA detected gamma ray showers with its above ground scintillators, while rejecting hadronic showers by detecting their muonic component with the veto array. In terms of lineage, the CASA-MIA team provided much of the expertise and leadership for the construction of Auger. Functionally, the detector is more akin to the High Altitude Water Cherenkov (HAWC) observatory, a gamma ray detector that uses direct particle detection air shower techniques. Through the CASA-MIA experiment, much was learned about how to set up a distributed array and how to design effective detection logic.

UHECR Physics Today

This more or less brings us to the beginning of the current era of UHECR physics, which is dominated by the Auger Collaboration (which is the primary focus of this work), and the Telescope Array (TA). A more detailed description of both experiments will be given in subsection 3.3.1 and subsection 3.3.2, respectively, but to give context, the Pierre Auger Observatory started development in the mid 1990's and began construction in 2000 in Malargue, Mendoza Province, Argentina, at the turn of the millennium. Operation began in 2004, and has continued through to today. The TA experiment began development around the same time as Auger began operation, and was originally located at Dugway Proving grounds, in a site expanded around the Fly's Eye experiment (for the full deployment it was moved to Delta, UT). Both apparatuses measure UHECRs of 10^{18} eV+ throughout the bulk of the array, and feature low energy infills for engineering tests and to expand their sensitivities to energies as low as 10^{16} eV.

1.2   Motivation and Goals

Broadly, the work in this document aims to advance the progress towards an upgraded Pierre Auger Observatory, dubbed *AugerPrime*. Auger has successfully shown that air shower arrays can cover massive areas to obtain exposures large enough to precisely fill in the highest energy portions of the cosmic ray spectrum. Its upgrade is primarily motivated by a two part case.

First, the composition of UHECRs is a vital measurement in moving the field of charged particle astronomy forward. By understanding the composition of the highest energy cosmic rays, we can work to infer the rigidity of each particle as it propagates to the Earth. With this information, Z dependent anisotropy studies may elucidate, or help elucidate, the mystery of the origin of UHECRs.

The other component of the motivation in upgrading Auger, is a modernization effort. Inexpensive and powerful electronics have developed over the past 15 years since Auger was built, and in order to ensure the continued operation, and to take advantage of increased performance, the entire electronics package is being upgraded. The scintillation and radio detectors that are to be integrated into the upgraded surface detectors will need additional data handling and processing abilities. This electronics upgrade also supports the composition measurement goal, by adding the facilities necessary to measure added channels of incoming data as well as a much more powerful onboard processing and programmable logic system. Adding a flexible programmable logic device will allow us the modify the functionality, especially the trigger logic, in-situ.

In particular, the work in this document will concern the design of the programmable logic, the operating system software, the time-tagging logic, and the hardware and software integration of the new GPS model. Supporting these design tasks, we will also review

verification data for the time-tagging system to determine the overall time resolution of the upgraded detector and show the results of testing GPS receiver models for the upgrade. Continuing along this path, results created using branches of the AugerPrime time-tagging hardware will also be included (e.g. quantification of spatial dependence of GPS timing errors, a miniaturized in situ timing system for AugerTA, etc.).

To round out the work included, we will report preliminary results from a follow-up to Auger's recent correlation of UHECR arrival directions with Starburst Galaxies. Towards this end, we will be using software written by the High Energy Astrophysics group at CWRU to execute and manage simulations of UHECR back propagation through the JF12 magnetic field model. We will simulate a meaningful portion of the available arrival direction data, and use a χ^2-like test to infer the significance of the distribution of SBGs over an isotropic distribution.

Chapter 2

Cosmic Ray Sky: The State of the Field

In this chapter, the all particle spectrum will be reviewed via the famous Swordy plot, which will lead us through a brief discussion of its main features and their phenomenology. After discussing cosmic rays of all energies, we will move our focus to the highest energies, begin the discussion on the basic models of acceleration to the highest energies. We then discuss the Fermi mechanism of acceleration, the Hillas Criterion, and power requirements of sources. We move into a discussion of specific sources and end on a overview of some of the observable quantities of UHECRs and.

2.1 Spectrum

Cosmic radiation in the broadest sense concerns all particles which enter Earth's atmosphere from space. The energy spectrum of cosmic radiation (or cosmic rays) as measured spans about 12 orders of magnitude in energy and many more in flux. At the lowest end of the spectrum, the sources are relatively clear from power concerns; in particular they are believed to be injected by supernovae and supernovae remnants [2]. Detectors of lower energy (below the knee) cosmic rays, see temporal modulation of the flux based on the solar magnetic fields, an effect commonly called "solar wind". Moving up in the energy, we see two distinct breaks in the all-particle spectrum (pictured in Figure 2.1). At slightly under
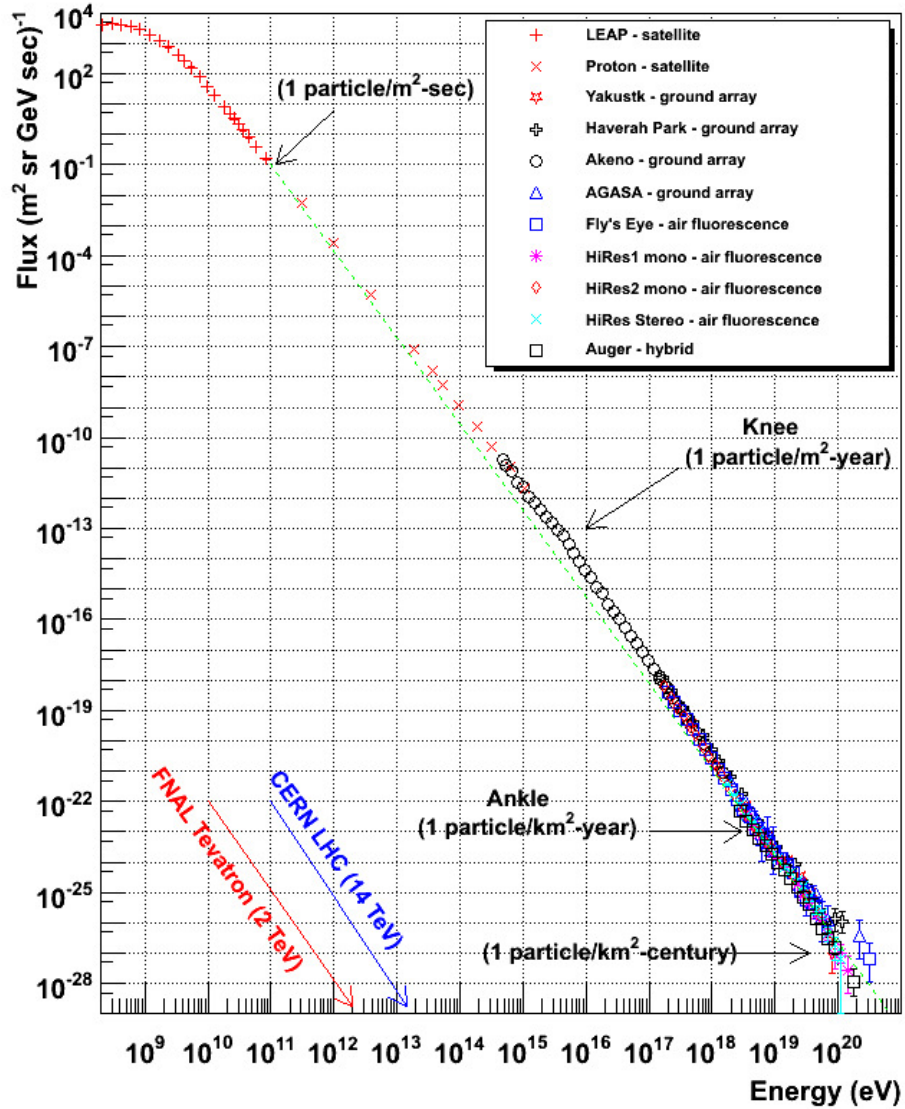
Figure 2.1: The all particle spectrum, often called the Swordy Plot [23], after Simon Swordy. This updated version is out of Hanlon (2008) [24].

13

10^{16}eV, we see a break at what is called the "knee" into a steeper (harder) spectrum, and then a softening at the "ankle".

Below the knee (lower than 10^{16}eV), it is widely agreed that supernovae and supernovae remnants make up the sources of the largest portion of cosmic rays [2, 19]. At these energies, charged particles are deflected far too much to be a reliable marker of sources, but gamma rays especially along with neutrinos and x-rays paint a fairly clear picture of supernovae remnants being the most probable source [2, 25, 26]. Around the knee but before the ankle, the sources are slightly less clear since this is above the cutoff of major gamma ray and neutrino experiment's sensitivities, but is below the sensitivities of UHECR detectors. In this regime, first and second order Fermi shock acceleration (often just Fermi acceleration, or the Fermi mechanism) provide good explanations for the shape of the spectrum and the particular physical mechanisms of acceleration, although these mechanisms are common to multiple proposed sources (see Table 2.2). Further explanation of Fermi acceleration will follow in subsection 2.3.1.

At the ankle and above, we enter the territory of Ultra High Energy Cosmic Rays, which the main body of work in this thesis concerns. It is worth mentioning at this point, the final feature in the spectrum, which is the approximate cutoff at the GZK limit of 5×10^{19}eV. The Greisen-Zatsupin-Kuzmin (GZK) limit, is a particle-physics-based upper limit on the energy of long traveling protons. Observationally, this limit is obfuscated by the measured mixed composition of cosmic ray primaries, at which point one can think about it naively as a "per proton" energy limit [27]. Sources in this energy regime become more mysterious, and are the topic of section 2.4.

2.2 Cosmic Rays at the Highest Energies

The field of Ultra High Energy Cosmic Ray physics focuses specifically on cosmic rays "above the ankle" or generally of energies >1 EeV. In this section, we will review the two basic source scenarios and discuss some of their merits in the context of observed phenomena. This section serves as an introduction to the next two sections, section 2.3 and section 2.4, where we will discuss the finer points of bottom up models.

2.2.1 Source Models: Top Down vs. Bottom Up

In the drive to understand the sources of UHECRs, there are effectively two over-arching categories of source proposals. There are those which involve the decay of supermassive relic particles, inflationary topological defects, or other Grand-Unified-Theory level particles, and those which invoke only the standard model of particle physics or less to explain acceleration via statistical or large astrophysical phenomena.

2.2.2 Top Down Models

Many top-down models are conceived as natural byproducts of exotic theories in cosmology and particle physics. In a top down theory, a particle or topological defect decays and produces a spectrum of particles. In many models these decays should release in the range of 10^3-10^7 EeV of energy [2].

Topological defects are proposed to have originated in fluctuations during phase transitions in the early universe. As the energy density of the early universe drops through inflation, models of topological defects predict it necessary to have regions where, in order to preserve causality, energy will be trapped. While these regions come in dimension 0-3, only magnetic monopoles of approximate dimension 0 and cosmic strings of dimension

1 should decay/break symmetry to produce UHECRs [2]. Some of the issues with such models include that the dynamics of the strings and monopoles become important, and that there should be modulation/suppression of the GZK cutoff. In some models, the cosmic strings and monopoles must actually meet (in the case of monopoles, it must meet its anti-particle), and given their rarity, this greatly suppresses the flux of UHECRs compared to what is observed [28]. It is also common in such models, to expect the topological defects to aggregate in the galactic halo. If this were true, then (in most models) the spectrum should not cutoff at extremely high energies, which is contrary to what is observed [28].

On the other hand, relic particles, a form of non-thermal cold dark matter, could have been formed in a freeze-out of the early universe, perhaps at the surface of last scattering, and would then wander the universe in a metastable state. At some later time, they would decay into some quantum-number-conserving mix of particles. The energies estimated here are generally around 10^{12} GeV, and the particles would be affected by gravitational pull [28, 2]. Given this, it is expected that they would agglomerate at the centers of galaxies, in which case distinguishing them from AGN lobes/flares as accelerates could prove difficult.

Most importantly for both models, is that decays from either would both be expected to produce large amounts of ultra high energy gamma rays and ultra high energy neutrinos. This is a serious problem for top down models given the current best constraints on the flux of ultra high energy gamma rays reported Auger, TA and Hi-Res [27, 26].

2.2.3 Bottom Up Models

Bottom Up models of UHECR acceleration invoke only known physical phenomena, and inside of that almost exclusively electromagnetic mechanisms. As is laid out in the recent review by Letessier-Selvon and Stanev (2011) [2], there are effectively only two physical mechanisms in this category: shock/statistical/Fermi acceleration and so-called "one-shot"

acceleration.

Shocks are outlined and discussed in subsection 2.3.1, but to review, any sort of abrupt region of fairly high magnetic field (compared to its surroundings) can act as a region of shock acceleration. Through stochastic processes, particles are accelerated by what are effectively considered point deflectors. While energy is not gained or lost in each interaction in the shock "cloud's" frame, the particle experiencing deflections can ultimately gain energy through cycles of such deflections.

One shot acceleration mechanisms (also sometimes called inductive acceleration) largely consist of "compact spinning magnetized" objects, which in effect is the long way of saying rotating young neutron stars and in particular magnetars. In this regime, an object simply produces a terrifyingly large electric field, which is not uniform by terrestrial accelerator standards, but is uniform enough, to take an entering particle and boost it up to UHECR energies in one shot (hence the name). There are also some scenarios where regions of radio-loud Active Galactic Nuclei (AGNs) can produce large enough electric fields to accelerate particles to ultra high energies, but these are oft constrained by the density of particles surrounding them (high particle density surrounding an accelerator implies likely collisions and therefore thermalization).

2.3   Bottom-Up Source Mechanisms and Constraints

This section is dedicated to discussing the Fermi mechanism, the Hillas Criterion and the power requirements of cosmic accelerators. We will begin with a discussion of Fermi acceleration, showing the general features of the model through a derivation of second order shock acceleration while discussing some of the features of first order shock accelerations. After this, we move into a discussion of the Hillas Criterion and the corresponding Hillas

plot as a constraint on bottom up UHECR source models. Finally, we will discuss the power requirements demanded by the all particle spectrum. This section serves as an outline of how power requirements work, and discusses an important injection mechanism which provides lower energy cosmic rays that may ultimately be fed into more powerful acceleration mechanisms.

2.3.1 Fermi Shock Acceleration

Fermi acceleration plays an important role in shaping the cosmic ray spectrum and provides an easy transition from below the knee to above it [19]. The critical phenomena of shock acceleration are moving regions of relatively high magnetic field which can turn around and effectively accelerate particles. In this section, we will review the first order Fermi acceleration as an example, although second order Fermi acceleration is explained in a very similar way. The two major differences are that the first order mechanism assumes a plasma cloud of magnetic deflectors and allows for decreases in energy, while the second order mechanism assumes a wall of magnetic deflectors (referred to simply as a *shock*) and always increases the energy of the incident particle.

Shock regions must be relatively sparse in terms of density as collisions in them will thermalize particles, and the Fermi mechanism is specifically an explanation for non-thermal particles. These magnetic mirrors can take a number of different forms, although they are believed to usually be the bow shocks of large objects. For example, shock acceleration can be seen in the bow shock of even such a mundane astrophysical object as the Earth [29]. In Fermi's original 1949 paper, he begins by considering a plasma cloud.

Before discussing some useful and convincing features of Fermi acceleration in the first-order, let us first go through and show the energy gain from such a scenario, following the procedure in Chapter 11 of Gaisser [19]. In the basic scenario of both types of Fermi
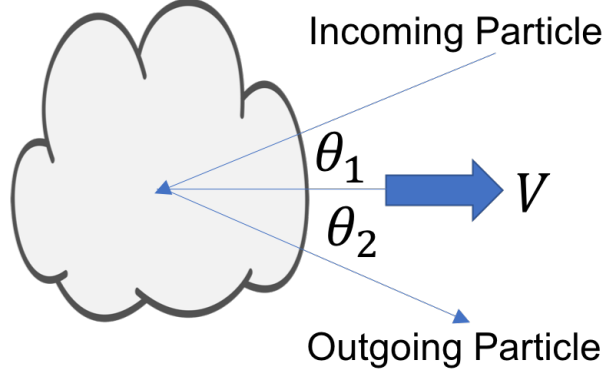
Figure 2.2: A simple diagram of the important quantities in Fermi acceleration.

acceleration, a particle comes in and is deflected in some way resulting in a net change in its velocity vector. To show this for second order Fermi acceleration, we start by looking at the energy of the incoming particle in the frame of the plasma cloud. This gives us:

$$E_1' = \gamma E_1 (1 - \beta \cos \theta_1).$$

where E_1 is the incoming particles energy in the lab frame, and γ and β are the relativistic factors as commonly used in the literature of special relativity, and θ_1 is the angle between the plasma cloud's velocity and the particle's velocity as measured in the lab frame. In what may seem like an obvious statement, magnetic fields do no work, or equivalently the magnetic field "collisions" are completely elastic, and therefore the total energy of the particle on the way out in the cloud's frame is the same as that on the way in, and so in the lab frame:

$$E_2 = \gamma E_2' (1 + \beta \cos \theta_2'),$$

where all is as above less that θ_2' is the particle's outgoing angle as measured in the frame of the cloud. Combining these two for a hyper-relativistic particle (i.e. assuming the dis-

persion relation for a photon), the change in energy can be written as:

$$\frac{\Delta E}{E} = \frac{(1 - \beta\cos\theta_1)(1 + \beta\cos\theta_2')}{1 - \beta^2} - 1.$$ (2.1)

From here, we need only average appropriately over both angles. In the case of $\cos\theta_2'$, considering the particle can exit at any angle, and in this model none are preferred, $\langle\cos\theta_2'\rangle = 0$. Phenomenologically, it is assumed that the probability of a deflection is proportional to the relative velocity between the cloud and the particle. This would come naturally by considering the ions within the cloud as little magnetic deflector dipoles. From this, the angular distribution of the deflected particle would be:

$$\frac{dn}{d\cos\theta_1} = \frac{c - V\cos\theta_1}{2c},$$

where we allow any incoming angle. This gives $\langle\cos\theta_1\rangle = -V/3c$, for a final result of:

$$\xi = \frac{1 + \frac{\beta^2}{3}}{1 - \beta^2} - 1 \approx \frac{4}{3}\beta^2.$$ (2.2)

Here, we give $\Delta E/E$ the convenient name ξ, for the fractional change in energy. This is perhaps the most basic but central result of the theory of second order Fermi acceleration. Given that β is strictly nonnegative, Equation 2.2 shows that simple interactions, which are elastic in the frame of the plasma cloud, can give an energy boost to particles in the lab frame. Here, we have made the assumption that the cloud will be non-relativistic, although the treatment can be extended to relativistic clouds. In fact, when this treatment is extended to relativistic sources, they can gain as much as a factor of γ^2 energy from their first interaction, which could be very large [29]. It is also common to do this exercise for sheets of magnetic deflectors, resulting in first order Fermi acceleration.

There is nothing in particular which stops a particle from undergoing this process multiple times, even within the same region, in which case we have the relation:

$$E_n = E_i(1 + \xi)^n,$$

with E_n the energy after the nth cycle, and E_i the initial energy. From here, we can go a number of ways. By assuming a time per acceleration cycle, we can begin to consider how the age of the plasma cloud affects the maximum energy of a particle accelerated by it. Furthermore, it is hardly a stretch to invoke a probability per unit time of re-acceleration, and from this a power law spectrum comes out.

2.3.2 Hillas Criterion

Examining the relationship between a source's ability to confine particles and the highest energy particles it can produce gives us the Hillas Criterion. First published in 1984, Hillas looked at the Larmor radius, i.e. the radius in which a particle of constant velocity is confined in a magnetic field, and considered the constraints on astrophysical accelerators that this gives [30]. This resulted in the Hillas plot or diagram, which has become a standard mechanism for displaying accelerators in the UHECR field [2]. Following the original treatment in his paper, the Larmor radius for a hyper-relativistic particle is:

$$r_L = 1.08 \times \frac{1}{Z} \left(\frac{E}{10^{15} \text{ eV}} \right) \left(\frac{B}{1 \, \mu\text{G}} \right) \text{ pc},$$

where r_L is the Larmor radius, Z is the atomic number, E is the energy of the cosmic ray in question and B is the magnetic field of the potential accelerator. We can put this expression in terms of the energy and account for the net movement of the magnetic centers by adding

a factor of β, representing their velocity. Taking the next step, we arrive at:

$$\left(\frac{B}{1\ \mu\mathrm{G}}\right)\left(\frac{L}{1\ \mathrm{pc}}\right) > \frac{2}{Z\beta}\left(\frac{E}{10^{15}\ \mathrm{eV}}\right),$$

in which L is the length of the accelerator, and β is the relativistic speed factor of its constituent shock producers. Finally solving this for the maximum energy, we produce what is known as the Hillas condition or criterion (which does not explicitly appear in the original paper, [30]):

$$E_{\mathrm{max}} \approx \beta Z e \left(\frac{B}{1\ \mu\mathrm{G}}\right)\left(\frac{R}{1\ \mathrm{kpc}}\right)\ \mathrm{EeV}. \tag{2.3}$$

In this particular version of it, out of [2], it is rewritten in EeV so as to be more useful in the context of UHECRs. From here, we can make a log-log plot in magnetic field versus radius and draw contours on it to represent the energy of particles produced along the contour. Such a plot is included in Figure 2.3.

Again, it is important to emphasize that this is a necessary but not sufficient criterion for a cosmic accelerator to reach ultra high energies in its output spectrum under the assumption of Fermi acceleration.

2.3.3 Power Considerations

By considering the density of cosmic rays through integrating the spectrum, we can put a constraint on their energy density in the galaxy. Let's take, for ease of calculation, Gaisser's rough estimate of an energy density of $\rho_E = 1$ eV/cm^3 [19]. We then need to know the volume of the galactic disk, which can be estimated as :

$$V_D = \pi R^2 d \approx \pi \times (15\ \mathrm{kpc})^2 \times 200\ \mathrm{pc} \approx 4 \times 10^{66}\ \mathrm{cm}^3.$$
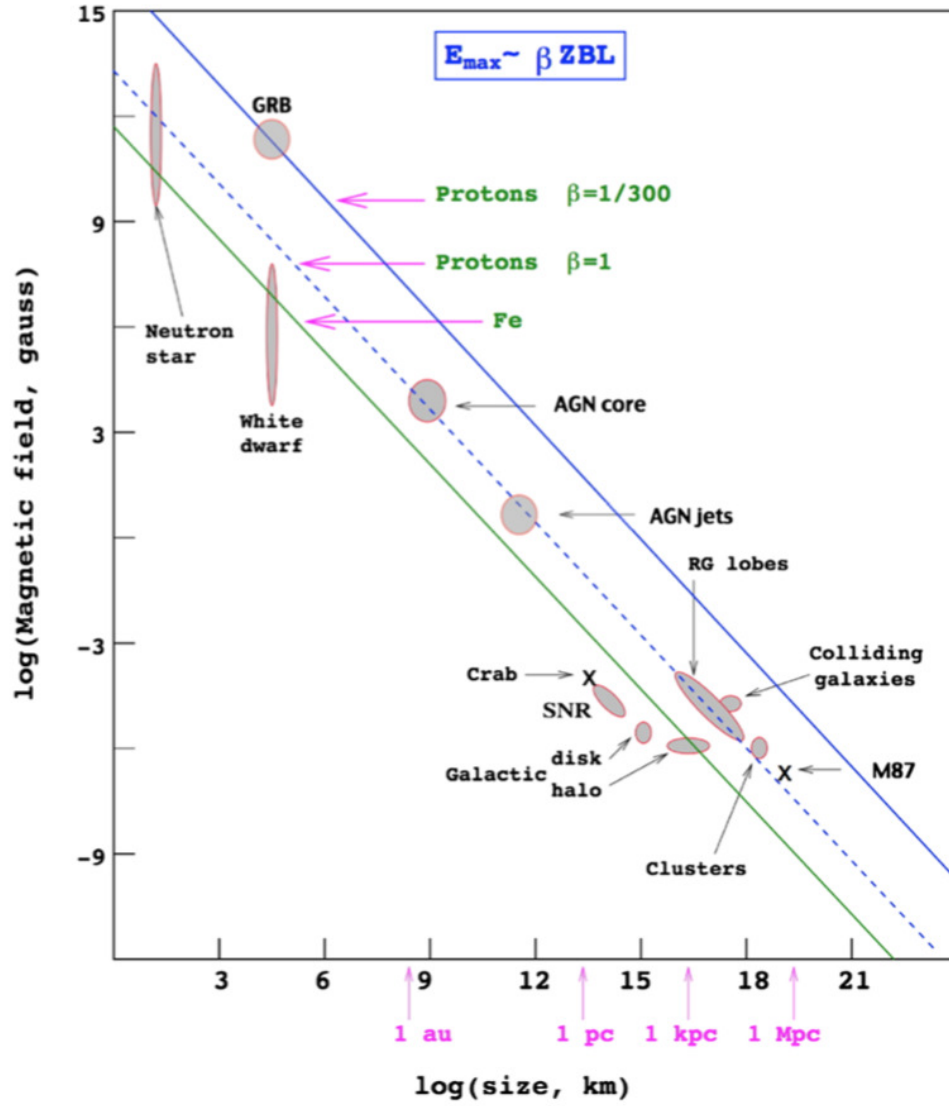
Figure 2.3: The Hillas Plot as interpreted in Letessier-Selvon and Stanev (2011) [2], showing possible UHECR acceleration candidates organized by their magnetic fields and radii. The contours drawn here represent 100 EeV and the βs are those of the diffuse shock acceleration fronts of which the accelerator would be composed.

Knowing this, we find the power by taking the total energy contained in cosmic rays, and

dividing it by the average confinement time (a result of the leaky box model), also known

as the mean residence time,

$$P_{CR} = \frac{V_D \rho_E}{\tau_R} \approx 5 \times 10^{40} \frac{\text{ergs}}{s} = 5 \times 10^{33} \text{ W}.$$

Here, we have τ_R as the mean residence time of a cosmic ray in the galaxy and I have broken the calculation out of the somewhat archaic ergs/s into watts. Continuing to follow Gaisser's treatment of the topic, a 10 solar mass ejection originating from a type II supernova with an average velocity of 5×10^8 cm/s and a frequency of 30 years gives a power of $P_{SN} = 3 \times 10^{42}$ ergs/s or 3×10^{35} Watts. With what would seem to be a fairly low efficiency, supernovae can easily power all of the observed cosmic rays below the knee. This, however, says nothing about the physical mechanisms of acceleration it simply gives a necessary but not sufficient condition. Further analysis shows that there is no obvious mechanism for supernova blast waves to accelerate cosmic rays to $>10^{18}$ eV, outside of their possible contribution to diffusive shock acceleration. Therefore, the calculation presented here pertains only to cosmic rays under approximately 100 TeV.

Looking at higher energies using the same technique, we see the requirements tabulated in Table 2.1. These are for rays of up to 10 PeV, and for higher energies, particularly those >1 EeV, a calculation beyond the scope of this work is required. That said, the point here is that power constraints can be an important requirement to rule out sources before looking for an actual physical explanation for their acceleration mechanism. This technique is commonly applied to low and mid energy cosmic rays, but can be used for UHECRs in more advanced computational models.

Energy	Power (erg/s)	Power (W)
> 100 TeV	~2×10^{39}	~2×10^{32}
> 1 PeV	~2×10^{38}	~2×10^{31}
> 10 PeV	~5×10^{37}	~5×10^{30}

Table 2.1: Here the power required by cosmic accelerators to account for portions of the spectrum with increasing energies are given. These calculations are out of [19].

2.4 Review of Specific Source Models

To simplify the wide range of source models, we have compiled Table 2.2. This table is particularly informed by the discussions in Letessier-Selvon and Stanev ([2], 2011), Berezinksy ([28], 1999) and Oikonomou ([26], 2014). A broader discussion of bottom up source models is found in subsection 2.2.3.

Source Type	Shock / Inductive	Matter Density Issues	Description	Main Issues
AGN (radio-quiet)	Shock	Yes	AGNs in general are a rather appealing possible source being relatively abundant and very energetic.	Radio Quiet AGNs rely on shocks at the end of their jets, however the matter density is too great for them to be likely UHECR producers.
AGN (radio-loud)	Both [31]	None in the extended radio lobes	Radio loud AGNs, and particularly those of type FR-II can have both shock acceleration sites and inductive acceleration sites where there is little matter density to stop acceleration.	There are fewer of these than radio-quiet AGNs, and so it should be harder for them to account for the total flux of UHECRs [32].
BL Lac Objects (subclass of Blazars)	Likely Shock, Possibly Inductive	Yes	Blazars are generally AGNs which are eating away at their accretion disks. Different configurations give different levels of variability in their emission. They are generally visible in the very high energy gamma-ray spectrum. A fairly recent paper by Murase et al. ([33]) points out that these can satisfy the Hillas Criterion but does not firmly pin down a physical mechanism of acceleration.	The two main issues are that BL Lac objects are bright because of the matter they accrete, and so there are some issues with thermalization here. Additionally, they only barely meet the Hillas Criterion to get to the highest energies.

25

Source Type	Shock / Inductive	Matter Density Issues	Description	Main Issues
Gamma Ray Bursts (GRBs)	Shock	Not significant	Gamma Ray Bursts are believed to be the product of a relativistic fireball in a super-luminous supernova. Since 1995, these have been pointed at as a tantalizing source of UHECRs, however they are surrounded in minor controversy as they were first introduced when the insufficient AGASA data came out, which indicated a flat spectrum and no cutoff at the highest energies [34, 2].	GRBs should produce large amounts of neutrinos in their interactions, since much of the relativistic fireball they are believed to consist of is leptonic. IceCube has, at this point, set constraints which make these somewhat less favorable [26].
Waves in the Interstellar Medium (ISM)/ Intergalactic Medium (IGM)	Shock	No	A variety of phenomena are capable of producing large shocks in the interstellar medium, or even in the relatively dense intergalactic medium around superclusters. Galactic mergers are considered a likely source of IGM shock waves, while supernovae and other star formation adjacent processes have been suggested to cause shocks in the ISM local to their Starburst regions [35].	Definitive proof of shocks that satisfy the Hillas Criterion has not yet been shown for the ISM in Starburst regions, although many sources say there will be qualifying shocks in the similar but less dense magnetohydrodynamic fluid of the general ISM. Murase et al. in 2008 [36] showed that the IGM can qualify for producing UHECRs around the ankle, although not at the highest energies.
Magnetars/ Pulsars	Both, but primarily inductive	No in most cases	Pulsars, a subclass of which are magnetars, are rapidly rotating neutron stars and are the primary candidates for inductive acceleration. From a theoretical standpoint, these are particularly promising as they can be shown to produce a spectrum similar to observations [37].	The viability of rotating neutron stars as UHECR sources is at the mercy of neutrino observatory data. In the near future, they could be ruled out if no high energy neutrinos are found [38].

Table 2.2: Here we have summarized some of the possible sources of UHECRs and discussed their advantages and disadvantages as models.

2.5 UHECR Observables

Here we endeavor to talk about some of the observable features of UHECRs. We will first speak generally about the spectrum and the models proposed to explain its suppression. Moving on, we discuss the composition of UHECRS, and we end the section looking at the anisotropy in their arrival directions.

2.5.1 Spectrum
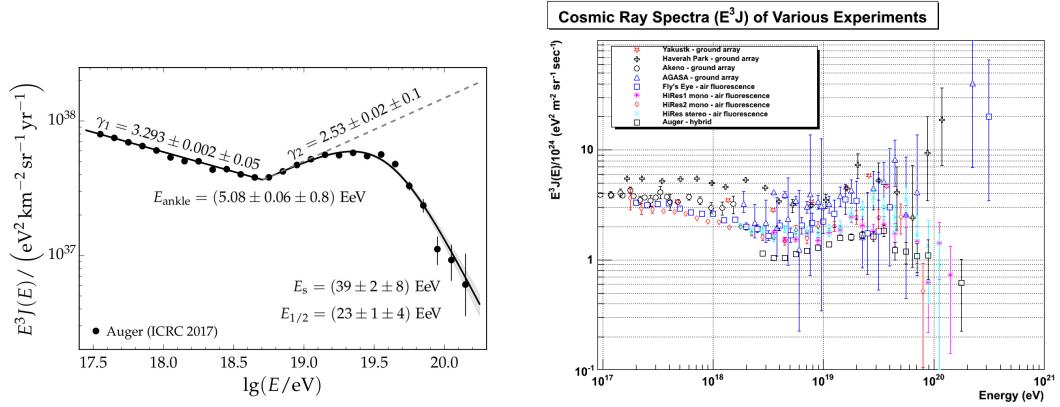
Figure 2.4: Left: The Auger high energy spectrum with cutoff model. E_s is the characteristic suppression energy and $E_{1/2}$ is the energy at which the measured flux is half of the projected power law flux (dotted line) after the suppression. The γs are spectral indices. Right: The high energy spectrum from multiple collaborations, multiplied by E^3 to allow for greater discernibility, from Hanlon (2008) [24].

In terms of their spectrum, UHECRs show a generally diminishing abundance going up to the very highest energies even compared to their expected power law spectra, as shown in Figure 2.4. At this time, there is sincere debate in the field about the nature of this diminishing abundance. One possible scenario explains this via the GZK limit and nuclear spallation, while another posits that the cosmic accelerators run out of the necessary power to reach higher energies [39, 40, 27, 41, 42, 43].

Physically, photo-disintegration in UHECRs is when the energy of a given photon

27

reaches gamma ray energies in the rest frame of the cosmic ray. Photo-disintegration applies to cosmic ray nuclei which can be broken up via the photon interactions through the Giant Dipole Resonance (GDR), while the analog for protonic primaries is the Greisen-Zatsepin-Kuz'min effect. The GZK effect is the name for the ultra-high-energy Δ^{++} resonance protons have with Cosmic Microwave Background photons creating pions and thereby shedding energy and changing direction [39, 40, 41].

The "end-of-steam" scenario starts with the observation that in both Fermi acceleration and inductive acceleration (see subsection 2.3.1 and subsection 2.2.3 respectively), we have Z dependence and that . The "end-of-steam" for a proton would then be lower than that for an iron nuclei, and if the spectrum for protons ended at, say $\sim$5 EeV, this would explain the heavier composition of UHECR primaries as discussed in subsection 2.5.2 [42, 43].

2.5.2 Composition

From a first principles perspective, the composition of cosmic ray primaries is one of the most important features of both sources and showers. However, the difficulties in determining the composition of primaries via the signal from surface detectors or even fluorescence telescopes are numerous. This will be discussed in greater detail in section 3.2. That said, by looking at the maximum of particle production, denoted X_{max}, a statistical determination can be made, in particular over many data points, to find the average composition per energy. Doing so has been one of the triumphs of the Auger Observatory, and the data as they stand are a point of contention amongst contemporary experiments. A trend worth noting in Figure 2.5 is the tendency for higher energy cosmic rays to be heavier in composition. If we look back at both the shock and inductive models of acceleration, both depend on Z, the atomic number directly, and so we expect this heavier composition at higher energies if these are correct.
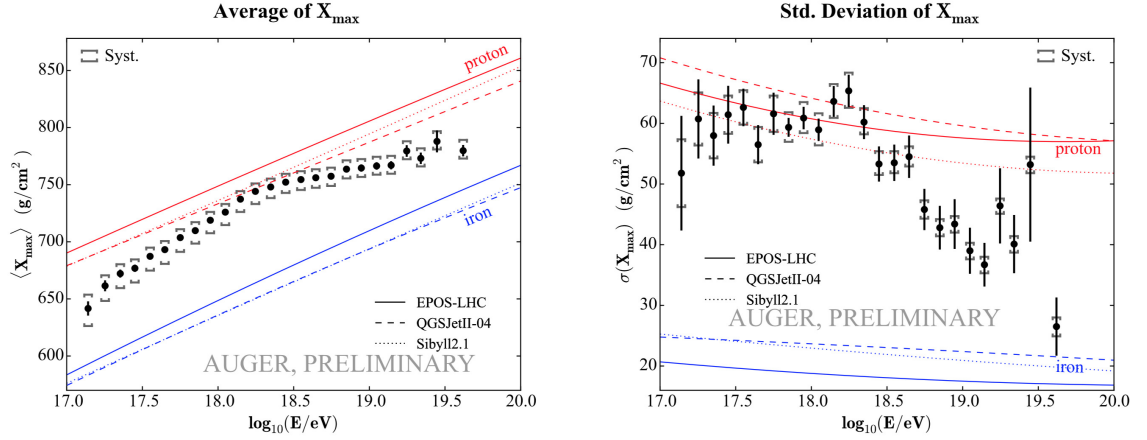
Figure 2.5: Left: the averages of $X_{\max}$ as a function of energy with contour lines drawn for their values based on 3 of the premier hadronic interaction models. Right: the standard deviation of $X_{\max}$ as a function of energy, with contours drawn similarly to those on the left. This figure is taken from an Auger Collaboration publication ([27])

One of the goals of the next generation of UHECR experiments is to better determine the composition of primaries especially at the highest energies. While Auger and TA currently attempt to find the composition on a per event basis, the errors are large enough that composition dependent anisotropy studies are not possible. Being able to determine composition on an event by event basis would be a major advantage since it gives the ability to select high energy proton events, which should be deflected less than higher Z nuclei in galactic and extragalactic magnetic fields.

2.5.3 Anisotropy
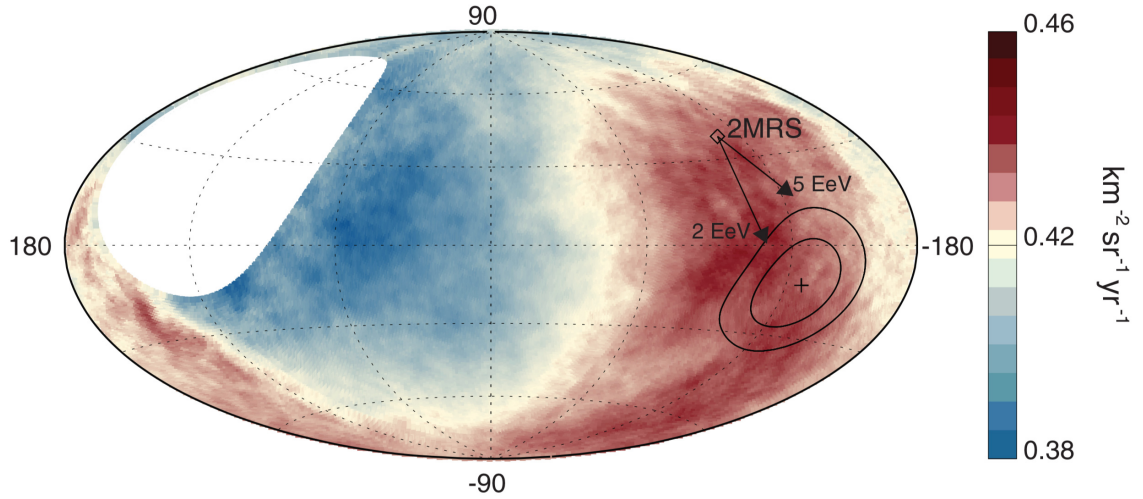
Figure 2.6: Pictured above is the sky map of UHECRs with energy > 8 EeV smoothed by convolving with a 45°top hat function. This figure is taken from an Auger Collaboration publication ([44])

Even since the first days of UHECR physics, one of the main questions has been whether the origin of UHECRs is galactic or extra galactic. In the late 90's, some conclusions were attempted with AGASA, Haverah Park, Fly's Eye and even Volcano Ranch data [28]. Recently, the Auger Collaboration has published findings in Science ([44]), which show that at the 5.2σ level, there is a dipole anisotropy in the arrival directions of UHECRs pointing away from the galactic center. This result represents the most definitive proof to date that the origin of UHECRs is extragalactic.

The analysis represented in Figure 2.6, essentially breaks down the sky along the direction of right ascension (i.e. direction of the Earth's rotation, effectively corresponding to galactic longitude) into a decomposition of cosines and sines of different frequencies. An example of how the data fits to these decomposing basis functions is given in Figure 2.7.

The important result here is that the maximum of the dipole is pointing away from the galactic center, which implies that the origin of most UHE-CRs is extragalactic. For all of the source models we have introduced, less the top-down ones, this is effectively assumed, however the extragalactic nature of UHECRs should not be taken for granted as it was a hotly debated topic for decades [30, 28].

Figure 2.7: Pictured here is the first harmonic in the anisotropy decomposition done by the Auger Collaboration in [44]

Additionally, as mentioned above, the "holy grail" of anisotropy studies would be a composition and energy dependent anisotropy study with high statistics, which is one of the probable final products of the AugerPrime upgrade.

Chapter 3

Extensive Air Showers and Cosmic Ray Detectors

Through this chapter, we give a description of some older and some more modern theories of how to quantify the important aspects of the air showers caused by Ultra High Energy Cosmic Rays. This will lead into a discussion of the various methods that have been employed to detect extensive air showers, and an overview of a few important experiments that employ them.

3.1   Extensive Air Showers

When an atomic nuclei or single proton of $10^{18}+$ eV (frequently called the primary) comes crashing into the upper atmosphere, particle physics at the very edge of humanity's current understanding ensues. After a violent disk of plasma created by slightly-beyond-the-verified-standard-model processes cools out into high energy Large Hadron Collider level interactions, the shower is carried down in phenomena quantified by a clever mix of particle physics and statistical mechanics. The first 10^{-15}-or-so seconds of this process (after the extrapolated standard model physics) are described by the basic principles of particles physics, and so the treatment in this chapter will start there. After this point, we typically become more concerned with how the particles "transport" from one species to another,

and how much energy is deposited in the nearby air. The rate at which particles are moved from pure kinetic energy of the prompt secondaries into matter in the later secondaries is a main parameter of cosmic ray showers and its maximum, X_{max}, is perhaps the best known predictor of the primary particle's composition.

After the shower undergoes its maximum of particle production, we usually become more concerned with the observable parameters on the ground, which can be seen by typical particle counting or calorimetric detectors. These observables are, in the case of the surface detector stations of Auger, derived from the signal recorded by one or more types of ground based detection. From this, a number of important observables can be extrapolated, however two of central importance are the muon-to-electron ratio (N_μ/N_e) and the lateral density function (LDF). N_μ/N_e can be derived from a single analog to digital converter (ADC) trace, or more accurately derived from multiple ADC traces from two or more types of detectors, ideally with different responses to muons and electrons, such as a scintillator and a Water Cherenkov Detector (WCD).

In the subsequent sections of this chapter, we summarize this process at each level. On the topic of particles physics, we present a treatment which discusses the steps for calculating decay widths and cross sections of the electroweak force, but does not delve into the group theoretic fundamental structure underlying the modern theories of particles physics. It should be noted, the mathematics used to describe these extensive air showers (or just "showers") was derived for lower energy primaries. After the initial higher energy processes, much of it is still entirely valid, but to tie the chapter into how this work is actually done, we include a discussion of the premier air shower simulation package, CORSIKA, which is almost always used instead of hand calculations for modern experimental design, verification and data analysis.

3.1.1 Particle Physics Basics

To start from the beginning of the story, particle physics as we know it today, stems from the 1928 equation by Dirac [6]:

$$i\slashed{\partial}\psi = m\psi. \tag{3.1}$$

Outside of i, the square root of negative one, and m, the mass of the particle, each of these objects will need an explanation. The first, $\slashed{\partial} = \gamma^\mu \partial_\mu = \gamma^\nu g_{\mu\nu} \partial^\mu$, is the four dimensional partial derivative operator contracted with a set of the "gamma matrices" (γ^μ). There are a number of bases we can put these in, but the two most common are the Dirac basis and the Weyl basis. The Dirac basis is used for most calculations, while the Weyl basis is typically used when discussing the chirality of particles, in which case it can be used to decouple the Dirac equation into two easily solved second order differential equations.

Going back to Equation 3.1, ψ is an object called a bi-spinor, and represents the particle's wavefunction. A spinor, is a vector with two complex entries which transforms under the SU(2) group. Frequently, these spinors are required to have magnitude 1, and so a bi-spinor is a four complex entry vector that mechanically looks something like:

$$\begin{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ \begin{pmatrix} c \\ d \end{pmatrix} \end{pmatrix},$$

a mathematical structure which is required by the Dirac equation. In general, we impose this normalization condition on the wavefunction ψ:

$$\bar{\psi}\psi = \psi^\dagger \gamma^0 \psi = 1.$$

34

Term: $\quad \mathscr{L}_{EW} \supset -2\sqrt{2}G_F e_L^\dagger \tilde{\sigma}^\mu \nu_{eL} e_L^\dagger \tilde{\sigma}_\mu^\dagger \nu_{eL}$

Operator Structure: $\quad e_L^\dagger \nu_{eL} * \nu_{eL}^\dagger e_L$

Figure 3.1: Here are two Feynman Diagrams and their corresponding terms in the electroweak Lagrangian density, along with the term's operator structure.

Above, the † represents the hermitian conjugate, and γ^0 is the zeroth member of Dirac's gamma matrices. In connection with Schrödinger's quantum mechanics, the entries of these bi-spinors are frequently simple plane waves and represent both the spatial and temporal behavior of the particles.

From a utilitarian perspective, these bi-spinors and the results of the Dirac equation become useful when seen through the standard model Lagrangians, Feynman diagrams, and ultimately the path integral formulation [19]. While the path integral formulation is certainly the most powerful tool for calculating cross sections and decay widths, a discussion of how to use it is outside the experimentally-focused nature of this work. Let's proceed by discussing some of the basic processes that occur in the detection of astrophysical particles from a basic particle physics view.

The Lagrangian is, at the lowest level of complexity, a statement of a classical systems total energy. When we apply this concept to particles of the standard model however, we see that it becomes a statement of the basic possible interactions available through each force [45]. For example, when we look at the Lagrangian density statement for the weak force as it pertains to interactions between electrons and neutrinos, we find a term for the charged current interactions of electron neutrinos and electrons, and its quantized operator structure like this: In Figure 3.1, the e and ν terms are bi-spinors, as discussed earlier. The

electroweak Lagrangian density, $\mathscr{L}_{EW}$, contains permutations of such interaction terms for each possible tree level (two vertex) interaction that can occur via the electroweak force. The interactions shown above involve neutrinos scattering off of electrons. These Feynman diagrams can be turned into a cross section calculation by tracing each vertex through to the end and multiplying each route together, and adding up the cross sections found by all possible topologically distinct permutations of the incoming and outgoing particles. Each vertex has a different term which we use to represent it, and this allows us to complete the calculation.

Once the expression for the diagram is assembled, we use the fact that the expression must produce a scalar quantity, and then invoke the Casimir trace theorem, which exposes the final result's sole dependence on the trace from the long expression of bi-spinors and gamma matrices. A number of these tricks are very well explained in Griffiths [46], which stands as a surprisingly good resource for many levels of students and researchers.

While walking through a calculation may be just outside the scope of this introduction, we will leave the steps summarized below:

1. Draw diagram

2. Write down scattering amplitude, $\mathcal{M}$ by tracing through each outgoing particles path backwards, including the appropriate vertices and propagators as you go

3. Multiply $\mathcal{M}$ by its Hermitian Conjugate to obtain $|\mathcal{M}|^2$

4. Insert $\gamma^0\gamma^0$ next to adjoint (†) operators to 'undagger' them

5. Commute γ matrices until you find a scalar term, i.e. $u(p_i, s)...\bar{u}(p_j, s)$

6. Permute the bi-spinors in the scalars ($u(p_i, s)$, etc.) to be next to each other, so you can sum over the spins to make the spin-summed square magnitude $\langle|\mathcal{M}|^2\rangle$

7. Take the traces and contract their results to produce dot and cross products of momenta (this can be done by FeynCalc[1] [47, 48])

8. Express the inner propagators in terms of outgoing momenta

9. Calculate kinematics, which directly translates to the arguments of the delta functions in $d\text{LIPS}_n$

10. Insert $\langle |\mathcal{M}|^2 \rangle d\text{LIPS}_n$ into the scattering amplitude

11. Calculate the integral numerically (this can be done with Mathematica).

3.1.1.1 Moving Particle Physics into Shower Physics

As we move from purely particle physics concepts into those that are applied to extensive air showers, we need to adopt a new formalism. Cosmics ray physicists studying showers, generally wrap up the particle physics aspects into various forms of inclusive cross sections, where we take a number of integral averages over quantities which are relevant to the particle physics, but less relevant to the transport of particles through the atmosphere.

As treated by Gaisser in [19], we start by effectively pulling apart the cross sections we can calculate or measure from particle and collider physics (respectively) into the quantity $\langle n_{ac}^{(b)} \rangle$, which is the mean number of particles of type c produced in an interaction between particles of type a and b. This is a function of the energy of the interaction between these particles, and from a deeper perspective, a function of the transverse momentum in the interaction, as well as the ratio of the energy deposited into the product to the energy of

---

[1] Additionally, I will note, as in the instructions above, I advocate using the FeynCalc ([47, 48]) package in Mathematica to complete these calculations if they are not for pedagogical purposes. It is certainly a rite of passage to complete a decay width or scattering cross section calculation by hand at some point, but for research purposes, modern symbolic computation system are advanced enough to save a researcher a large amount of time. The experimental values of many cross sections can be easily looked up in the Particle Data Group webpage (`http://pdg.lbl.gov`, Tanabashi et al. [49]).

the primary particle. Such averages exist for different numbers of particles interacting to create different multiplicities of products.

To begin working towards the statistical portion of extensive air showers, in the form of transport equations, we need to finish wrapping the particle physics into convenient quantities. The next step after taking cross sections and turning them into $\langle n_{ac}^{(b)} \rangle$, is to define the quantity $F_{NN}(E, E')$ as:

$$F_{ac}(E_c, E_a) \equiv E_c \frac{dn_c(E_c, E_a)}{dE_c},$$

which is the unit-less inclusive cross section for a particle of species a to interact and produce dn_c particles of species c within an energy bin the size of dE_c around an average energy deposited into the outgoing particle E_c with incident energy E_a. This quantity has already been appropriately integrated over transverse momentum. These F_{ac} cross sections are useful in a number of calculations, and it is here where it becomes important to introduce the concept of correlated and uncorrelated fluxes.

In the history of cosmic ray physics, people have been concerned with particles of energies where air showers are induced and observable, but also with energies where a detection method would never be able to tell whether two particles stemming from the same interaction were related in any way. The former, air showers, are considered a *correlated flux*, where an observer on the ground would be able to tell that the particles they are seeing come from the same primary interaction. However, many interesting phenomena and hints of fundamental physics are contained in the interactions from particles that are not high enough energy to cause correlated fluxes. In these uncorrelated fluxes, we consider the production that a portion of a species' spectrum induces. Each interaction that produces final state observable particles is not important, it is the total spectrum that we observe on

the ground that makes the observable quantity in uncorrelated fluxes.

In general, we can extend the treatment covered herein from correlated fluxes to uncorrelated fluxes by invoking the spectrum weighted moment, which effectively tells you how a spectrum of particles of a particular type produce how many particles of another type. To this end, the transport equations for correlated and uncorrelated fluxes are closely related by whether or not we treat $F_{ac}(E_c, E_a)$ via it's own contribution integral in a transport equation, or if we further wrap it into a *spectrum weighted moment*. These moments make what Gaisser calls "approximation A", which is effectively that all of the relevant quantities, but in particular the unit-less inclusive cross section, do not depend on energy. This is, of course, untrue over large energy scales, but it is easily believable over the energies at which uncorrelated fluxes are relevant.

The spectrum weighted moment is calculated as:

$$Z_{ac} = \int_0^1 (x_L)^{(\gamma-1)} F_{ac}(x_L) dx_L,$$

where $x_L = E_c/E_a$, is frequently used in the transport equations we are about to discuss, when extending them to uncorrelated fluxes.

3.1.2   Transport Equations

Once the particle physics has been appropriately handled, the production of particles through the atmosphere is determined by transport equations, which effectively describe the production of particles from each species to the others as a function of depth. For example, a basic transport equation which describes the production of pions from nucleons follows:

$$\frac{d\Pi}{dX} = -\left(\frac{1}{\lambda_\pi} + \frac{1}{d_\pi}\right)\Pi + \int_0^1 \frac{\Pi(E/x_L)F_{\pi\pi}(E_\pi, E_\pi/x_L)}{\lambda_\pi(E/x_L)}\frac{dx_L}{x_L^2} + \int_0^1 \frac{N(E/x_L)F_{N\pi}(E_\pi, E_\pi/x_L)}{\lambda_\pi(E/x_L)}\frac{dx_L}{x_L^2}.$$

(3.2)

Here, we have a number of characters, but the relevant ones for an explanation of transport equations are Π, the number of pions, and N the number of nucleons, along with the $F_{ac}(E_c, E_a)$ style inclusive cross sections [19]. While this is far too simple of a model to describe what happens in UHECR air showers, it stands to illustrate how these calculations can be handled analytically or numerically. The leftmost term on the right hand side of the equation represents the two modes in which the pion would exit the shower, through either an interaction that would effectively take it out of the shower, or that the pion would decay. The middle term represents the possibility that the pion will scatter elastically (in the particle physics sense) and remain a pion, while the right term represents nucleons colliding into anything (but probably a nitrogen molecule, i.e. other nucleons) and producing a pion.

To extend this system fully, we would write a term for each particle in the shower (after we exit the LHC-level-interactions portion) which can produce pions, and then do the same for each other species of particles. In simple cases, we can compute these integrals and complete a computation of the abundance profiles of various species through the shower. In more complex cases, and perhaps in the most general case, we would apply a numerical solver to the set of differential equations to find such profiles.

An example applied to uncorrelated fluxes, is the electron neutrino production spectrum as measured by the IceCube experiment [50]. Through this process, and the extended

process for Kaons, neutrinos are produced:

$$\pi^\pm \to \mu^\pm + \bar{\nu}_\mu / \nu_\mu$$

$$\searrow$$

$$\mu^\pm \to e^\pm + \bar{\nu}_e / \nu_e + \nu_\mu / \bar{\nu}_\mu$$

for which we can write the muon production spectrum as:

$$\mathcal{P}_\nu = \frac{\epsilon_\pi}{X \cos\theta (1 - r_\pi)} \int_{E_\nu/(1-r_\pi)}^\infty \frac{\Pi(E,X)}{E} \frac{dE}{E} + \frac{\epsilon_K}{X \cos\theta (1 - r_K)} \int_{E_\nu/(1-r_K)}^\infty \frac{K(E,X)}{E} \frac{dE}{E}.$$

(3.3)

The objects in this equation are analogous to those in Equation 3.2, except here we have a term for the production of neutrinos from Kaons. After substituting in the integrals, we can use this to compute an energy spectrum using spectrum weighted moments ([19]):

$$\frac{dN_\nu}{dE_\nu} = \frac{N_o(E_\nu)}{1 - Z_{NN}} \left(\frac{\mathcal{A}_{\pi\nu}}{1 + \mathcal{B}_{\pi\nu} \cos\theta E_\nu/\epsilon_\pi} + 0.635 \frac{\mathcal{A}_{K\nu}}{1 + \mathcal{B}_{K\nu} \cos\theta E_\nu/\epsilon_K} \right),$$

where,

$$\mathcal{A}_{i\nu} \equiv Z_{Ni} \frac{(1-r_i)^\gamma}{\gamma + 1} \text{ and } \mathcal{B}_{i\nu} \equiv \left(\frac{\gamma + 2}{\gamma + 1} \right) \left(\frac{1}{1 - r_i} \right) \left(\frac{\Lambda_i - \Lambda_N}{\Lambda_i \ln(\Lambda_i/\Lambda_N)} \right).$$

Where γ is the spectral index less one, r_i is the Molière radius of a given species and the Λ terms are interactions lengths. Frequently, a more easy to fit form is used in experimental applications, such as this one from an IceCube report on atmospheric neutrinos [50]:

$$\Phi(E_\nu) = C E_\nu^{-\alpha} (w_\pi + w_K),$$

That said, only the transport equations which do not use spectrum weighted moments,

41

i.e. those for uncorrelated fluxes, are relevant to the propagation of UHECR extensive air showers in the atmosphere (or elsewhere). While these equations were important when researchers truly started to understand the nature of extensive air showers, we should revisit the fact that today careful and vast Monte Carlo simulations are the primary tool for understanding extensive air showers.

3.1.2.1 Grammage, Interaction Length and Radiation Length

After touching on the particle and statical physics processes involved in air showers, we need to discuss how the particles in the air shower interact with the atmosphere. Before discussing specific interactions, we first introduce some important quantities dealing with height in the atmosphere and how particles transport through it.

How height is measured in UHECR physics may seem counter intuitive to the outside observer. In terms of how particles interact with the atmosphere, the density of the atmosphere is of primary importance, however, the density as a function of physical position (for example, altitude above sea level) can change based on upper atmospheric weather. To account for this, the height in the atmosphere is measured by *vertical depth* (also called grammage, a quantity originally used in papermaking), which is the integrated density of the atmosphere along a line. This quantity can be thought of as how much material one would catch if one dropped a box straight down to a particular height. Deriving further off of arguments presented in Chapter 3 of Gaisser [19], we find this relation:

$$X_v = \int_h^\infty \rho(h) \, dh$$

Now if we take a look at the US Standard Atmosphere of 1976 ([51]), which is still used today, the density profile (Figure 3.3), is not well described by any easy analytical func-

tion, and so we rely on approximations to describe it. One such approximation, from
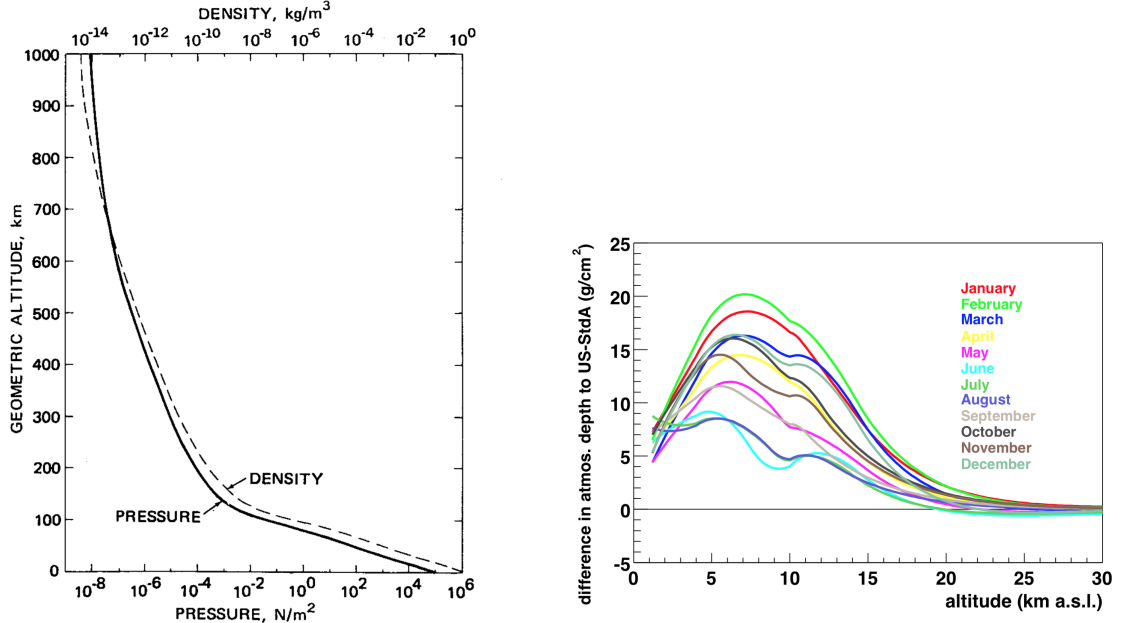
Figure 3.2: Left: Here we have the pressure and density of the standard atmosphere. This could be integrated as a function of height to find the grammage at a given layer of the atmosphere. This comes from the 1976 Standard Atmosphere [51]. Right: Differences in measured vertical depth at the Auger Observatory from the Standard Atmosphere model, from the Auger Collaboration in Keilhauer et al. [52].

Gaisser, Engel, and Resconi [53], is plotted in Figure 3.2. This figure gives the basic properties of vertical depth, namely that it monotonically increases as a cosmic ray approaches the ground. As shown in Figure 2.5, the depth of maximum particle production, the best available predictor of primary composition, is measured in grammage. Equipped with this concept, we can begin talking about interaction and decay lengths.

The interaction length takes into account the probability of a species of particle interacting with anything in the atmosphere, and represents the length at which 1-1/e of a population would have interacted. We may also define interaction lengths for different processes, such as brehmstrahlung, Compton scattering of photons, or ionization. These are

43

calculated as:

$$\lambda_j = \frac{\rho}{n_A \sigma_j^{\text{air}}} = \frac{A m_p}{\sigma_j^{\text{air}}}$$

where ρ is the density of the air, n_A is the number density of nuclei of mean mass A, and m_p is the mass of a proton. Note that much of the physically interesting part of the interaction length is contained in the cross section, which is the part we modify if we are concerned with a particular process.

The decay length of a particle that has available states to decay into, commonly muons or neutrons, is written as:

$$d_j = \rho \gamma c \tau_j,$$

with γ as the Lorentz factor of the particle, c the speed of light, and τ_j the particle's decay time. Using this idea of defining lengths over which particles can undergo certain processes, we can continue on to defining radiation parameters relevant to cosmic ray showers.
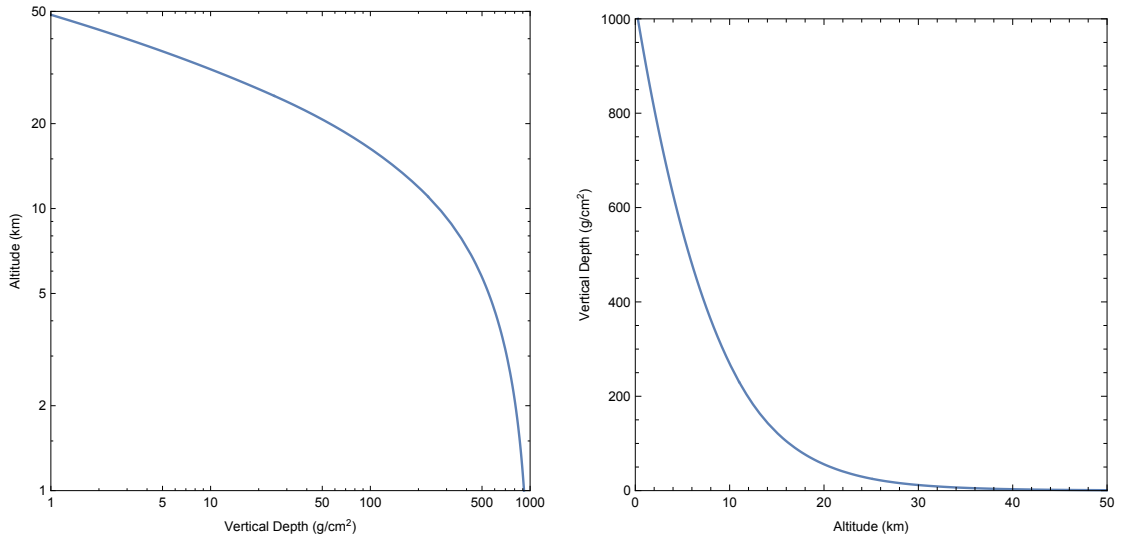
Figure 3.3: Based on the description and approximation in Gaisser, Engel, and Resconi [53], these are some simple plots of vertical depth versus altitude on logarithmic axes, and the inverse on regular axes. These are presented to give some intuitive context to the discussion.

3.1.3 Atmospheric Losses

In the energy regime that we are concerned with regarding cosmic ray secondaries ($\approx$100 MeV - TeV), there are two types of losses for charged particles that we must concern ourselves with, namely ionization and brehmstrahlung. Ionization does not scale nearly as strongly as brehmstrahlung. Brehmstrahlung is the basic interaction process of two charged particles interacting and ejecting a photon in the recoil. Cherenkov light is a much lower energy loss process. Both ionization and brehmstrahlung can be accounted for in a simple differential equation from Gaisser, Engel, and Resconi [53]:

$$\frac{dE}{dX} = -\alpha - \frac{E}{X_0},$$ (3.4)

whose general solution is:

$$E(X) = Ce^{-X/X_0} - \alpha X_0.$$

Above, α is the ionization loss in MeV/(g cm^{-2}), and X is the depth traversed. X_0 is the radiation length, which is material specific. The left term in Equation 3.4 represents ionization and approximates no energy dependence in the process while the right term represents brehmstrahlung. The additive term that arrises in the solution is the *cutoff energy*, $E_c = \alpha X_0$.

The analogous and most prominent loss for high energy photons is pair production, through which a photon becomes a positron and an electron, which continue to feed the shower while moving the mean energy per particle down. The probability of pair production in each photon is 7/9 per radiation length. At lower energies, such as medical x-rays, photoelectric losses and Compton scattering dominate [49].

3.1.4 Shower Parameters

On its way through the atmosphere, a shower goes through a number of phases, as mentioned in the introduction to this chapter. At the *point of first interaction* the particle undergoes interesting very high energy physics, which can have effects on the shower's dynamics, but on average these effects largely wash out. As the shower leaves it's first couple of interactions, the primary particle species of importance will have been determined, in part at random and in part due to the composition of the primary. The terminology used in the field of extensive air showers (EAS) to refer to the different constiutents of the shower, known as components is:

- *Electromagnetic Component* (frequently just *EM component* or *EM cascade*): the component of the shower composed of electrons, positrons and photons. Generally, by the time the shower reaches surface detectors such as those in Auger most of the positrons have annihilated in the atmosphere, leading to a net charge in the shower, an essential feature for radio detection [54].

- *Muonic Component*: from the initial hadronic interaction at the beginning of the shower, the hard quark-heavy collisions produce charged pions, whose primary decay channel is into muons and neutrinos. This makes the muonic portion of the shower an essential observable for determining the primary's composition. The muonic component effectively serves as a proxy for the number of quarks, and therefore nucleons which initiated the shower.

- *Hadronic Component*: while it is rarely a direct contributor to the final detectible signal, the hadronic component of the shower, largely composed of pions and some K_L (K-long) mesons, is important in some aspects of the shower formation before shower maximum. Interestingly, the largely unobserved hadronic component has

been shown to be perhaps the only tell tale sign of Quark-Gluon Plasma in EAS by LaHurd [55]

3.1.4.1 Heitler-Matthews Model

In the commonly used toy model by Heitler [56], and it's extensions such as [57], Particles come from the initial interaction and either pair produce or undergo brehmstrahlung at each "interaction length", which does not necessarily correspond to an actual interaction length as introduced previously. As these generations continue, they reach the cutoff energy where their losses come primarily from lower energy processes that do not contribute to the shower. This is the same E_c as in subsubsection 3.1.2.1.
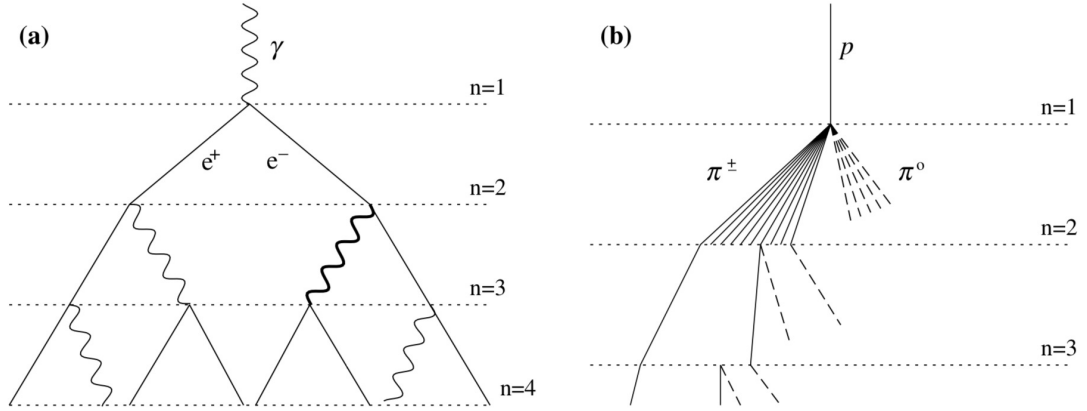
Figure 3.4: Left: a photon-initiated shower in the Heitler model. Right: the extension to hadronic processes by Matthews [57] (also the source of this diagram).

In the simple Heitler model, diagrammed in frame (a) of Figure 3.4, the maximum of particle production, X_{max}, is given by ([58]):

$$X_{max} = \lambda \ln\left(\frac{E_0}{E_c}\right).$$

The classical model does not take into account pionization from hadronic processes, how-

ever Matthews [57] extends the simple model to account for this, and is often used as a moderately more accurate account of shower physics, which is much more applicable to UHECR physics. To walk through a form from the Heitler-Matthews model [57], we start by assuming the average height of first interaction for a proton initiated shower:

$$X_0 = \lambda_I \ln 2,$$

where λ_I is the interaction length of the proton, which is a function of energy effectively proportional to $-\ln E_0$. Through arguments about how the hadronic portion of the shower feeds the EM cascade, we can write the production maximum for such a proton shower as:

$$X_{max} = X_0 + \lambda_r \ln \left(\frac{E_0}{3N_{ch}E_c} \right).$$

Here, N_{ch} is the charged pion multiplicity of hadronic production, i.e. how many charged pions are created per high energy hadronic interaction. The three in the denominator comes from the fact that, on average, 1/3 of the produced pions will be neutral, and therefore feed the EM component of the shower. This form can be conveniently rewritten as:

$$X_{max} = X_{max}^{\gamma} + X_0 - \lambda_r \ln(3N_{ch}). \tag{3.5}$$

This form isolates the pure EM shower maximum, X_{max}^{γ}, from the hadronic contributions, which necessarily decrease the depth of interaction, i.e. hadronic showers will always reach maximum production higher in the atmosphere than an equal energy photonic shower. This form is rewritten with elasticity K_{el} accounted for in Stanev [58] as:

$$X_{max} = X_0 \ln \left(\frac{2(1 - K_{el})E_0}{3N_{ch}E_c} \right).$$

The elasticity is the portion of the energy of the leading hadrons (the particles coming out of the initial very high energy interactions) that gets translated into pions. This effectively lowers X_{max} by stealing energy away from the EM component of the shower and hoarding it until after it is too late to contribute at the shower maximum.

Another important feature of EAS that the Heitler-Matthews model reveals is the *superposition model* or *superposition principle*. Effectively, it is postulated that while Quantum Chromodynamics is extremely difficult to model, we can ignore the inter-nucleon interactions of atomic nuclei primaries and view each nucleon as starting its own shower. In this way, when we consider the showers caused by a particle of atomic mass A and energy E_0, we assume, in almost every possible way, that this is the same as a shower consisting of A subshowers of energy E_0/A. The modified expression for X_{max}, again from Matthews [57], is (with everything else as in Equation 3.5):

$$X_{max}^A = X_{max} - \lambda_r \ln A.$$

This tells us that, as it is in the case of photons to protons, the shower maximum of nuclei should be strictly lower (i.e. higher altitude) than protons, up to statistical fluctuations. A final quantity of historical and intuitive (although perhaps not mechanical) importance to the topic of EAS, is the *elongation rate*. Early in the study of EAS, Linsley introduced this quantity as the base-10 logarithmic derivative of X_{max} with respect to energy [2, 53, 57]. Generally, this quantity is used in today's literature to give an intuitive understanding of how much more penetrating a higher energy primary is compared to a lower energy primary, but in the beginning of EAS studies, it was considered to be one of the few quantifiable observable quantities.

3.1.4.2 Shower Universality

In recent years, the ideas behind *shower universality* have become increasingly common and important. Shower Universality dictates that the *age parameter*, s (a function of vertical depth in the shower), and the lateral distance from the core are the only relevant parameters for the EM component of a shower, up to scaling for energy. It is, in some sense an approximation, and in another a statement, that showers ultimately end up having the same production profiles as a function of depth. The definition of the age parameter, which will also show up in the discussion about LDFs (subsubsection 3.1.4.3), is [53]:

$$s = \frac{3}{1 + 2\frac{X_{max}}{X}}. \tag{3.6}$$

Shower Universality is motivated by the observation that shower profiles, in particular electron and photon spectra, are approximately the same at X_{max} for any shower [59]. This is shaped by the processes of production and absorption in each spectra reaching equilibrium (as they must for it to be a maximum of particle production).

There is some relatively widespread misunderstanding in the community about the extent to which shower universality is a fact, and to what extent it is a convenient approximation. Of course, it is a mathematical fact that in order to have a maximum of particle production, the competing processes of particle creation and destruction/annihilation (and ultimately removal from the shower), must reach equality at some point [53]. Since the spectra of EM cascades are shaped by absorption and multiplication, it is in this way shower universality is a fact. That, however, does not guarantee that fluctuations induced by longer-lived hadronic processes can't shape the profile somewhat, and that universality completely neglects the production of muons, a tell-tale sign of primary composition. In these ways, it is a useful approximation, which allows for simple but deep analyses in many

situations [59].

3.1.4.3   Lateral Distribution Functions

Some detection methods, especially those based on particle detection and counting like water Cherenkov detectors and scintillators only record a shower's footprint on the ground, while others, namely air fluorescence and air Cherenkov give lateral profiles of the shower's development [54]. That said, particle detection methods tend to have very desirable operational qualities, especially near 100% uptime, lower failure rates, and a lack of moving parts.

To generally describe the lateral distribution functions (LDFs) used in UHECR physics, they tend to depend on s, the age parameter (Equation 3.6), $x = r/r_i$, the distance in Molière radii and some constants, usually related to fluctuations or the total number of particles. A rather complex pair of parameterizations for the electronic and muonic components of the shower are given in Gaisser, Engel, and Resconi (2016) [53], based on work in the 60's by Kenneth Greisen amongst others. These are:

$$
\begin{aligned}
\rho_{e^\pm}(s, x) &= C_1(s) x^{s-2} \left(1 + x\right)^{s-4.5} \left(1 + C_2 x^d\right), \\
\rho_\mu(r) &= \frac{\Gamma(2.5)}{2\pi\Gamma(1.25)^2} \left(\frac{1}{320}\right)^{1.25} N_\mu r^{-.75} \left(1 + \frac{r}{320}\right)^{-2.5}.
\end{aligned}
\tag{3.7}
$$

The derivation for the muonic component shown above, assumes that no muons are lost in the shower. Counter intuitively, however, in order to adjust this LDF for different heights of the shower, we adjust N_μ using any number of empirical formulae. Additionally, it is worth noting that the muon LDF cannot, in this form, be adjusted for differently inclined showers. Frequently, before analysis is done using an LDF like this, the shower is effectively "rotated" so it can be considered vertical. The reconstruction used by the Auger Collaboration

uses the "Auger LDF", which is given by:

$$S(r) = \tilde{C} \left(\frac{r}{r_s} \right)^{-\beta} \left(1 + \frac{r}{r_s} \right)^{-\beta},$$

where r_s is a characteristic distance used (500m or 1000m in most applications), $\tilde{C}$ is the effective normalization constant and $\beta = a + b \sec \theta$ is a shape factor, which can be used to characterize fluctuations and directional arrival of the shower.

Such LDFs exist in numerous forms and for numerous applications. One interesting aside, is the current development of Radio detection LDFs. Since radio detection is a relatively young field, LDFs motivated from a theoretical perspective have not been widely developed. In Schröder (2017) [54], an in depth discussion of this is given. Futhermore, Haungs, Rebel, and Roth (2003) [60] gives a rather excellent account of various LDFs, their parameters, and many other aspects of detection and detector engineering for UHECRs.

3.1.5   Air Shower Monte Carlo Methods: CORSIKA

At time-of-writing, almost all of the aspects of shower and particle physics mentioned in this chapter are, at the professional level, handled by Monte Carlo simulations, and the most consummate package to do so is called CORSIKA [61]. This package is used for showers from the low TeV range to the high EeV range and effectively bypasses the need for statistical accounting of species transport by invoking detailed models of particle physics at each step. CORSIKA uniquely takes into account certain aspects of the atmospheric transport and kinematics, but outside of these, relies on invoking the best model of interactions at each level. These include EPOS-LHC, Sybil and QGSJET at the absolute highest energies, down through GHEISHA and FLUKA, and then uses GEANT for its modeling of interactions with matter.

Studies for most major detectors in VHE and UHE astrophysics have been done with CORSIKA. Going into the mid 2000's, CORSIKA cam to dominate the field and is now the gold standard for EAS simulations. Notably, it was used in validation for Auger (Risse, Heck, and Knapp (2001) [62]) and design work for CTA (Bernlöhr (2008) [63]).
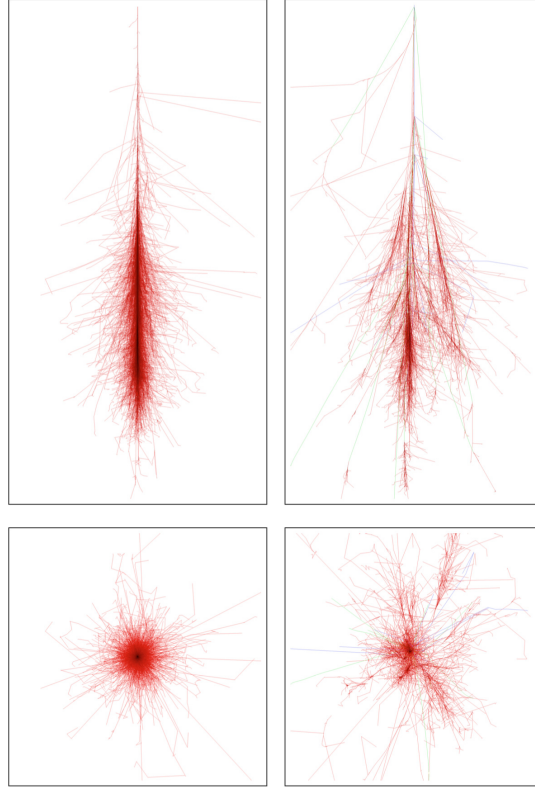
Figure 3.5: Left: A gamma ray shower of 100 GeV, longitudinal profile above, lateral profile below. Right: A proton shower of the same energy and same cross sections in each panel. Note the spread of the hadronic shower is much larger, and if you zoom in more closely, you can see the blue hadron and green muon tracks in addition to the red EM component. These images were published in Steinke (2012) [64], but are originally from Fabian Schmidt.

3.2 Detection Methods

Over the years since Pierre Auger began studying strange torrents of particles from the sky, many different techniques have been used to detect and identify extensive air showers with ground based techniques. Some of the first versions of these are covered in section 1.1, while the section below will exclusively cover techniques used successfully by modern experiments. We start by discussing the techniques of greatest importance to the work done for this dissertation: water Cherenkov and scintillation detection. The discussion will then move into the fluorescence technique for air shower detection, given its importance to surface detector energy estimation. We will then discuss air and ice Cherenkov to give context to multi-messenger neutrino detections, to better methods of air shower composition measurements and to set the stage for a discussion of Cherenkov telescopes. Finally, we will discuss radio detection methods, since the AugerPrime upgrade will be adding radio detectors to each station. There are a number of techniques currently under development, including radar imaging of the plasma disk left after the first interaction and molecular fluorescence, but for the sake of brevity, discussions of these will be left out of this work.

3.2.1 Water Cherenkov

Detection of the Cherenkov light made by particles traversing water at speeds greater than the speed of light in the medium is a technique only a couple of decades older than particle physics itself. The first use of Cherenkov light for detecting extensive air showers was by the team led by N.A. Porter at the Harwell Laboratory in England as part of the Atomic Energy Research Establishment (one of the many advances that came out of the greatly increased research funding of the early '50s [10, 65]). He and his team were the first to figure out how to stop bacterial and fungal growth long enough to create a stable and viable

detector [65].

This technique allowed for measurement of a direct proxy of the showers total energy at a point making it a very favorable technique, not to mention that water is much cheaper than a comparable volume of scintillator, and lends itself to detectors with a wide aperture for many zenith angles. It was also N.A. Porter who figured out the fruitful application of highly reflective liners to keep a maximal amount of light in the detection volume, where it might scatter into the photomultiplier tube for detection.

The Haverah Park team used Water Cherenkov detectors in their first proper air shower measurement application, which then served as a stepping stone on the way to deploying the Auger Observatory. As documented in Watson (2011) [65], many of the techniques that Porter found effective for Water Cherenkov Detectors (WCDs) were pivotal to the decision to use them in Auger.

Physically, Cherenkov light arrises from the interactions of fast particles with their medium. As mentioned above, the naive explanation for Cherenkov light, is that it occurs when a charged particle is traveling faster than the speed of light in the medium it is traversing. Many texts and educators like to make the analogy that it is like the shockwave which causes sonic booms in air, and although this analogy is useful, it is truly completely different dynamics which govern these physical processes.

As a charged particle moves through water at any speed, its interactions with the molecules cause slight changes in the polarization of the medium, however when the phase velocity of the light is greater than that allowed by the medium, the medium emits light in spherical shells. This light, however, destructively interferes within what would be referred to as the particles "Mach cone" and, from a naive point of view, constructively interferes in the forward facing Cherenkov cone. In depth discussion of this is given in chapter 23 of Zangwill (2013) [66] and a diagram of some of the relevant quantities is shown in Figure 3.6. The

angle of the Cherenkov cone in the laboratory frame is:

$$\cos \theta_c = \frac{1}{n\beta} \approx \frac{1}{n} \tag{3.8}$$

We make the above approximation as the particles we are discussing, even at low energies
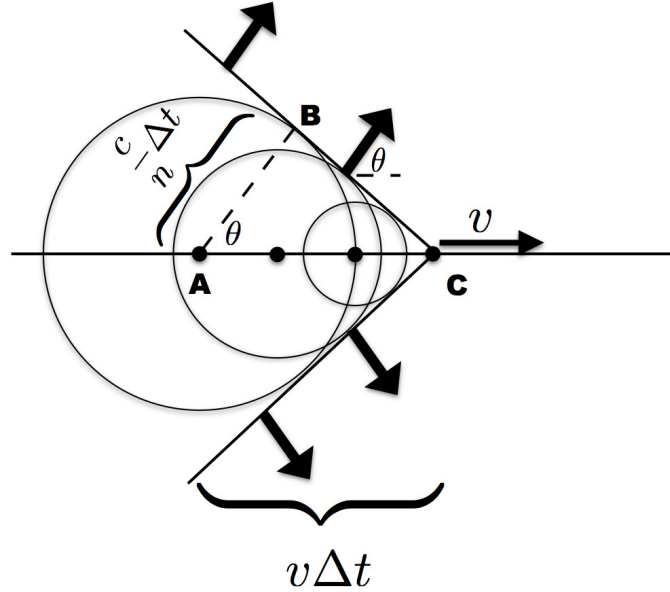
Figure 3.6: A diagram of some of the relevant quantities for Cherenkov light from [67]. Note that the length of the chord AB divided by the length of AC gives us the cosine of the Cherenkov angle. In this diagram, θ is the aforementioned θ_c.

compared to the scale of UHECR showers, are moving at very nearly the speed of light[2]. This emission angle can be very small in the laboratory frame, but luckily for the case of WCDs, this light is emitted inside the detector enclosure and reflected upwards by the liner (the air Cherenkov technique is not so lucky in this regime). This Cherenkov emission rate is constant as the particle traverses the detector and so longer traversals will leave larger signals to be seen by the photomultiplier tubes.

The Auger WCD is diagrammed in Figure 3.16 and with a radius of 1.6m on the top and

---

[2]Of course, for some phenomena how close their speed is to c is important, here it is not.

a height of 1.2m on the sides, it holds about 12 metric tons of water. The most successful technique that has been used to purify the water in Auger tanks is simply to start with relatively pure water, let the microbes and fungi grow in it, and then wait until they have exhausted their food supply. In a light tight tank, there will be no algae or otherwise to fuel an ecosystem and therefore the "gunk" will eventually die, falling to the bottom harmlessly.

Finally, before moving on to scintillation technologies, we must dwell on an important technical point around calibration of a WCD. In Auger, we use a unit called VEM or Vertical Equivalent Muons as a standard measure of the signal produced in an air shower. By making a histogram of pulse heights over many events, one can determine the minimum energy deposited by a muon, which corresponds to it going through the shortest possible path in the detector, or vertically.

3.2.2 Scintillation

Scintillation counters or detectors, frequently referred to simply as scintillators, are perhaps the most venerable and longest lasting detection technique. In early experiments, going all the way back to Anderson and Neddermeyer (1936) [8] and earlier, scintillators were used to trigger the cameras researchers used to photograph the events they were detecting. The first prominent use of scintillators for very high energy air shower detection was by Linsley at Volcano Ranch, where they were used in much they same way they currently are by a number of experiments.

Scintillators are usually made of crystals or plastics, and are luminescent, or emit visible light, particularly when exposed to ionizing charged particles [68]. The scintillators that are useful for particle detection (versus, say, medical or metrology applications) operate on the principles of fluorescence. When a charged particle excites the scintillator, the molecules in the material enter a very shortly held metastable state which is usually lower energy than the

initial excitation, and then decay releasing a photon of a wavelength outside of the materials absorption bands. This allows light created by the scintillator to travel through it without having to worry about unnecessary attenuation from processes like the photoelectric effect, which would be the dominant loss in this regime. The minimum energy of the excitation determines the Minimum Ionizing Particle or MIP energy, which is the correct analog of VEM in water Cherenkov.

While scintillators have a number of nice properties, perhaps the best of these for astrophysics is that they are compact and relatively robust, while maintaining the ability to distinguish individual particles under the right conditions. On the other hand, most scintillators degrade over time, and should not be exposed directly to ambient light if avoidable. They also, depending on their for factor, are prone to internal cracking.

Organic scintillators, which AGASA, TA and AugerPrime all use, have short nanosecond or sub-nanosecond response times to though-going particles [69, 70, 68]. They are extruded from what is, in effect, appropriately doped styrofoam [71]. Due to their extrusion process, they can be manufactured in a number of form factors. Additionally, both inorganic and organic scintillators come in liquid forms, and can be as simple as chemically simple as liquid Xenon (e.g. direct detection of dark matter experiments).

In the AugerPrime implementation, many of the same principles as water Cherenkov are used in the design of the scintillator housing. In particular, a highly reflective titanium oxide layer is used to coat the sides of the scintillator panels, which allows reflection of pulses increasing the fairly short ($\approx$10 cm) attenuation length of the material. Birefringent wavelength shifting optical fibers are used to transport the light around the edges of the panel to the photomultiplier tube. The detectors are to be called SSDs or Scintillator Surface Detectors, and will consist of 4 m^2 of aperture atop each of the 1660 stations in the array [69].
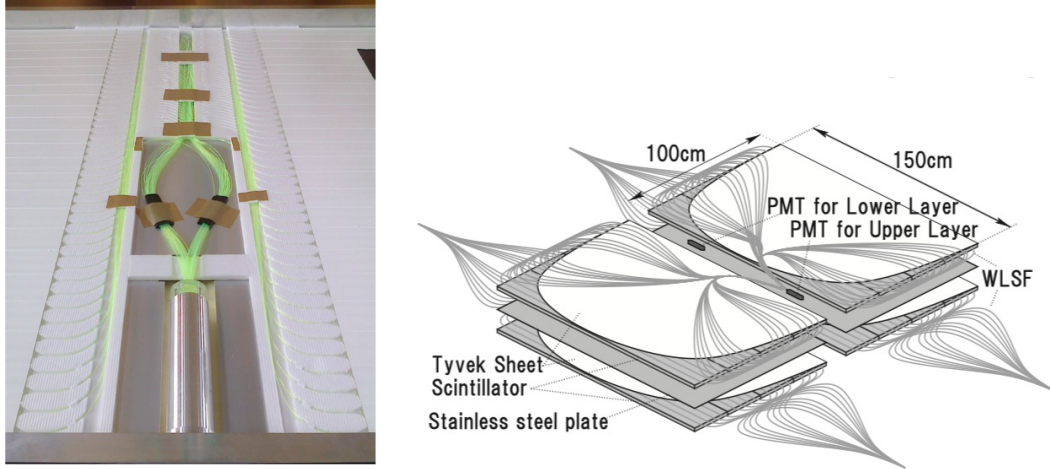
Figure 3.7: Left: A photo of the internal fiber structure and scintillator panels of the SSD for the AugerPrime Upgrade from Cataldi (2017) [72]. Right: A diagram of the two layer structure of the TA surface detector's scintillator panels, from Abu-Zayyad et al., 2012 [70].
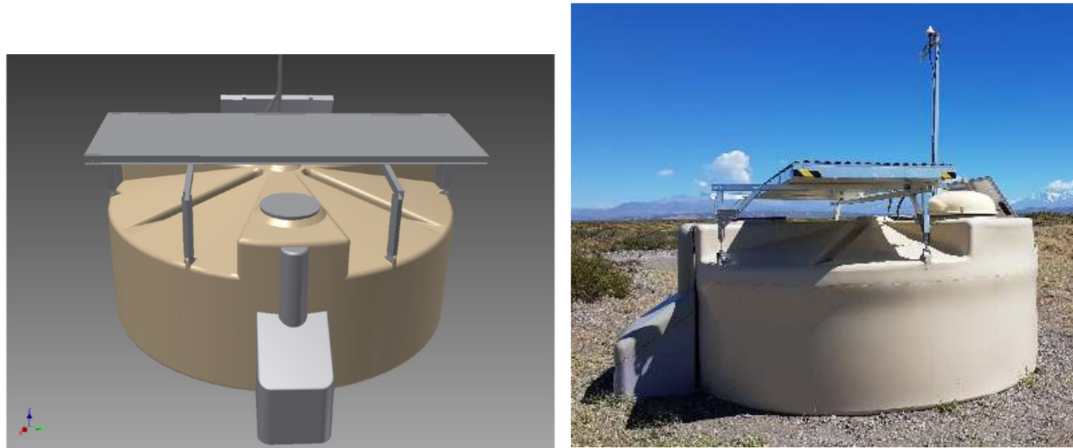
Figure 3.8: Left: a 3D mock-up of the AugerPrime station with scintillators prominently featured on top, from [72]. Right: A picture of the first SSD in the field from [69].

3.2.2.1 Comparison and Compatibility of Water Cherenkov and Scintillation Detectors

Scintillators and water Cherenkov detectors (WCDs) operate at the highest level in very similar ways. The detector volume, either the water or the plastic scintillator, is connected to a PMT (for the scintillator, it is usually mated using an optical glue or grease, and for the water it is partially submerged). While scintillators do not carry the issues of water

59

purity and PMT suspension, they are more expensive and fragile. Additionally, they are relatively two dimensional thin sheets and often require a lead layer on top to encourage pair production from the photonic portion of an incoming extensive air shower. WCDs, on the other hand, have a favorable aperture vs. zenith profile, allowing faithful detection and reconstruction of high zenith angle showers (Figure 3.9 shows a plot of the WCD from Auger and the SD from TA's apertures). In defense of scintillators, the deployment process for WCDs frequently involved trucks of purified water, flatbeds which can only carry two or three stations and sometimes even helicopters. Logistically, WCDs can be challenging and expensive.

Perhaps Figure 3.9 may appear as a relatively unfortunate fact for scintillators, but looking forward, there is a good argument to be made that radio detection and scintillation detection are complimentary techniques. As the radio detection technique advances, upward facing scintillators with high aperture to vertical showers, and radio antennae with high aperture to inclined showers (more on this in subsection 3.2.5) provide similar measurements, i.e. high EM and low muon sensitivity. In this way, perhaps the best detector one could build would utilize all three of these techniques to accurately determine composition with high aperture at all angles and relatively low cost for the performance. Modern electronics have made handling this data rate possible and the only seriously needed technology is cheap radio repeaters, which exist but have not been readily configured for use in air shower arrays.

Historically, water Cherenkov and scintillation detectors have not been used in concert. However, in considering ways to better measure the composition of UHECR primaries, the Auger Collaboration has decided to pursue the use of both detector types in tandem. This allows one to use the different muon and electron efficiencies of water and scintillator detectors to determine the number of muons and electrons that hit a station during a shower
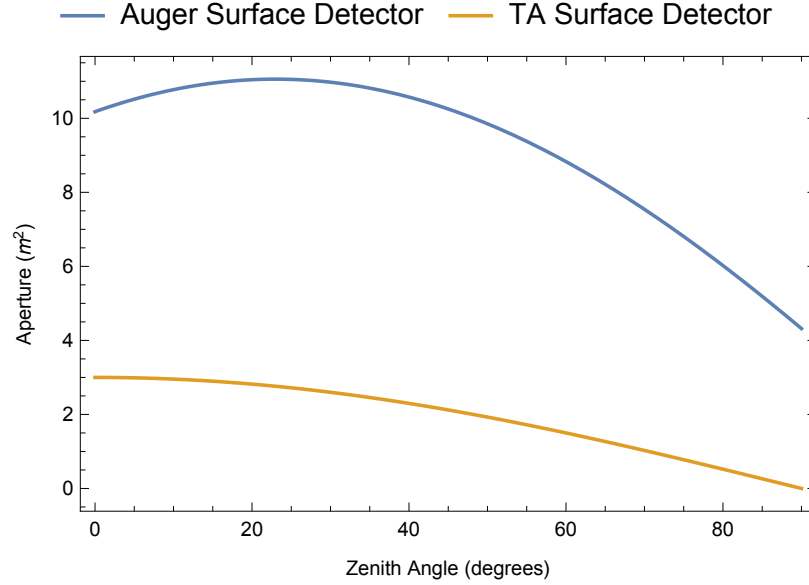
Figure 3.9: Here is a comparison of the aperture as a function of zenith angle for the Auger and TA surface stations. Water Cherenkov detectors tend to allow higher apertures for lower cost, but are logistically inconvenient.

(see section 4.1 for a more in depth discussion of this). As mentioned previously, it is this quantity, N_μ/N_e that is perhaps the best estimator of composition that is available to surface detectors.

3.2.3 Atmospheric Fluorescence

Originally pioneered by the Fly's Eye experiment (in conjunction with Volcano Ranch), the very successful Nitrogen Fluorescence technique provides perhaps the most visually relatable detection of UHECR showers. Besides Fly's Eye and its successors, Auger and TA also operate fluorescence telescopes [14, 73].

As charged particles from an air shower traverse the atmosphere, they excite the metastable states of the Nitrogen atom, which in turn emit light in the visible spectrum. The key detail of this process, and the fact which makes the fluorescence technique successful is that the emission of this light is isotropic. Because of this, an array of telescopes can be set up

to view a fiducial volume of atmosphere. Since the amount of fluorescence is effectively
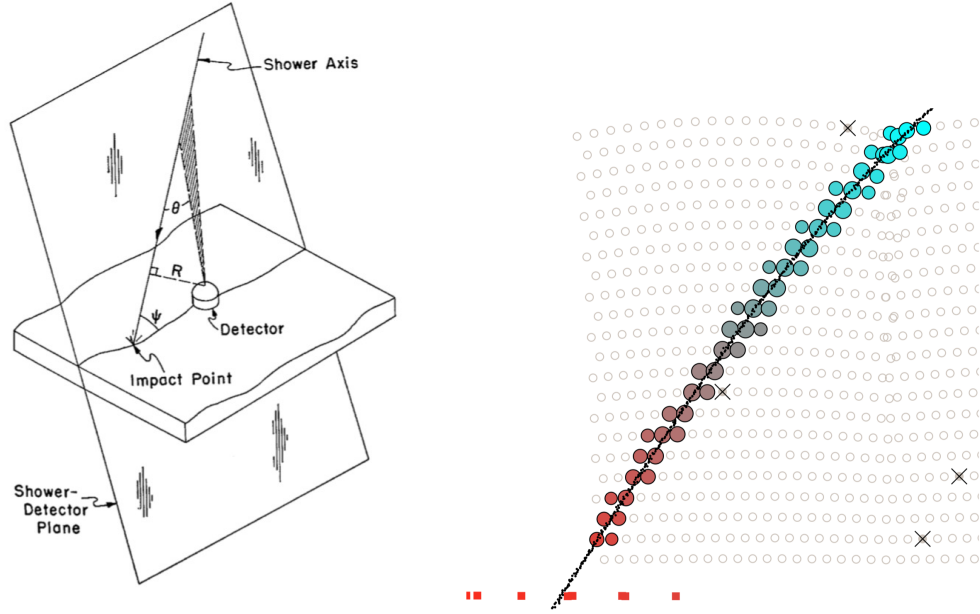
Figure 3.10: Left: A diagram of the principle at work in FD reconstruction. This is originally from the doctoral thesis of a student of Greisen named Brunner, reproduced in [14]. Right: A reconstructed track from the Auger FD, taken from [74].

proportional to the high energy part of the EM component of the shower, the fluorescence technique allows for direct observation of X_{max}. In fact this was one of the driving forces behind the construction of the Fluorescence Detector (FD) at Auger; accurate measurements of X_{max} by the FD gives the approximate composition of primaries (averaged over many showers), while the 100% uptime of the WCDs fills out the spectrum.

Technically, fluorescence telescopes work on a very basic reflector design. Usually a large mirror of some sort focuses the light onto the focal plane, which is packed photomultiplier tubes. These are almost universally in a hexagonal tiling [73]. In modern applications, many times the primary mirror of the telescope is segmented [74]. This allows for greater field of view and more light collection, while avoiding many of the complications of segmented telescopes since the PMT pixels are so large the point spread function (PSF) of

the telescope does not have to be nearly ideal. Accordingly, no adaptive optics or other corrections are needed.

A pitfall of fluorescence telescopes is that, much like optical telescopes, they can only be operated on nights with a clear moonless sky. This brings their typical duty cycle down to the range of 5-10% [73, 74, 14]. While this sounds bleak, the lack of exposure from time is made up for with the increased instrumented volume. Furthermore, the aperture of the FD array is a function of the energy of showers. In the Auger FDs, low energy showers are not even collected throughout the entire array due to insufficient light, however $10^{19}+$ eV showers are collected even outside of the array in some cases. At $10^{17.5}$ eV we have 900km^{2}sr and at 10^{19} eV, we have 7400km^{2}sr of aperture [74].
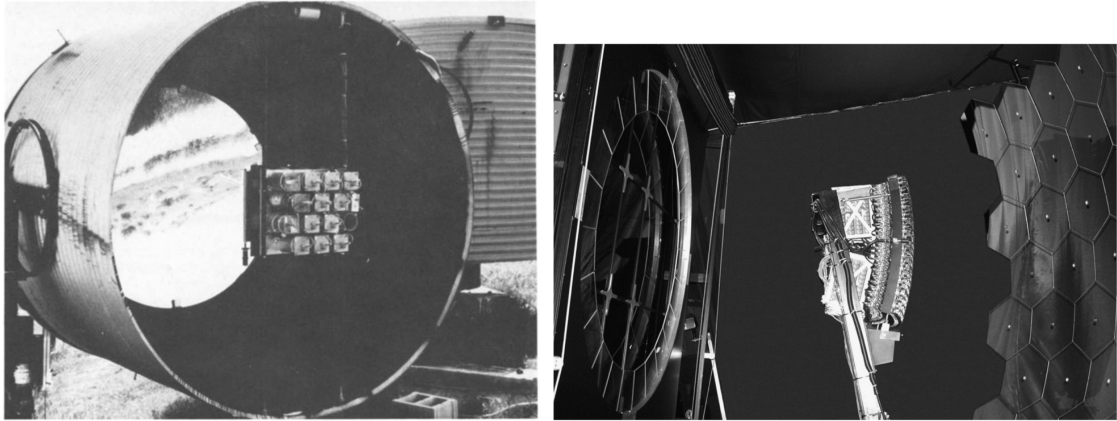
Figure 3.11: Left: A telescope from the original Fly's Eye detector, published in Baltrusaitis et al. [73]. In the center, we have the camera module contained in the stainless steel drum. Right: One of the Auger Fluorescence detectors. In the middle of the picture between the aperture on the left and mirror on the right is the honeycomb lattice of PMTs.

3.2.4 Direct and Ice Cherenkov

As we discussed in the Water Cherenkov section, Cherenkov light is emitted any time a charged particle traverses a medium faster than the local speed of light. Since the particles in an air shower are all effectively moving at the speed of light, we need not wait for the

shower to hit a relatively high index of refraction medium like water, we can detect the Cherenkov light produced from the air. We have also put Ice Cherenkov in this category since, from a hardware point of view, they are very similar techniques.

The technique was first pioneered in the same era as N.A. Porters work at the Harwell laboratory. John Jelley and Bill Galbraith went off of work by P.M.S. Blackett, a Nobel prize winning cosmic ray physicist, and looked for pulses from the night sky with an oscilloscope. Over time, their experiments became more complex and the eventually proved conclusively that they were detecting Cherenkov light produced in air. Interestingly, Blackett was convinced that we would be able to lay back on a dark night and view the Cherenkov pulses with his bare eyes [65]. The technique was later used with mild success in air shower physics.

Ultimately, the first very successful application of air Cherenkov was in Imaging Atmospheric Cherenkov Telescopes. In this configuration, gamma rays enter the atmosphere and interact at relatively high vertical depth (i.e. low altitude). They produce almost exclusively electromagnetic showers (as seen in Figure 3.5), wherein the electrons, which have a relatively low threshold for producing Cherenkov light in air, create a large and relatively bright cone which hits the telescopes optics and is then focused into an array of photomultiplier tubes. Some of the experiments that make use of this are VERITAS, MAGIC and HESS. The original Whipple observatory operated into the new millennium as well. These experiments will be superseded by the Cherenkov Telescope Array, or CTA, in the next decade.

In UHECR physics, a number of experiments have toyed with the idea of directly detecting the Cherenkov radiation from extensive air showers. The issue with this technique, and the reason it is not more widely adopted, is firstly that the Cherenkov cone from vertical showers is relatively small (<a few degrees), while the Cherenkov light from oblique
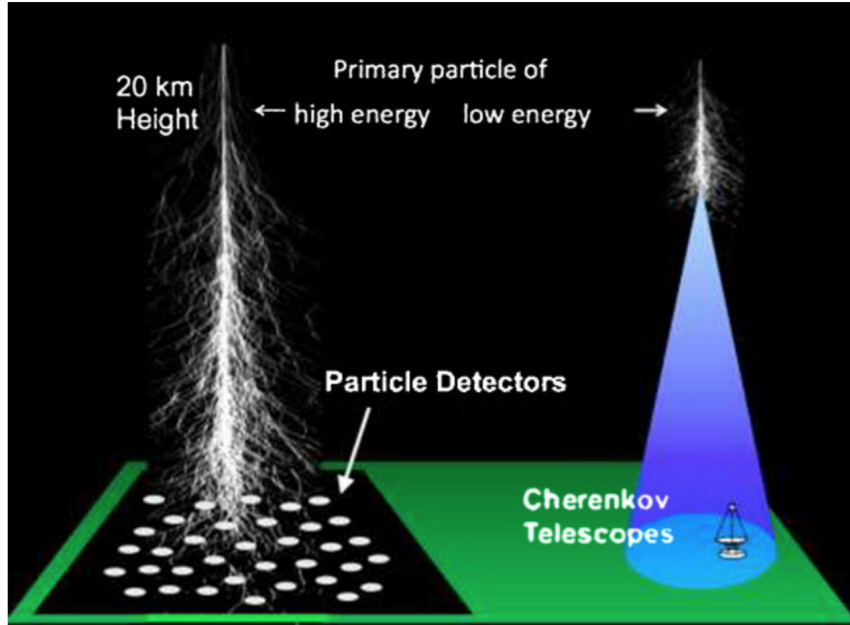
Figure 3.12: This diagram from Kampert and Watson [14] shows the basics of the IACT method of gamma ray detection in comparison to the standard method of UHECR detection. To be fair to HAWC, Milagro and CASA-MIA, there are a number of surface detector arrays for high energy gamma rays.

showers is usually too attenuated to be detected. Furthermore, this technique effectively involved exposing bare or lightly filtered photomultiplier tubes to the sky. This tends to be a fairly risky proposition since weatherproofing is non-trivial and requires moving parts to allow for remote operation. On top of all this, these sky facing PMTs can only be operated on clear moonless nights[3].

One benefit of the direct air Cherenkov technique, is that an array of this style would be able to faithfully distinguish the Z number of the primary to $\approx$5% as long as it has a small enough time resolution ($<$2 ns) [75]. This type of experiment has been deployed, although the only one currently operating is TAIGA (formerly Tunka-133). Briefly, it is an array of upward facing photomultiplier tubes that view the night sky when they can.

The form of this technique used by IceCube, which could be described as an ice Cherenkov

---

[3]Much of this discussion is informed by Schröder [54], which serves as an excellent review of all of UHECR physics, and not just radio detection.

experiment, takes strings of photomultiplier tubes with onboard electronics more than 1 km long and immerses them in the Antarctic ice. The tubes face down towards the center of the Earth (a useful quality for background rejection) since the neutrinos they are looking for are more likely to interact on their way through the Earth than through the atmosphere. In a naive sense, IceCube is an array of Tunka-133s in layers, facing downward. Since the index of refraction is higher in the ice, the particles excited and created by the neutrino interactions in the ice will produce wider Cherenkov cones and therefore it will make detection easier.

3.2.5 Radio

While the possibility of detecting air showers via radio antennae had existed at least by the 90's (it was considered as the technique for the Auger surface detectors [65]), this method of detecting UHECR showers has really come to fruition in the last decade. At time of writing, a number of experiments including LOFAR, KASCADE-Grande, Auger and IceTop (IceCube's surface detector) are adding radio components to their detectors [54]. Much of this work as it pertains to UHECRs was initially done in the Auger Engineering Radio Array (AERA) [76]. The physical basis of the technique arises from two effects, the geomagnetic effect and the Askaryan effect. As the shower progresses, the large numbers of positively charged and negatively charged electrons are separated over large distances by Earth's magnetic field. This phenomena of charge separating is the same principle as the operation of a simple Hertz dipole antenna, however it has no limit to how separated it will become. This means that a wave packet of increasing wavelength per time is created. This effect is detectable in and of itself so long as the shower is perpendicular enough to the geomagnetic field and creates linearly polarized light.

Joining in concert with the geomagnetic effect, the Askaryan effect arrises from positron
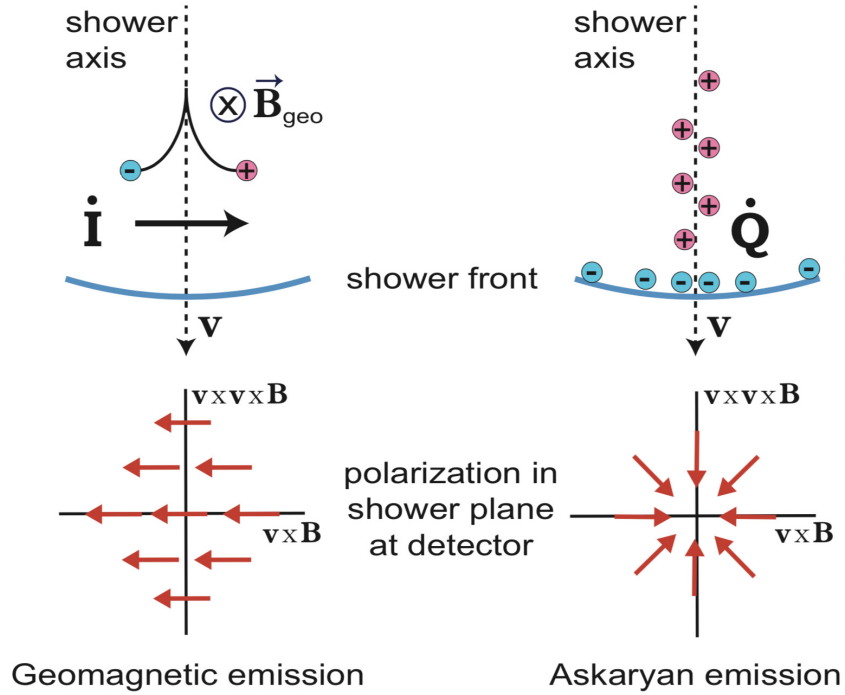
Figure 3.13: This diagram shows, in the top left, the kinematics of the geomagnetic effect and under it the polarization. On the right it shows the Askaryan effect and below it, the radial polarization it produces. Sourced from the excellent review by Schröder [54].

depletion in the shower. The essence of the Askaryan effect is positrons annihilating in the medium as the shower progresses. This leaves an excess of electrons in the shower and as the total charge of the shower changes ($\dot{Q}$), they create radially polarized waves in radio wavelengths. In solid matter, the Askaryan effect dominates dramatically over the geomagnetic effect, while in air the lower density makes it such that the Askaryan effect is still prominent but not as important as the geomagnetic effect [54]. A diagram of both effects is shown in Figure 3.13.

Similarly to air Cherenkov, radio only emits within and near the edges of the Cherenkov cone of the shower. The technique is saved in the UHECR regime (in contrast to air Cherenkov), because of the geometrically expanding cone of highly inclined showers. In other words, for inclined showers the longer time of propagation allows the Cherenkov

cone to spread, therefore making it detectable by multiple Auger stations. This, coupled with the radio-transparency of the atmosphere, has made radio detection a fairly successful technique.

3.3 Operating UHECR Experiments

Here we give an overview of currently operating experiments relevant to this work. We will discuss each of Auger, TA and CTA with importance placed on describing Auger and AugerPrime, the central foci of this work. The discussion of CTA paves the way for subsection 6.3.5, while the description of TA corresponds to subsection 6.3.4 and the description of Auger begins the work described in chapter 3.3.3.1 and chapter 4.7.3.2.

3.3.1 Pierre Auger Observatory
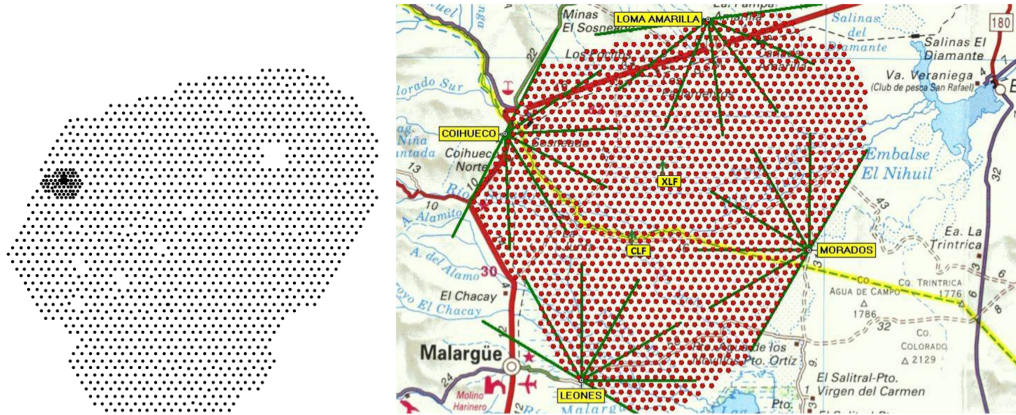
Figure 3.14: Left: A dot map of my own creation showing the positions of the surface detectors. Note that some areas are not instrumented due to terrain. This dot map makes the position of the infill and engineering arrays obvious by the bunchin towards the top left. Right: The standard mapping diagram used by the collaboration, showing the FDs and SD positions as well as local geography, from [77].

The Pierre Auger Observatory is a large UHECR observatory in Malargüe, Mendoza Province, Argentina. It consists of 1660 surface detectors and 24 Fluorescence detectors instrumenting an area of >3000 km^2 (maps presented in Figure 3.14). The fluorescence telescopes have a duty cycle of about 15%, while the surface detectors have >99% uptime, interrupted primarily by solar power outages from the rainy season. A rather large spacing of 1.5 km between stations in a hexagonal grid ensures maximum aperture in the desired

energy range of $10^{18.5}$+ eV. From the very large instrumented area and near 100% uptime of the array, the observatory has a total exposure of 40,000 km^2 sr yr [77]. This makes the observatory the largest and highest exposure in the world.

Design began in 1995 after the idea was conceived by Jim Cronin and Alan Watson. Prototyping ran until 2002 when construction officially began. This ran until 2008, when the full array was commissioned, however the array started taking data in 2004. While a number of designs for the surface detectors were considered, including water Cherenkov, scintillation, RPCs or radio detectors, ultimately the water Cherenkov technique was found to be the best in terms of cost per aperture per angle.

3.3.1.1 Auger Surface Detector

The Auger Surface Detector consists of a large tank of about 12 metric tons of water, 3.6 m in diameter and 1.2 m tall (1.2 meters is chosen as it is optimal for Cherenkov production versus attenuation), made of polyethylene with a Tyvek reflective liner. It has 3 photomultiplier tubes mounted on top, each housed in an air tight container (Figure 3.15), slightly submerged in the water thereby optically coupling them to their detection volume. The photomultiplier tubes are 9" with eight dynodes (model Photonis XP1805/D1). They have

Figure 3.15: This is a digram of the housing that the PMT is situated in to be partially immersed (i.e. optically coupled) into the water in the WCD. Figure from Collaboration [77].

their preamplifiers and high voltage supplies built into their base. Digging in to the technical details, photomultiplier tubes are sampled at two different gains by a 40 Mhz, 10 bit flash Analog-to-Digital Converter (ADC, Analog Devices model AD9203). Their output is
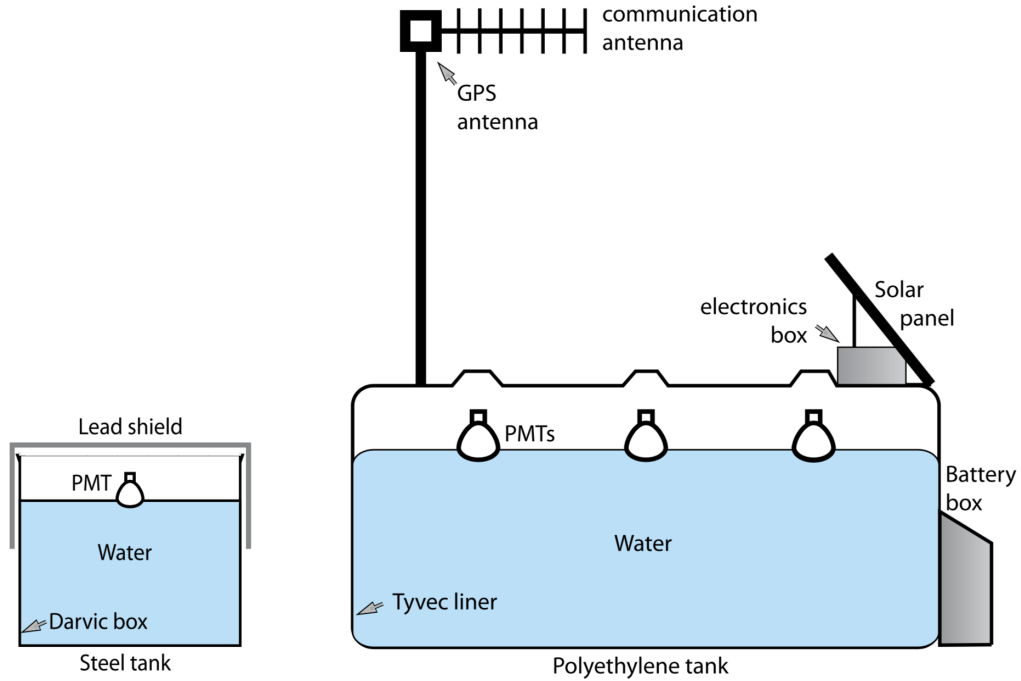
Figure 3.16: Above is a schematic diagram of the WCDs used by (left) N.A. Porter, the leader of the team that created the water Cherenkov technique, and (right) the WCD used by Auger. In some sense, the Auger WCD is the grandchild of the Porter version. From Kampert and Watson [14].

fed into either two Altera ACEX FPGAs, or one Altera Cyclone, depending on the date of construction. Later models have the single FPGA; there is also a block ram that is moved into the Cyclone in the newer version. Trigger logic is handled in the FPGA, and the appropriate outputs are made available from there to feed data into the onboard computer. The processing system is a IBM PowerPC 403 GCX at a clock rate of 80 MHz with 32MB of RAM, 2MB firmware storage.

For timing, the stations employ a custom ASIC with a 100Mhz counter which takes in the sawtooth granularity correction from the GPS units. The model chosen for the GPS timing receiver is the Motorola Oncore UT+ which features a 20Mhz clock capable of 25 ns RMS accuracy with the clock granularity message (sometimes called the sawtooth, after the shape it makes over time).

Communications is handled by a custom radio from the University of Leeds electrical engineering team which runs in the hundreds of bits per second range. Power is handled by a solar panel and lead acid batteries. Once the stations are set up, they are fairly robust and the main sources of maintenance are standard electronics errors, weather effects (in some situations, the SD is the tallest metal object in a large area, which is to say lightning strikes do destroy them sometimes), and the occasional issue with the local goats climbing onto the top of the tanks and knocking over solar panels and antennae.

3.3.1.2   Auger Fluorescence Detector

Some of the details of the FD has already been discussed in subsection 3.2.3, but I will reiterate and expand on these here. The collection of 4 sites with 6 telescopes each are usually collectively referred to as the Fluorescence Detector, although sometimes the term FDs is used to refer to the individual telescopes. Each telescope has a $30° \times 30°$ field of view, and each station with 6 telescopes then has a $180°$ field of view around it.

The telescopes are equipped with rectangular aperture mirrors, with a set of corrective lenses in front of them. Although the mirror's diameter is 3.4 m, the effective aperture is close to 1.5 m diameter due to the geometry of the housing [78], shown in Figure 3.17. Each telescope is equipped with a garage-door style shutter, which is operated automatically, but monitored at night to ensure faithful operation, which is shown open for maintenance in Figure 3.17. There is also a backup cloth shutter that can be employed in case of emergency. Most weather issues are also handled automatically, and the corrective lenses and filters at the aperture work as a window to keep the cool air from the air conditioners inside. Precautions are taken to keep the facility clean as well [77]. Scientifically, the FD works as a very important calibration component for the SD, as well as the science that it accomplishes alone. The events caught by the FD and SD simultaneously are called "golden
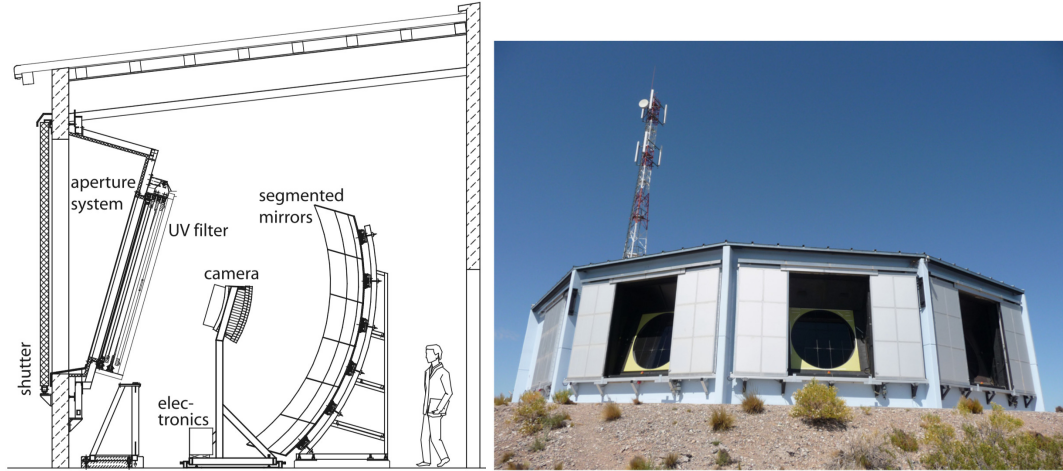
Figure 3.17: Left: A diagrammatic representation of the components of the fluorescence detector. Right: A photo of the fluorescence detectors with their shutters open for maintenance. Both from [77].

hybrids" and are used to determine the correct SD energy estimators to use. This is due to the fact that, at time of writing, the errors and energy calibration of fluorescence detectors are the best understood amongst all of the UHECR detection techniques [54]. In this way, the energy is determined from the SD and from the FD and then used to recalibrate the energy estimators used for the SD.

One rather interesting aspect of the FD's construction, is that each station (i.e. each 6 telescope bundle) is controlled by a single computer called an *EyePC*. The readout electronics for each telescope are connected via firewire to what is called an MPC. These MPCs are basically single board computers which bring the firewire data in, buffer it, and then send it across a Base-100 ethernet connection to the EyePC [77].

3.3.1.3 AugerPrime Upgrade

AugerPrime, an upgrade to Auger that is currently underway and the main focus of the original work in this document, will provide new electronics to each of the 1660 stations. This will include new ADCs, new trigger processing, a new central FPGA and processing

73

system on the same chip and a number of other upgrades. These will be discussed in depth in chapter 3.3.3.1. In finality, the upgrade will add radio and scintillation detectors to each station in addition to the improved electronics.

3.3.2   TA

Near Delta, Utah (not far from the Dugway Proving Grounds, the historic site of CASA-MIA and Fly's Eye), the former AGASA group operates the Telescope Array (TA), a detector which is similar in nature to Auger. Schematically, it has FD and SD components which can operate in hybrid mode, like auger. There are a handful of differences and we will summarize a few key points below [79].

- TA uses double-stacked scintillators, while Auger uses water Cherenkov detectors.

- Auger views the southern hemisphere's sky; TA views the northern hemisphere's sky.

- Outside of trigger processing, almost all in-situ analysis for the Auger SD is done on the processing system, while TA moves much of this to the FPGA. This gives faster event registration, but lends itself to operational opacity.

- TA employs a more rigorous data-blinding process, while Auger data are available to collaboration members on a daily basis.

- Auger does not section its surface array, meaning that all coincidences of any subset of stations will be recorded, while TA sections its array (see Figure 3.18) into three parts, making data collection easier but preventing a small number of detections on the edges of the sections.

The surface detector for TA uses 507 stations with a 1.2 km hexagonal spacing. As with Auger, these detectors have 100% uptime and are used to measure the spectrum with rel-
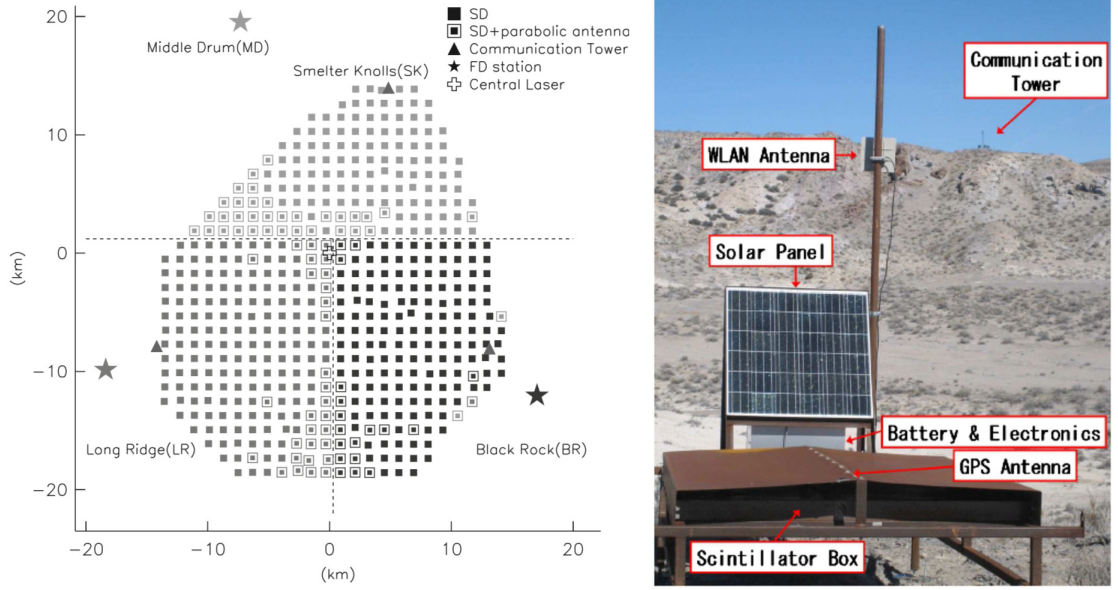
Figure 3.18: Left: A map of the telescope array located near Delta, Utah. Right: The TA surface detector with components diagrammed. Both from [80].

atively high statistics [80]. These detectors instrument a total area of 700 km^2, of which the area covered by their FD is a subset. The electronics consist of a CPU board which is fed data by an FPGA. The CPU is a Renesas SH-4 running at 266Mhz, which is interestingly the same unit that powers the 1998 Sega Dreamcast. The programmable logic is provided by a Xilinx SPARTAN3 FPGA, and the FADC is a 50 MHz 12-bit Analog Devices AD9235.

3.3.2.1 Work Towards a Cross Calibration of Auger and TA

There are a number of unresolved questions in the field of UHECR physics, and a handful of these revolve around tensions between the data collected by Auger and the data collected by TA. In particular, Auger has shown that in its spectrum, as the energy increases, the composition gets heavier [81]. TA reports the opposite. Both collaborations use fluorescence measurements as the standard of measuring showers, although Auger is uniquely equipped

to measure composition with the SD array. Going forward, the AugerPrime upgrade will allow a much more accurate determination of composition from Auger's SD array.

Additionally, TA was originally built to answer a question that haunted Auger, which is the issue of different energies as reconstructed by the FD and SD. Fluorescence techniques are considered to be powerful enough that the SD at Auger is calibrated based off measurements made with it, however when the SD reconstruction is done from first principles, it does not agree with the FD reconstruction [79].

In an attempt to resolve these tensions, the High Energy Astrophysics group at CWRU along with the Auger group at the Colorado School of Mines have undertaken the task of setting up an Auger WCD in the TA facility for direct cross calibration [25]. Ultimately, this involved preparing most of the equipment to run a detector of our own design in the TA, since the TA stations do not have accessible local triggers.
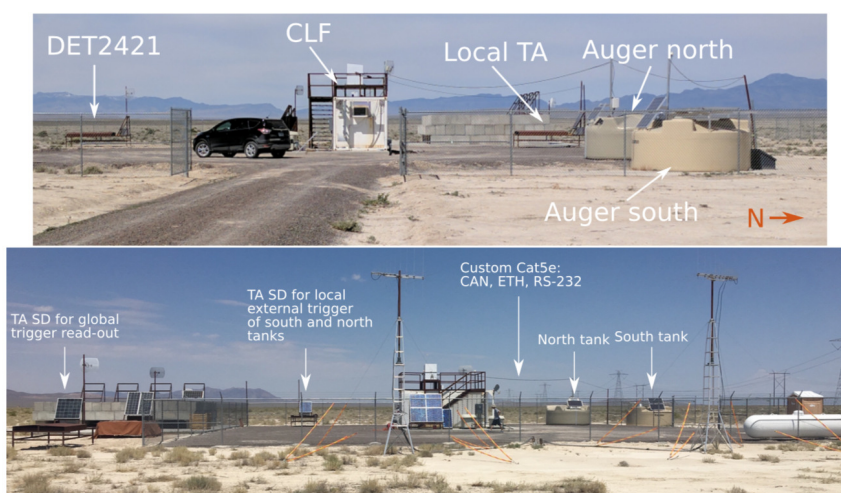
Figure 3.19: Here are two photos of the Auger@TA site with components labelled. From Quinn [25].

In all, our detector, an Auger WCD of the same design deployed in Argentina, and Auger North WCD (a prototype for a cancelled project) and two TA detectors are running at the Extreme Laser Facility (panoramas shown in Figure 3.19). One of the TA stations

reads out global TA triggers, while the other is effectively used for it's detection electronics and hooked up to all custom processing, including discriminators, coincidence gates and time tagging electronics. The construction of the time tagging electronics will be discussed in subsection 6.3.4.

3.3.3 CTA

For the past 30 years, the Air Cherenkov has been the primary method for Very High Energy (VHE) gamma ray astronomy (with significant contributions from others such as HAWC and CASA-MIA). As many US and European groups began thinking about the next generation of Imaging Atmospheric Cherenkov Telescopes (IACTs), it was realized that gamma ray physicists and astronomers from around the globe would have to pool resources to build a bigger array of telescopes. Hence, the Cherenkov Telescope Array (CTA) was conceived. In this subsection, we aim to introduce CTA and the pSCT MST to give context to subsection 6.3.5.

Passing on the legacy of HESS, VERITAS and MAGIC, CTA will consist of two main sites, one in the Canary Islands, at La Palma on the Roque de Los Muchachos site, and another in Chile at the European Southern Observatory's Paranal site [82]. The sites have been chosen to maximize the performance per unit time, and minimize the cost of land and logistics [83]. The array will consist of 3 or 4 telescope designs, with large and medium telescopes at the north site, and all three sizes of telescopes at the southern site. In air Cherenkov detectors, larger telescopes focus more light and thereby make lower energy events detectable, while multiple smaller telescopes allow for a wider field of view with less light collecting area, making them effective for high energy events.

Of the current designs for CTA telescopes, there is one model in the Large Sized Telescope (LST) category, two in the Medium Sized Telescope category (MST) and 3 in the

Small Sized Telescope (SST). Ultimately, the collaboration will likely decide on 1 SST, while employing two MSTs of one type per site. The LST, one MST and one SST use the tried and proven Davies-Cotton optics design, while the rest use forms of two-mirror optics. In general, two mirror optics are more delicate, but allow for a wider field of view, and therefore more showers from gamma rays detected. A "family photo" of the potential designs is shown in Figure 3.20.

Figure 3.20: In this figure, from Ong [82], we can see all six of the candidate telescope designs[5].

In CTA, the workhorse telescopes will be the Medium Sized Telescopes (MST). This is also the only design range in which deploying two different types of telescopes, with different optics and cameras, is being considered. To comment on this briefly, using two instruments can be more difficult than having a unified approach due to differing development, maintenance and analysis issues, but having them developed relatively independently

can give important cross checks. Furthermore, from a political perspective, a country will understandably be more likely to fund the construction of an apparatus if scientists from the country played an important role in its development. From this perspective, accepting two MST designs can provide more infrastructure funding for both groups and sites.

To briefly take a broader view, in many of the Cosmic Accelerator candidates that I've discussed in section 2.4, large leptonic populations are expected to be accelerated in addition to UHECRs. These leptons generally interact in the source and should produce VHE to UHE gamma-rays. With powerful enough optics, and a long enough exposure CTA and other gamma ray observatories should be able to find events at the top of their energy spectra which perhaps point back to UHECR sources. More data collection and higher statistics should aid in any future multi-messenger analysis involving VHE gamma-rays, UHE neutrinos and UHECRs.

3.3.3.1 Schwarzchild-Coudèr Telescope

The northern MST is a design originally spawned by Vladimir Vassiliev of University of California, Los Angeles. This design was under consideration well before it became a candidate in CTA. Originally, this work was to culminate in the Advanced Gamma-ray Imaging System, but the team saw a more feasible funding opportunity by joining CTA [84]. With a mostly US and partly Italian team, the prototype of this telescope has been set up at the Fred Lawrence Whipple Observatory (which is, perhaps confusingly, not the site of the original Whipple Telescope). On Jan\u{}uray 31, 2019, this telescope saw first light.

The optics and camera design make the Schwarzchild-Coudèr Telescope an advanced and unique design. More pixels will be active in the focal plane of this camera than every IACT currently in use (at time of writing) in the world, combined.

The two mirror Schwarzchild-Coudèr have been known about since the turn of the 20th
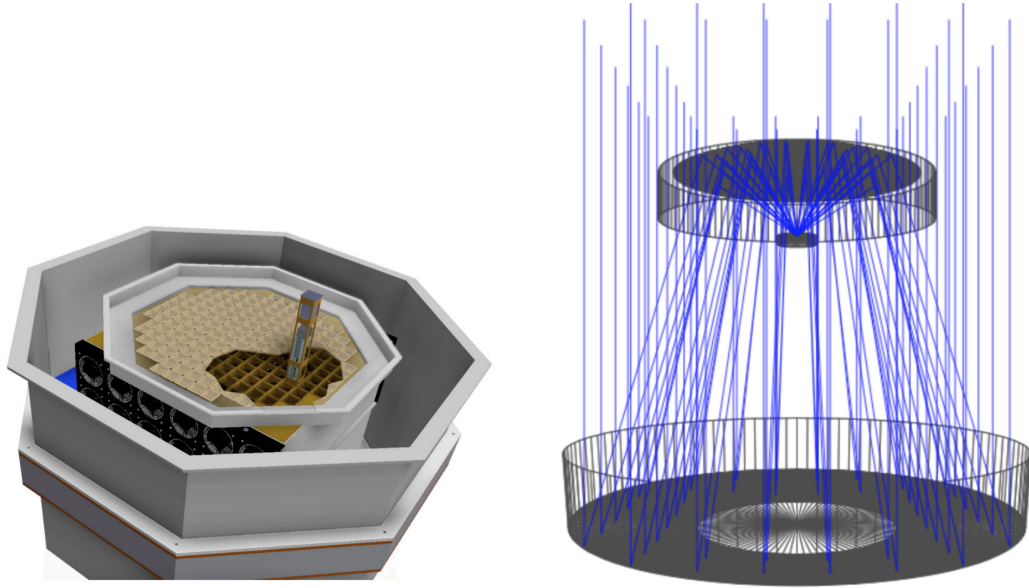
Figure 3.21: Left: The SCT camera design rendered, from al. [85]. Note the detector modules being inserted into the lattice and the muffin fans mounted below for cooling. Right: A diagram of the camera's optics with a ray tracing simulation overlaid [86].

century, but mirror finishing technology has only recently reached the needed precision-at-cost to make this telescope a viable option. The optics, shown in Figure 3.21, give this telescope an 8°field of view in each dimension, allowing a larger collecting area over the night sky than any previous or planned IACT [86]. The optics also focus the image onto a focal plane that is narrower than previous designs, ruling out the use of PMTs in the camera.

The camera, due to optical constraints, must then use Silicon Photomultipliers (SiPMs), which are small silicon-based avalanche photodiodes, that effectively rely on the breakdown of their charge binding structure when put under a relatively high voltage (high for nature, low for PMTs). These allow us to fit 11.2 kPixels in the focal plane. At this resolution, it should be (and from the first couple of events, is) possible to accurate determine the morphology of the events, and therefore very efficiently reject cosmic ray events. Currently, the telescope is operating and needs a time-tagging system to allow for cross calibrations

and meaningful event determination. This project is underway and will be discussed in subsection 6.3.5.

Chapter 4

Electronics and Firmware for the Auger Prime Upgrade

We will open this chapter discussing the scientific motivations for the AugerPrime upgrade, which in turn dictate the station electronics requirements. Then, a description of the science and mathematics behind decoupling the electron and muon signals in the SD is given. Subsequently, we will have an overview of the Upgraded Unified Board (UUB) which will control the upgraded SD stations, and describe its main features. The organization of work packages for the upgrade will be discussed briefly, and any features not covered in the board overview will be discussed as individual work packages, less GPS and Timing which has its own section towards the end of the chapter. After the work packages, we will discuss the operating system of the UUB and how it was brought-up. The chapter continues into a discussion of the GPS software and hardware integration. Finally, we end on a walk-through of the boards operating system and bring up procedure, focusing on the author's documentation on the topic.

4.1   Scientific Motivations for AugerPrime

Since the days of John Linsley, much of the early work in UHECR science has been motivated by measurements of spectra; while successful methods for measuring primary com-

position directly using balloon-borne and space-based instruments have been developed (e.g. Hams et al. (2016) [87]), inferring composition from air shower measurements has been more challenging. However, at this point the spectrum is resolving (Figure 2.4) and primary composition measurements on a shower-by-shower basis can lead to important constraints on source models, thereby improving searches for anisotropy and efforts to identify specific astrophysical sources.

Auger's Fluorescence detector already measures X_{max}, and therefore composition, but its reach is limited by its 10% duty cycle: a consequence of needing clear skies to operate. At a fixed energy, we expect the relative muon and electronic component abundances to be a strong indicator of primary atomic composition. Therefore, by separating the muonic and hadronic component of showers on the ground, we have an opportunity to employ the entire surface detector array with its 100% uptime towards event-by-event composition measurements. This is the basis of the AugerPrime program[1].

In order to deconvolve the muonic and hadronic component of the shower, new hardware for the surface detector is required. Towards this end, each station in the array will be equipped with a Scintillation Surface Detector (SSD) and a radio antenna (preliminary details in [88]). Each of these instruments on an upgraded SD station have a different response to the electromagnetic (EM) and muonic components of the shower, and we can exploit this to separate the muon and electron components of the shower (see subsection 4.5.1 and [89, 90]). Finally, since a shower's muon content depends on the number of initial hadrons and therefore the atomic number of the primary, we can use the muon to electron ratio, N_μ/N_e, to find the composition. In practice, this actually is accomplished by finding the depth of shower maximum via shower universality (see subsubsection 3.1.4.2 or [90]). Ultimately, this is correlated with Z via air shower Monte Carlo simulation results such as those shown

---

[1]Prime, although stylized lower case, is an acronym for Primary Ray Identification through Muons and Electrons

and discussed in subsection 2.5.2, specifically Figure 2.5 or those shown in Figure 4.1.

As an additional consequence of the ability to better determine the muon content of showers, we will be able to cross-check models of particles physics based on extrapolations from LHC results. In particular, the ultra high energy models that CORSIKA (subsection 3.1.5, [61]) employs predict more muons than are observed in the Auger SD [91]. It is unclear if this is an issue with the muon production in the models or if it is a problem with the current muon content measurement of the SD. An upgrade of the surface detectors to more accurately measure muon content is required to answer these questions, as well as make shower-by-shower composition measurements. This upgrade takes the form of AugerPrime.

4.2 Decoupling the Shower Components: SD Composition Measurements

As mentioned above, the primary goal of the AugerPrime upgrade is to determine the composition of primaries with the Surface Detector. The effective workflow here is to find the muon content of the shower, which we can relate to X_{max} (see Figure 4.1), which can in turn be used to give an estimate of the primary composition [89, 72, 92, 93]. Looking at the observables on the ground, we can consider the question of how we unravel the components of a shower using the efficiencies of two different detector types. As we think about the different responses of the new detectors added to the SD, we should note that the water Cherenkov detector will have a stronger response to muons than the scintillation detectors. Knowing this, we can imagine using some type of signal addition or subtraction to find the showers components from the detector's traces. This is formalized in a classic linear algebra problem: matrix inversion. For Auger, this method was developed in Letessier-

Selvon et al. (2014) [94], and furthered in The Pierre Auger Collaboration et al. (2016) [92][2].
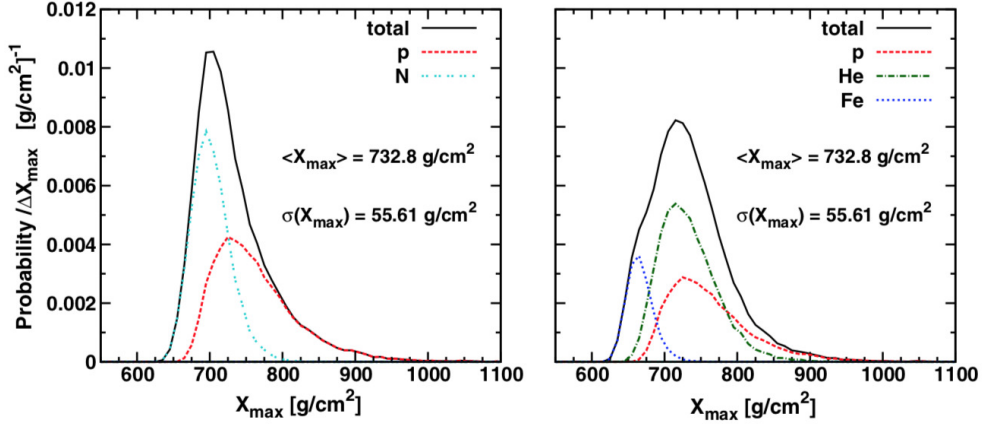
Figure 4.1: These plots out of Aab et al. (2014) [93] give some visually simple intuition for the width and mean of the X_{max} distributions for different primaries. It should be clear from these that a precise statement of primary composition is out of reach using this technique, but we can assign probabilities of an event's composition.

The Pierre Auger Collaboration et al. (2016) [92] states the problem as follows:

$$
\begin{pmatrix} S_{SSD} \\ S_{WCD} \end{pmatrix} = \begin{pmatrix} \lambda A_{SSD} & A_{SSD} \\ \beta A_{WCD} & A_{WCD} \end{pmatrix} \begin{pmatrix} f_{EM} \\ f_{\mu} \end{pmatrix},
\tag{4.1}
$$

where S_i is each detector's respective signal, A_i is each detector's respective area, λ is the MIP/VEM/$\cos\theta$ ratio, $\beta = \cos\theta + 2h/(\pi R)\sin\theta$ with h and R the height and radius of the WCD, respectively, and f_i the flux of the respective shower component (see subsection 3.2.1, subsection 3.2.2 and subsection 4.5.6). Solving this linear algebra problem to give the fluxes is then straightforward.

---

[2]This publication ([94]) was originally meant as a proposal for the AugerPrime new detector technology, in which the tank would be split into two parts with an extra rubber liner. The liner would have significant stopping power for the EM component of the shower, thereby effectively creating two different detectors in the same tank. This was rejected because, somewhat unintuitively, it costs much more to overhaul and refill the water in 1600 stations than it does to simply mount a scintillator on top. The electronics costs between the two plans were comparable.

One thing to note about this method is that we still have to put it through the statistical filter of X_{max} to composition correlations. Another is that it does not depend on better timing. Ultimately, this method depends on shower universality (subsubsection 3.1.4.2) in order to be valid, and specifically it relies on the separability of the shower components [92, 90].
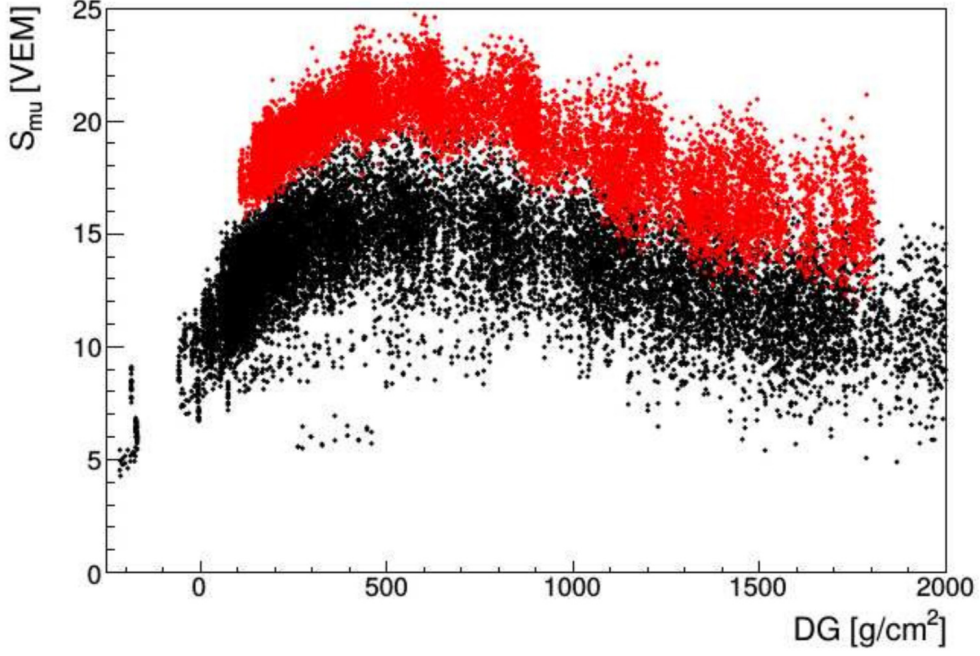
Figure 4.2: This plot, originally from [89], shows the signal strength of muons (S_μ, as in Equation 4.1) as a function of depth and primary composition. The black markers correspond to protons and the red markers are for iron. DG is the difference in vertical depth between the detector and X_{max}.

After discussing this fairly solid and relatively straightforward technique, it is necessary to mention that there *should* be a better way to relate the number of muons in the shower directly to the primary composition, without relying on the statistical variability of X_{max}. Hints of this are shown in Figure 4.2, where we can see that the muon signal is not precisely related to X_{max}. A number of techniques have been looked at and successfully simulated

to find the number of muons in a shower, and the better timing does help with this ([95, 89]) but techniques to take the observables and turn them into a measure of the primary composition are still under development. In principle, improved detector timing for the WCD could lead to better particle identification since the signals from individual muons being detected in the tank should appear as short bursts relative to the more dispersed electromagnetic particles [95].

4.3   Upgraded Unified Board Overview

The Upgraded Unified Board (UUB) serves as the "nerve center" of an upgraded surface detector, controlling its primary scientific functions. The UUB is a custom board, designed by Auger collaborators as a drop-in replacement for the Unified Board. It is made to fit in the same electronics enclosure and connect to the same radio, GPS antenna, power system and photomultiplier tubes [92]. The board is powered by a Zynq^{TM} 7Z020 SoC/FPGA (System-on-a-Chip/Fully Programmable Gate Array), and consists effectively of three layout sections: the Zynq^{TM} and interfaces, the front end inputs and digitizing, and slow control [96]. In this section we will discuss the context and general features of the UUB, paving the way for a detailed investigation of its subsystems including the Board Bring-up and GPS Integration documented in section 4.4 and section 4.7, respectively.
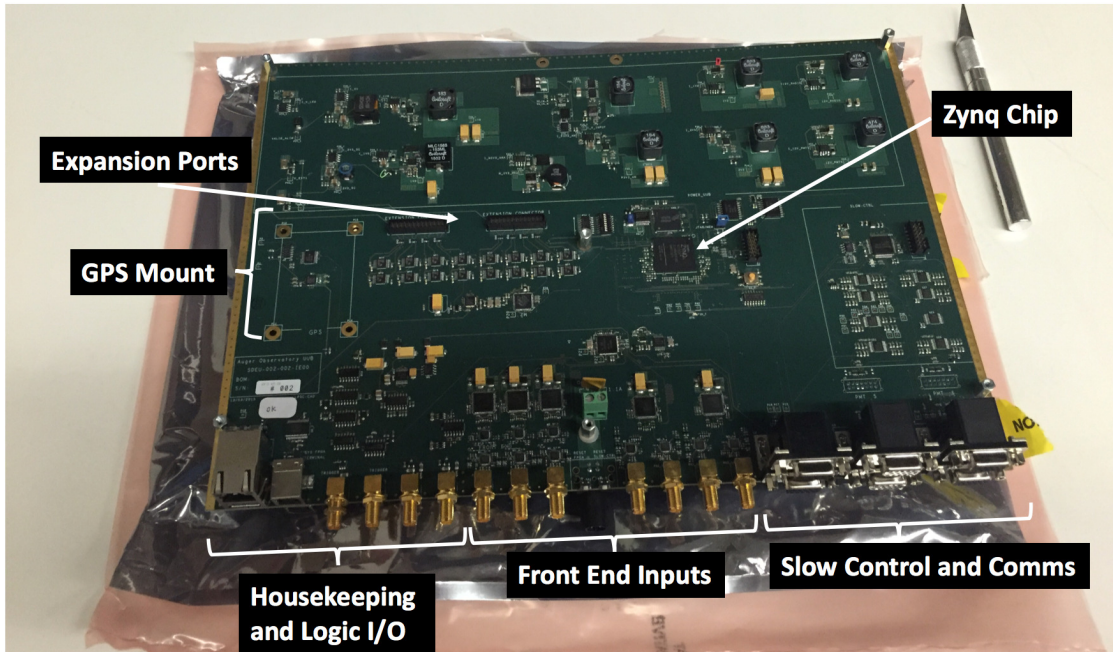
Figure 4.3: A Picture of the UUB with some of the basic parts and connections labeled.

The Zynq^{TM} chip provides most of the processing functionality of the UUB, enclosing event encoding including trace recording, time-tagging, trigger logic, communications packet creation, housekeeping, monitoring and more. The Zynq^{TM} chip is split into two essential parts: the Processing System (PS) and the Programmable Logic (PL). In the diagram in Figure 4.4, the PL is not explicitly shown because of it's undetermined nature, but what is made clear is the architecture of the PS. There are both "hard" and "soft" peripherals, such as GPIOs (General Purpose Input Output) and UARTs (Universal Asynchronous Receiver/Transmitter). These devices live both in hard silicon on the processing side, where some of their parameters are customizable, and in the programmable logic, where they are completely customizable.

The CPU in the processing system consists of two ARM Cortex^{TM} A9 cores running at 333Mhz. This effectively allows an appropriate operating system to invoke parallel processing to decrease dead time and increase reliability of the detector. An example where

this is especially useful is interpreting GPS messages through the receivers rather slow baud rate. At a 9600 baud rate of communication, the processing system must wait for 150 or more bytes to come in, corresponding to more than a microsecond of dead time. In this situation, the second core will handle all other functions, in particular critical data handling. That said, the PL firmware is designed to save all important information, i.e. anything that is science-critical, until it is ready to be read out. Furthermore, interrupts have been implemented in the processing system for both shower and muon triggers, and Direct Memory Addressing (DMA) is used to transfer the data from the trigger buffers to the PS with minimum processing power and maximum speed.
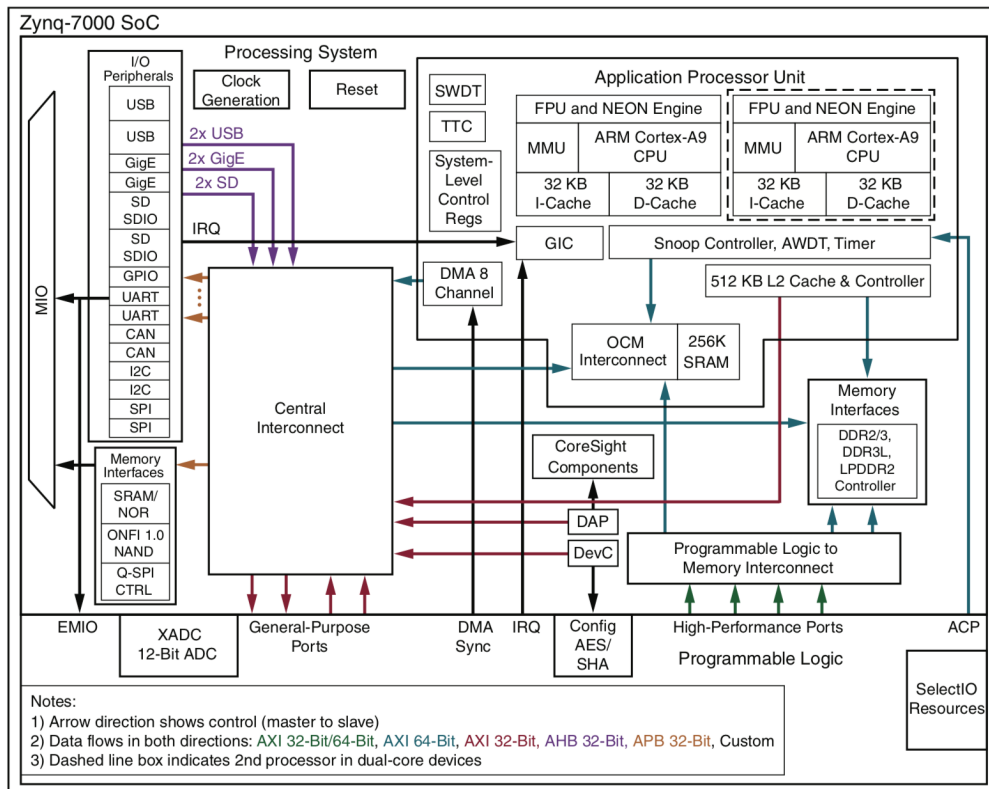
Figure 4.4: A schematic of the internals of the PS in the Zynq$^{\text{TM}}$ chip, from Xilinx [96].

The external features of the physical board layout are shown in Figure 4.3, while the

high level logical and signaling layout is shown in Figure 4.5. A quick examination of Figure 4.5, shows that the Zynq$^{\text{TM}}$ chip is truly the the hub of all of the activity on the board, in fact even external clocks are distributed through it. The main external external clock is the Abracon ABLJO-V-120.000MHZ-T2 ([97]), a VCXO (Voltage Controlled Crystal Oscillator) whose control voltage is set by an SPI bussed DAC. This VCXO was chosen for its remarkably low jitter of .1ps, however a number of options have been discussed for future versions of the UUB[3]. The speed of the oscillator was chosen to match the speed of the AD9628 120MS/s ADCs.

Signal processing-wise, the analog portion of the front end is largely ported over from the Unified Board (UB, the previous control board for Auger), and is effectively designed to ensure pulse integrity from the photomultiplier tubes. Additionally, the UUB includes amplifier electronics for each channel and an anti-aliasing filter, to ensure faithful reconstruction of relatively high frequency features. Finally, the board also has a number of power delivery features. The UUB serves as a hub for the various voltages that need to go to its detectors and other peripherals[4].

4.4 Board Operating System and Bring-Up

To accomplish the goal of finding the composition of primaries through the SD, we need to add new detectors to each station. Handling the data from these new detectors requires a new board, the UUB, and that board requires an operating system. In this section, we explain how that operating system has been set up for the UUB, a project which the author headed in the beginning phases. This section is structured to first give a brief overview

---

[3]These include temperature compensated and oven compensated oscillators to deal with the relatively rapid temperature changes (see Brandt (2015) [98] for more about the desert temperatures).

[4]As a note, it is worth mentioning that the power delivery has caused some amount of engineering headaches as it can induce voltages in other lines that are routed above, below or even just nearby. This can be a scientific show-stopper if the affected line happens to be a signal input.

of the boot stages and PetaLinux, and finally an overview of the documentation we created for the task, which was propagated through the AugerPrime working team. Complete documentation written by the CWRU group is given in **??**.

4.4.1 Boot Stages

Four stages are required for the UUB to start up. They are listed and defined below:

1. Bootstrapper: Initializes the Zynq^{TM} and is hardware specific but can be made trivially from a board support package.

2. First stage bootloader: Provides resources for second stage bootloader. This is made for Zynq^{TM} chips in general, not for a specific board.

3. Second stage bootloader: The most commonly used solution and the one we will employ here is a program called U-Boot, which is available preconfigured for Zynq. U-Boot is effectively a stripped down Linux kernel which gives a large enough command set to do some debugging and launch the main kernel.

4. Operating system: Here we need a true Linux operating system, like PetaLinux or Debian. The OS will be configured using a device tree which is available for the UUB through the Xilinx Design Tools. It will be cross compiled on an external Linux platform using the provided Xilinx toolchain (Code Sourcery).

4.4.2 PetaLinux

To fulfill our operating system needs, the UUB makes use of the Xilinx recommended PetaLinux operating system. PetaLinux is easily compiled for ARM systems and gives all of the needed functionality in a package smaller than 10MB, although this can be expanded

for greater functionality. The PetaLinux Source Developer's Kit has a number of useful options for all manners of communication, including a lightweight SSH server, drivers for many commonly used peripherals, and an open source configuration, allowing even kernel level modifications. Ultimately, however, the decision to use PetaLinux over any other operating system was made because it is integrated into the Xilinx ecosystem, making development much easier.

4.4.3 Bring-Up Documentation

As part of our responsibilities for prototyping the board bring-up, we wrote detailed documentation on the process. For the sake of brevity, this has been put in Appendix C. The basic steps are to download and install the Xilinx design suite on a 32-bit Linux machine, then install the PetaLinux SDK package. We then load our hardware specification into the PetaLinux SDK, select our Operating System options, and finally compile the binary executable, which is on the order of a handful of megabytes. The documentation then walks through the subtleties of booting the kernel over JTAG, paves the way for booting from an SD card or SPI Flash, and finally discusses how to implement software in the PetaLinux environment.

4.5 Relevant Subsystems

Before proceeding on to a discussion of the electronics and GPS integration, we should first discuss some of the subsystems of AugerPrime with which we will need to interact. We start with a discussion of the scintillator, which is critical to the science case for AugerPrime as noted in section 4.1. After discussing this new detector element, we move on to a discussion of the trigger logic. This feeds into a subsection about the trigger hierarchy, and

finally we will discuss slow control and calibration. A full schematic of the UUB is shown in Figure 4.5.
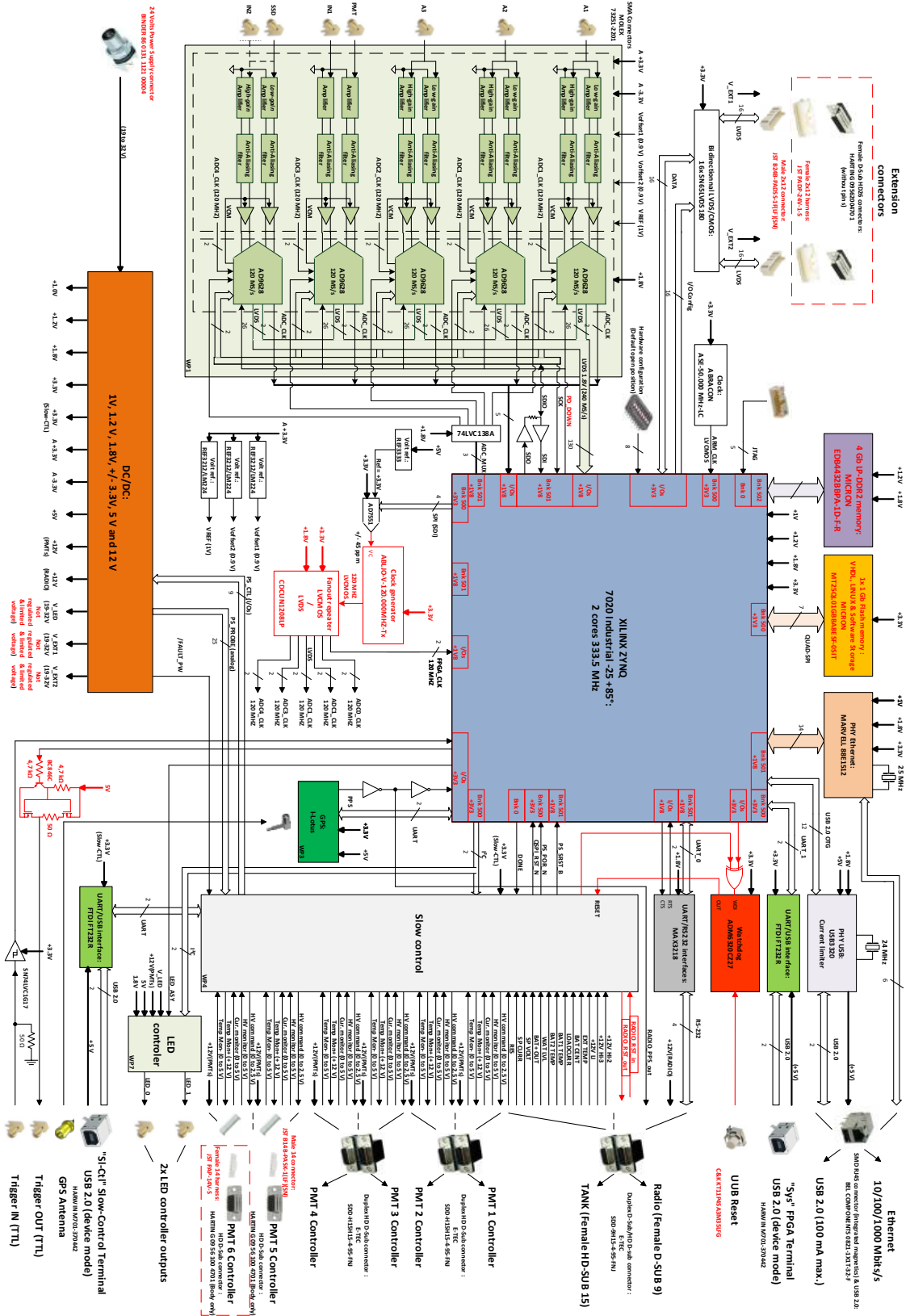
Figure 4.5: A block schematic of the UUB, displaying the major components and their interconnections, from Suomijärvi [99].

4.5.1 Scintillation Surface Detector

While some details of the SSD have already been discussed in subsection 3.2.2, we will review them here and provide a few more technical diagrams and pictures. The Scintillation Surface Detector is an organic extruded plastic fluorescent scintillator of about 3.8m^2 manufactured by Fermilab. As mentioned above, the addition of this detector should give us a handle on the muon content of the shower, in conjunction with the water Cherenkov portion of the detector.

In the final stage of AugerPrime, each SD station will be fitted with a scintillator which sits on top of the station in a specially designed mount, pictured in Figure 4.6. As pictured in the left panel of Figure 3.7, Wavelength Shifting Fibers (WSFs) pickup and transport the scintillation light to the photomultipliers. To determine the best choice for photomultiplier tubes, both silicon photomultipliers and photomultiplier tubes (Hamamatsu R9420) were evaluated, with the PMTs being the final choice ([69, 100]).

Each scintillator bar, of which the larger panel is comprised (48 in total), is coated with titanium oxide on the top for handling and for reflection of usable signals back into the bulk of the scintillator. on top of the panels is a corrugated steel heat pipe structure which allows air to flow, and reflects the sunlight, thereby keeping the SSD at ambient temperature. The fibers are bundled together at the center of the SSD and aggregated at the 'cookie' in front of the PMT. The first SSDs were commissioned to the Engineering Array on Sept. 16, 2016.
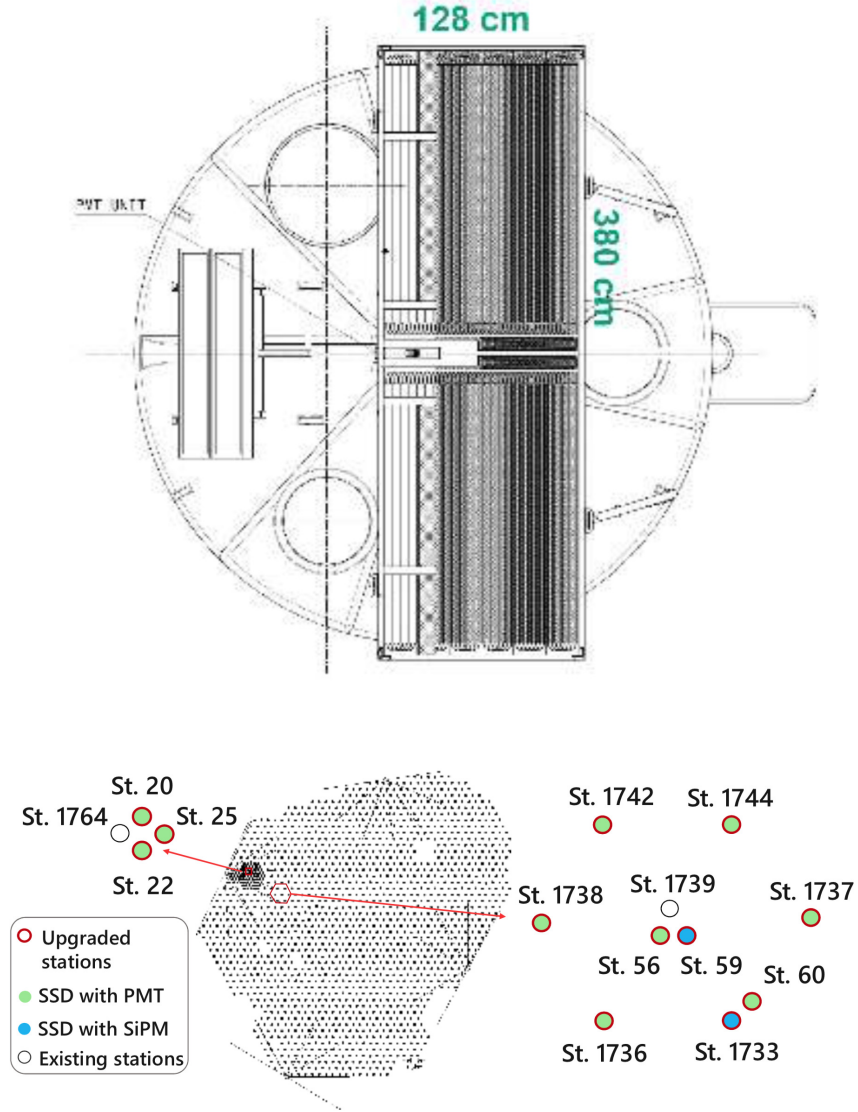
Figure 4.6: Top: A schematic of the SSD mount originally shown in Šmída [69]. Bottom: The positions of the SSD carrying stations from Zong [100].

4.5.2 Trigger Logic

The trigger logic is implemented in the Xilinx Zynq[TM] chip, on its Artix-7 FPGA, and is fed by the analog front end through the digitizers (discussed and named in section 4.3)[5]. In the early days of Auger, a number of triggers were tested, but the most successful was the

---

[5]In large part, the trigger logic for AugerPrime was designed by Dave Nitz of Michigan Technological University.

Time-Over-Threshold (TOT) trigger. Single bin (i.e. threshold) and MoPs (Multiplicity of Positive Steps) triggers are also available. Some of the triggers have compatibility modes, in which they are downsampled to 40Mhz and the frequency response of the UB ADCs is modeled and applied to them. Since much of the work done for timing in chapter 4.7.3.2 involves triggering or using the signal chain from the triggers, we will briefly review the nature and functions of the triggers available on the UUB.

We can divide up the triggers into the Compatibility, Current and Calibration categories. The compatibility triggers allow old analysis routines to be used on the data. The current triggers are the state of the art triggers that have been developed for the UUB. These use more advanced techniques than the compatibility triggers, and take advantage of the digital signal processing capabilities of the Zynq^{TM}. Finally, the calibration trigger looks for single through going muons which can then be used in the charge histogram to set the VEM level (see subsection 4.5.6). These descriptions and more technical information about the UUB's PL can be found in Nitz (2018) [101], and a less technical description of the triggers is available in Appendix D.

The two basic trigger types on the UUB ported form the UB and currently in use are single bin triggers, in which the station triggers if a threshold in a single bin is exceeded and Time-Over-Threshold triggers, in which the station looks for its ADC traces to be above a certain threshold for a predetermined number of bins. For the work we will discuss in section 5.3, we will only be looking at single bin and TOT compatibility triggers. The Time-Tagging module discussed in Table 5.1 has separate timing registers for each of the single bin and TOT triggers (see Appendix D for more information).

Triggers at the station level are designated T1 if they pass a first set of conditions; upon passing a second set, they are elevated to T2 which corresponds to 'real showers'. T1s can be used for calibration, both for checking trigger rates and for forming MIP and VEM

97

measurements as discussed in subsection 3.2.2 and subsection 4.5.6. T2s are sent across the radio to the Central Data Acquisition System, which sends a T3 request if certain larger trigger conditions are met. A detailed description with references is available in appendix section D.1

4.5.3 Slow Control

When deploying autonomous detector stations over such a wide area as Auger, it is of the utmost importance to ensure that there are redundant solutions to solve as many problems as possible without physical maintenance. According, the UUB has three lines of defense: (1) Software controls, (2) Slow control and (3) Watchdog system. Many issues can be avoided by careful and clever software, but for those that can't, having a redundant Linux system on board can save many otherwise untenable situations. This is one of the primary functions of the slow control module, largely developed at the University of Wuppertal by Karl Heinz-Becker and his team. The other main functions of slow control are to monitor various voltages, the temperature and other easily measured conditions, as well as communicating with the radio, setting the PMT parameters and taking care of some of the housekeeping tasks. We discuss this here since GPS telemetry and PPSs are used by the slow control module to help monitor and save the UUB in problematic situations.

To ensure the integrity of the system, the slow control module itself makes sure that the Zynq^{TM} chip itself is still running well. By communicating with it and monitoring its reference voltages, slow control can determine if operation is proceeding smoothly, or if there are issues. If there is something that needs to be addressed, the slow control can send a reset pulse to the Zynq^{TM} chip. It also helps manage the correct boot order of the Zynq^{TM}; in particular, the external block RAM must be reset on startup to ensure proper functionality. The Watchdog chip (to the left of slow control in Figure 4.5), monitors the

voltage parameters of both the Zynq$^{\text{TM}}$ and slow control, and is programmed to send a reset pulse to either or both as a last line of defense against failure.

4.5.4 Hardware Design and Integration

Integrating all of the subsystems produced and creating a board layout that facilitates the necessary interconnections of each component is a critical part of designing new detector electronics. The design effort for the UUB involves several international working groups each with specific responsibilities for UUB subsystems. To coordinate integration, each working group packaged their system in the form of an "IP" module. Working groups then coordinated to link IP into the master block diagram.

Among the constraints in board layout and integration, we have layering and induction concerns, thermal and mechanical stress concerns, crowding, power and finally FPGA usage concerns. The Xilinx tool set which is provided to work on the Zynq$^{\text{TM}}$ chip gives tool to manage power and FPGA concerns, while layering and induction can be simulated through SPICE or modeled in ORCAD. Mechanical and thermal stresses are simulated before board production, where designs are iteratively improved on as successive board versions are fabricated and tested. Hence, there are several pre-production runs of boards so that the non-obvious issues can be worked out.

4.5.5 Software Development

For the UUB, much of the software could be ported from the old UB as it was largely written in C. This involves using the PetaLinux command line utilities to instantiate programs to be cross-compiled, and then checking that the port worked[6]. While the software for the

---

[6]This work was done by Patrick Allison of OSU and later by Ricardo Sato who works directly for Auger as our science coordinator, along with contributions from the author and others towards the software pertaining to our specific work packages, i.e. GPS communications software.

UB is rather complex, only certain components needed to be ported for mission-critical tasks. Some parts were rewritten and others were simply transferred over. For the GPS integration, both options were presented and we modified the UB code for possible use but ultimately decided that the differences between PetaLinux (section 4.4) and OS9000 would allow us to write a simpler more reliable version of the GPS code.

4.5.6 Calibration

A number of calibrations must be done for AugerPrime, and as with the initial Auger detectors they must be done autonomously in-situ (i.e. while deployed in the field). From a science perspective, the chief amongst these is the MIP/VEM calibrations. Since triggers are based on a station's VEM measurement, or for the SSD, the MIP measurement, a poor calibration can directly affect a station's science performance. The general method of
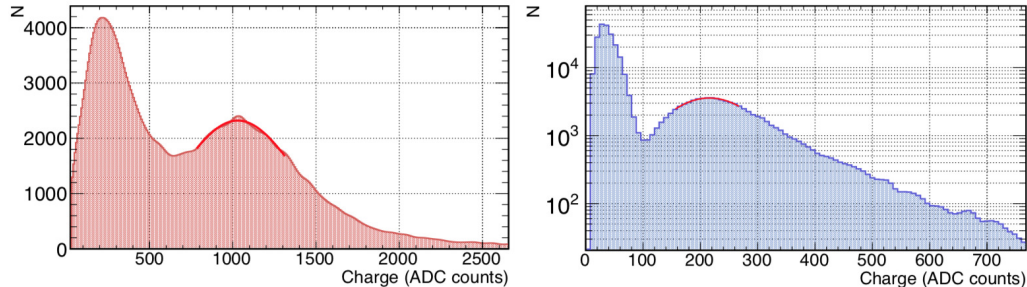
Figure 4.7: This charge histogram, from Zong [102], shows counts from the muon trigger.

finding calibrations for the SSD and WCD is to look at traces whose recording is initiated by the muon trigger. These are stored and analyzed at the station level, and used make a histogram of the charge deposited (calculated from integrating the traces). This reveals two distinct peaks. The left is the noise pedestal, which is used to find the appropriate ADC count offset. The right is the MIP/VEM peak (depending on whether you are looking at a WCD or an SSD). The second peak is fit with a Gaussian and the center is used as the VEM

value in all of the stations internal trigger calculations. Examples of the charge histograms for an upgraded station running in the EA are shown in Figure 4.7.

There are other important calibrations, such as the high voltage calibration. This is done by looking at the noise level in the ADC traces and handing histograms of that off to slow control. This is a largely straightforward process which, usually goes well; when it doesn't succeed stably, it usually indicated an issue with the PMT which will have to be fixed manually. Additionally, in order to determine important parameters from deconvolving the different components of the shower, MIP vs. VEM diagrams must be made (see section 4.1 for more details). We have shown one of these in Figure 4.8.
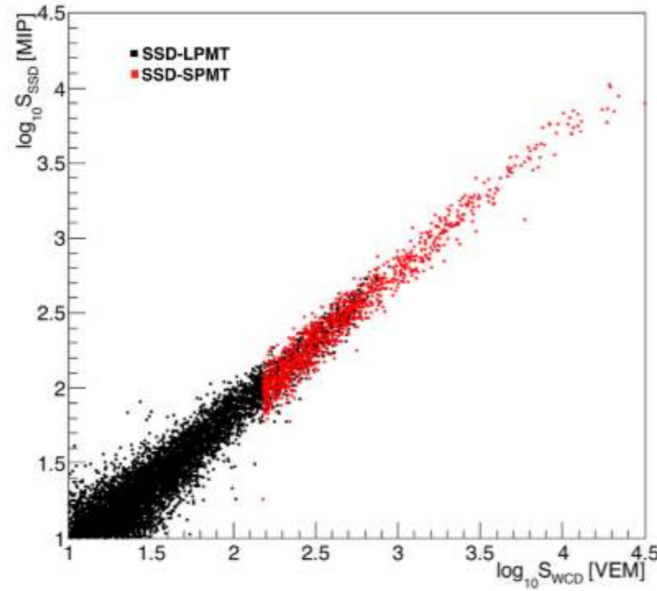
Figure 4.8: This diagram from Castellina [103] shows a typical MIP vs. VEM calibration curve. While it looks linear with a slope near unity, a closer inspection shows that the VEM signals are slightly higher than the MIP signals, and this makes sense because a muon will deposit more energy over its >1m traversal of the WCD than it will in its short traversal of the SSD.

4.6 GPS Timing

For the AugerPrime upgrade, the main responsibility of the HEA group at CWRU was timing and GPS integration. In order to accurately reconstruct the properties of the primary cosmic ray in an air shower event, each station must know what time it is according to some shared standard [7]. For Auger, connecting each station with a wired connection is not technically nor financially feasible, and so time must be distributed in some other way. The cheapest and most accurate solution is to use the constellation of atomic clocks that float over our heads 24/7, delivering accurate timing information for free: the Global Positioning System (GPS).

In this section we will discuss the essentials of GPS timing starting with the mathematics of GPS, then move into a discussion about the GPS constellation. We will discuss some of the early concerns which have shaped the way the GPS is used in Auger and the coordinate system used by GPS insofar as we need to understand it to calculate distances between sets of GPS coordinates. At this point we move into the more technical aspects of GPS, discussing messaging protocols, message selection for AugerPrime and finally the sawtooth granularity correction.

4.6.1 GPS Basics

To inform the coming sections, namely the GPS integration in section 4.7, the receiver testing in section 5.2, and the spatial correlations in section 6.2.6, we will provide an overview of the relevant facets of the GPS.

Underlying the operation of GPS is the idea of broadcasting the position and current

---

[7]In smaller experiments, it is sufficient to simply collect all data at a central location using the same cable lengths and electronic delays in each detector. This was, for example, the configuration of Volcano Ranch and AMANDA. In medium sized experiments, where stations can still be connected via wires, but data transfer must be asynchronous, time is often distributed over networks or through distributed protocols. Examples of this include the Large Hadron Collider and IceCube.

time of a set of satellites to a receiver on the surface of the Earth. The receiver collects the information broadcast from several satellites and compares the time of transmission to the time received to produce a time-of-flight for each collected signal. These time-of-flight values are then converted into what are referred to as pseudoranges,

$$\rho_i = c(t_r - t_i),$$

where t_r is the time on the receiver's clock when the signal was received and t_i is the time on ith satellite's clock, in its own reference frame at time of transmission. We can then compare each pseudorange with equations measuring the distance between the satellite and receiver to create a system of equations which will herein be referred to as the **GPS equations**,

$$|r - r_i| = \rho_i, \ i \in \{1, 2, 3...\}.$$

Here i is an index that runs over the available satellites, of which there must be at least four to fix position and time. More explicitly, in the case of four satellites these equations are:

$$(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = c^2(t_r - t_1 + \Delta t)^2$$
$$(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = c^2(t_r - t_2 + \Delta t)^2$$
$$(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2 = c^2(t_r - t_3 + \Delta t)^2 \qquad (4.2)$$
$$(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2 = c^2(t_r - t_4 + \Delta t)^2,$$

where Δt is a correction term to be applied. We can see that these equations can be manipulated in various ways. If we are only concerned about position, we can forget solving for the time coordinate and use only three satellites (i.e. three equations). If we want to

103

know only our current time t_r, we can fix our last known position and only need to use one equation to find the time. It is common practice to re-parametrize the problem to use pre-corrected pseudoranges and solve for the unaccounted-for error. In equation 4.2, we can think of this as moving the Δt correction term outside and solving for it, assuming we have the correct set of values for t_i.

Of course, in practice we do not want to throw any information out; furthermore, we are dealing with a non-linear system of equations. These do not always admit an easy solution and can have somewhat unpredictable behavior. This is why, in any receiver's electronics, the GPS equations are always solved using either a least squares method, or an iterative solution allowing the receiver to use as many satellites as it can see. These methods allow use of an inconsistent (over constrained) set of equations, meaning that we can use more equations than degrees of freedom, and obtain a more accurate solution. As noted in Langley (1991) [104], almost all receivers use iterative methods to solve these equations because they can do so quickly and to a higher degree of accuracy.

4.6.2   GPS Constellation

The Global Positioning System was fully brought into service in 1995 with a constellation of twenty four satellites. The original geometry of the GPS constellation had the 24 satellites at $55°$ inclination in six different orbits, each azimuthally separated by $60°$ (see figure 4.9). The constellation has since been expanded to 32 satellites for greater worldwide coverage. With the original geometry, a user should always have at least six satellites in view, allowing for redundancy in the constellation and giving more data for the numerical solution of the positioning equations.

Initially, the specification for the GPS system was largely dictated by the military's requirements, and so a slow rolling local error was added in to the civilian channel, to
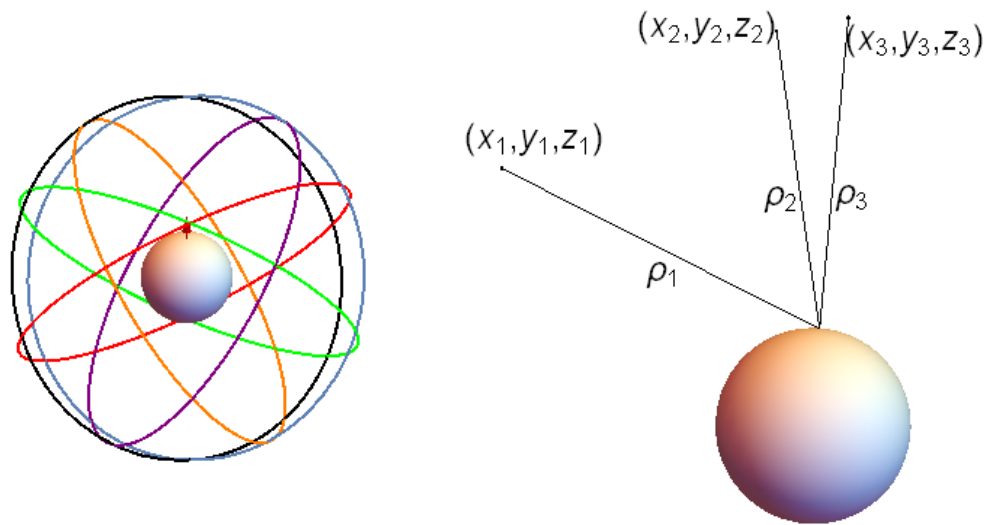
Figure 4.9: Left: A diagram of the available GPS constellation orbits of the original 24 satellites. The red arrow indicates the North Pole. Right: this diagram shows the coordinates and quantities used in GPS calculations.

ensure that only the military was able to get the initially specified minimal $10 - 15$ m spherical error probable (SEP: the radius from the mean of a collection of points within which 50% of the points lie). Civilian users received this slow rolling error, officially called Selective Availability resulting in inaccuracies as large as 100 m [105]. In 2000, Selective Availability was turned off by executive order, and the next generation of satellites will be built without this feature.

4.6.3 World Geodetic Survey of 1984

To calculate distances and understand altitudes and positions as needed both in Auger and for a number of other HEA group activities (e.g. subsection 6.3.4, section 6.2.6), we will need an introduction to the useful details of GPS coordinates.

The coordinate system used by GPS is called WGS-84, short for World Geodetic Survey of 1984. While there are plenty of technical details to discuss, they are largely outside the

purview of this work. WGS-84 represents a detailed model of the Earth including tidal and tectonic effects. For us, the only relevant portions is referred to as the WGS-84 Ellipsoid, and it tells us three very important parameters. From Imagery and Agency [106], the semi-major axis of the coordinate system the Earth is embedded in via this standard is a=6378137 m, and the semi-minor axis is c=6356752.3 m. The radius as a function of latitude is:

$$r = \left(\sqrt{\frac{\cos(\theta)^2}{a^2} + \frac{\sin(\theta)^2}{c^2}} \right)^{-1}, \tag{4.3}$$

where θ is the latitude. This is a simply derived fact of 2D geometry. For the purposes of this work, this is what we need to know about WGS-84.

4.6.4 GPS Receiver Communications Protocols

Since we will soon discuss the integration of the new GPS receivers for AugerPrime, we will give a general overview of the communication protocol used by our GPS receivers.

Due to Motorola's early dominance in the GPS market, many receivers that are not for maritime applications use the "Motorola Binary" (MB) communications protocol. [8] The Motorola Binary protocol specifies that all connected devices will operate at 9600 baud[9]. The way it is written uses bit and byte packing techniques to absolutely maximize the amount of information contained in a message. In general, MB is not human readable and even has to be interpreted to fit into standard memory structures[10]. In general, the fields in MB are character, short, integer or sub-byte, i.e. 1, 2, 4 or <1 bytes. In the cases where they are less than one byte, each bit generally represents a boolean field.

---

[8]For the stations running in Auger currently (less the EA), the Motorola Oncore UT+, is the integrated GPS unit. Information on the GPS units we tested for AugerPrime is in subsection 5.2.1.

[9]Also bits per second, in the case of a single bit communications interface

[10]Alternatively, 'Maritime Protocol 2000' from the National Maritime Electronics Association (NMEA) takes an approach of simplicity in interpretation and integration. This protocol uses plain ASCII letters and packet spacing to make message relatively human readable. It specifies a faster baud rate to make up for the lower information density.

An example of the type of messaging that would be sent in MB is shown in Figure 4.10. In this specification, we can see how the bytes of different fields are packed in. It should be noted that MB is a big-endian format, i.e. the first byte (left-most if you like to hearken back to the tape machine formalism) represents the highest value. Each message is on the order of 100 or so bytes, and so simple math tells us that the 154 byte Time, Position and Status @@Ha message will take $(154 \times 8)/9600 = .12$s or more than a tenth of a second, and therefore properly handling this long time-of-transmission in software in a non-blocking way is important and will be discussed in both the context of the CWRU HEA group's TIM instrument (section 6.2) and the Auger UUB (section 4.7).

@@Hamdyyhmsffffaaaaoooohhhhmmmmaaaaoooohhhh
mmmmVVvvhddttimsidd (repeat 'imsidd' series for remaining 11 channels)
ssrrccooooTTushmvvvvvvC<CR><LF>

Date
    m = month   1..12
    d = day   1..31
    yy = year   1998..2079
Time
    h = hours   0..23
    m = minutes   00..59
    s = seconds   0..60
    ffff = fractional second   0..999,999,999 nanoseconds

Position (Filtered or Unfiltered following Filter Select)
    aaaa = latitude in mas   -324,000,000..324,000,000 (-90°..+90°)
    oooo = longitude in mas   -648,000,000..648,000,000 (-180°..+180°)
    hhhh = GPS height in cm   -100,000..+1,800,000 (-1000..+18,000m)*
    mmmm MSL height in cm   always 0,000,000 with M12M

Position (Always Unfiltered)
    aaaa = latitude in mas   -324,000,000..324,000,000 (-90°..+90°)
    oooo = longitude in mas   -648,000,000..648,000,000 (-180°..+180°)
    hhhh = GPS height in cm   -100,000..+1,800,000 (-1000..+18,000m)*
    mmmm = MSL height in cm   always 0,000,000 with M12M

Figure 4.10: Pictured here is the @@Ha message from the i-Lotus M12M users manual [107]. We include this as an example of how the Motorola Binary format is packed and interpreted.

MB as implemented on the receiver for AugerPrime offers a number of commands

which are used to provide the essential timing functionality desired. For an experiment like Auger, satellite coverage for continuous operation is important[11] Figure 4.9 shows that the inclination of the orbits. The fact that they have Half Sidereal Day (HSD) periods, means that they will precess around the Earth and their even spacing will effectively provide coverage everywhere on Earth. A simulated plot of coverage is shown in Figure 4.11.
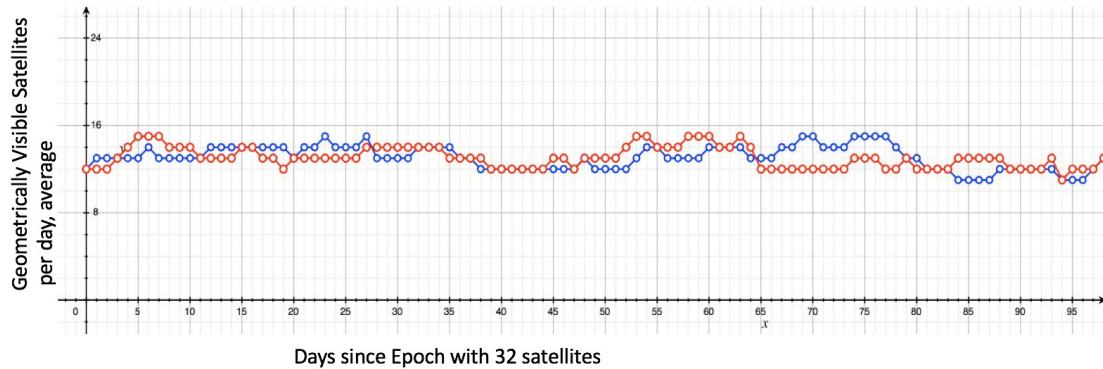
Figure 4.11: This plot, created by then-undergraduate Peter Fedrizzi under my supervision, shows the average number of satellites visible over Malargüe (i.e. Auger, for our purposes) and Cleveland in a given day.

4.6.4.1 Message Selection

The aforementioned concerns led the designers of the UB software to choose three UT+ messages to be recorded in each SD station: @@Ea, @@En and @@Bb. Respectively these are the 8-channel: Time/Status, T-RAIM, and Satellite Status messages. @@Ea is in parity with @@Ha, where both give all of the information the receiver can provide about its positioning and timing solutions, as well as all operational information about the receiver, including temperature, satellite (Space Vehicle, SV) health and clock speed amongst others. @@Hn and @@En are the Time-Receiver Autonomous Integrity Monitoring messages. These messages also contain estimates on the accuracy of the timing solution, as well

---

[11]In the early days of Auger, the period of Selective Availability was still at the forefront of the community's thinking and there were still valid questions about how well the GPS constellation would work in South America.

as the very important Clock Granularity Message (see subsubsection 4.6.4.2), and some further information on the health of the GPS constellation. Finally, the @@Bb message, implemented similarly in all of the relevant receivers, gives relevant signal processing and local orbital parameters (elevation and azimuth from the receiver).

For the UUB, we[12] looked at the actual science goals and realized that local stations had been needlessly wasting considerable processing power in parsing and saving GPS telemetry messages that did not contribute to the physics we are trying to accomplish. Specifically, the @@Ha message can be replaced by the approximately 75% shorter @@Hb message, which does not give irrelevant information about the receiver's operation, and instead relays only critical time and positioning data. Furthermore, the @@Bb message could be removed entirely since concerns about the operation of GPS have more or less subsided (see section 6.2.6 for some that remain). These changes have improved our required data load from 324 bytes per second to parse, down to 132 bytes total. All of the messages we have mentioned above and more are implemented in the UUB software and can be enabled remotely (see section 4.7).

4.6.4.2   Sawtooth Correction: The Clock Granularity Message

If we consider the M12M GPS receriver's data sheet, it reports a <2 ns PPS (Pulse-per-Second) accuracy, this despite the fact it sports only a 40Mhz oscillator with 25ns coarse grained timing accuracy. The mechanism through which the receiver makes its PPS accuracy more precise than the coarse granularity of the oscillator is by accurately calculating the phase of a clock cycle in which the PPS should truly land in comparison to the leading edge of the clock cycle when the PPS is sent. The receiver communicates this "sawtooth correction" via its serial connection in the @@Hn message. This is called the clock gran-

---

[12]Ricardo Sato and the author.

ularity message, which takes the form of a bipolar sawtooth pattern, hence its colloquial name, showing the precession of the true PPS around the receiver's clock.
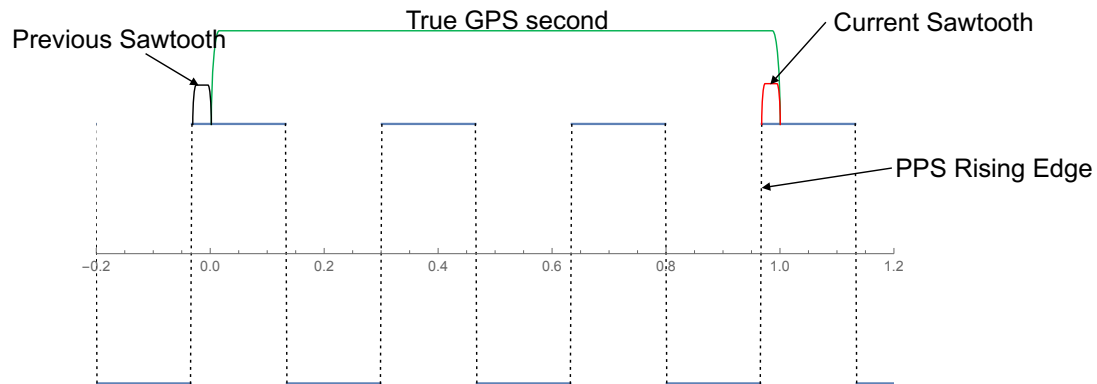
Figure 4.12: Here we have a schematic diagram showing how the sawtooth correction works. By finding the phase difference between the leading PPS edge and the actual beginning of a GPS second, the receiver can deliver a serial message which allows the user to obtain a very accurate beginning-of-second time. In this diagram, we have used 4Hz for the square wave to represent the GPS receiver's clock.

The mechanics of this precession are, at surface level, relatively straightforward (see Figure 4.12): the receiver's oscillator is not perfectly trained to the GPS constellation, and we need not require it to be. If we ask about the specific form of the sawtooth over a given interval, such as that shown in Figure 4.13, more complicated dynamics come into play. Some of these will be discussed in section 6.2.6, but to summarize, these include sky conditions and temperature modulations.

4.7   GPS Timing Integration in the UUB

Integrating a GPS receiver into a board like the UUB requires hardware, firmware and software to work in concert to relay information and signals to a number of destinations
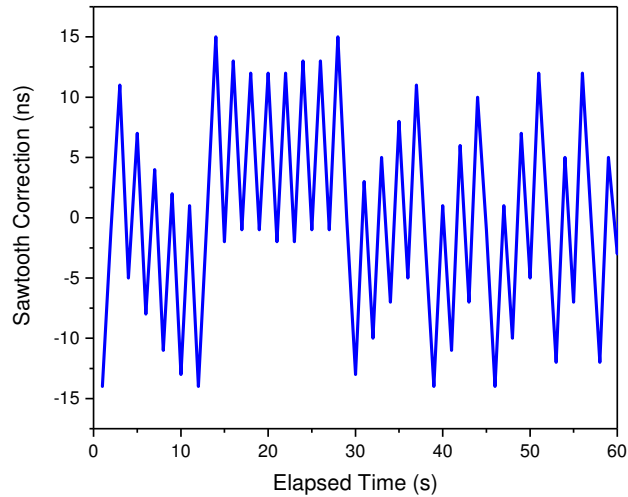
Figure 4.13: Provided here is a plot of 60 seconds of M12M sawtooth corrections. This plot shows the sawtooth behavior, as well as a phase alignment shift around 15 seconds, where, because of the precession of the true PPS around the output PPS, we see a change in the sawtooth pattern and then a change back. These occur on the time scale of anywhere from hours to minutes, depending on the sky conditions and how long the receiver has been running.

in the board. The description of these elements, the design choices made around them and their mechanisms of operation will be the topic of this section and represent a main pillar of the original work described in this document. To begin, we will discuss the firmware design tools, and then we will move into a discussion of the actual software architecture[13].

4.7.1 The Purpose of GPS in AugerPrime

To reiterate, by implementing GPS receivers for Auger and the AugerPrime upgrade, we give the mission-critical ability of synchronizing time across the vast distances covered by the SD. The time differences in stations for moderately inclined showers will be on the order of small numbers of microseconds, but for near vertical showers, accuracy on the

---

[13]We would like to note that the i-Lotus M12M receiver and the SSR-6Tf are entirely interchangeable outside of minor performance differences which will be discussed in section 5.2.

order of tens or hundreds of nanosecond is required to obtain sub-degree reconstruction. This accuracy is already marred by the variable width of the shower front, and so we can keep our aggregate error down by having the best timing and time synchronization possible.

Beyond timing, the AugerPrime upgrade gives the collaboration the opportunity to implement GPS receivers which have positioning capabilities that are more accurate than those of human surveyors. With the new SSR-6Tf receivers, excellent positioning data should be immediately available and a survey campaign can be used to obtain positions to much greater accuracies than human surveyors [14].

4.7.2 Firmware Design for GPS Timing

The first step in implementing the AugerPrime GPS receiver is to build a firmware that allows communication with it. The prime character here is the UART or Universal Asynchronous Receiver and Transmitter. A UART allows serial communication without clock synchronization between two devices. The M12M and SSR-6Tf receivers utilize a slightly modified version of the RS-232 communications protocol. Namely, they use a logical-high voltage of 2V$<$V$_{high}$$<$3V. Precariously, this lands between the commonly used 1.8V and 3.3V logic standards, however we have extensively tested it at 3.3V and it works as expected, even for years on end. In choosing a UART and implementing it, we have to consider a few criteria:

- Reliability- How long will the UART run for without resets or software-based housekeeping?

- Simplicity- How much FPGA resources will the UART take up?

---

[14]Upon construction of the array, trepidations related to Selective Availability and the relatively poor accuracy of the UT+ led the collaboration to employ surveyors to find accurate positions for all 1660 SD stations. These were then uploaded to the UT+ models. Ultimately, it was found that the UT+ models, by taking week long survey campaigns, could find their position better than the human surveyors anyway.

- Flexibility- Can the UART settings be changed, or are they hard coded?

Ultimately, the choice of UART boils down to the latter two conditions, since both options that we looked at were extremely reliable as far as our testing could discern. Since, as was mentioned before, in the UUB design, Zynq^TM slice space is at a high premium, whichever UART is chosen should be evaluated based on being able to accomplish the task at hand with an absolute minimum of resource utilization.

4.7.2.1 Brief Overview of FPGA Resources

From a logical point of view, an FPGA is simply a device which connects sets of logical gates (AND, OR, XOR, etc.) to emulate actual hardware functions. A known principle at work here is that any piece of digital hardware or software can be broken down into a logical construct using logic operations. FPGA resources boil down into units of slices. A slice contains three high level types of resources which are used to create the requested set of logical gates. These are LUTS, registers and multiplexers. The Look-Up Tables, or LUTs, give the FPGA the ability to handle functions which need to take a semi-continuous input and map it to a semi-continuous output. Registers and flip-flops (at the level of this task, these are effectively the same thing) allow the storage of information, and the synchronous passing of it along a data path. Finally, the multiplexers, which are effectively invisible to the firmware programmers, are used by the design software to make the connections between various gates. In some sense they act as hardware relays, but without any electromechanical movement.

4.7.2.2 UART Selection

For AugerPrime, we will be using the UART solely to communicate with one device at a fixed baud rate. Therefore, flexibility is the less important criteria, and simplicity dictates

our UART choice. In Table 4.1, the resource utilization for the UART 16550 ([108]), a

long standing industry standard, and that for the UART Lite ([109]), a Xilinx original IP,

are shown. Further on, we have a top-level diagram for the UART Lite in Figure 4.14.

Based on the lower resource utilization, we made the recommendation that we use the

UART Lite.

The HEA group at CWRU was responsible for both UART selection and validation.

Shakedown tests were performed and the UART Lite was, in finality, used in the UUB and

all of the GPS Timing projects the HEA group took part in.

Device	Slices	Slice Registers	LUTs
UART 16550	118	308	347
UART Lite	49	79	119

Table 4.1: Here are the different resource utilizations for the relevant configuration of the two Xilinx UART IPs in question [109, 108].
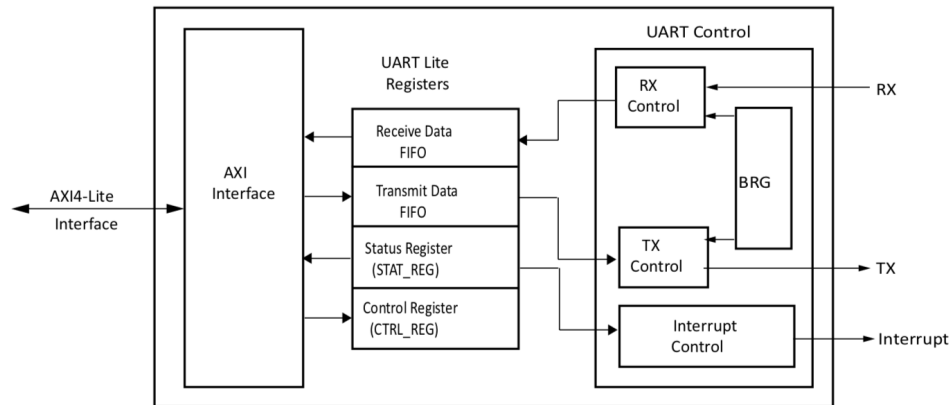

Figure 4.14: From Xilinx [109], we have the top-level digram for the UART Lite IP. This is identical to the top-level digram for the UART 16550, but its resource utilization is lower since the BRG (Baud Rate Generator) is locked in during firmware synthesis and implementation.

4.7.2.3 Introduction: Xilinx Vivado^{TM} Design Tools

In general, FPGAs are programmed in hardware verification languages (sometimes called hardware design languages) to produce what is called a *bitstream*. The design suite that was used for AugerPrime comes from Xilinx and consists of two important products, Vivado^{TM} and the Source Developer's Kit (SDK). Vivado^{TM} is used to create the bitstream, and allows both graphical block diagram editing as well as hardware design language editing. To stymie confusion, we will use 'hardware design language' to refer to any available hardware verification language whether HDL, Verilog or VHDL.

The general work flow in Vivado^{TM} is to start a project, which prompts the user to select the type of hardware to be used. Vivado^{TM} knows the descriptions for all Xilinx programmable devices and imports a specification describing the devices abilities and port structure. The port structure essentially lists all of the available pins on the device. In some cases such as when working with their Zynq^{TM} evaluation board, the ZedBoard^{TM}, the software already knows how the port structure maps onto physical ports, and can take advantage of this. In the case of custom boards such as the UUB, all ports must be manually mapped. This involves the use of a *design constraints file*, in the .xdc format. These files allow specification of a logic standard and an external port to be mapped to a physical pin on the Zynq^{TM} chip.

All of the other functions of Vivado^{TM} pertain specifically to what is happening on the FPGA. This is where the *block diagram* comes in. In Vivado^{TM}, there are 3 interconnected units you have to manage. The block diagram is the portion where one can invoke IP, which stands for Intellectual Property, but effectively these are firmware versions of hardware devices, i.e. ICs. The block diagram allows users to spatially arrange and interconnect IPs and is where most of the work in Vivado^{TM} is done. Most standard electronics have a corresponding IP and utilizing it involves invoking the IP, connecting its buses and ports

to the other elements of the block diagram, customizing the IP in whatever way is needed, and finally, if needed, mapping it to an AXI (Advanced eXtensible Interface) address. AXI is the primary communication bus in many ARM devices, and certainly is the most widely used in the UUB.

Beyond the constraints file and the block diagram, the last facet of the design we must implement is the top level wrapper. This is a piece of hardware design language (Vivado™ supports both Verilog and VHDL generation) which relates the external connections of the constraints file to the internal connections of the block diagram. It is in here that we can, for example, connect a signal line to an external port. We can also map such a signal line to split off into an LED for monitoring, or connect it to two ports to allow a signal to be used by multiple IP in the block diagram.

4.7.2.4   Implementing the UART

As elucidated in subsubsection 4.7.2.3, we have to look after three tasks to implement the UART in the UUB. First, we instantiate the block diagram and make our customizations, then we add a connection to the top level wrapper, and finally we make a external connection with the constraints file. Adding the UART Lite is as simple as right clicking on the block diagram in Vivado™ and searching for it. Upon adding it, we are greeted with Figure 4.16. Here we can select the baud rate and parity bits we want, which are respectively 9600 and 0 for the SSR-6Tf. We can then accept these parameters and move to the AXI Mapping tab, where Vivado's™ built-in address mapping tool does a fine job of taking care of this part of the task automatically. Next, we have to make the connection from the PL to the top level wrapper. Vivado™ can do this automatically, but it will destroy any custom
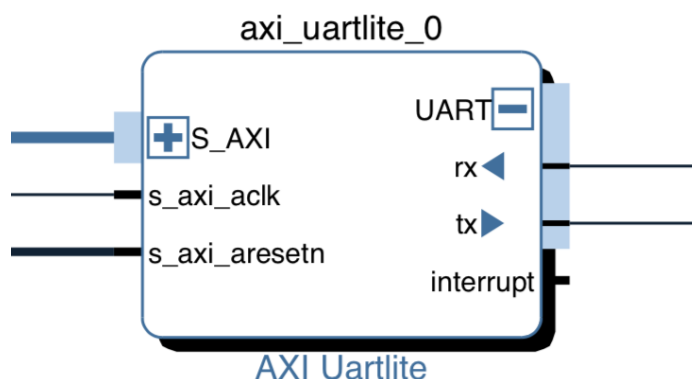
Figure 4.15: This is the UART Lite's block on the block diagram. It is shown here to give some idea of what the IPs in the block diagram look like and to show the ports that must be connected for it.
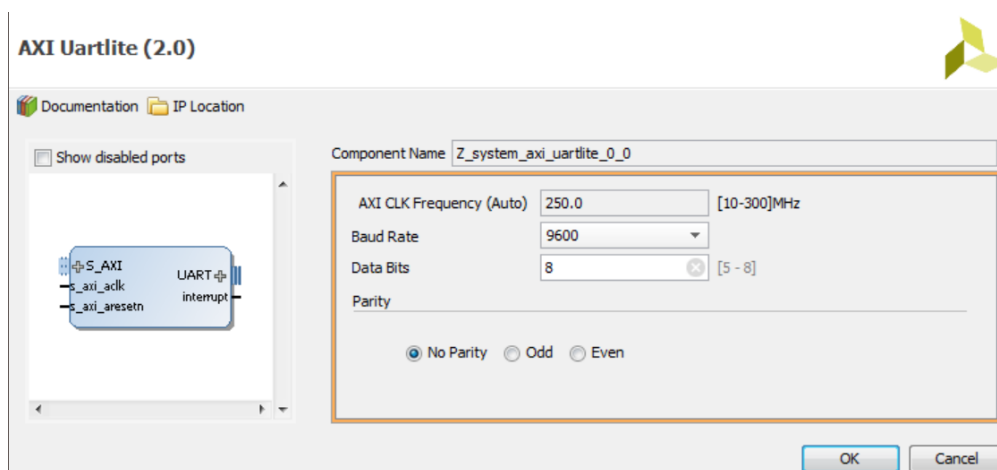
Figure 4.16: In this Vivado™ window, we are given our rather sparse options to customize the UART Lite. Those shown here are the correct configuration for the SSR-6Tf GPS receiver.

firmware written into the wrapper, so care must be taken here[15]. In figure Figure 4.17, we

have shown a snippet of the top level wrapper for a custom Vivado™ project [16].

In the UUB's wrapper, the UART Lite's I/O signals are called *GPS_RX* and *GPS_TX* for

---

[15] We would endorse saving your wrapper elsewhere, allowing Vivado™ to automatically generate the wrapper, and then re-inserting your custom PL code appropriately. This can sometimes be a moderately confusing and lengthy task and in my experience, it must be done in one sitting or you will certainly forget some detail of what you have done. This is also an excellent place to take good notes on what you have done.

[16]The UUB's top level wrapper is not publicly available, although the hardware description file which contains binarized information from the top level wrapper is available through the uub-firmware github, `https://github.com/auger-prime-sde/uub-firmware`, under work package 5.

GPS receive and transmit, while in Figure 4.17, the analogous signals are *sin* and *sout*. To clarify, *signal* is the correct hardware design language term for this object, which represents a physical signal carrying wire. Once we have the wrapper taken care of, we need only attend to our final task: correctly constraining the signals to physical ports or pins on the Zynq^{TM} chip. This is as simple as adding two lines per signal to the constraints file (usually just called *constraints.xdc*). One line specifies the signaling standard, e.g. 1.8V or 3.3V, and the other specifies the "package pin" or physical site on the Zynq^{TM} chip that the signal is to end up in. In the case of integrating the SSR-6Tf into the UUB, these lines are:

```
set_property PACKAGE_PIN V10 [get_ports GPS_TX]

set_property PACKAGE_PIN V9 [get_ports GPS_RX]

set_property IOSTANDARD LVCMOS33 [get_ports GPS_TX]

set_property IOSTANDARD LVCMOS33 [get_ports GPS_RX]
```

With this, we can produce firmware for a working UART Lite in the context of GPS receiver communications as implemented in the firmware for AugerPrime. Specifications of the ports and pins are publicly available in Bouvier (2014) [110].

```
  Diagram  X    Address Editor  X    Z_system_wrapper.v  X    sysgpsc.xdc  X
  C:/TIM/CTA_messingAround3/CTA_messingAround3.srcs/sources_1/imports/hdl/Z_system_wrapper.v
181      wr_en);
182  inout [14:0]DDR_addr;
183  inout [2:0]DDR_ba;
184  inout DDR_cas_n;
185  inout DDR_ck_n;
186  inout DDR_ck_p;
187  inout DDR_cke;
188  inout DDR_cs_n;
189  inout [3:0]DDR_dm;
190  inout [31:0]DDR_dq;
191  inout [3:0]DDR_dqs_n;
192  inout [3:0]DDR_dqs_p;
193  inout DDR_odt;
194  inout DDR_ras_n;
195  inout DDR_reset_n;
196  inout DDR_we_n;
197  inout FIXED_IO_ddr_vrn;
198  inout FIXED_IO_ddr_vrp;
199  inout [53:0]FIXED_IO_mio;
200  inout FIXED_IO_ps_clk;
201  inout FIXED_IO_ps_porb;
202  inout FIXED_IO_ps_srstb;
203
204  input REF1PPS;
205  input TEST1PPS;
206  input GRES;
207  output LED;
208  output LED2;
209  output LED3;
210
211 //  input [31:0]din; //might need to comment these back in, really not sure what
212 //  input [1:0]gpio_1_tri_i;
213 //  input [31:0]gpio_2_tri_i;
214 //  output [0:0]gpio_tri_o;
215  input sin;
216  input sin_1;
217  output sout;
218  output sout_1;
219  input rd_en;
220  input wr_en;
221
```

Figure 4.17: This Vivado^{TM} window shows the Verilog wrapper used in a custom project. As mentioned in the text, the UUBs wrapper is not publicly available, however it is similar with a bit more going on.

4.7.2.5   Physical Connection

On the UUB and in the various testing enclosures, the GPS unit is physically mounted on a 60mm×40mm pad, with a 20 pin connector. The technical drafting diagram is shown in Figure 4.18 and the physical board layout in Figure 4.19. The specification of most importance to the work described above is in Table 4.2 [17]. While these images referenced above are for the M12M, the analogous images for the SSR-6Tf are only cosmetically different.

---

[17]Ultimately the routing of these pins is handled by the integration team at LPSC Grenoble, headed by Eric Lagorio, one of the lead electrical engineers on AugerPrime.

The pinout is the same and so are the features up to minor performance differences [18].
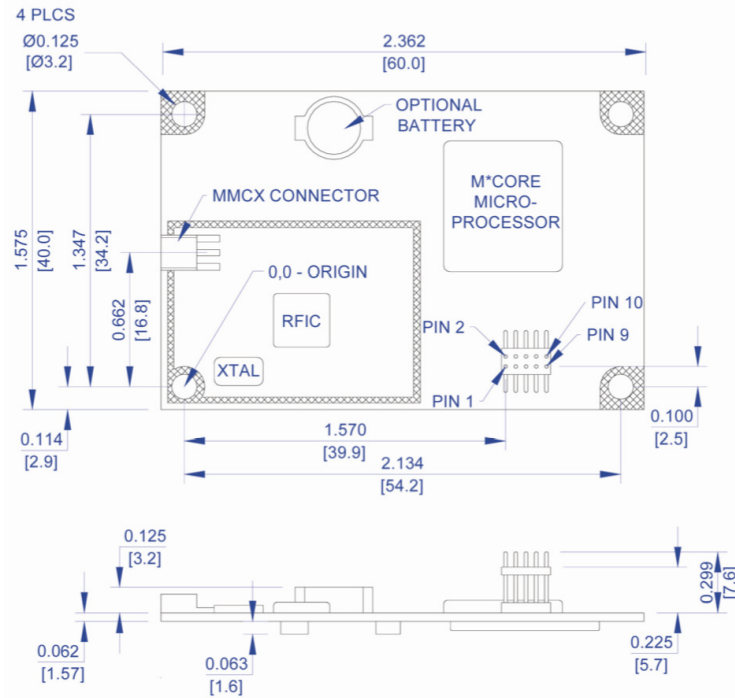
Figure 4.18: Here is a technical drafting of the M12M and SSR-6Tf's shared footprint from i-Lotus [107].
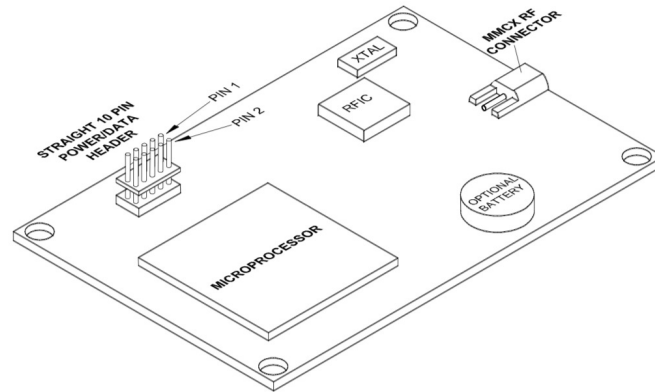
Figure 4.19: This rendering shows the physical layout of the M12M receiver, but more importantly for our purposes, it shows the pin ordering which corresponds to Table 4.2. From i-Lotus [107].

---

[18]The SSR-6Tf users manual is still in draft form and not ready to be released to the public, so we cannot show the corresponding diagrams for the SSR-6Tf. We in the HEA group at CWRU have been granted a copy for use, but not distribution.

Pin #	Signal Name	Description
1	TxD1	Transmit Data (3V logic)
2	RxD1	Receive Commands (3V logic)
3	+3V PWR	Regulated 3Vdc Input
4	1PPS	1 pulse-per-second output
5	Ground	Signal and Power common
6	Battery	Optional External Backup
7	Reserved	Not currently used
8	RTCM In	RTCM correction input
9	Antenna Bias	3V-5V antenna bias input
10	Reserved	Not currently used

Table 4.2: This table from i-Lotus [107] shows the all-important pinout. These pins are then routed, some into the FPGA and some into other components on the UUB.

4.7.3 GPS Communications Software

After implementing the hardware to communicate with the GPS, we need to write the software[19]. The software really consists of 2 main parts: the initial message handling is operated through what we call a *parser*, and then the data is put into classes using *packing functions*[20].

4.7.3.1 Parser

Essentially, the parser performs two functions; it figures out when a message has been received and then it figures out what type of message it was. These functions are accomplished in the UUB in the main loop of a file called *gps_main.c*. The *select* command, shown in-situ in Figure 4.20 is the real workhorse of the parser. This command takes five arguments: the first is an integer describing the place in the second argument, which is a file descriptor with read privileges. The second and third arguments are unused in the AugerPrime software as they wait for a file descriptor to become available for writing and to monitor a file descriptor for exceptions, respectively. The final argument gives a timeout, which in our case is set to 10s. When the *select* command is activated, it passes the point

```
while(1){
  aux=readfds;
  timeout.tv_sec=10;
  timeout.tv_usec=0;
  if(select(gps_rec_fd+1,&aux,NULL,NULL,&timeout)==1){
    usleep(80000);
    read_flag=1;
    do{
  msg_buff_len=gps_receiver_read_msg(gps_rec_fd,msg_buff,read_flag);
```

Figure 4.20: Shown here is the main loop of the AugerPrime GPS message parser. Note that the select statement effectively does all of the work of monitoring for changes in the file descriptors.

---

[19]This was a joint effort between Ricardo Sato and the author.

[20]We should note now, that a careful reading of subsection 4.6.4 is an absolute prerequisite for understanding what will be said in this subsection.

in the readable file descriptor to *gps_receiver_read_msg*, which takes in the message until it encounters a valid message. This is specified by a carriage return and line feed, proceeded by a correct XOR checksum of the entire message. Due to the nature of XOR, it is sufficient to take the XOR of the message from the very beginning, before the first "@" and go all the way through the checksum. This should result in a value of 0x00, indicating that the message XOR matched the checksum sent by the receiver. Once this is true, *gps_receiver_-process* is activated, with the data from the most recent message. This function, shown in Figure 4.21, infers the message type from the available messages implemented. Finally, the type is passed back into the parser loop where a set of if/else statements activate the appropriate packing function and extract the needed data. The ultimate product of this loop is the *gps_process_info* function, which globalizes the information parsed and packed here for the rest of the UUB software to make use of.

```c
int gps_receiver_process(unsigned char *msg,int len)
{
  if(GPS_RECEIVER_MIN_MSG_LEN<=len){
    if(memcmp((msg+2),"Hb",2)==0){
      return(1);
    }
    if(memcmp((msg+2),"Hn",2)==0){
      return(2);
    }
    if(memcmp((msg+2),"Ha",2)==0){
      return(3);
    }
    if(memcmp((msg+2),"Bb",2)==0){
      return(4);
    }
    if(memcmp((msg+2),"As",2)==0){
      return(5);
    }
    if(memcmp((msg+2),"Cj",2)==0){
      return(6);
    }
    return(0);
  }
  return(1);
}
```

Figure 4.21: Here is the process which decides what type of message has been received. It simply asks if the first two letters match any of the pre-assigned messages. If not, it passes back to the *gps_main* process and resets the receiver's messaging parameters.

4.7.3.2 Classes and Packing

Once the *gps_receiver_process* function has decided on the correct message type, the message gets passed back to the parser process in *gps_main* where it is packed into a struct for readout. Packing these data into structs allows sensible access to any of the parameters contained within, and is widely considered the "correct" way to handle serialized data. From a programming point of view, creating these structs and packing functions is relatively straightforward.

To create a struct and packing function for a particular message, we need only look in the M12M User's Guide [107], and create the appropriate fields for each piece of data we desire to store. In keeping with good practices, we have made all of the data from any of the possible messages available, except for bit fields, which are generally not collected in housekeeping. As an example of making a struct, we can look back at Figure 4.10, and compare it with Figure 4.22. In Figure 4.10, we see that each parameter has a big-endian set of bytes that it corresponds to. The largest of these is an int, and the smallest is a char, although as mentioned before, some chars represent bit fields. The reasons why having various field types forces us to use a packing function are detailed in Appendix E.

```
struct ha_msg
{
  char month;
  char day;
  short year;
  char hours;
  char minutes;
  char seconds;
  unsigned int frac;

  int lat;
  int lon;
  int height;
  int msl;
```

```
int gps_receiver_get_info_Ha(struct ha_msg *p,unsigned char *msg)
{
  p->month=msg[0];
  p->day=msg[1];
  p->year=msg[2]<<8 | msg[3];
  p->hours=msg[4];
  p->minutes=msg[5];
  p->seconds=msg[6];
  p->frac=(msg[7]<<24) | (msg[8]<<16) | (msg[9]<<8) | msg[10];
  p->lat=(msg[11]<<24) | (msg[12]<<16) | (msg[13]<<8) | msg[14];
  p->lon=(msg[15]<<24) | (msg[16]<<16) | (msg[17]<<8) | msg[18];
  p->height=(msg[19]<<24) | (msg[20]<<16) | (msg[21]<<8) | msg[22];
  p->msl=(msg[23]<<24) | (msg[24]<<16) | (msg[25]<<8) | msg[26];
```

Figure 4.22: Left: This is the code snippet where the struct for an @@Ha message is contained, note the parity in field with Figure 4.10. Right: here we have the corresponding packing function, which takes the contiguous data in the serialized Motorola Binary message, and maps it to the correct fields in the struct for easy data access.

To handle repeated fields in a message, the repeated data types are instantiated in their

own struct (an example is given in Figure 4.23), which becomes a substructure of the greater message. These are packed by a *for* loop, with the offset from the beginning of the message noted (an *int* was used for convenience; it is now a *const*).

```
Channel Data
       i = SVID                    0...37
       m = mode                    0...8
             where:                0 = Code Search        5 = Message Sync Detect
                                   1 = Code Acquire       6 = Satellite Time Available
                                   2 = AGC Set            7 = Ephemeris Acquire
                                   3 = Freq Acquire       8 = Available for Position
                                   4 = Bit Sync Detect
       s = signal strength         0...255
       l = IODE                    0...255
       dd = channel status (16 bits)
```

```
struct ha_sat
{
  char satid;
  char trakmode;
  unsigned char sigstrength;
  unsigned char iode;
  short chanstat;
};
```

```
int i;
int offset=52;
for(i=0;i<12;i++){
  p->sats[i].satid=msg[offset+1+6*i];
  p->sats[i].trakmode=msg[offset+2+6*i];
  p->sats[i].sigstrength=msg[offset+3+6*i];
  p->sats[i].iode=msg[offset+4+6*i];
  p->sats[i].chanstat==(msg[offset+5+6*i]<<8) | msg[offset+6+6*i];
}
```
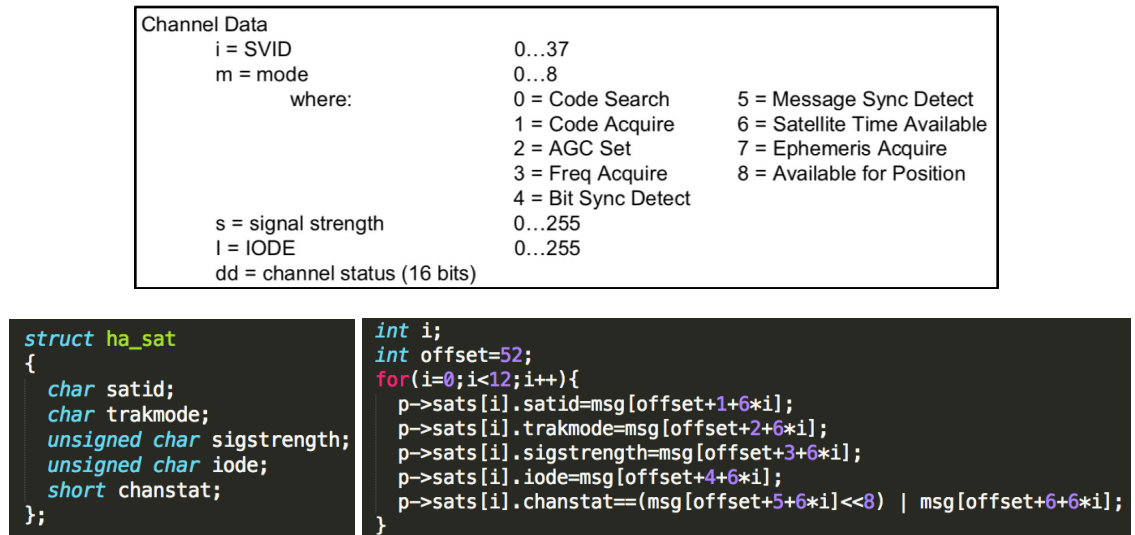
Figure 4.23: Top: the manual entry from i-Lotus [107] for the repeated field of channel-wise satellite data. Bottom Left: the struct corresponding to one channel. Bottom Right: the packing function which iterates over the contiguous data, putting it into the channel structs.

Taking a look at Figure 4.22, we can see in the right panel, the packing function corresponding to the struct in the left panel and the message in the User's Manual shown in Figure 4.10. Each of the allowed messages has such a corresponding struct created in the same way: (1) examine the manual (i-Lotus [107]), (2) create a struct corresponding to the data in the message, (3) write a packing function for the message, (4) add the message header to *gps_receiver_process* and finally (5) add the packing function and any relevant readout to the *while* loop in *gps_main* (under the select block). The messages implemented currently are likely the only ones which will be relevant for future use, and are shown in Table 4.3.

Message Header	Content Summary
@@Ha	Time/Position/Receiver Status, Detailed
@@Hn	T-RAIM + Sawtooth Information
@@Hb	Short Time/Position
@@Bb	Satellite Statuses and Information
@@As	Position Hold Parameters
@@Cj	Receiver Identification Message (all ASCII, no packing needed)

Table 4.3: Above are the message types implemented in the finalized version of the UUB GPS code.

4.8 Conclusions: Reliability of the Integrated GPS and Operating System

At time of writing, the PetaLinux operating system and the GPS software and firmware have been running in the Auger Engineering Array (EA) for more than two years. Beyond having no reported issues with either system, we have not received any request for further functionality. Recently a Critical Design Review was initiated for the complete preproduction UUB. Although a number of concerns and issues remain to be resolved prior to production runs, no concerns were raised regarding either the GPS or its integration into the UUB. Based on these results we can say that the integration and board bring-up were successful.

Chapter 5

Time-Tagging and GPS Performance for AugerPrime

In this chapter, we will investigate the timing performance of two aspects of the Auger-Prime upgrade. To set the stage, we review some timing basics. After this, we move on to the topic of GPS receiver selection for AugerPrime. We motivate this endeavor, discussing the need for a new receiver selection, and then describe the absolute, relative and temperature timing tests done in the lab at CWRU for the M12M and SSR-6Tf receiver models. To conclude this section, we justify our recommendation of the SSR-6Tf, which has been accepted and acted on by the collaboration.

The measurement for the time resolution of the AugerPrime SD prototypes in the engineering array is then described. We give our methods for two different measurements of the time resolution, one of which requires us to calculate a physical error factor for arrive directions of triggering showers, and the other which involves using a cable to time synchronization pulses between two detectors. We find the time resolution to be ∼8ns and that

both methods are in agreement.

5.1   Time-Tagging Basics

In order to keep time in any context, we must first start with a clock. Even back to the days of Galileo or John Harrison, clocks have been based on the simple principle of an oscillator, keeping time by 'ticking' at a precise rate. In today's context, most clocks are kept by electronic oscillators with a variety of methods employed to generate the signal. In the case of the UUB, this is an Abracon ABLJO-V-120.000MHZ-T2 ([97]), which uses the 3rd overtone of a quartz crystal to keep time. By modulating the voltage across this VCXO, we can control its vibrational frequency.

The field of air shower physics usually requires its time-tagging hardware to accomplish two tasks. First, it must time the occurrence of events, which can be broadly described in terms of timing the rising edge of a logic pulse. Generally, the main task of trigger hardware and firmware is the generation of a logic pulse which corresponds to the time of the trigger, often offset by some known amount. The second task of a time-tagging module in an experiment such as Auger, is to accomplish frequency distribution, i.e. synchronizing clocks at a very fine level across many stations.

To synchronize stations for time-tagging, Auger uses the GPS constellation's space-borne bank of atomic clocks. Using the techniques outlined in subsection 4.6.1, the GPS receiver can calculate the time of a GPS second to a handful nanosecond's accuracy. These GPS seconds are timed by the receiver's 1PPS, which comes in the form of a logic pulse. By counting the ticks of an oscillator from when a 1PPS comes in to when an event occurs, and then from the event to the next 1PPS, we can interpolate to find the point in the second

when the event occurred. This process is encapsulated in Equation 5.1 below.

$$t_e = \frac{N_e}{N_{i+1,PPS} - N_{i,PPS}}.$$

(5.1)

Here, we have N_e, the number of counts up to an event, $N_{i,PPS}$ the counter value at the preceding PPS, and $N_{i+1,PPS}$, the counter at the current (i.e. subsequent) PPS. This gives us t_e, the time (in seconds) of an event within a GPS second, and then we can simply use the telemetry messages from the receiver or the counter from the time tagging module to put together the GPS second[1]. With these pieces of information, we can put together a full decimal time-tag of when an event occurs, allowing reconstruction of air showers and comparison with data from other experiments.

To get the highest possible accuracy in time tagging, we also need to apply the clock granularity message. This has two effects on Equation 5.1: first, the clock granularity messages (subsubsection 4.6.4.2) from the proceeding and current PPS events give the information that we need to adjust how many seconds the number of calibration counts, $N_{cal} = N_{i+1,PPS} - N_{i,PPS}$, represents. The preceding sawtooth message then tells us how the end of second represented by the PPS needs to be shifted relative to GPS time to be correctly aligned.

We address the first problem by applying a correction to N_{cal}:

$$N'_{cal} = (N_{i+1,PPS} - N_{i,PPS}) \left(1 + 10^{-9}(\Delta t_{curr} - \Delta t_{last})\right),$$

where the Δt_i correspond to the preceding and current saw tooth (*last* and *curr*, respectively) in nanoseconds, as it is delivered by the receiver. The second problem is fixed by

---

[1]There is some difficulty in finding the number of leap seconds to date, but this can be requested from the receiver periodically.

simply adding the preceding sawtooth (Δt_{last}) to the event time. This gives us a new t'_e of:

$$t'_e = \frac{N_e}{(N_{i+1,PPS} - N_{i,PPS})\left(1 + 10^{-9}(\Delta t_{curr} - \Delta t_{last})\right)} * 10^9 + \Delta t_{last}. \qquad (5.2)$$

In this formulation, we have added the factor of 10^9 to put t'_e into units of nanoseconds, which is the usual convention for the type of timing accuracy measurements we will be doing in this work. This is in line with the framework set forth by the original group who worked on GPS verification for the initial Auger construction (documented in Meyer and Vernotte (2001) [111]). This concludes the introduction to time-tagging relevant for this chapter however, the discussion on timing standards and accuracy is continued in section 6.1.

5.2 GPS Receiver Selection Testing

As mentioned above, Auger requires precise timing synchronization amongst the SD stations. This is accomplished via GPS timing boards, which are available from the manufacturer as OEM products. Several manufacturers offerings were considered, and based on work completed at CWRU in the early 2010's, the i-Lotus M12M [112] was initially chosen (before the author began working with the Auger Collaboration). Unfortunately, when pressed for a quotation and delivery schedule in 2017, i-Lotus was unable to meet the volume of our request. Accordingly, we required a new GPS unit. We ultimately selected the Synergy SSR-6Tf after the testing described in this section.

For this section, we will first briefly discuss of the two models of receivers. After this, the first test, measuring absolute timing, is outlined and then results are shown. We then discuss the method for the relative timing and temperate dependence tests, and show their

results. Finally, we describe our recommendation to purchase the SSR-6Tf and justify it from the data and context.

5.2.1   Specifications: i-Lotus M12M vs. the SSR-6Tf

The SSR-6Tf is a GPS receiver made specifically for accurate timing applications. The unit is made by Synergy Systems (the same company that acts as the reseller for i-Lotus in the US). The SSR-6Tf is designed to optionally operate in "compatibility mode" that provides nearly identical functionality to that of the i-Lotus M12M with comparable or better timing performance. In compatibility mode, the SSR-6Tf is designed to function as a "drop in replacement" for the M12M.

Both units specify operating temperatures in the range -40C to 85C. The i-Lotus draws 123 mWatts power while the SSR-6Tf draws 155 mWatts. Both receivers have the same form factor, pinout and antenna connectors.

By definition, timing accuracy here relates to the accuracy of timing for 1 Pulse-Per-Second (1PPS) output pulses. For every measurement reported here, we apply granularity corrections (so-called 'sawtooth', as explained in subsubsection 4.6.4.2) which are reported for each 1PPS pulse via internal serial line from the GPS receiver.

5.2.2   Initial Bench Tests: Absolute GPS Timing

Our first tests were conducted on the bench. GPS antenna signals were routed into the laboratory from a roof-top antenna. Both the SSR-6Tf and the M12M have been exercised for many 100's of hour in our laboratory without fault.

We find that the SSR-6Tf as delivered by Synergy cold starts into Motorola compatibility mode. For both cold and warm starts, the SSR-6Tf is much faster than the M12M (usually seconds vs. minutes).
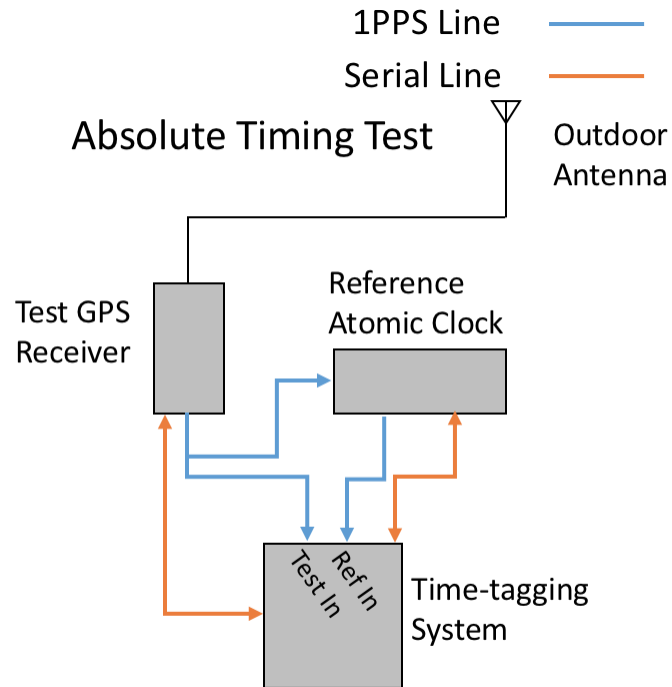
5.2.2.1 Test Stand for Absolute GPS Timing

Figure 5.1: Bench test schematic of time-tagging system to measure absolute timing of the Synergy Systems SSR-6Tf vs. the i-Lotus M12M GPS receivers.

For our initial tests, we compare absolute time-tagging performance of the SSR-6Tf against that of the M12M over a range of time-scales using a GPS-disciplined Atomic Clock. Before each test, the atomic clock is trained for several hours to the GPS constellation's timing. The clock we are using is a FS725 Rubidium frequency standard from Stanford Research Systems, which is trained (disciplined) by the GPS constellation. Individual 1PPS produced by the disciplined FS725 provide an accurate time standard to better than a few 100s of picoseconds over timescales from seconds to days.

Figure 5.1 shows a schematic diagram for the absolute timing test stand. Our bench test timing test stand, TIM (see section 6.2 for more details), is based on a ZedBoard™ and runs our own 750 Mhz time-tagging system firmware through a GPIO interface. Operation of the board is controlled by a Standalone linux operating system script. This script controls

the time tagging system firmware, data logging, and serial communications with both the GPS and the atomic clock. On a pulse-by-pulse basis we measure the timing differences between the 1PPS from the atomic clock and the corrected 1PPS from each receiver. The final result will be a plot of the standard deviation of the pulse arrival times as a function of timing window scanned over to calculate the variance. The derivative of this is directly related to the Allen Deviation.
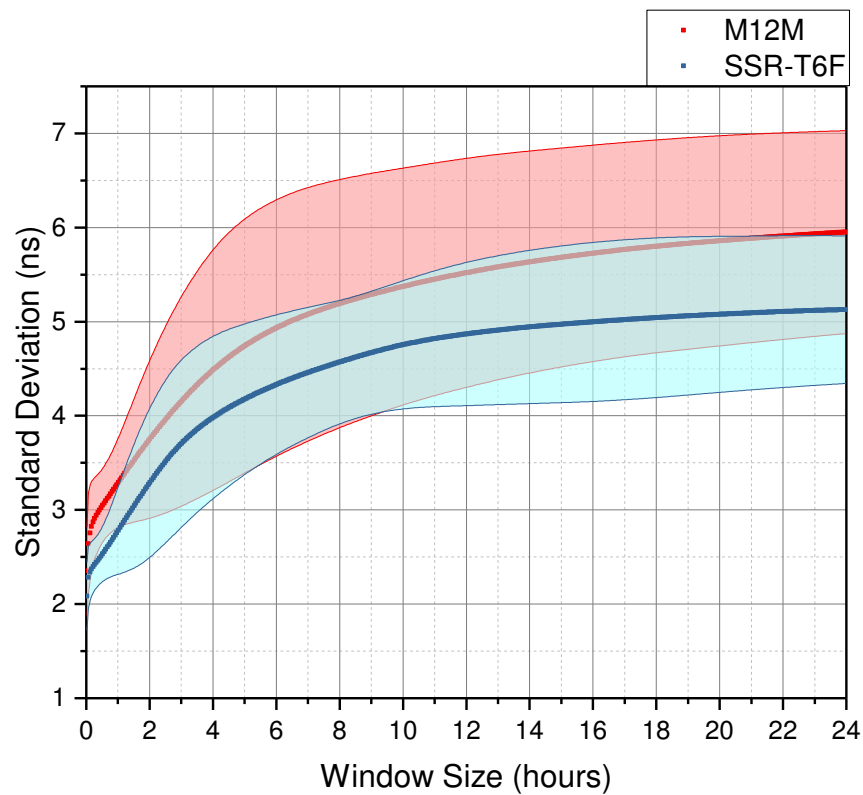
5.2.2.2   Results for Absolute GPS Timing

Figure 5.2: Absolute time-tagging: The standard deviation of each receiver's 1PPS compared to the 1PPS of the FS725 atomic clock. The i-Lotus M12M is shown in red. The Synergy Systems SSR-6Tf is shown in blue. The shaded areas correspond to a 1-σ error region for each receiver.

Figure 5.2 shows our results for absolute time-tagging of sawtooth-corrected 1PPS GPS timing for the M12M vs. the SSR-6Tf. We show the RMS difference between the measured 1PPS for each receiver vs. the atomic clock standard as measured on timescales ranging from a few seconds to 24 hours.

In terms of short-term timing accuracy (timescales less than a few minutes), the SSR-6Tf reports a timing precision of 2.3 ns while the M12M reports a timing precision of 2.8 ns. Absolute timing errors gradually increase for both receivers over timescales of a few hours, presumably due to drifts in the electron content of the ionosphere. Over many hours, the long-term absolute timing resolution of the SSR-6Tf is generally better than about 5 nanoseconds while the M12M is closer to 6 nanoseconds. We find that on any time-scale from seconds to hours, the SSR-6Tf outperforms the M12M by approximately one nanosecond for absolute timing accuracy. We note that in terms of the performance specifications required for the the AugerPrime upgrade (e.g., better than 8 nanoseconds absolute timing) both the SSR-6Tf and the i-Lotus M12M meet the required specification.

5.2.3   Relative GPS Time-Tagging

Although absolute timing accuracy is an important performance parameter for GPS tim-ing, in the field at Auger the *relative timing* between two receivers is the more important quantity, since only timing differences between receivers will impact the reconstruction. To verify relative timing accuracy, we developed a second test stand configured to accept telemetry data from two GPS units and to compute the difference in arrival times of their (sawtooth corrected) 1PPS time signals. We also use this configuration to explore possible temperature dependence of the arrival time of the 1PPS on the order of nanoseconds.

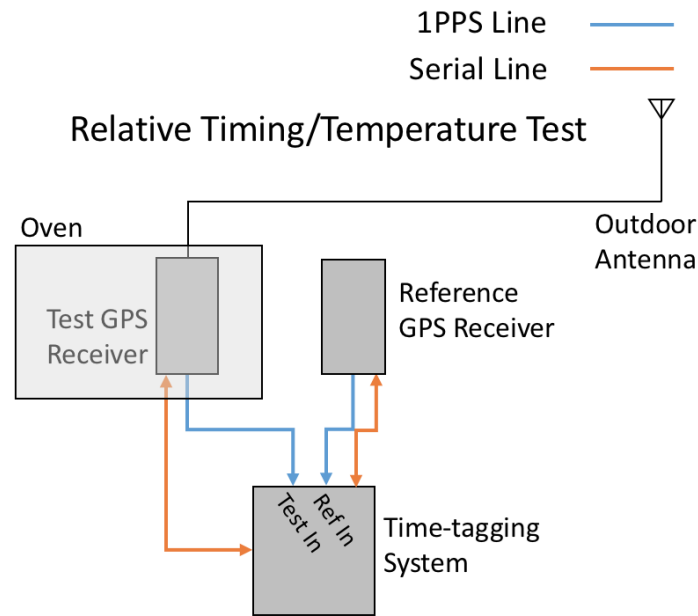### 5.2.3.1 Test Stand for Relative GPS Timing and Temperature Dependence

Figure 5.3: Bench test schematic of time-tagging system to measure relative timing of the Synergy Systems SSR-6Tf vs. the i-Lotus M12M GPS receivers, including temperature dependence.

Figure 5.3 shows the schematic setup of our test for relative timing and temperature dependence. The configuration is closely matched to that used for previous time-tagging calibration and temperature dependence measurements conducted and reported by the CWRU group [98]. For relative GPS time-tagging we select two GPS receivers of the same model and then measure the relative arrivals of their sawtooth-corrected 1PPSs. These measurements provide a series of time differences from which timing precision (standard deviation) can be computed over long-duration tests.
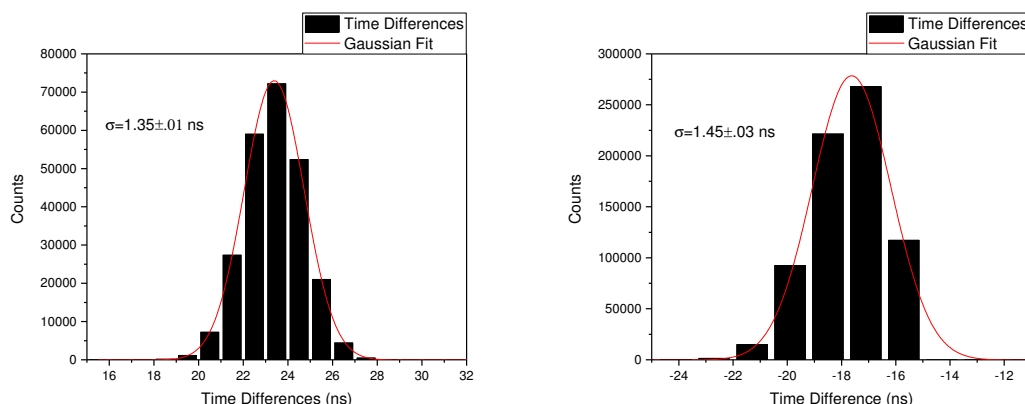
5.2.3.2 Temperature and Relative Timing Test

Figure 5.4: Left: the relative timing test for the SSR-6Tf. Right: the relative timing test for the M12M. These are the time differences of the sawtooth corrected

In comparison to the ~1.45ns standard deviation in the time differences of the M12Ms, the SSR has a standard deviation of 1.3 ns, which is the instrumental uncertainty of the 750 Mhz time-tagging system used for this testing [98]. It is worth noting that it's possible the SSR-6Tf may have an even better timing accuracy, but this result reaches the limits of our test stands ability to time-tag it. To reiterate, these time differences represent two identical receivers timed against each other with the TIM unit (section 6.2). The data have been filtered for counter rollover and sawtooth rollover outliers which account for less that 1% of the second by second time differences recorded. These measured timing precisions are in line with the manufacturer's presented figures [113, 112].

For the temperature dependence testing, the profile of temperatures is derived from a study of the weather at the Auger site [98]. The M12M shows some temperature dependence during extreme temperature ramps, but the mean of its 1PPS arrival time does not depend on temperature, In contrast, the SSR-6Tf does show a small direct dependence on the temperature. In spite of this, the performance of the SSR is still better than that of the M12M when put under thermal variations.
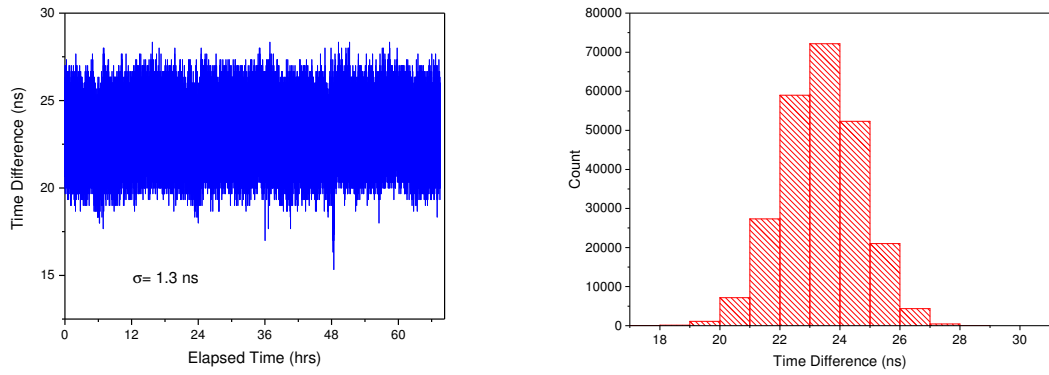
Figure 5.5: Left: A time series test of the SSR-6Tf's performance with sawtooth correction measured against another sawtooth corrected SSR-6Tf. The data has been corrected for counter and sawtooth rollover errors. Right: A histogram of the time series.
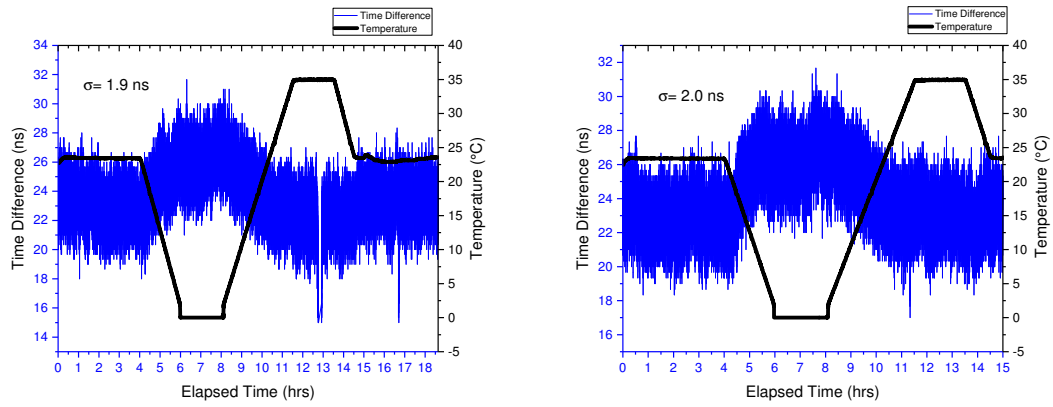

Figure 5.6: Two runs from the temperature testing of the SSR-6Tf.
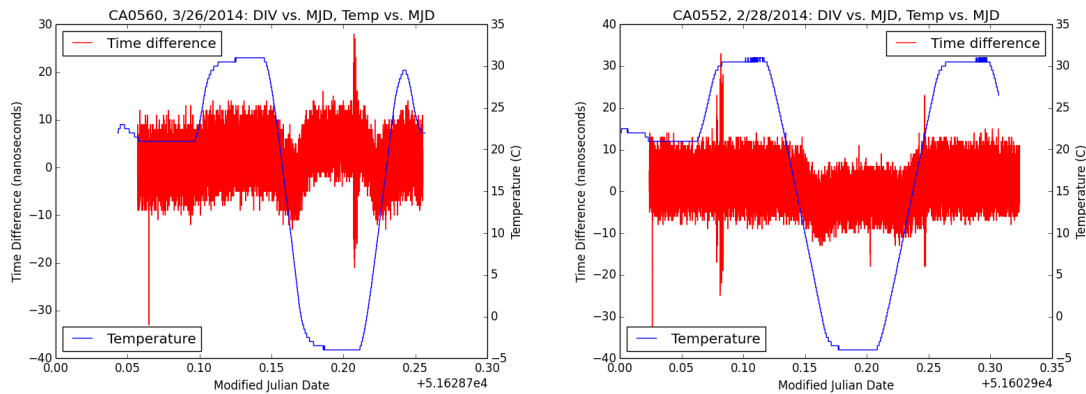

Figure 5.7: Plots from Dan Brandt's original analysis of the M12M's temperature dependence, available in Brandt [98].

5.2.4 Conclusions

We have tested the M12M and SSR-6Tf receivers using two different methods, first looking at the absolute timing accuracy over different time scales and second looking at the timing of each receiver relative to another receiver of the same model under temperature variation. The absolute accuracy test shows that, while both models perform satisfactorily for use in AugerPrime, the SSR-6Tf has slightly better performance over all time scales. According to our relative timing tests, the SSR-6Tf outperforms the M12M; where the M12M has a standard deviation in time differences of ~1.45 ns, the SSR has a relative timing accuracy of ~1.3 ns or better. The temperature dependence testing showed that M12M receivers have a temperature dependence during large temperature changes but stabilize when the temperature does. On the other hand, the SSR-6Tf shows a direct temperature dependence where the mean of the time differences depends on the temperature. If the SSR was less accurate, this could cause issues, but even with this direct dependence the SSR is more accurate than the M12M.

All things considered, the SSR-6Tf is a newer receiver which outperforms the M12M in all relevant parameters. Both receivers have now been tested in the field for >3 months without fault, and so our recommendation for the AugerPrime upgrade is to purchase the SSR-6Tf receivers when Synergy is ready to deliver their final version.

5.3 AugerPrime Time-Resolution

An essential parameter of shower reconstruction for Auger is the time resolution of its detectors. This affects directional reconstruction amongst other observables, and if the time-resolution is low enough it may contribute to helping determine the muonic content of incident air showers (see Espadanal [95] and section 4.1). With the integration of Auger-

Prime completed, the time-tagging module has been added in a similar way to other IP (see section 4.7). At this point, it has been running continuously for over two years in the engineering array without issue, thereby confirming its reliability. This then leaves us to confirm the timing performance of the whole system, end-to-end.

In this section, we will first outline the basic facets of the time-tagging system. We will then move on to a discussion of each of two methods to measure the time resolution of the upgraded stations and the results of the measurements. First we will discuss a method using only air showers, and then a method using a synchronization cable. Finally, we will compare the results of these two tests and their uncertainties.

The timing tests will be performed in the stations Trak Jr., Clais Jr. and Peteroa Jr., which have been dubbed the 'timing triplet' and are some of the engineering array stations most frequently used for UUB verification and testing. A map of their positions is shown in Figure 5.8.
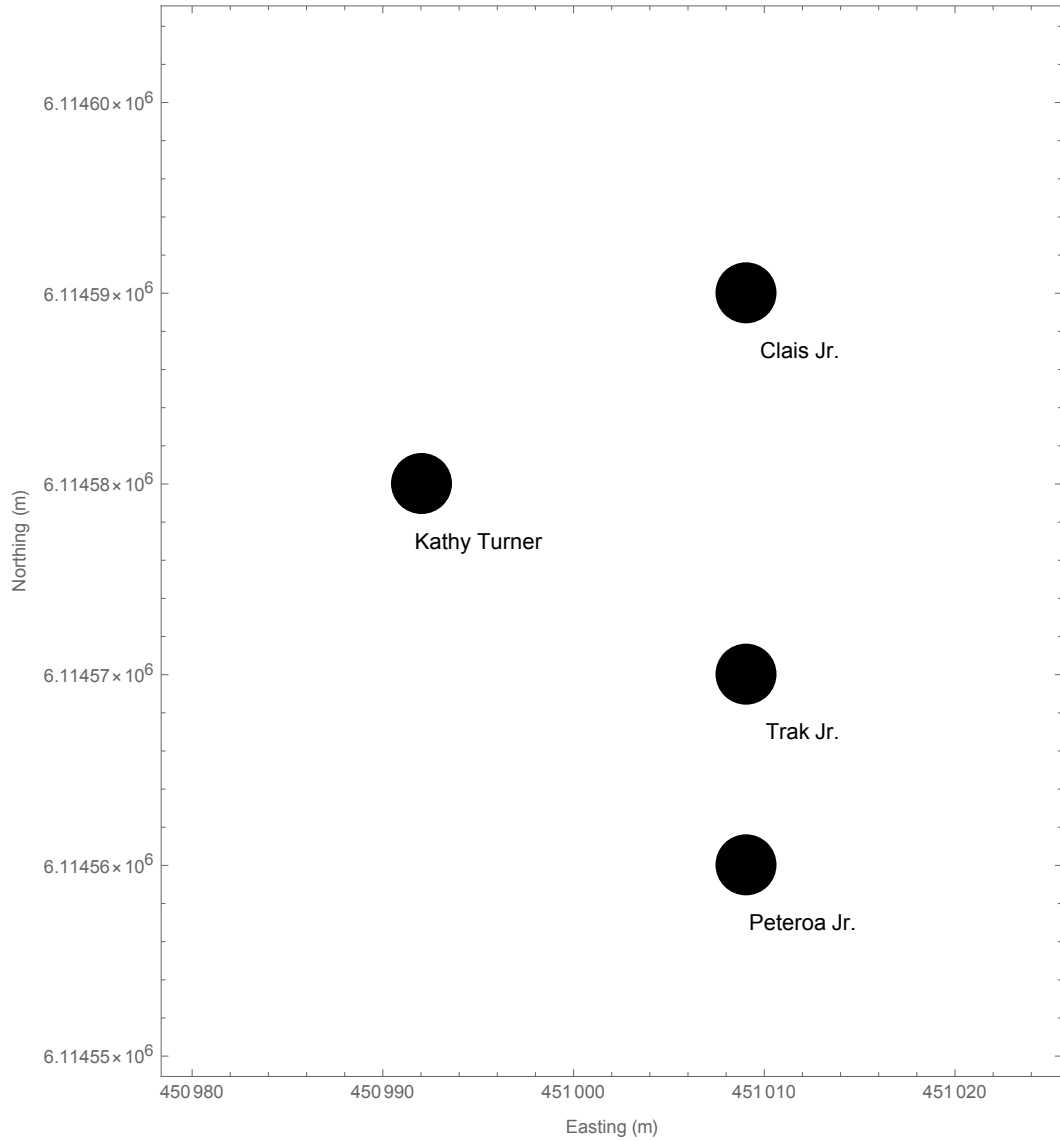
Figure 5.8: Here we have a map of the relative positions of the stations in the EA. The distance between Trak Jr. and Clais Jr. is ~20m and the distance between Trak Jr. and Peteroa Jr. is 11m (this is somewhat obscured since we have used the global Auger coordinate system on these axes).

5.3.1 Time-Tagging Specifications and Ports

In comparison to the time-tagging module for TIM, shown in Figure 6.4, the final time-tagging module for AugerPrime integrates all of the GPIO connections into internal wires and registers and makes some of the calculations that are done in post-processing for TIM.