Sensitivity of AugerPrime to the masses of ultra-high-energy cosmic rays

For the attainment of the academic degree of

Doctorate in Science

from the
Karlsruher Institut für Technologie (KIT)
and the
Universidad Nacional de San Martín (UNSAM)

the accepted

Dissertation

of

M.Sc. David Schmidt

of Fürth

Day of the oral examination: 29.06.2018

Referee: Prof. Dr. Dr. h.c. Johannes Blümer

Co-referee: Prof. Dr. Alberto Etchegoyen

Advisor: Dr. Markus Roth

Sensitivity of AugerPrime to the masses of ultra-high-energy cosmic rays

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)
und der
Universidad Nacional de San Martín (UNSAM)

genehmigte

Dissertation

von

M.Sc. David Schmidt
aus Fürth

Tag der mündlichen Prüfung: 29.06.2018

Referent: Prof. Dr. Dr. h.c. Johannes Blümer

Korreferent: Prof. Dr. Alberto Etchegoyen

Betreuer: Dr. Markus Roth

Sensitivity of AugerPrime to the masses of ultra-high-energy cosmic rays

Tesis presentada para optar por el título de

Doctor en Astrofísica

del Instituto de Tecnología "Prof. Jorge A. Sábato" de la
Universidad Nacional de San Martín (UNSAM)
y del
Karlsruher Institut für Technologie (KIT)

por

M.Sc. David Schmidt
de Fürth

Fecha de la defensa oral: 29.06.2018
Director: Prof. Dr. Dr. h.c. Johannes Blümer
Co-director: Prof. Dr. Alberto Etchegoyen
Colaborador: Dr. Markus Roth

Abstract

A number of conundrums endure regarding the very highest energy cosmic rays measured at Earth. For one, their sources remain shrouded in mystery with attempts at back-tracing these particles to their origins confounded by deflections in galactic and extra-galactic magnetic fields. The region in energy for the emergence of the extra-galactic contribution remains unclear, and each hypothesis is accompanied by an additional set of complications to address. The suppression of the flux at the highest energies, once explainable by photo-pion production through interactions with the cosmic microwave background, has been placed under scrutiny by measurements of an increasingly heavier composition above the ankle, a region which was thought to be dominated by protons just a little over a decade ago. Relatedly, the relative roles of photo-disintegration and the maximum rigidity achievable by nearby extragalactic sources in defining features of the energy spectrum are unclear. A deficit in the number of muons predicted by simulations of extensive air showers using the LHC-tuned hadronic interaction models also remains unexplained. Knowledge of the energy dependencies of the elemental abundances of cosmic rays would be highly beneficial in answering each of these questions and numerous others.

The Pierre Auger Observatory is the largest of the two experiments currently measuring ultra-high-energy cosmic rays and aims to address the aforementioned outstanding issues in the field. In this pursuit, it is performing a large-scale detector upgrade, known as AugerPrime, in order to reconstruct the mass of cosmic rays on a per event basis by means of measuring the correlated magnitude of the muonic component of extensive air showers. The placement of a scintillator atop each water-Cherenkov detector is the highlight feature of the detector upgrade with the principle concept being the exploitation of the differing responses of the two sub-detectors to the electromagnetic and muonic shower components. The work presented in this dissertation includes (i) analyses used to inform the design of the scintillator, (ii) the upgrades to the observatory's software framework in order to accommodate the detector upgrade, (iii) the development and tuning of comprehensive scintillator surface detector simulations, (iv) the assessment of a matrix-based algorithm used to reconstruct the mass of cosmic rays from measurements of the two surface sub-detectors, and (v) an application of this algorithm in selecting protons from the background of two different astrophysical scenarios for the purpose of performing charged particle astronomy.

The tools and methods developed in this work are of pivotal importance for future analyses in the context of AugerPrime.

Zusammenfassung

Eine Reihe wichtiger Fragen zur Natur der kosmischen Strahlung höchster Energien sind nach wie vor unbeantwortet. Zum einen sind die Quellen dieser Strahlung weitgehend unverstanden, da Versuche, die Teilchen zu ihrem Ursprung zurückzuverfolgen, von galaktischen und extragalaktischen Magnetfeldern erschwert werden. Der Energiebereich des Spektrums, in dem der Anteil der extragalaktischen Strahlung dominant wird, ist bisher unklar, und jeder Erklärungsversuch bringt eine Reihe von Problemen mit sich. Die Unterdrückung des Flusses bei höchsten Energien, die einst durch Wechselwirkungen der kosmischen Strahlung mit Photonen der Kosmischen Hintergrundstrahlung erklärt wurde, wird dadurch in Frage gestellt, dass Experimente bei Energien über der Knöchelregion des Spektrums eine Verschiebung der Zusammensetzung hin zu schwereren Kernen messen. Vor nur etwas mehr als einem Jahrzehnt dachte man, diese Region sei von Protonen dominiert. In diesem Zusammenhang sind auch der relative Einfluss von Photodisintegration und der maximalen Rigidität, die von nahen, extragalaktischen Quellen erreicht werden kann, auf die entscheidenden Charakteristika des Spektrums unklar. Zudem ist der Mangel an Myonen in Luftschauersimulationen, die hadronische Wechselwirkungsmodelle verwenden, die an Daten des LHC angepasst wurden, nach wie vor unverstanden. Kenntnisse darüber, wie die Zusammensetzung der Kosmischen Strahlung von der Energie abhängt, wäre äußerst hilfreich, um diese und weitere Fragen zu beantworten.

Das Pierre Auger Observatorium ist das größere zweier Experimente, die aktuell in Betrieb sind und deren Ziel es ist, die kosmische Strahlung höchster Energien zu messen und die zuvor genannten Fragen zu beantworten. Zu diesem Zwecke wird momentan ein groß angelegtes Upgrade durchgeführt, das unter dem Namen AugerPrime bekannt ist. Das Ziel von AugerPrime ist es, die Massenzahl des Teilchens, das einen Luftschauer erzeugt, für jeden Luftschauer einzeln bestimmen zu können. Um dies zu erreichen, wird die korrellierte Myonenkomponente des Luftschauers gemessen. Kernstück des Upgrades ist die Anbringung eines Szintillators über den Wasser-Tscherenkovdetektoren. Mit dessen Hilfe kann die Tatsache ausgenutzt werden, dass die elektromagnetische und die myonische Luftschauerkomponente unterschiedliche Signale in den zwei unterschiedlichen Detektoren hinterlassen. Die Arbeiten, die in dieser Dissertation beschrieben werden, umfassen (i) Untersuchungen zum optimalen Design des Szintillators, (ii) Anpassungen in der Software, die das Upgrade nötig macht, (iii) die Entwicklung und Anpassung umfassender Simulationen des Szintillator-Oberflächendetektors, (iv) die Entwicklung eines matrixbasierten Algorithmus zur Rekonstruktion der Masse des Primärteilchens aus den Messungen der zwei Detekoren, und (v) eine Anwendung dieses Algorithmus, bei der Protonen vom Untergrund zweier unterschiedlicher astrophysikalischer Szenarien unterschieden werden, um Astronomie mit geladenen Teilchen zu ermöglichen.

Die Werkzeuge und Methoden, die in dieser Arbeit entwickelt wurden, sind von entscheidender Bedeutung für zukünftige Analysen im Zusammenhang mit AugerPrime.

Resumen

Un cierto número de enigmas aún perduran respecto a los rayos cósmicos de muy alta energía medidos en la Tierra. En primer lugar, el misterio acerca de las fuentes que los causan, debido a intentos frustrados de rastrear estas partículas hacia su origen por culpa de las desviaciones en campos magnéticos galácticos y extra-galácticos. Además, sigue sin ser claro el rango de energías donde inicia la contribución extra-galáctica, y cada hipótesis planteada es acompañada por un cierto número de complicaciones adicionales a resolver. La supresión del flujo en las energías más altas, alguna vez explicada por la producción fotón-pión por interacción con la radiación del fondo de microondas, ha sido puesta bajo escrutinio por mediciones de composición progresivamente más pesada arriba del codo, una región que se pensó dominada por protones desde hace poco más que una década atrás. Relacionado con esto, tampoco son claros aún los roles relativos de la foto-desintegración y la rigidez máxima alcanzable por fuentes extragalácticas cercanas en la definición de las características del espectro de energía. También sigue sin ser explicado el déficit en el número de muones predichos por simulaciones de lluvias atmosféricas extendidas usando modelos de interacción hadrónicos ajustados por el LHC. Por todo esto, conocer la dependencia con la energía de las abundancias de elementos de rayos cósmicos sería muy beneficiosa para responder cada una de estas incógnitas y muchas otras.

El Observatorio Pierre Auger es el experimento más grande de los dos que actualmente miden rayos cósmicos ultra-energéticos y pretende abordar los temas anteriores aún no respondidos. En esta búsqueda, se está implementando una actualización y mejora de detectores, llamada AugerPrime, para poder reconstruir la masa de los rayos cósmicos evento por evento, por medio de la medición de la magnitud correlacionada a la componente muónica de la lluvia de partículas atmosférica extendida. La colocación de centelladores encima de los detectores Cherenkov es la característica distintiva de la actualización, cuyo concepto es el aprovechamiento de la discrepancia de señales medidas entre estos dos sub-detectores respecto a la componente muónica y electromagnética de las lluvias. El trabajo que se describe en esta tesis incluye (i) el análisis del diseño del centellador de superficie, (ii) las mejoras hechas al software de análisis de datos del Observatorio para poder incluir estos detectores, (iii) el desarrollo y ajuste de la simulación completa de los centelladores, (iv) la evaluación de un algoritmo basado en una matriz para la reconstrucción de la masa del rayo cósmico a partir de las mediciones de los dos sub-detectores de superficie, y (v) la aplicación de este algoritmo para seleccionar protones del fondo, para dos escenarios astrofísicos diferentes, con propósito de realizar astronomía de partículas cargadas.

Las herramientas y métodos desarrollados en este trabajo son fundamentales para futuros análisis en el contexto de AugerPrime.

Contents

CHAPTER 1

Cosmic Rays and their Detection

Cosmic rays are primarily relativistic charged particles originating elsewhere in the universe but which find their way to Earth's atmosphere at a rate of approximately 1000 per square meter per second. The flux of these particles is dominated by protons, but heavier ionized nuclei contribute to a degree which varies depending on the energies in question.

1.1 Cosmic Ray Overview

The differential flux of the very highest energy cosmic rays, the most energetic of which are the focus of this work, is shown as a function of energy in Fig. 1.1. A number of spectral features may be observed, which relate to the the nature and distribution of galactic and extragalactic sites of particle acceleration in addition to interactions occurring during propagation.

The steepening of the spectrum at an energy of approximately 3×10^{15} eV is referred to as the "knee" and is most often attributed to galactic supernova remnants (SNRs) reaching their maximum acceleration potential [2]. The maximum acceleration energy of proton is reached first and corresponds to the onset of the softer spectrum at the energy of the knee. Increasingly heavier nuclei then progressively reach their maximum acceleration potential, which is proportional to mass, resulting in a maximum energy for iron that is 26 times higher than that of proton, or around 8×10^{16} eV.

The next clearly discernible feature in the spectrum is the "second knee" at an energy of approximately 10^{17} eV. Results of the KASKADE experiment showed a softening in the spectrum of heavier elements at this energy and a subsequent hardening in the spectra of lighter elements [3]. It is tempting to interpret the former as the SNR iron knee and the latter as the onset of the extragalactic component of cosmic rays; however, the existence of an additional population of galactic accelerators capable of accelerating protons to 10^{17} eV (and subsequently iron to 2.6×10^{18} eV), referred to as Population B, has also been postulated [4].

At approximately 3×10^{18} eV, the energy spectrum becomes harder at a feature referred to as the "ankle." If the extragalactic flux of cosmic rays is dominated by

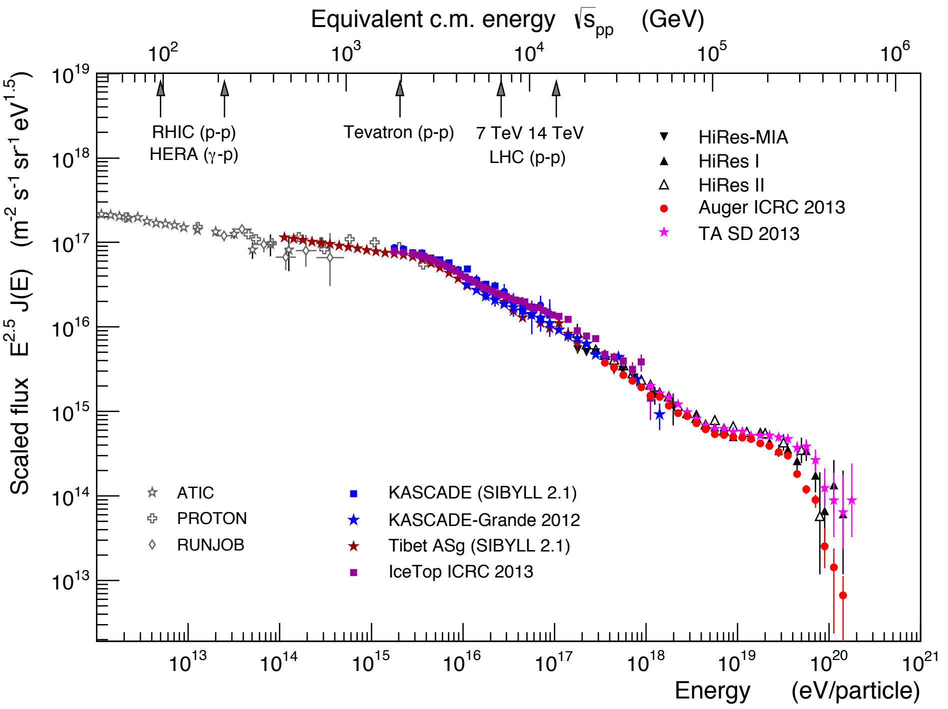


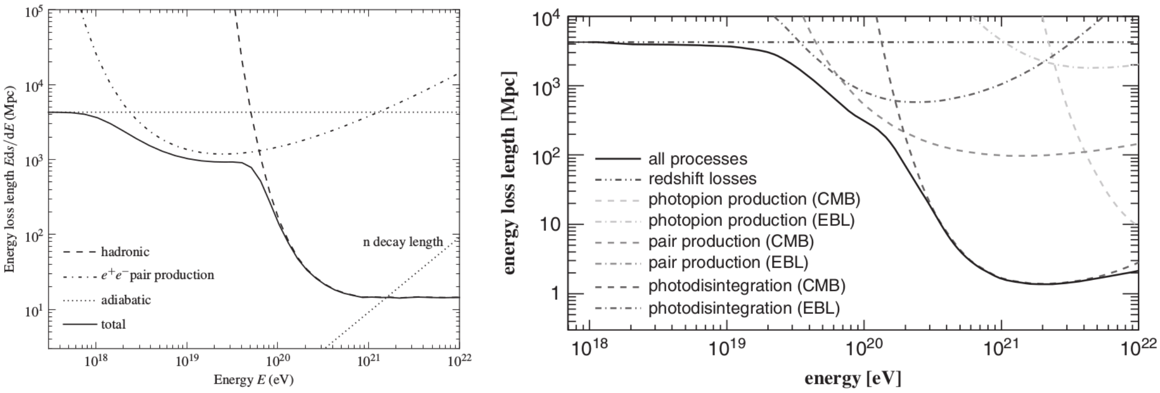
Figure 1.1: Differential energy spectrum of cosmic rays [1]. The flux is scaled by a factor $E^{2.5}$ to increase the visibility of spectral features.

protons, this can be interpreted as a pile-up effect due to electron-positron pair production [5], which becomes the dominant process for proton energy losses above approximately 2×10^{18} eV. Alternative theories suggest that the ankle marks the onset of the extragalactic component of the cosmic ray flux, which requires the existence of the aforementioned Population B of galactic accelerators. Regardless, it is widely assumed that the flux above the ankle derives from extragalactic sources due to the absence of any directional evidence for production of these cosmic rays in the galactic plane and the apparent absence of sources with sufficient acceleration potential within our galaxy.

Above 4×10^{19} eV, the spectrum is strongly suppressed. It remains unclear if this suppression can be primarily attributed to propagation effects or to extragalactic sources reaching their maximum acceleration potential. The former case would involve a significant impact of energy loss processes including photo-disintegration of nuclei via delta resonance and photo-pion production for protons, both due to interactions with photons of the cosmic microwave background radiation (CMB). The energy loss length of protons is depicted in Fig. 1.2a, and the same is shown in Fig. 1.2b for iron nuclei.

1.1.1 Sources and Propagation

General characteristics of sources and their distributions can be surmised with theory constrained by the measured energy spectrum and an understanding of propagation processes; however, individual sources for very high energy cosmic rays have never been directly observed. This can be largely attributed to the fact that cosmic rays are heavily deflected by galactic and extragalactic magnetic fields during



(a) **(b)**

Figure 1.2: (a) Energy loss length for protons for various processes [6]. (b) Energy loss length for iron nuclei for various processes [7].

transit from their point of origin to Earth. This renders it exceptionally difficult to back-trace individual cosmic rays to potential sources. Nonetheless, for very low charge primaries, such as proton, trajectories begin to approach rectilinearity near the upper end of the energy spectrum. Unfortunately, such high energy particles and nuclei rapidly undergo the discussed energy loss processes during transit. What results is referred to as the "GZK horizon," referring to the fact that beyond approximately 100 Mpc, any protons with energies exceeding approximately 10^{20} eV at their source will have lost the majority of this energy. Additionally, most nuclei with such energies will have undergone photo-disintegration within this radius. All of this effectively results in an opacity of the universe to the highest energy cosmic rays and a limited distance from Earth in which sources could be directly observed before the trajectories of cosmic rays have been excessively scrambled by magnetic fields.

Aside from attempts at directly observing the unknown sources of ultra-high energy cosmic rays (UHECRs), significant efforts have been made to classify astrophysical objects capable of accelerating charged particles to such extreme energies. On the first order, the maximum energy to which a particle can be accelerated by a specific type of source is considered roughly proportional to the radius of the acceleration region R, the strength of the magnetic field B, and the charge Z of the particle or nuclei. Various potential sources classified by this criterion are shown in a so-called "Hillas plot" in Fig. 1.3.

1.2 Extensive Air Showers

Due to their low flux, which is on the order of one primary per square kilometer per century in the region of the flux suppression, direct detection of ultra-high-energy cosmic rays is not feasible. Instead, their properties are reconstructed from the observation of the extensive air showers they induce in Earth's atmosphere.

An extensive air shower (EAS) most frequently begins when a cosmic ray hadronically interacts with a nitrogen or oxygen molecule high in the atmosphere. What follows is a shower that can be broken down into three components, namely the hadronic, electromagnetic, and muonic. The core of the shower consists of high en-

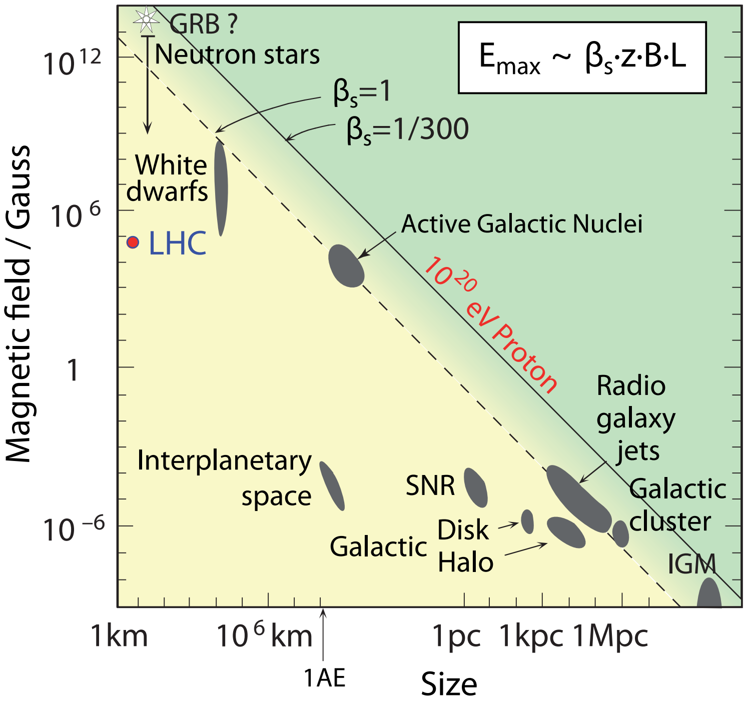


Figure 1.3: Hillas plot depicting potential astrophysical sites for the acceleration of particles and nuclei [8]. Strong candidates for 10^{20} eV protons have yet to be identified.

ergy hadronic interactions, which feed the other two shower components. Neutral pions and eta particles produced in the hadronic core can decay to two photons, each of which initiates an electromagnetic sub shower propagated through alternating stages of electron positron pair production and bremsstrahlung. Charged kaons and pions can decay to produce neutrinos and muons, the latter of which comprises what is referred to as the "muonic component" of EASs. For a schematic depiction of the three shower components, see Fig. 1.4.

Existing methods for detection of EASs from UHECRs include the observation of fluorescence emission from nitrogen molecules excited by passing air showers as well as measuring the energy deposit of the particles reaching Earth's surface. Both of these methods will be discussed in the context of the Pierre Auger Observatory (Auger) in Section 1.3. Measurement of the radio emission of showers has also been recently validated as a viable detection method.

1.3 Pierre Auger Observatory

The Pierre Auger Observatory [10] is an UHECR observatory located at an elevation of approximately 1400 m near Malargüe, Mendoza Province, Argentina. At the heart of the observatory is a hybrid detection mechanism consisting of the observatory's Surface Detector (SD) [11] and Fluorescence Detector (FD) [12]. The former consists of a 3000 km^2 surface array of 1600 water-Cherenkov detectors (WCDs), which samples the lateral distributions of EAS particles arriving at ground in order to obtain an estimate of the "shower size." The FD consists of 24 fluorescence telescopes overlooking the SD array from four sites and observers the longitudinal development

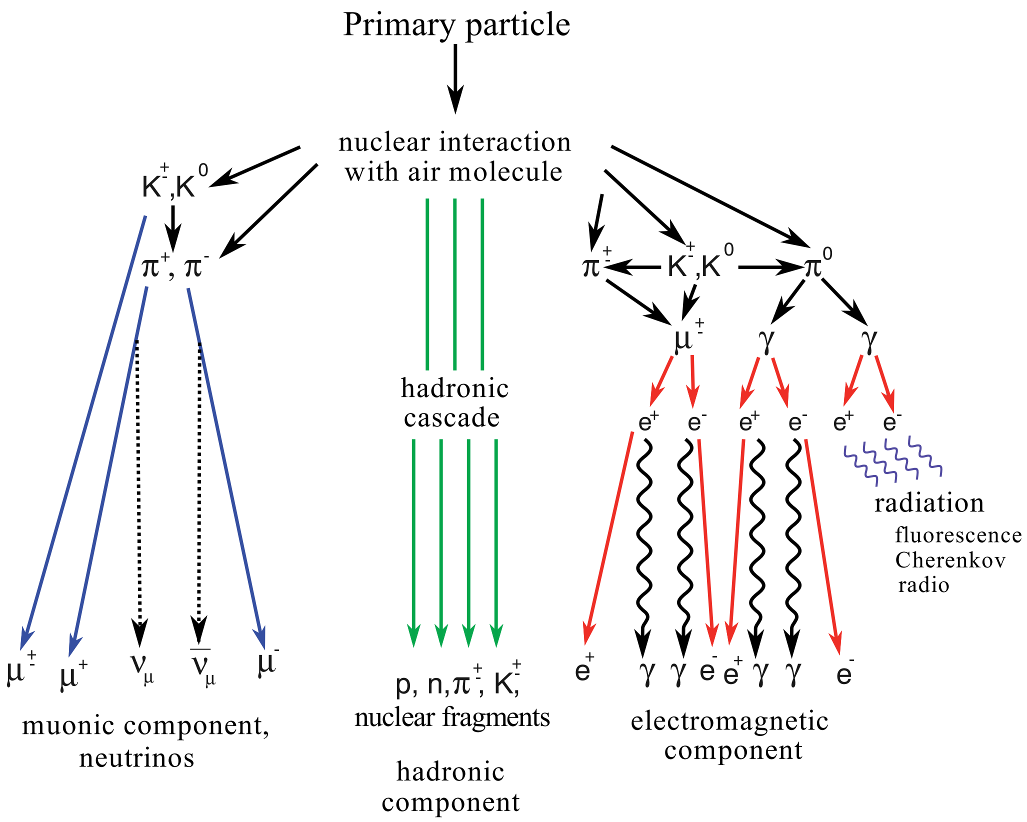


Figure 1.4: Schematic depiction of the hadronic, electromagnetic, and muonic components of EASs (taken from [9]).

of showers via their fluorescence emission to obtain a calorimetric estimate of the primary cosmic ray energy. The energy estimates of the FD are used to calibrate the SD such that the energy of primary particles can be reconstructed by the SD alone without extensive dependence on simulations. The observatory is schematically depicted in Fig. 1.5.

1.3.1 Surface Detector

The 1600 WCDs that comprise the SD of Auger are tessellated hexagonally such that each WCD is equidistant from six neighboring stations with a grid spacing of 1500 m. The array is fully efficient in the detection of cosmic rays above approximately 3×10^{18} eV for events with zenith angles of up to 60°. Each WCD and its associated electronics and communication systems is referred to as an SD "station," one of which is depicted in Fig. 1.6a. Each station consists of a cylindrical tank with a radius of 1.8 m and a diffuse, highly-reflective liner inside. These tanks are filled with ultra-pure water up to a depth of 1.2 m, making for a total active detection volume of slightly over 12 m^3. Relativistic charged particles from the electromagnetic and muonic components of EASs produce Cherenkov photons inside this water volume, which are collected by three 9-inch photo-multiplier tubes (PMTs). This signal is amplified and digitized resulting in a time trace of ADC counts. The integral of this distribution is calibrated to the average signal produced by a vertical muon passing through the center of the station, a quantity derived from its measured relationship

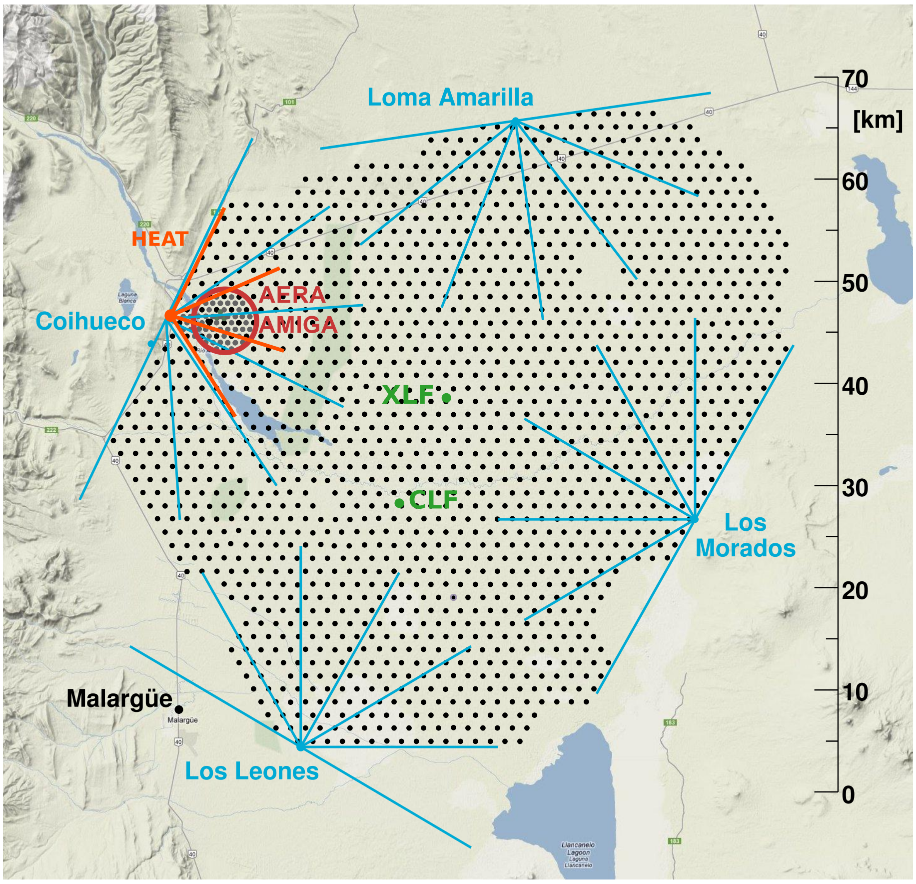


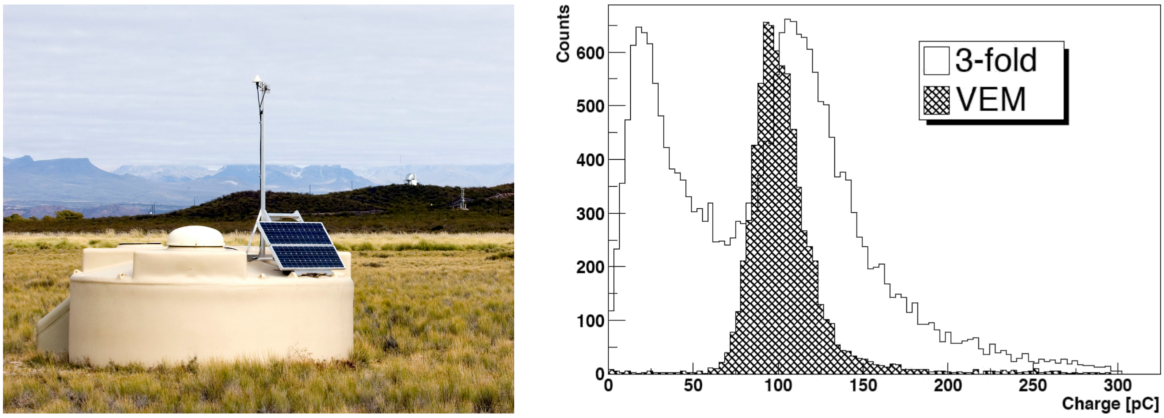
Figure 1.5: Schematic depiction of the Pierre Auger Observatory. Each black dot represents one SD station. The blue lines represent the fields of view of the four sites housing fluorescence telescopes. The orange lines represent additional high-elevation telescopes for observation of showers induced by lower energy primaries. The locations of the observatory's two laser facilities used for calibration of the fluorescence telescopes (the CLF and XLF) are also depicted. The denser array with a spacing of 750 m between surface detector stations is also depicted. Each of these stations in the denser array will be accompanied by a 30 m^2 buried muon counter for the direct detection of the muonic component of EASs [13].

with the peak of the histogram of signals from atmospheric particles continuously measured in one minute intervals. Such a histogram is depicted in Fig. 1.6b.

The particle footprint at Earth's surface is sampled for each EAS, and the resulting station signals are fitted as a function of their distance from the reconstructed shower axis. Once this is performed, the signal at a given reference distance, referred to as the "shower size," is corrected for the shower's arrival angle, and the energy of the primary cosmic ray is obtained via the calibration with the calorimetric measurements of the FD. More details on this reconstruction procedure will be given in Section 1.4.

1.3.2 Fluorescence Detector

The FD consists of 24 telescopes split evenly between 4 locations on the perimeter of the SD array. One such location is pictured in Fig. 1.7a. Each location has a 30°



(a) (b)

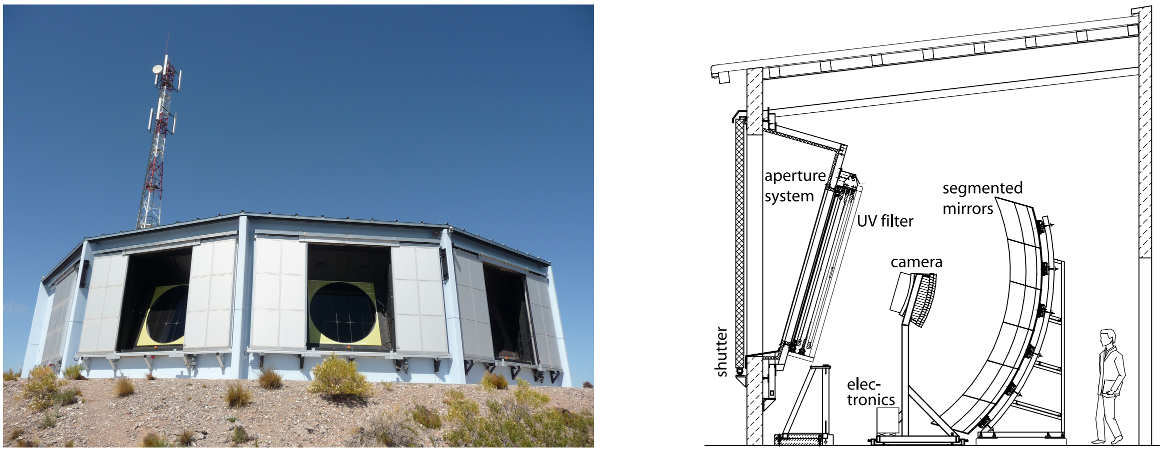
Figure 1.6: (a) Picture of an SD station. (b) Sample calibration histogram of an SD station [14]. Signals are calibrated to the response of one vertical, through-going, muon (hatched "VEM" histogram) via the relationship with the histogram of signals from all the atmospheric particles triggering the detector in a one minute time span ("3-fold" histogram).

by 30° field of view in azimuth and elevation. Photons produced via fluorescence emission by an EAS as it traverses the atmosphere enter a telescope bay and are reflected by a 13 m^2 segmented mirror into the telescope camera, which consists of a 22 row by 20 column grid of PMTs. A schematic of a telescope bay layout is given in Fig. 1.7b. The traversing shower leaves a down-going track in the camera from which the longitudinal energy deposit of the shower can be reconstructed. The integration of this longitudinal profile effectively provides the energy of the primary cosmic ray. The maximum of this profile is referred to as the *depth of shower maximum*, X_{max}, which is a composition sensitive parameter. Heavier nuclei develop higher in the atmosphere and exhibit less fluctuations in their depths of shower maximum than lighter primaries.

As the FD can only perform measurements on dark, clear nights, its duty cycle in limited to approximately 15%. Since it uses the atmosphere as a calorimeter, precise atmospheric monitoring is required. Each of the four FD sites is equipped with a lidar, infrared cloud cameras, and a weather station. Additionally, there are two laser facilities, the Central Laser Facility (CLF) and Extreme Laser Facility (XLF) centrally located in the SD array. The known power and intensities of these lasers allows for calibration of each FD telescope that views the up-going tracks they produce. Additionally, information from the Global Data Assimilation System (GDAS) is used to inform atmospheric conditions at the event time during reconstruction.

1.3.3 Upgrades

A number of upgrades to the basic hybrid detection design of Auger have already been performed or are currently under development. One such upgrade is a 27 km^2 portion of the array which is instrumented with an additional 49 SD stations such that the grid spacing in this region is a denser 750 m. This denser array reaches full detection efficiency at 3×10^{17} eV for showers with zenith angles of less than 55°.



(a) (b)

Figure 1.7: (a) Fluorescence detector site in which individual telescope bays may be observed [10]. (b) Schematic of a telescope bay [10].

30 m^2 scintillator detectors buried at a depth of 2.3 m will be also be deployed at 61 of the 71 positions (23.5 km^2) of the denser array, allowing for the direct detection of the muons in EAS up to energies of about the ankle of the cosmic ray energy spectrum. Together, the 750 m SD array and the buried muon counters are referred to as Auger Muon Detectors for the Infill Ground Array (AMIGA), a station of which is depicted in Fig. 1.8. In the same denser region of the array, antennas have also been deployed to measure the radio emission of air showers, which is a newly flourishing detection technique that theoretically has a 100 % duty cycle and does not suffer from the uncertainties in hadronic interaction models relating to the muon component of EAS. This denser array is also overlooked by three additional high-elevation fluorescence telescopes, which lower the energy threshold of the FD. Given these upgrades, the potential exists and is being realized to measure, for the same EASs, the footprint at ground with the SD, the muon content with the buried scintillation counters, the radio emission with the aforementioned antennas, and the longitudinal profile with the high-elevation extension of the FD. This promises to be a rich data set for multi-hybrid measurement and calibration at and below the ankle of the cosmic ray energy spectrum.

1.3.4 Results

Auger has produced a number of results in the field of UHECR physics. It has unequivocally confirmed the suppression in the energy spectrum above 4×10^{19} eV [16]. It has also set leading photon [17] and neutrino [18] limits at the highest energies which rule out a significant effect of top-down processes in the production of cosmic rays in extragalactic sources. The observatory has also discovered a dipole in the arrival directions of cosmic rays with energies greater than 8×10^{18} eV with over 5σ post-trial significance [19]. The lack of significant anisotropy at lower energies around the ankle also provides reasoning against a transition from galactic to extragalactic origins at the ankle, as a stronger anisotropy would be expected from cosmic rays produced within the galaxy at these energies, especially considering the mea-

Figure 1.8: Installation of one of the 30 m^2 buried, scintillation-based muon counters. These counters will directly measure the muonic component of EASs approximately at and below the energy of the spectrum's ankle. Picture from [15].

sured dominance of a light component of the energy spectrum in this energy region. Auger also measures a spectrum which grows increasingly heavier above the ankle [20]. Additionally it has measured proton-air and proton-proton cross sections at the highest energies [21], and demonstrated a muon deficit in EAS simulations using the latest LHC-tuned hadronic interaction models when comparing with measurements [22].

1.3.5 Open Questions

A number of conundrums remain for UHECRs, which Auger seeks to resolve. Sources of UHECRs have yet to be definitively identified, although some indications exist as to their origin. Additionally, it is unclear whether the suppression at the highest energies can be attributed primarily to propagation effects, such as photo-pion production and photo-disintegration of nuclei, or if it arises predominantly from extragalactic accelerators reaching their maximum acceleration potential. It also remains to be clearly demonstrated whether the ankle is an indicator of the transition from galactic to extra-galactic cosmic rays, which would require a Population B of galactic accelerators, or if this transition occurs at a lower energy.

Knowledge of the elemental composition of the total cosmic ray flux, and, if possible, mass estimation at the per event level, would be of great value in answering each of these questions. Due to its limited duty cycle, the FD cannot provide reasonable estimates of composition above energies around the onset of the flux suppression. A detector upgrade allowing for mass reconstruction with a 100% duty cycle is called for and will be further discussed in Chapter 2.

1.4 Surface Detector Reconstruction

The standard SD reconstruction of Auger emphasizes a data-driven approach that minimizes dependence on simulations. The steps of this reconstruction are briefly summarized in this section starting from station-level hardware triggers and ending with the reconstructed energies and arrival directions of the UHECRs.

1.4.1 Triggers

A chain of hardware and software triggers exists to reject background composed largely of secondaries from significantly lower energy cosmic rays than those of interest to Auger. The first step in this chain is the hardware-level T1 trigger which may be satisfied in one of two ways.

1. *Threshold*: All three PMTs measure a signal crossing the threshold of 1.75 VEM above baseline in coincidence.

2. *Time-over-Threshold*: At least two of the three PMTs measure 0.2 VEM above baseline for at least 13 time bins in a window of 120 bins.

If the T1 trigger is satisfied, the local station software processes the T2 trigger conditions. These are identical to those of the T1, except the T2 *Threshold* trigger requires a signal of 3.2 VEM as opposed to 1.75 VEM. At the T2 level, the *Threshold* trigger is designated with the code *Thr2* and has a frequency of approximately 20 Hz. The primary aim is to detect the signatures of muons in highly inclined showers. The *Time-over-Threshold* trigger is designated with the code *ToT* and has a frequency of 1 Hz - 5 Hz. It is optimized for background discrimination for air shower events below 60° in zenith angle. If either T2 trigger is satisfied, the station notifies the central data acquisition system (CDAS), which examines time coincidences of T2 triggers to identify the signatures of extensive air showers.

In explaining the demands for satisfying the conditions of the T3 triggers, it is helpful to define the concept of the NCN sub-trigger schema. The first N in this schema corresponds to the number of stations with a reported T2. The second N corresponds to the maximum "crown" around a central station (which may be any of the stations reporting a T2). The first crown corresponds to the hexagon of stations at the elementary array distance away from a central station (approximately 1500 m). The second crown corresponds to the stations two elementary array distances away from the central station (approximately 3000 m). There are two primary ways to satisfy a T3 trigger with combinations of NCN sub-triggers, namely:

1. *3-fold coincidence (T3-3ToT)*: satisfaction of 2C1 and 3C2 triggers with stations reporting a *ToT* T2 trigger. This is aimed primarily at events with zenith angles below 60°.

2. *4-fold coincidence (T3-4T2)*: satisfaction of 2C1, 3C2, and 4C4 triggers with stations reporting any type of T2 trigger (*Thr2* or *ToT*). This trigger is intended to recover highly inclined events.

An event may also be promoted to the T3 level if a trigger is received from the FD. A final pre-reconstruction T4 physics trigger also exists, which has stricter requirements than those of the T3 and can be satisfied in one of three ways:

1. *T4-3ToT*: 3 neighboring stations satisfying the *ToT* T2 trigger, which form either an equilateral triangle (where the three stations are separated from one another

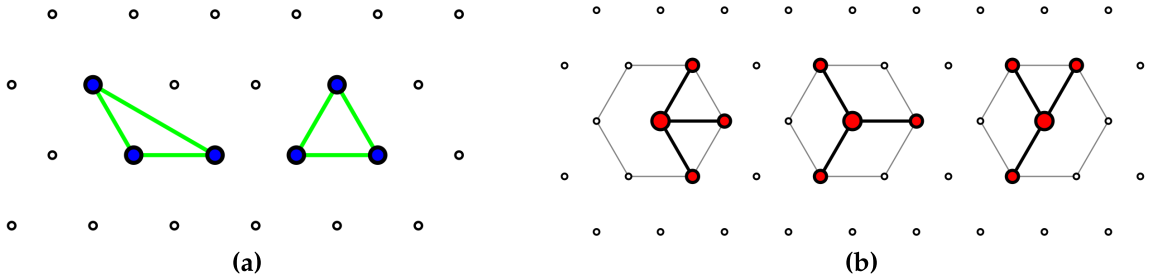


Figure 1.9: (a) Schematic depiction of the two possible configurations satisfying the T3 3ToT trigger conditions. (b) Schematic depiction of three possible configurations satisfying the T3 4C1 trigger conditions.

by the elementary array distance) or isosceles triangle. The schematic depiction of instances of these two possible configurations is given in Fig. 1.9a.

2. *T4-4C1*: 4 stations satisfying either of the T2 trigger conditions (*Thr2* or *ToT*), where three of the stations are in the first crown of a fourth station. A schematic depiction of different instances is given in Fig. 1.9b.
3. *T4-A3ToT*: 3 stations satisfying the *ToT* T2 trigger form a line.

With the T4 triggers in hand, an initial "bottom-up" selection may be performed in order to reject stations geometrically and temporally incompatible with a rudimentary first reconstruction of the event geometry. Highly isolated triggering stations are also rejected.

A T5 trigger also exists and may be applied prior to or following reconstruction. Designated the NT5 trigger, it requires that N neighbors of the station in the event with the largest signal were functioning, independent of whether or not they triggered. This trigger is intended to filter out events at the very edge of the array or in which stations vital to the quality of reconstruction were not online. For analyses of the cosmic ray energy spectrum, a 6T5 trigger is employed, whereas this requirement is often relaxed to a 5T5 trigger for anisotropy studies in order to increase the data set at the highest energies.

1.4.2 Shower Geometry

The first estimate of the shower geometry is obtained with the assumption of the a plane for the shower front. First, a signal-weighted barycenter $\vec{b}$ is obtained, which is used to define the weighted bary-time origin t_0. With this spatial and temporal coordinate system in hand, the plane front moving at the speed of light c along an axis $\hat{a}$ can be written as

$$\hat{a}(\vec{x} - \vec{b}) = c(t - t_0), \tag{1.1}$$

which is visualized in Fig. 1.10a. The time at which the shower plane passes through any given point $\vec{x}$ can thus be calculated and a chi-squared minimization of the differences between predicted and observed times can be performed.

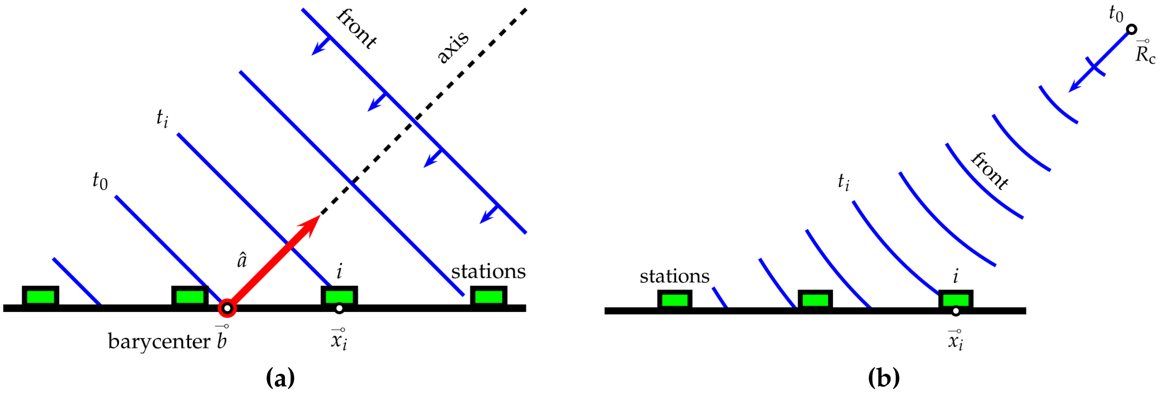


Figure 1.10: (a) Schematic depiction of the definition of the planar front assumption. (b) Schematic depiction of the spherical shower front assumption.

In the case of at least four triggered stations and a successful fit of the lateral distribution of signals (see Section 1.4.3), a spherical approximation of the shower front may be applied, which may be expressed as

$$c(t_i - t_0) = \left|\vec{R}_c - \vec{x}_i\right| \tag{1.2}$$

where $\vec{R}_c$ serves as the virtual spherical origin of the shower and may be related to the axis $\hat{a}$ of the shower by

$$\vec{R}_c = \vec{c} + R_c\hat{a} \tag{1.3}$$

where R_c is the shower's radius of curvature at the impact point $\vec{c}$. For a schematic depiction of this model, see Fig. 1.10b.

1.4.3 Lateral Distribution

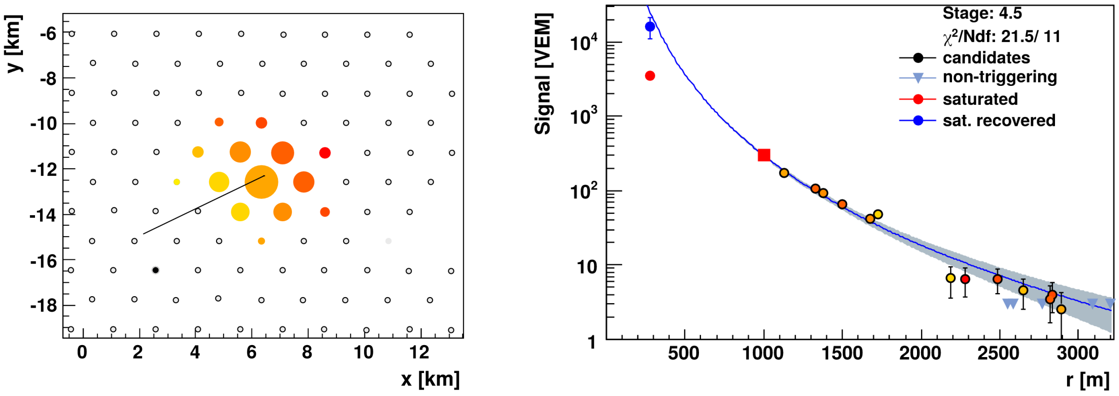
The lateral dependence of signal, that is the dependence of signal on a station's distance from the shower axis in the shower detector plane (SDP), is modeled with what is known as the lateral distribution function, which may be written as

$$S(r) = S_{r_\mathrm{opt}} f_\mathrm{LDF}(r) \tag{1.4}$$

where the adopted form of LDF is a modified Nishimura-Kamata-Greisen (NKG) type function [23, 24], written as

$$f_\mathrm{LDF}(r) = \left(\frac{r}{r_\mathrm{opt}}\right)^{\beta} \left(\frac{r + r_\mathrm{s}}{r_\mathrm{opt} + r_\mathrm{s}}\right)^{\gamma}, \tag{1.5}$$

The parameter r_opt is the distance at which the uncertainty in the shape of the LDF (largely defined by β) is minimized. This parameter has been shown to be largely defined by the geometry of an array and is approximately 1000 m for the triangular grid of Auger [25]. The parameters r_s and γ provide flexibility in the shape of the lateral distribution function (LDF), particularly at larger distances from the shower axis where the fraction of signal due to the muonic component of air showers becomes increasingly relevant and eventually dominant. A depiction of the lateral



(a) **(b)**

Figure 1.11: (a) Top-down schematic view of the footprint of an EAS on the SD array. The size of the circles corresponds to the magnitude of the signals whereas the coloring corresponds to signal start times. The black line represents the shower axis. (b) Fit of the lateral distribution function for the same event. The station closest to the shower axis was saturated, but the signal was recovered. Non-triggering stations, which may be observed at the larger distances, are depicted with downward facing equilateral triangles.

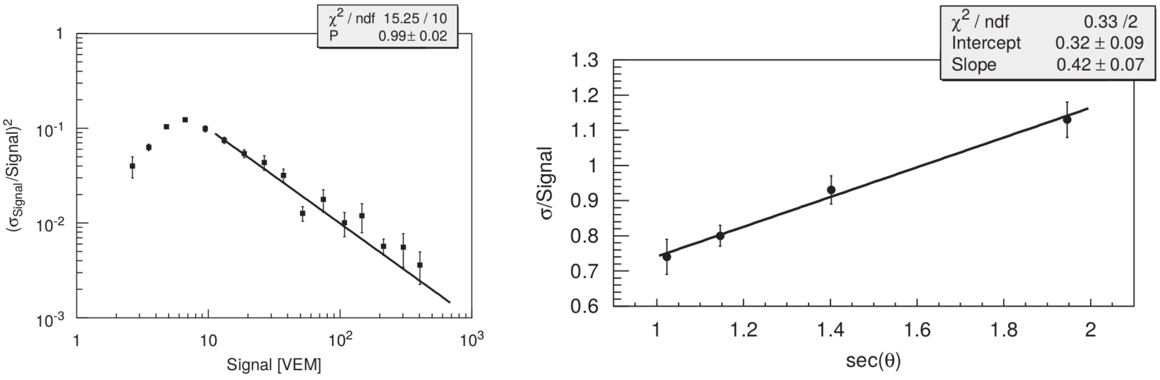
distribution of a sample event is provided in Fig. 1.11a along with a fit of this modified NKG functional form to this distribution in Fig. 1.11b. In this sample event, the station closest to the shower axis received a signal beyond the capabilities of the dynamic range and is therefore designated as "saturated." A signal recovery procedure exists for such instances [26].

Signal Uncertainties

In order to perform the likelihood fit of the lateral distribution function, an uncertainty associated with each measured signal is required. Uncertainties in signal can be attributed to sampling fluctuations of the ground particle densities, detector response, and uncertainty in the signal calibration. The signal uncertainties for the considerable majority of the dynamic range are dominated by sampling fluctuations.

A model for the uncertainty of signals measured by WCDs was derived from coincident measurements of two or more stations placed approximately 10 m from one another. Such configurations are referred to as multiplets. For any given shower, the two (or more) stations, although located at nearly identical positions in the SDP, measure entirely different particle samples. From many measurements in the range of signals within the dynamic range of the WCDs, it was confirmed that the signal uncertainties could described by the underlying assumption of Poisson statistics as the uncertainty in signal scales with its square root. Fig. 1.12a shows this explicitly. A zenith dependence of this proportionality was observed, however, and could be modeled linearly in secθ with

$$\sigma_S = f_s(\theta)\sqrt{S} = (a + b\,\text{sec}\theta)\,\sqrt{S}. \tag{1.6}$$



(a) **(b)**

Figure 1.12: (a) Signal accuracy for WCDs of the SD array. The linear fit at higher signals corresponds to the assumption that $\sigma/S = P/\sqrt{S}$. At lower signals, distortions due to the triggering efficiency are observed. (b) Dependence of proportionality constant between observed signal uncertainty and $\sqrt{S}$ on $\sec\theta$. Plots from [27].

This zenith dependency, shown in Fig. 1.12b, is largely due to the increasing fraction of energy deposit in the WCD from muons as zenith angle increases and the electromagnetic component of showers is relatively more attenuated at ground. Differences in detector response, partially attributed to geometry, also contribute. For the data set of the multiplet stations, no strong dependence on distance or shower size were observed; however, it should be noted that the data sample for the multiplets is dominated by low energy events. A concept which naturally follows from the underlying Poisson statistics is that of an effective particle number, which would, for an ideal particle counter, results in the same signal uncertainties observed in data. This effective particle number is defined as

$$n_{\text{eff}} = p\,S \tag{1.7}$$

where p is referred to as the “Poisson Factor” and is given defined as

$$p = \max\left(1, f_s(\theta)^{-2}\right) \tag{1.8}$$

where $f_s(\theta)$ is taken from Eq. (1.6). The constraint of a minimum factor of 1 is based on physical assumptions as to the maximum signal particles should be able to produce on average. This will be readdressed in Section 4.5.1.

Likelihood

A maximum likelihood method is used to perform the LDF fits. Prior to constructing the terms of the likelihood, each non-saturated signal is converted into an effective particle number using Eq. (1.7). Thereafter, stations are classified and treated in four different ways:

1. **Small Signals**
   Stations with an effective particle number of less than 30 are treated with a Poisson probability distribution, namely

$$f_{\mathrm{P}}(n_i, \mu_i) = \frac{\mu_i^{n_i} \mathrm{e}^{-\mu_i}}{n_i!} \tag{1.9}$$

   where n_i is the effective particle number calculated for a given station and μ_i is the expectation from the LDF.

2. **Large Signals**
   A Gaussian approximation is used for stations with an effective particle number of 30 or more, namely

$$f_{\mathrm{G}}(n_i, \mu_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left(-\frac{(n_i - \mu_i)^2}{2\sigma_i^2}\right) \tag{1.10}$$

   where σ_i is obtained from the signal uncertainty model derived from doublets such that $\sigma_i = p\sigma_S$. Recovered signals from saturated signals are also treated in this way.

3. **Saturated Signals**
   If a signal is saturated and can not be recovered, it is treated as a lower limit. As such, it's contribution of the likelihood consists of integrating the Gaussian approximation over all values greater than n_i, namely

$$F_{\mathrm{sat}}(n_i, \mu_i) = \int_{n_i}^{\infty} f_{\mathrm{G}}(n, \mu_i)\,\mathrm{d}n = \frac{1}{2}\,\mathrm{erfc}\left(\frac{n_i - \mu_i}{\sqrt{2}\,\sigma_i}\right), \tag{1.11}$$

4. **Zero Signals**
   For non-triggering stations, the probability of obtaining an effective particle number n_i less than the threshold of observance (trigger threshold) of n_{th} serves as the contribution to the likelihood. This requires integrating the Poisson probability distribution from $n = 0$ to n_{th}.

$$F_{\mathrm{zero}}(n_{\mathrm{th}}, \mu_i) = \sum_{n=0}^{n_{\mathrm{th}}} f_{\mathrm{P}}(n, \mu_i) \tag{1.12}$$

The likelihood constructed from stations with their various classifications may be written as

$$L = \prod_i f_{\mathrm{P}}(n_i, \mu_i) \prod_i f_{\mathrm{G}}(n_i, \mu_i) \prod_i F_{\mathrm{sat}}(n_i, \mu_i) \prod_i F_{\mathrm{zero}}(n_i, \mu_i) \tag{1.13}$$

where what is actually minimized is the negative log likelihood such that each station acts as a sum in its construction, namely

$$-\ell = \sum_i \ln f_{\mathrm{P}}(n_i, \mu_i) + \sum_i \ln f_{\mathrm{G}}(n_i, \mu_i) + \sum_i \ln F_{\mathrm{sat}}(n_i, \mu_i) + \sum_i \ln F_{\mathrm{zero}}(n_i, \mu_i), \tag{1.14}$$

1.4.4 Fit Stages

The fit for the shower geometry and size are performed in the following stages. First, an initial estimation of the geometry is obtained with the plane approximation of the shower front. Next, an estimation of the shower size S_{1000} is performed using the station-weighted barycenter as the estimate for the core position and estimations for the LDF shape taken from the parameterizations of β and γ. After this, the core and shower size are fit again, but together, followed by the optional inclusion of the LDF shape parameters depending on the satisfaction of conditions relating to the multiplicity and relative positioning of stations in the SDP. Finally, the core and shower size fit is performed once more considering non-triggering stations in the likelihood.

1.4.5 Energy

Once an estimate of the shower size S_{1000} is obtained, a normalization for zenith angle is performed, namely

$$S_{38} = \frac{S_{1000}}{1 + a\,x + b\,x^2 + c\,x^3} \tag{1.15}$$

where S_{38} is the shower size had it arrived at a zenith angle of 38°. The parameters a, b, and c are obtained through a fit to the event intensities at different zenith angles, under the assumption of an isotropic flux in $\cos^2\theta$. This method is referred to as the Constant Intensity Cut method [28].

The energy of the shower is then calculated from the zenith-independent estimator S_{38} by

$$E = A\,(S_{38})^B\,\mathrm{EeV} \tag{1.16}$$

where the parameters A and B are obtained through a calibration of high-quality events measured by both the FD, from which an energy estimate is obtained, and SD.

CHAPTER 2

Design of AugerPrime

Given the open questions in UHECR physics discussed in Chapter 1, the upgrade to Auger, named AugerPrime, bears the responsibility of equipping the observatory with the means to accomplish a number of objectives. To repeat, these include investigating whether the dominant cause of the flux suppression at the highest energies is due to effects of cosmic ray propagation through the universe or to nearby sources reaching their maximum acceleration capacity. Exploration of physics beyond the energies of the Large Hadron Collider and addressing the apparent inability of the current LHC-tuned hadronic interaction models in predicting the number of muons produced in EAS are also vital objectives. Determining the origins of ultra-high-energy cosmic rays remains an aspiration. Each of these objectives, among others, either requires or would greatly benefit from an improved estimation of the mass of cosmic rays on an event-by-event basis. With that said, the primary technical objective of AugerPrime is to enhance and augment Auger's existing detectors with hardware that can provide the information necessary for the reconstruction of cosmic-ray mass. As the lifetime of the Observatory has been extended to the end of 2024, this will provide an enhanced data set roughly comparable in size to that already acquired as well as allow for the re-analysis of existing data with an improved understanding of the existing detector and of extensive air showers.

A principle piece of information in reconstructing cosmic-ray-primary mass is the number of muons produced and reaching Earth's surface in extensive air showers. With this in mind, a number of different upgrades to Auger's SD array were proposed for AugerPrime, each intended to aid in the disentanglement of the electromagnetic and muonic components of air showers in order to reconstruct a muon number. These included a detector based on resistive plate chambers (RPCs) to be situated below the WCDs, a segmentation of the existing WCDs into two independent water volumes, and full-array coverage with buried scintillators. In the end, the solution of placing a scintillator on top of each existing WCD detector was chosen as it was a non-invasive, cost-effective proposal shown to fulfill the physics requirements set forth.

In addition to the choice of scintillators for the full SD array as the focal point of AugerPrime, a number of other upgrades are planned with implementation already underway. These include:

- **Small Photo-multiplier Tube (SPMT)**
  A fourth, “small” PMT, the Hamamatsu-R8619-22 [29], with a diameter of less than 30 mm will be added to each of the WCDs. The smaller light collection combined with the gain settings will extend the dynamic range of WCDs to match the upper limits planned for the scintillator detectors such that both sub-detectors saturate at approximately the same distance from the shower axis. For showers at energies of around 10^{20} eV, this distance is approximately 200 m.

- **Upgraded electronics**
  The existing front end and unified board (UB) have 6 channels, a sampling frequency of 40 MHz, and 10 bits alloted to ADC counting. All of the channels are being used for high and low gain signals of the WCD’s 3 PMTs. The addition of a scintillator and SPMT requires additional channels. The upgraded unified board (UUB) design accommodates this demand with 10 channels in addition to a faster sampling frequency of 120 MHz and 12 bit ADC counting, which offers the option of either extending the dynamic range or more finely digitizing the detectors’ analog signals.

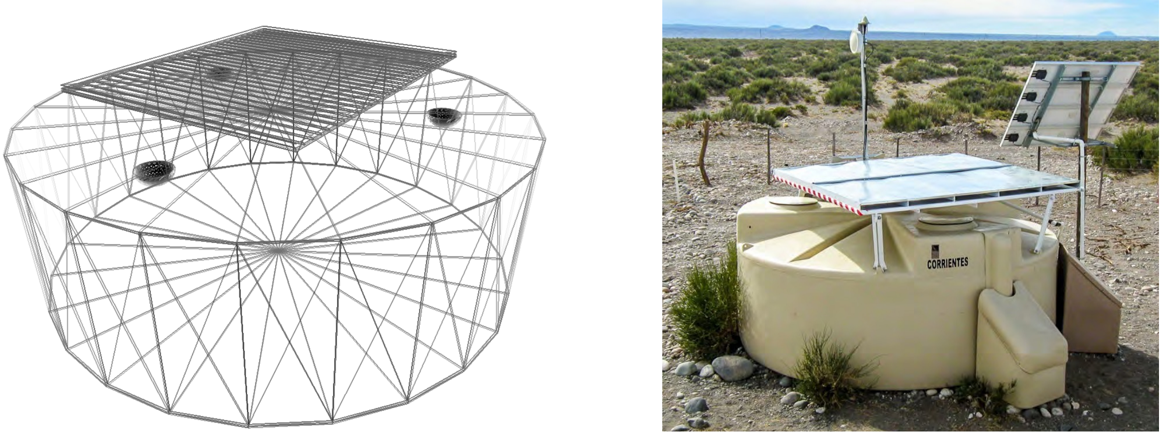
- **Extended Fluorescence Detector uptime**
  The FD is being equipped with an additional mode of operation in which the PMT gains are reduced by approximately a factor of 10. This allows for measurement by the FD earlier in the evenings and later in the mornings thereby increasing its uptime by around 50% resulting in a total uptime of approximately 20%. The procedure has been tested and shown to work technically while still collecting data of sufficient quality for FD reconstruction. The effects of aging on the PMTs have also been examined and found to be acceptable for the expected lifetime of the Observatory.

- **Underground Muon Detector (UMD)**
  In the denser-spaced 750 m array, 30 m^2 scintillator-based UMDs are being deployed at a depth of 2.3 m for 61 of the WCD positions. These will significantly enhance composition reconstruction on an event-by-event basis in the region of the energy spectrum’s ankle as well as provide a direct calibration of the methods used to reconstruct muon numbers from the upgraded SD’s scintillator and water-Cherenkov sub-detectors.

2.1 Design Decisions

With a scintillator as the choice detector of the AugerPrime upgrade, a number of decisions as to its basic design needed to be made. Key questions in this process were, simply stated,

1. How big should it be?
2. Where should it be located relative to the WCD?
3. In what range of signal size should it be able to measure?



(a) (b)

Figure 2.1: (a) Visualization of the GEANT4 volumes of the prototype detector simulations used to inform scintillator design decisions. (b) Picture of one of the first scintillator prototype detectors, called *Auger Scintillator for Composition - II*, or "ASCII" for short.

Each of the following three sections aim to answer one of these interrelated questions. The accompanying analyses make use of a preliminary simulation, where a scintillator of variable size is located above a WCD at 1.5 m above ground as visualized in Fig. 2.1a. Some parameters of this preliminary simulation were informed by actual prototypes conceptualized, developed, and constructed by the Auger group at the Centro Atómico Bariloche (Comisión Nacional de Energía Atómica) in Bariloche, Argentina, which received technical advisement during construction from experts at the Instituto de Tecnologías en Detección y Astropartículas (ITeDA) in Buenos Aires, Argentina. A photo of such a prototype is shown in Fig. 2.1b.

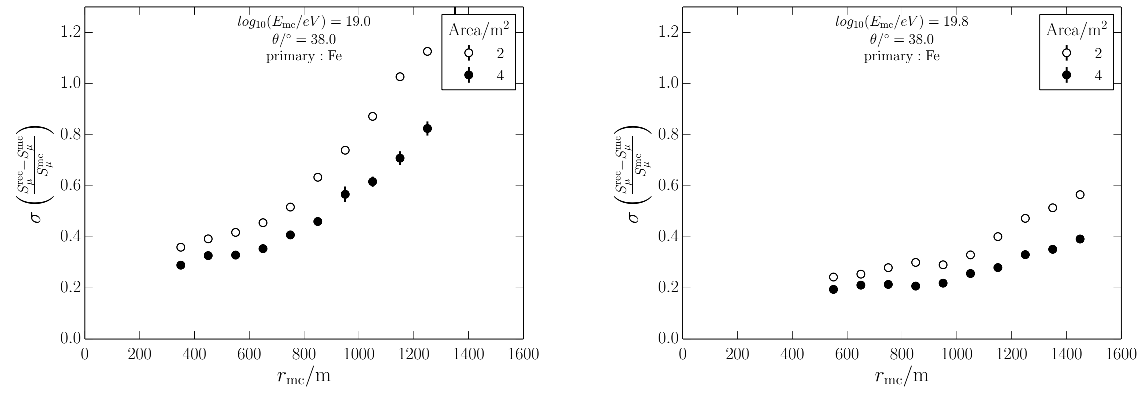
In these analyses, a simple reconstruction algorithm was employed in which the muonic signal in WCD detectors is written as a linear combination of the total signals of the scintillator and WCD, that is,

$$S_{\mathrm{wcd}}^{\mu^{\pm}} = g S_{\mathrm{scin}} + h S_{\mathrm{wcd}} \,, \tag{2.1}$$

where $S_{\mathrm{wcd}}^{\mu^{\pm}}$ is the signal in vertical-equivalent muon (VEM) produced by muons in the WCD, and S_{scin} and S_{wcd} are the total signals respectively measured by the scintillator (in MIP) and WCD (in VEM). A much more detailed formulation and assessment of such an algorithm is presented in Chapter 4.

2.1.1 Size

In determining the size of the scintillator, it was important to consider how it will be used in analysis. It serves as a complimentary detector to the WCD in the SD array. The general idea is to use the different responses of the two detectors to backtrack the number of particles from different shower components at ground. The size of the scintillator should thus be decided upon in order to minimize the limitations of its resolution on the resolution of variables reconstructed using the information from both sub-detectors. Scintillators with areas of 2 m^2 and 4 m^2 were simulated to



(a) **(b)**

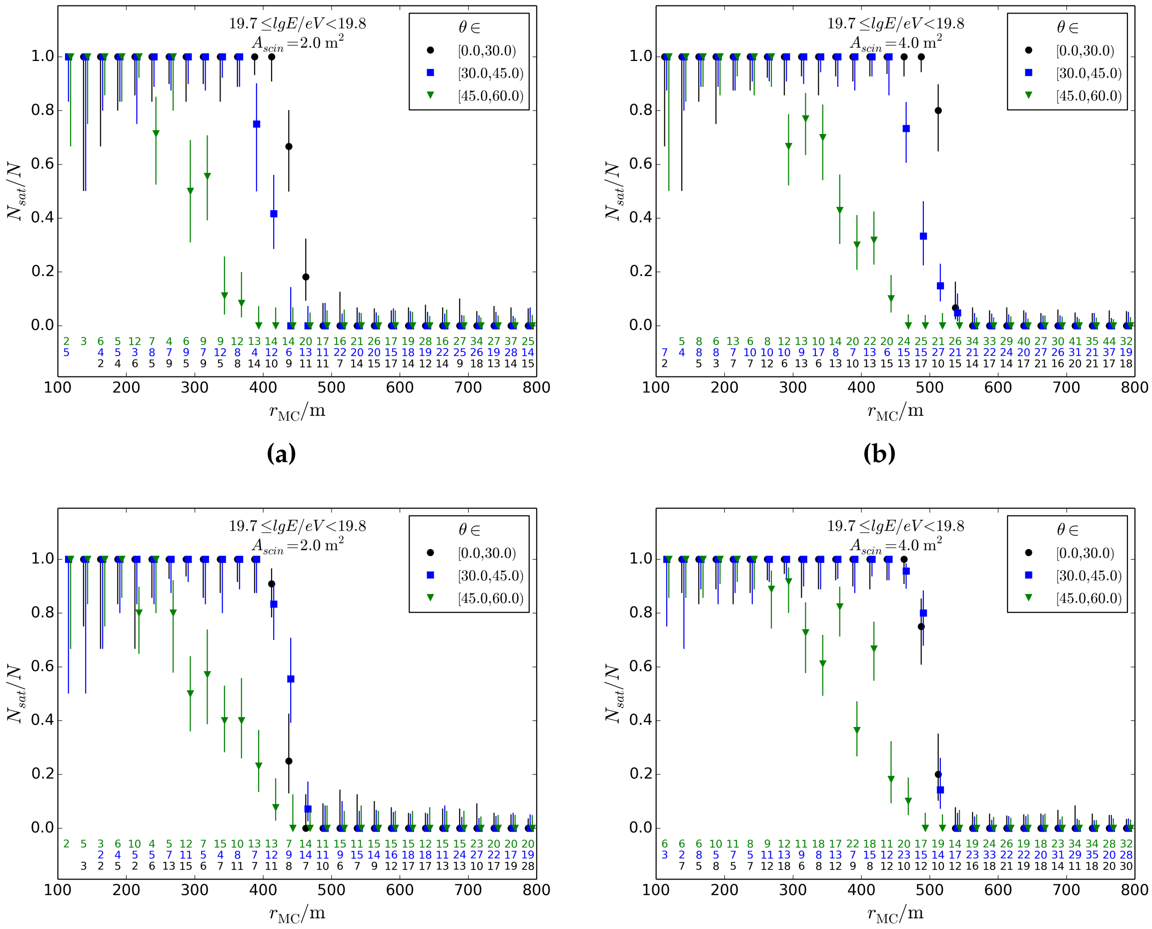
Figure 2.2: Resolutions of the reconstructed muonic signal in the WCD as a function of distance from the shower axis for primaries of (a) $10^{19.0}$ eV and (b) $10^{19.8}$ eV. The lack of data closer to the shower axis is due to the limitations of the dynamic range of the prototype scintillator detectors used to tune the prototype detector simulations.

gain an idea as to the impact of size on resolution, where 4 m^2 was the approximate financial size limit for scintillation material. The results of the reconstructed muonic signal in WCDs shown in Fig. 2.2 demonstrate that the two fold increase in the area of the scintillation material corresponds to a 20-100% improvement in resolution, depending on the distance of the detectors from the shower axis, and extends the distance from the axis for which measurements with usable resolutions are possible.

Naturally, an increase in the size of the scintillator corresponds to a proportional increase in the size of signals, and therefore PMT and electronics saturation already occur at larger distances from the shower axis. At an energy of $10^{19.8}$ eV, doubling the active scintillation area results in saturation approximately 100 m further from the shower axis (see Fig. 2.3).

Conclusions

All indications suggested significant benefits for a scintillator with an area of 4 m^2 or larger without massive decreases to the distance range for which the dynamic range of a scintillator would result in reliable measurements. The 4 m^2 scintillator also results in a signal size of approximately the same magnitude as that of the WCD. Under a naive Poisson assumption, this would result in signal uncertainties of approximately the same order of magnitude. Considering the roughly equal weight of the two detectors' uncertainties on the reconstructed muonic signal in the WCD according to Eq. (2.1), this would, in theory, indicate that neither of the two detectors is a dominant limitation in reconstruction. The reality is a bit different and is further discussed in Chapter 4. Taking financial considerations, which do not permit a considerably larger area, into account, the decision for a 4 m^2 scintillator was made.



(a) **(b)**

(c) **(d)**

Figure 2.3: Fraction of stations with saturated electronics as a function of distance from the shower axis for primaries with energies between 19.7 eV and 19.8 eV. (a) and (b) respectively depict the saturation curves for 2 m^2 and 4 m^2 scintillators for iron primaries. (c) and (d) depict the same for proton primaries. The numbers at the bottom of each plot provide the number of simulated stations used in the calculations for the given bin.

2.1.2 Position

The position of a scintillator relative to the WCD determines to what degree the two sub-detectors share the same sample of particles. By placing the scintillator at a distance of a few meters from the WCD, the two sub-detectors would measure entirely different batches of particles, whereas placing it directly on top of the WCD would result in many of the same particles entering both detectors. This is depicted in Fig. 2.4. There are hypothetical advantages and disadvantages to different arrangements. Separated detectors would provide two independent samples of the lateral distributions of particles at effectively the same position in the shower detector plane. Having the detectors share the same particle sample does not boast this benefit, but does imply that any algorithms making use of both sub-detectors at the station-level will not suffer as much from fluctuations in particle distributions on the order of meters in the shower footprint at ground.

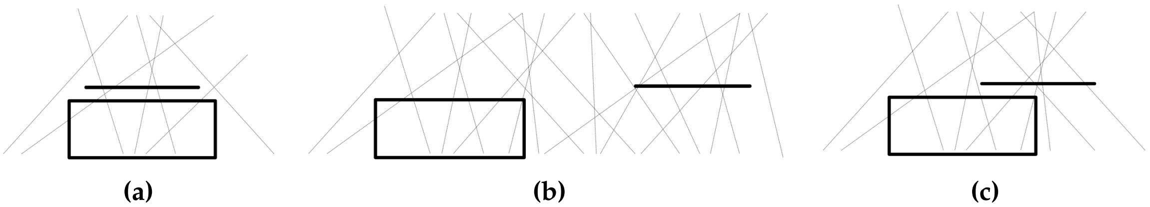


Figure 2.4: Depictions of three possible relative positions of a scintillator and WCD. In the arrangement depicted in (a), the two sub-detectors will share the same particle sample, therefore minimizing fluctuations between the two sub-detectors, whereas in the arrangement of (b), they will measure entirely different particle batches, providing independent samples of the lateral distribution of particles at effectively the same point in the shower detector plane. The schematic of (c) simply serves to illustrate that many configurations where different fractions of the particle flux at ground are shared exist and that this fraction can depend on the shower geometry.

Table 2.1: Dense ring simulation library used for positioning and correlation studies. The quantity given for "Showers per bin" refers to the number of unique showers available in each energy and zenith bin. Non-zero zenith angles appear in gray as they could only provide a limited contribution to the results of the analyses due to azimuthal signal asymmetries.

Model	QGSJetII-04
$\theta/°$	0, 12, 25, 36, 45, 53
$\lg(E/eV)$	18.6, 19.0, 19.5, 20.0
Dense ring distance / m	600, 800, 1000
Stations per distance	12
Primaries	proton
Showers per bin	60

The resolutions of each individual sub-detector will not significantly differ between the two arrangements. Only the resolutions of quantities derived from combined use of both measurements will exhibit differences. An analysis of simulations was performed in order to estimate the magnitude of the correlations between the two sub-detectors' signals and demonstrate the impact of these correlations on the resolutions of the muonic signal in the WCD reconstructed via the linear combination described at the start of this chapter and written in Eq. (2.1). The simulation libraries used for the subsequent analysis are summarized in Table 2.1.

These simulations make use of "dense rings," which are stations placed in regular intervals of azimuth at a given distance from the shower axis in the shower detector plane. Once defined in the shower detector plane, the station positions are projected onto Earth's surface, where they are simulated. A schematic depicting such a dense ring may be observed in Fig. 2.5a.

Scintillator and WCD Signal Correlations

The correlations between scintillator and WCD signals are assessed independently for each distinct combination of energy, zenith angle, and distance. First, for each unique instance of a dense ring, the WCD and scintillator signals are independently normalized by the average signal of all detectors in the ring via

$$\widetilde{S}^i_{\mathrm{det}} = \frac{S^i_{\mathrm{det}}}{\langle S_{\mathrm{det}} \rangle}, \tag{2.2}$$

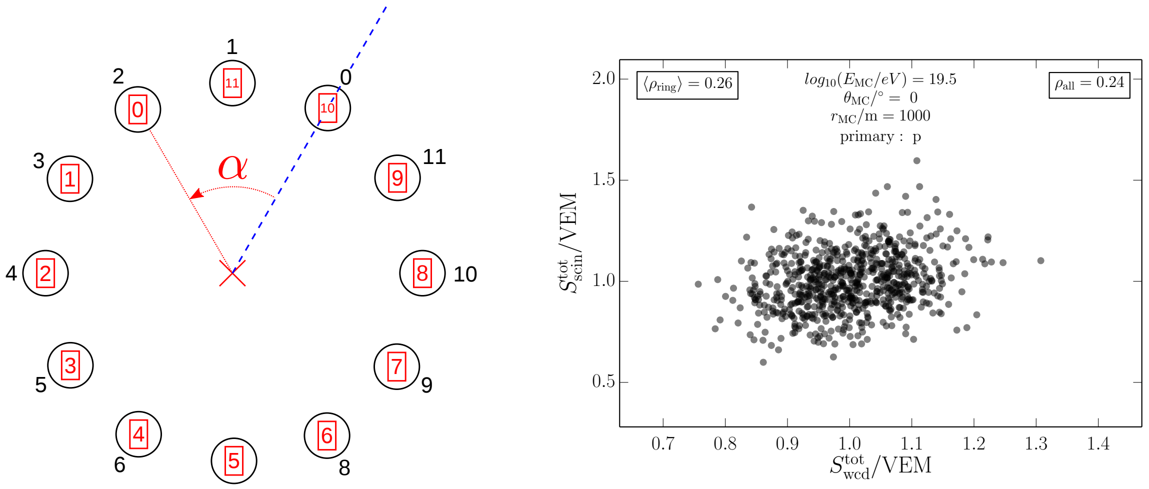
where i refers to individual signals in the dense ring and “det” may refer to the scintillator or WCD. This normalization is performed such that all stations for a given energy, zenith angle, and distance may be grouped without inducing an artificial correlation due to the effects of shower-to-shower fluctuations on the signals measured for individual showers. Once this normalization has been performed, all normalized signals are combined into one set for a given energy, zenith angle, and distance and the correlation between scintillator and WCD signals is calculated by dividing the covariance by the product of the standard deviations of the signals of the two sub-detectors, namely

$$\rho(\widetilde{S}_{\mathrm{wcd}}, \widetilde{S}_{\mathrm{scin}}) = \frac{\mathrm{COV}(\widetilde{S}_{\mathrm{wcd}}, \widetilde{S}_{\mathrm{scin}})}{\sigma_{\widetilde{S}_{\mathrm{wcd}}} \sigma_{\widetilde{S}_{\mathrm{scin}}}}. \tag{2.3}$$

A sample plot depicting the normalized signals from which a correlation coefficient is calculated is shown in Fig. 2.5b, and the correlation coefficients for different energy and zenith bins at 600 m are given in Fig. 2.6a. Correlation strengths at a zenith angle of 0° range from approximately 0.2 at an energy of $10^{18.6}$ eV to 0.4 at $10^{20.0}$ eV. The obvious manner in which to simulate WCD and scintillator pairs where the two detectors do not share the same particles but measure the same point in the shower detector plane would be to place them a few meters apart in the simulations. This method encountered difficulties, however, due to limitations in the implementation of the prototype scintillator simulation. Additionally, as the particles injected in the shared volume of the two detectors are re-sampled from the same area of the shower footprint, an artificial reduction in the fluctuations between the sub-detector signals could arise.

With that said, an alternative method for the analysis was invoked. The scintillators in the dense ring were rotated by an angle α (see Fig. 2.6a for a visual depiction) relative to the WCDs with which they were simulated. In doing this, the newly paired detectors did not share the same particles during re-sampling in their simulation, and should thus exhibit no correlation in their signals once the discussed normalization is performed. The cost of this method is that any real correlations in particle distributions at the ground on the order of a few meters to a few tens of meters would not be included in this analysis. Without un-thinned CORSIKA showers, the magnitude of this correlation, if it exists, is difficult to estimate, although a few indications will be discussed in Chapter 4 this work.

A limitation to this analysis on dense rings exists in that with increasing zenith angle, differences are apparent in “early” and “late” station signals due primarily to geometric reasons but also to differences in attenuation, especially at large zenith



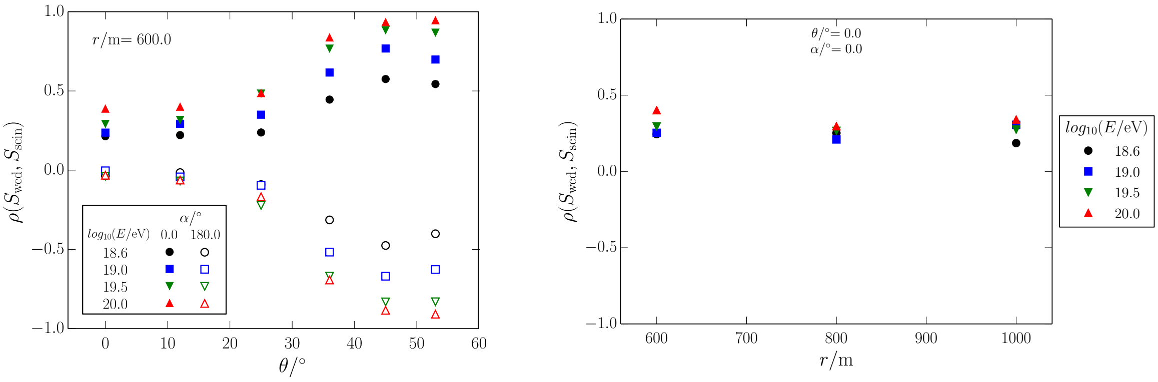
(a) **(b)**

Figure 2.5: (a) Depiction of a dense ring of scintillators and WCDs. For this schematic of a nearly vertical shower, evidenced by the fact that the projected dense ring on ground is nearly circular, the dashed blue line corresponds to the shower axis. The black circles and numbers respectively correspond to the WCDs and their IDs. Analogous for the scintillator are the red rectangles and numbers. WCDs and scintillators with the same ID were simulated together. For analysis, the WCDs remain "stationary" while the corresponding scintillator signals are rotated by the angle α (in this case, rotation along the arc length of two adjacent stations or $\alpha = 60°$. (b) Normalized scintillator and WCD signals at 1000 m for the set of showers of energy $10^{19.8}$ eV and zenith angle 0°. The correlation is calculated both using all normalized signals at once and also by averaging correlations calculated for individual rings. Only minimal difference are observed between the two results.

angles. Such azimuthal asymmetries in dense ring signals are not accounted for in the normalization procedure previously described and begins to significantly contribute to the calculated correlation coefficient even at modest zenith angles ($\theta > 12°$), which may be observed in Fig. 2.6a. In principle, this asymmetry should be parameterized out of this analysis or a different normalization procedure should be employed. In this analysis, however, the influence of the asymmetry effect is briefly investigated at all zenith angles, but only completely vertical showers ($\theta = 0°$), where azimuthal asymmetries are absent, are used when examining the resolution of the reconstructed muonic signal via the methods briefly described at the start of this chapter. Naturally, a more complete analysis at all zenith angles should be performed with the full simulations developed for the final scintillator design for AugerPrime.

Effects on Station-level Muonic Signal Reconstruction

Given the influence of azimuthal asymmetries, only showers with a zenith angle of 0° were used to examine the influence of signal correlations. Station-level resolutions of the reconstructed muonic signal in the WCD were calculated for scenarios where a signal correlation does and does not exist, which respectively correspond to positional arrangements where the scintillator is on top of the WCD or some



(a) **(b)**

Figure 2.6: (a) Correlation coefficients calculated for sub-detectors sharing the same particle sample ($\alpha = 0°$) and different particle samples ($\alpha = 180°$). For completely vertical showers ($\theta = 0°$), the correlation is compatible with zero at all non-zero values of α, as expected. A clear influence of the azimuthal asymmetry's influence on calculated correlation coefficients comes into play at $\theta \geq 12°$, however, where even scintillators and WCDs simulated with different samples of particles exhibit a correlation. At $\alpha = 180°$, where "early" and "late" scintillator signals have essentially been swapped, one might expect a correlation arising from azimuthal asymmetries to mirror those at $\alpha = 0°$ if the asymmetry's effect on signal correlation is dominant. For large zenith angles this is appears to be the case. (b) Correlation strength as a function of distance for the dense rings simulated at 600, 800, and 1000 m. In this distance range, there is no apparent dependence.

meters off to its side. In order to minimize rotation of the scintillator dense ring, $\alpha = 0°$ and 30° were used for the two scenarios. Results are shown in Fig. 2.7a. The resolutions calculated for scenarios with correlated signals exhibit improvements of up to around 20 % over uncorrelated signals. The difference between the two does not converge with increasing energy of the primary particle and is more-or-less of the same magnitude for the distances examined as shown in Fig. 2.6b.

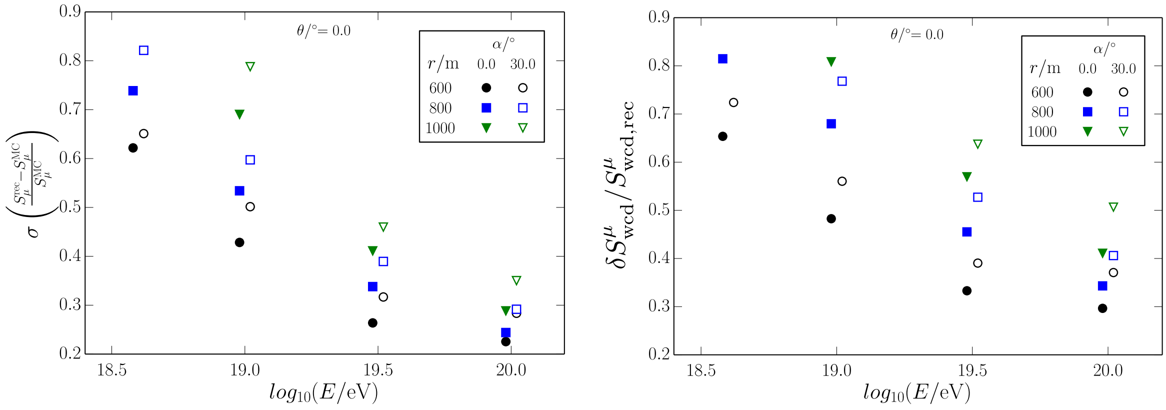
Analytical Uncertainty

For comparison, analytical calculations of the uncertainty in the WCD's reconstructed muonic signal were performed. Considering signal correlations between the scintillator and WCD, the uncertainty of the muonic signal in the WCD reconstructed via Eq. (2.1) may be expressed, under a Gaussian assumption, as

$$\delta S^{\mu^\pm}_{\rm wcd} = \sqrt{\left(\frac{\partial S^{\mu^\pm}_{\rm wcd}}{\partial S_{\rm scin}}\delta S_{\rm scin}\right)^2 + \left(\frac{\partial S^{\mu^\pm}_{\rm wcd}}{\partial S_{\rm wcd}}\delta S_{\rm scin}\right)^2 + 2\rho(\delta S_{\rm scin}\delta S_{\rm wcd})\left(\frac{\partial S^{\mu^\pm}_{\rm wcd}}{\partial S_{\rm scin}}\frac{\partial S^{\mu^\pm}_{\rm wcd}}{\partial S_{\rm wcd}}\right)} \tag{2.4}$$

where ρ is the correlation coefficient and signal uncertainties are indicated by the corresponding variable name preceded by a δ. This reduces to

$$\delta S^{\mu^\pm}_{\rm wcd} = \sqrt{(g\delta S_{\rm scin})^2 + (h\delta S_{\rm scin})^2 + 2gh\rho\delta S_{\rm scin}\delta S_{\rm wcd}} \tag{2.5}$$



(a) **(b)**

Figure 2.7: (a) Resolutions of the reconstructed muonic signal of the WCD for instances where the sub-detector signals are correlated (filled markers, $\alpha = 0°$) and uncorrelated (open markers, $\alpha = 30°$). These instances approximately correspond to scenarios where the scintillator is positioned, respectively, on top of and beside the WCD. (b) Analytically propagated resolutions of the matrix formalism for $\alpha = 0°$ and $30°$ (approximately corresponding to instances where signals are respectively correlated and uncorrelated). These resolutions are systematically greater than those calculated through direct comparison of reconstructed and Monte Carlo quantities. A number of possible explanations exist for this offset. One such explanation is the use of an average resolution derived from all signals within an energy, zenith, distance bin instead of a yet to be determined single station resolution. Further investigation is required. Nonetheless, the general trends in the differences between resolutions with and without correlated signals are similar.

For interpretation, it is import to note that the parameters g and h are opposite in sign and nearly equal in magnitude. With that in mind, it can analytically be seen that

1. the two sub-detectors uncertainties have equal weight in the uncertainty of the reconstructed muonic signal in the WCD and

2. positive correlations between the scintillator and WCD signals will improve the resolution of the reconstructed muonic signal.

This matches what was observed with simulations. Resolutions on $S^{\mu^\pm}_{\mathrm{wcd}}$ were explicitly analytically calculated using Eq. (2.5) and numerically derived estimates of WCD and scintillator uncertainties. These uncertainties were obtained by taking the standard deviations of dense ring signals. The results, which are depicted in Fig. 2.7b, show analytically calculated uncertainties up to approximately 35 % higher than those obtained numerically (although this number varies strongly as a function of energy), which indicates something is awry or missing in the computation. However, the differences between resolutions with and without correlated signals are relatively well reproduced, which supports the effect of scintillator position on station-level muon signal reconstruction observed in prior analysis.

Decision

Although reasonable improvements in resolutions for the station-level reconstruction of muonic signal were observed, it is important to note that these analyses do not investigate the positive impact of having two independent samples of the lateral distribution could have on a more global reconstruction of the shower muon multiplicity. This statement should be deeply considered.

Taking these analyses and considerations into account along with more practical mechanical and logistical considerations of one arrangement as compared to the other, the decision was confirmed to place the scintillator on top of the existing WCDs. This arrangement allows the WCDs to be used as a basis to mount the frames that will support the scintillators and also have the scintillators at a height where the local animals, mostly cows, which wander the pampa are less likely to interfere with the equipment. Put simply, modest gains in station-level reconstruction resolutions do not justify the cost of an independent support system, increased complexity in setting up signal routing and station electronics, and the larger risk of damages incurred by cows deciding to take a nap on a warm, sun-heated scintillator roof. Nonetheless, the demonstrated effect of correlations demands that they be properly taken into account in the reconstruction process, especially in the calculations of uncertainty in reconstructed parameters.

Partial Overlap

Having made the decision to place the scintillators on top of the WCDs, questions arose as to the impact of the amount of positional overlap between sub-detectors on reconstructed muonic signal resolutions. Additionally, it should also be noted that different shower geometries will result in a different amount of overlap when projecting the areas of the two detectors into the SDP. An analysis was performed to investigate the relationship between the amount of overlap between the detectors and their signal correlations in conjunction with the resolution on the reconstructed muonic signal in the WCD. A method similar to the dense ring rotation method described near the start of Section 2.1.2 and depicted in Fig. 2.5a was used, but instead of integer rotations of scintillator signals relative to those of the WCD, a statistical method simulating fractional rotation was employed.

The basic idea is to construct scintillator signals by summing a fraction of the signal simulated with the corresponding WCD (“correlated” signal) with the remaining fraction taken from a scintillator simulated with an independent batch of particles (“uncorrelated” signal), naively written as

$$S_i \to f S_i + (1-f) S_{i+1} \tag{2.6}$$

where f is the fraction of the scintillator area overlapping that of the WCD. This naive expression fails to properly manage signal uncertainties, however, in that it has an averaging effect, which artificially reduces the total scintillator signal resolution. This can be corrected for. Thinking of each scintillator as a sum of the signal measured by an arbitrary number portions of its surface, one may write

$$S = S_1 + S_2 + \cdots + S_N. \tag{2.7}$$

Under the assumption that the signals in each of these portions is independent from the others, one may then express the uncertainty in the total signal as

$$\delta S = \sqrt{\sigma_1 + \sigma_2 + \cdots + \sigma_N} \tag{2.8}$$

which under the assumption of equal size portions reduces to

$$\delta S = \sqrt{\sum_{i=1}^{N} \sigma_f^2} = \sqrt{\frac{1}{f}\sigma_f^2} \tag{2.9}$$

where f is the fraction of the total scintillator area that one portion covers. Solving for the uncertainty in one portion, one has

$$\sigma_f = \delta S \sqrt{f}. \tag{2.10}$$

The uncertainty of a portion of the total scintillator signal that is obtained purely by propagating error in Eq. (2.6) is

$$\sigma_p = f \delta S \tag{2.11}$$

which implies that by simply applying Eq. (2.6) to obtain a signal comprised of a mixture of correlated and uncorrelated signal will result in a quantity with artificially low fluctuations. To resolve this, a new quantity Δ is defined as what must be added in quadrature to the uncertainties already inherent to the calculated σ_p of Eq. (2.11) in order to obtain an uncertainty in the composite signal (comprised of correlated and uncorrelated signals) equal to the uncertainty in the original total signals, δS. Expressed mathematically,

$$\sigma_f = \sqrt{\Delta^2 + \sigma_p^2} \tag{2.12}$$

which gives the result

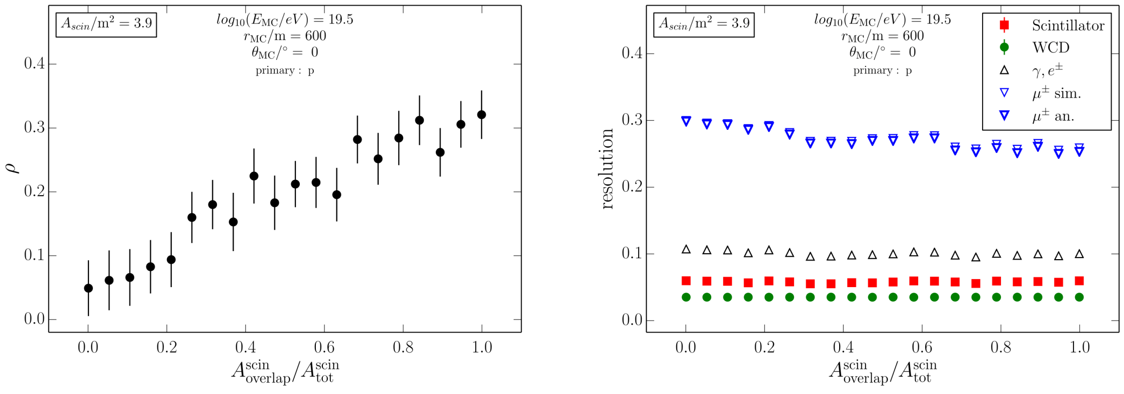
$$\Delta(\mathbf{f}) = \delta S \sqrt{\mathbf{f}(1 - \mathbf{f})} \tag{2.13}$$

and thus, Eq. (2.6) may be re-expressed as

$$S_i \to [f S_i]_{\Delta(f)} + [(1-f) S_{i+1}]_{\Delta(1-f)} \,. \tag{2.14}$$

where the subscripts to the quantities in square brackets indicate the fluctuations that must be applied to these quantities. These missing uncertainties for various fractional signals are added by randomly sampling a Gaussian distribution around a mean of $f S_i$ and with width taken from $\Delta(\mathbf{f})$ of Eq. (2.13).

Using this procedure, simulated fractional overlap of the scintillator area with the WCD was investigated. The correlation strength and resolutions of both the reconstructed muonic and electromagnetic signals in the WCD may be observed a function of the fractional overlap in Fig. 2.8a and Fig. 2.8b. It is clear that correlation strength and resolution change in an approximately linear manner with the fraction of detector overlap. As a result, it can be stated that small differences in overlap are not expected to have a large effect on the resolutions of reconstructed quantities.



(a) **(b)**

Figure 2.8: (a) Correlation strength increases approximately linearly as a function of scintillator area overlap with the WCD. (b) Resolution of the reconstructed electromagnetic and muonic signals in the WCD as a function of the fraction of the scintillator area that overlaps with the area of the WCD in the projection onto the SDP. Resolutions obtained via analytical error propagation are also shown, and are compatible with those obtained numerically.

Conclusions

The placement of the scintillator relative to the WCD has been shown to result in non-negligible signal correlations between the two sub-detectors. These signal correlations can moderately improve the resolution of quantities reconstructed at the station level using information from both sub-detectors, but come at the cost of not having two independent samples of the lateral distribution. Once the decision was made to place the scintillator atop the WCD, primarily for practical reasons, it was shown that the amount of overlap between the WCD and the scintillator in the shower detector plane results in only small differences in the resolution of reconstructed quantities. Nonetheless, in order to get a non-biased analytical estimate of the resolutions of reconstructed quantities, these signal correlations must be handled correctly under different shower geometries and thus warrant further study and modeling to avoid introducing biases and mis-estimated uncertainties in quantities reconstructed at the shower level. Such studies could greatly benefit from measurements with doublet stations, where at least one doublet is equipped with both sub-detectors. However, proper parameterization of signal correlations for high energy primaries will either require other methods or require the demonstration of a dependence of correlation strength only on signal size and not on primary energy, since the data sets obtained with doublet configurations are dominated by extensive air showers initiated by lower energy primaries.

It is also interesting to note here that the strength of signal correlations is likely intimately related to the energy spectra of the incident particles. Knowledge of the correlation strength as measured with real data could therefore open the possibility of gaining insight into these energy spectra in other analyses.

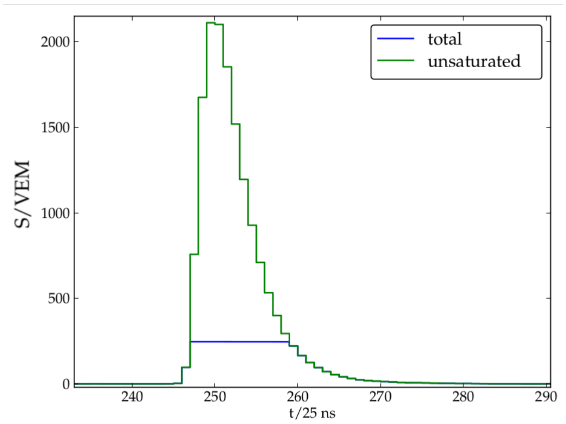


Figure 2.9: Sample trace to illustrate "saturation multiplier" statistic. The multiplier is equivalent to dividing the peak value of the "unsaturated" trace (what would be measured if there were no limits on the dynamic range) by the maximum of the trace which may actually be measured under a fixed set of constraints, or the maximum value of the "total" trace represented by the blue line.

2.1.3 Dynamic Range

It is desirable to measure as close to the axis of extensive air showers as possible. Particularly for primaries at the highest energies of around 10^{20} eV, no clear measurements have ever been performed closer than approximately half a kilometer from the shower axis, as this is the point at which the electronics of the WCD saturate. The lower limit of the dynamic range of the scintillator is fixed by the gain necessary to measure the charge deposited by individual minimum ionizing particles (MIPs) with a signal to noise ratio high enough to allow for accurate calibration of the detectors to this quantity. With this limitation, the upper limit of the dynamic range is then placed as high as possible such that there are no gaps in the distribution of signals where reliable measurement is possible.

In order to investigate the demands placed upon the hardware-dictated dynamic range by measurements at various distances from the shower axis, a statistic known as the "saturation multiplier" was defined, namely

$$\text{saturation multiplier}(C\sigma) = \frac{\langle q_{\text{s.b.}} \rangle + C\sigma(q_{\text{s.b.}})}{q_{\text{s.b.}}^{\text{sat}}} \tag{2.15}$$

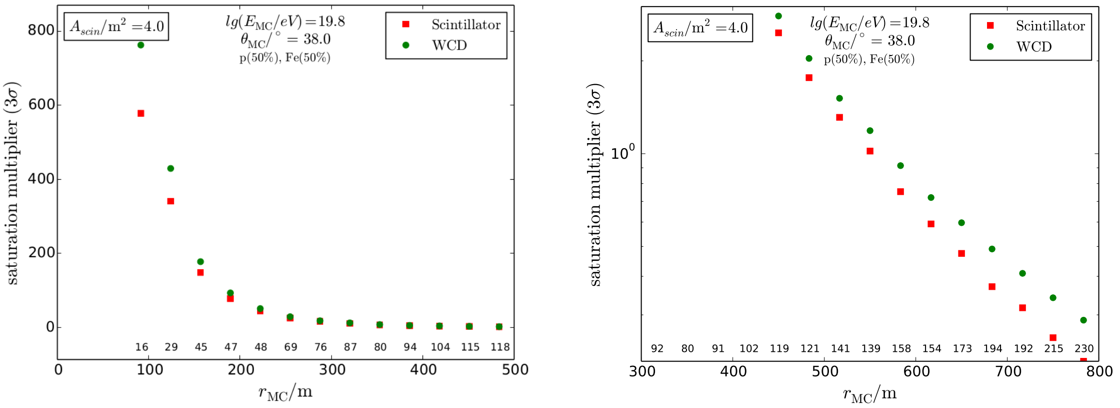
where

$$q_{\text{s.b.}} = \max(\text{ADC counts in single trace bin}) - \text{baseline}, \tag{2.16}$$

or the maximum number of ADC counts in a single bin of a signal trace minus the baseline number of ADC counts, and

$$q_{\text{s.b.}}^{\text{sat}} = 2^{N_{\text{bits}}} - \text{baseline} \tag{2.17}$$

where N_{bits} is the number of bits in the electronics for the lowest gain channel. A saturation multiplier of $C\sigma$ gives the factor by which the assumed dynamic range of a scintillator's electronics must be extended on it's upper end such that only a fraction of stations equal to one minus the cumulative distribution function (c.d.f.) of a Gaussian distribution at the value $C\sigma$ above its mean will saturate. A more



(a) **(b)**

Figure 2.10: The saturation multiplier plotted as a function of distance from the shower axis in (a) linear and (b) logarithmic scale. From the linear scale, it is clear that the demands on the upper limit of the dynamic range begin to rapidly increase at around 200 m from the shower axis. The logarithmic scale shows that the rate at which the demand increases is not exactly equal between the WCD and scintillator, which is important to consider when determining the hardware and gain settings for the two sub-detectors if they should be able to measure up to the same distances from shower axes. This difference can likely be attributed to differing slopes in the two sub-detectors' lateral signal distributions and the compositional makeup of their respective signals. The numbers at the bottom of the plots give the number of simulated stations used in the given bin for calculations.

intuitive way of explaining this is by using the example of a saturation multiple of 2σ. The value of this multiplier tells you the factor by which the upper end of the dynamic range must be extended such that only 5% of the scintillators will saturate. For a helpful visual depiction of this statistic, see Fig. 2.9 and its caption.

For this analysis, saturation multipliers of 3σ were calculated at various distances from the shower axis and are shown plotted in Fig. 2.10. The values correspond to the factor by which the upper limit of the dynamic range must be extended such that 99.7% of scintillators are unsaturated. At a distance between 250 m and 200 m from the shower axis for the energy of $10^{19.8}$ eV, the saturation multiplier begins to rapidly increase. This is due to the rapidly increasing slope of the lateral distribution of particles at ground. At larger distances, an extension of the dynamic range on the order of a factor of 0-10 is required. At distances closer than 200 m, this factor quickly eclipses 100 and grows increasingly faster. This distance range of 200 m thus represents the limits of extending the upper limit of the dynamic range under the given constraints of two channels, continuity, and a fixed lower limit. 200 m at the highest energies was thus set as the objective for the scintillator design.

Conclusions

Measurement at a distance of 200 m for the highest energy showers of approximately 10^{20} eV was set as the objective for the dynamic range of the scintillators based on the rapidly increasing slope of the lateral distribution of particles in the region closer to the shower axis. Slight differences in the evolving demands on the upper limit of the dynamic range between the WCD and scintillators should be considered when aiming to match their saturation distances. To measure much closer than 200 m from the shower axis would either require many more channels with different gains, giving up the constraint of a continuous dynamic range from the lowest to the highest signals, and potentially developing new calibration procedures.

2.2 Scintillator Surface Detector

The decision was made for a scintillator with an area of approximately 4 m^2 to be placed on top of the each existing WCD with a dynamic range enabling measurement up to 200 m from the shower axis even for the highest energy events. With these specifications in mind, a prototype design was developed and named the Scintillator Surface Detector (SSD). A brief description of the aspects of this design relevant to this work is provided in the following sections. For a more detailed description, see [30]. In September 2016, 12 prototypes were deployed in an engineering array (EA), which is situated in Auger's SD array.

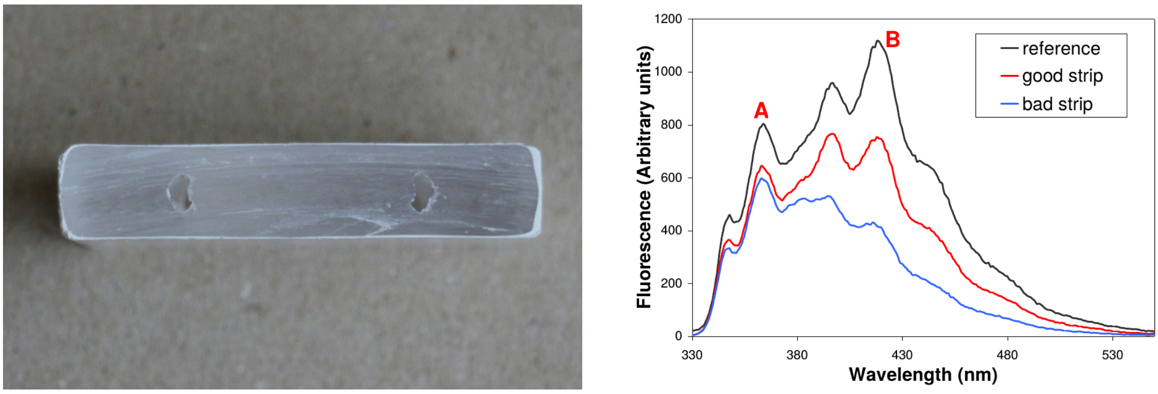
2.2.1 Design

The components of the SSD primarily dominating its response to EASs are the scintillator bars, optical fibers, and PMT.

Each of the scintillator bars is 50 mm in width, 1600 mm in length, and 10 mm thick and has two kidney-shaped holes running along its length as pictured in Fig. 2.11a. The bars, produced by Fermi National Accelerator Laboratory (FNAL) by extruding liquid polystyrene (Polystyrene Dow Styron 663 W), contain two wavelength-shifting dopants, PPO (1%) and POPOP (0.3%), whose respective emission peaks of 365 nm and 420 nm may be observed in the fluorescence spectra of the scintillators in Fig. 2.11b. Each bar is also co-extruded with an outer layer of a 0.25 mm polystyrene TiO_2 (15%) mixture, which has diffusive reflective properties reducing light losses.

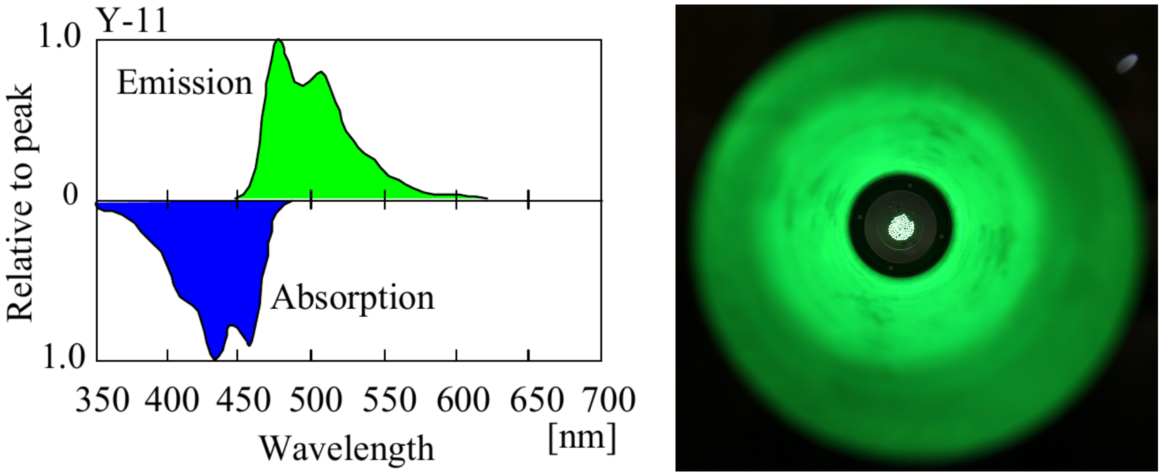
The optical fibers chosen for the SSD are the Y-11 (300) MSY from Kuraray [32] with a 1.0 mm diameter. With a multi-cladding structure, photons are trapped inside the fiber at multiple surfaces resulting in a higher photon yield. As S-type fibers, they also possess a molecular orientation of the polystyrene which provides higher durability and flexibility resulting in a minimum bending radius of 100 mm as opposed to that of 200 mm for non S-type fibers. These fibers are also doped with the wavelength shifting dye K-27, the absorption and emission spectra of which may be observed in Fig. 2.12a. Photons with wavelengths below approximately 490 nm are absorbed and re-emitted as green light. The manufacturer quotes an approximate attenuation length of 3.5 m.

The chosen PMT for the SSD is the Hamamatsu R9420, which has an 8 stage dynode system and a bialkali photocathode with a diameter of 38 mm. It has a high quantum efficiency in the green portion of the visible light spectrum, which matches the specifications of the chosen fibers.



(a) (b)

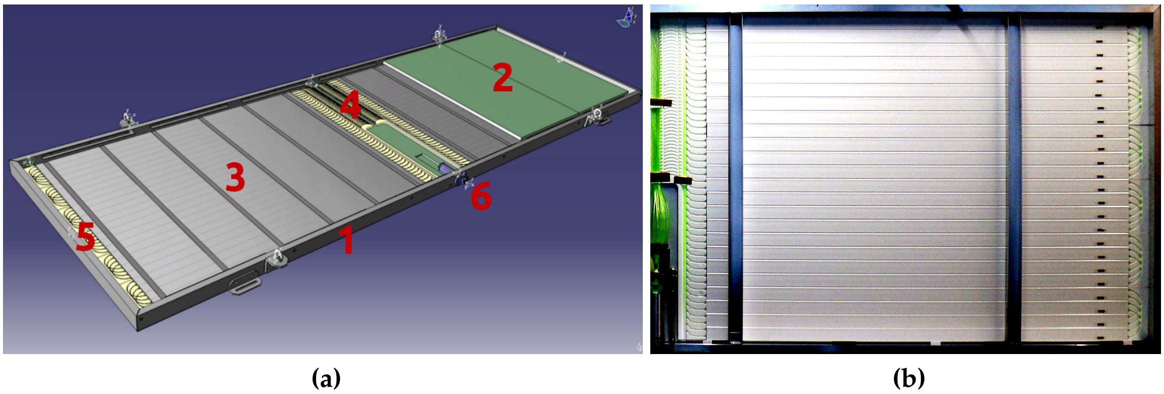
Figure 2.11: (a) Cross section of a scintillator bar. The kidney shaped holes for fiber routing and the outer layer composed of the polystyrene and TiO_2 mixture may be observed. (b) Fluorescence spectra of scintillators produced by FNAL [31] and chosen for the SSD. The peaks marked with "A" and "B" respectively indicate the emission peaks of the two dopants, PPO and POPOP. The black curve represents the reference values, whereas the red and blue curves represent samples of high and low response, respectively, which could be associated with to what degree the desired concentrations of the dyes are achieved during production.



(a) (b)

Figure 2.12: (a) Absorption and emission spectra of the K-27 wavelength shifting dye [33] in the Y-11 (300) MSY fibers produced by Kuraray [32] and chosen for the SSD. (b) Optical fiber bundle inside a PMMA cookie as seen by looking down the readout device tube of the SSD.

Naturally, there are many other materials, properties, and methods involved in SSD construction. Perhaps the most comprehensive overview of the SSD design and construction at present may be found in [34]. A schematic depicting the layout of an SSD may be observed in Fig. 2.13a. In this schematic, the aluminum frame of the detector may be observed (designated with a red "1") along with the filling material (designation "2") and scintillator bars (designation "3") inside the detector



(a)

(b)

Figure 2.13: (a) Schematic depiction of an open SSD. (b) Top-down picture of one side of an open SSD. The contrast on this image has been significantly increased to increase the visibility of the optical fibers.

enclosure. The central and side optical fiber routing (respectively designated with "4" and "5") may also be observed in addition to the covered hole in the frame in which the readout device may be inserted through an aluminum tube (designation "6"). A ventilation hole covered is drilled opposite the hole for the readout devices on the inside wall of the frame and covered with sintered metal for light tightness.

The beams that comprise the frame have a 80 mm by 30 mm profile with 2 mm thick walls. They span 3800 mm and 1280 mm on the longer and shorter sides of the detector, respectively. A 22 mm layer of extruded polystyrene between two 1 mm aluminum sheets composes the filling below the scintillator bars, whereas expanded polystyrene is used above. Together, the filling materials reduce the volume of air inside the scintillator to approximately 10 L. 24 scintillator bars laid tightly side-by-side comprise each of the SSD's two modules resulting in an active scintillating area for the complete detector of 3.84 m^2. The two scintillation modules are located on either side of the detector's central axis on which the readout devices (including the PMT) are located at one end. The optical fibers are cut in 5.85 m segments. Each of these segments is routed away from the SSD's central axis through a hole in one scintillator bar, bent with a diameter of 100 mm and routed back towards the central axis through another scintillator bar. The two ends of each fiber segment are routed along the central axis into a fiber bundle inside a PMMA cookie (transparent for wavelengths larger than approximately 45 nm), shown in Fig. 2.12b, which is coupled with the PMT. A real image of this setup may be observed in Fig. 2.13b. This U-bend fiber configuration increases signal uniformity as photons may travel both directions along the fiber and still reach the PMT.

2.2.2 Engineering Array

During September 2016, 10 prototype detectors were deployed to form a complete hexagon including multiplet arrangements towards the southeast of the more densely filled 750 m area of the SD array. Additionally, 2 prototype detectors were placed inside the 750 m array in a doublet configuration. This set of detectors, which is

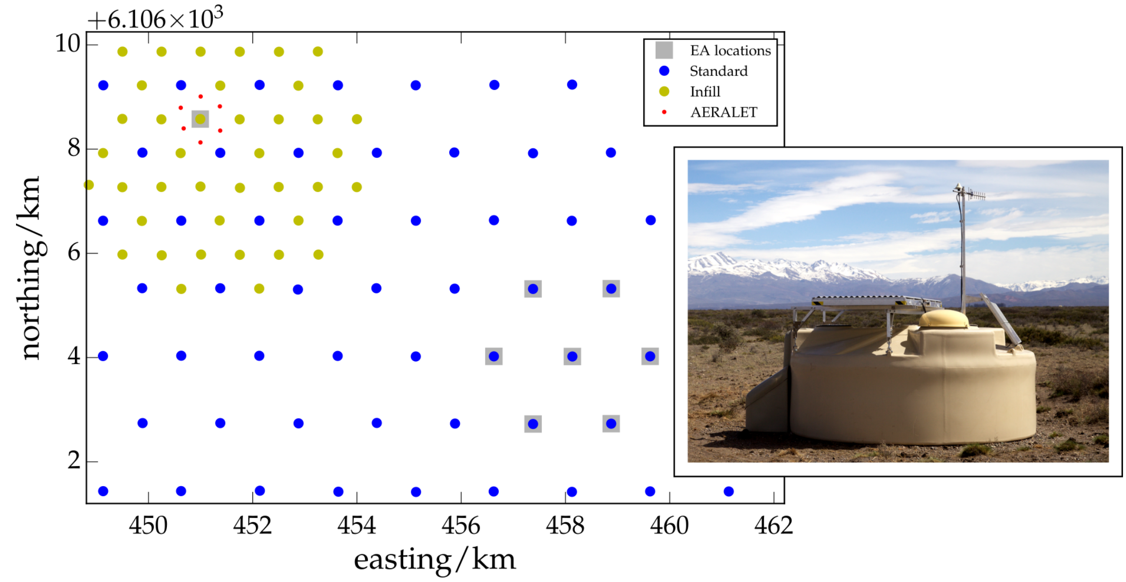
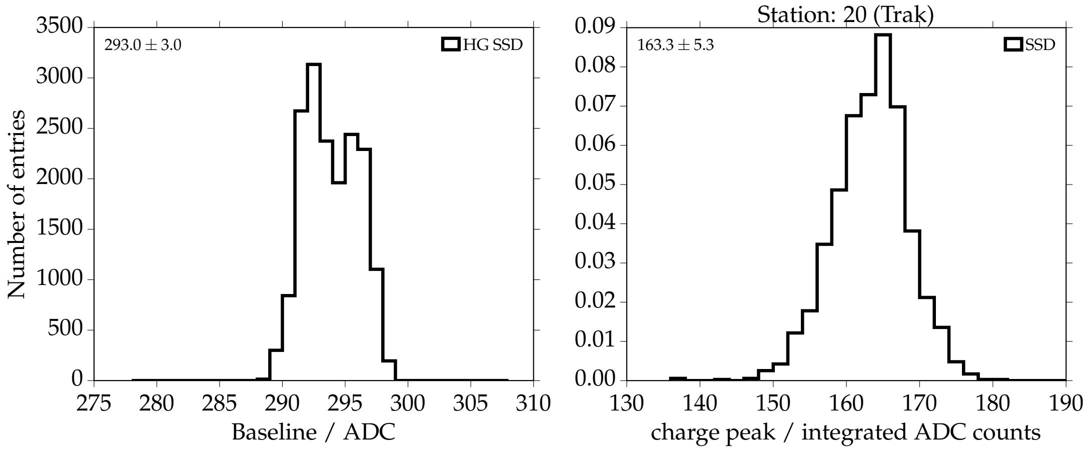


Figure 2.14: Schematic map of the portion of the SD array housing the SSD prototype detectors, which comprise the AugerPrime EA. A photograph of one such prototype SSD atop a WCD is also shown.



(a) **(b)**

Figure 2.15: (a) Distribution of baseline for the high gain channel of an SSD for a station of the AugerPrime EA. (b) Distribution of measured charge per MIP for different calibration periods in an EA SSD.

depicted in Fig. 2.14, has been collecting data as slaves to their partner WCDs for nearly a year and half and continue to do so.

Estimations of electronics properties as well as calibration information were taken from two of the EA detectors for both channels of the 9-inch WCD and the SSD PMTs. These values include baseline, variance therein, and the amount of charge associated with the calibration quantities of a VEM (WCD) and MIP (SSD). Two examples of distributions from which these values are obtained are shown in Fig. 2.15 and the values for the quantities used in simulations are provided in Table 2.2. The specific implementation of these quantities is discussed in detail in Chapter 3.

Table 2.2: Values taken from measurements of electronics and calibration properties of two AugerPrime EA stations for use in simulations are given by the columns indicated with "(UUB)". The analogous values used in WCD simulations using UB electronics are also provided for comparison. Uncertainties are not given for these quantities as, at this point in time, the systematic uncertainties in these quantities between stations of the EA are considerably larger than the statistical errors which can be obtained from individual station measurements. These values are to be taken as first-order, general estimates of what can be expected for the production design SSDs and are subject to change.

Quantity / ADC	WCD (UB)	WCD (UUB)	SSD (UUB)
HG baseline	50	292	300
σ(HG baseline)	0.5	2.0	2.3
VEM or MIP charge	200	1575	197
LG baseline	25	233	226
σ(LG baseline)	0.5	2.5	1.5

CHAPTER 3

Scintillator Surface Detector Simulations

Simulations of the Scintillator Surface Detector enable both an assessment of proposed reconstruction algorithms and the interpretation of real shower measurements. At the core of the SSD simulations lies the use of real detector measurements and parameterizations obtained therefrom wherever possible. Where measurements of a desired quantity are not possible or unavailable, for instance the energy deposit of individual particles inside the scintillating material during runtime, the latest version of the field-standard simulation software, GEANT4 [35] has been employed. The response of the detector is simulated in detail. This includes energy losses of shower particles traversing the polystyrene bars, decays occurring in atoms of the bars and fibers, photon attenuation along the fiber, photoelectrons produced at the PMT's photo cathode, electrical current at the base of the PMT, and the response of the station electronics. Given this level of detail and especially the use of GEANT4 in propagating and calculating energy losses of particles in the detector material, the simulations can be very computationally expensive. Therefore, a station-level particle thinning procedure was also developed and validated to reduce the time required for simulation productions, such as those performed for the GRID, without significant losses in the quality of the output in parameter spaces of interest. Auger's Offline simulation and reconstruction framework was the natural choice to house the simulation software. As such it was modified and augmented to incorporate the SSD simulations as well as the other components necessary in the accommodation of a large-scale detector upgrade. This chapter begins with this upgrade to the Offline software, continues with a detailed description of the SSD simulation procedure and tuning to measurements performed with a prototype detector, and concludes with a summary of the thinning algorithm aimed at reducing computational expenses.

3.1 Offline Upgrade

Auger's Offline software framework [36] has served as a simulation and reconstruction backbone for the collaboration for over a decade and has been adopted or used

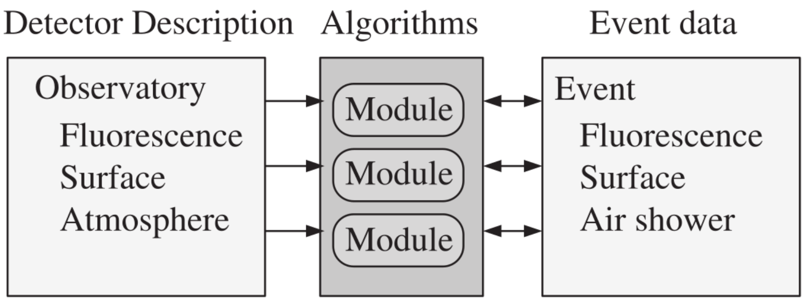


Figure 3.1: Structure of detector description and event interfacing with simulation and reconstruction algorithms (from [36]). While the detector description contains static or slowly changing information, the event contains data specific to the air shower in question. Modules may read from both the detector description and the event and write newly calculated information to the event. Since its conception, the detector description has also been augmented to include Auger's radio and underground muon counter extensions.

in some manner by a handful of other collaborations. In order to house the storage, access, processing, and simulation needs of the AugerPrime hardware, a number of modifications and augmentations of the software were performed.

From its conception, the $\overline{\text{Off}}\underline{\text{line}}$ software framework was designed to take proper advantage of the benefits of C++ and object oriented design in its underlying framework whilst maximizing the simplicity and ease of use of its high-level interfaces. This allowed for a low entry barrier to the development, inclusion, and use of physics simulation and reconstruction algorithms, thereby making the framework accessible to users with varying degrees of computational knowledge. The changes and extensions of the framework to accommodate for AugerPrime adhere to these same principles.

The software framework is broken down into three parts, which are schematically depicted in Fig. 3.1. One part is dedicated to the *detector description* and contains static or slowly changing properties of the surface detector, fluorescence detector, and the atmosphere in addition to the observatory's more recent radio and underground muon counter extensions. The second part stores *event data* and contains information collected by the respective detectors in addition to reconstructed information pertaining to the air shower in question. The third part consists of simulation and reconstruction *algorithms*, which are housed in sequentially processed *modules* capable of reading data from both the *detector description* and *event data* and writing newly derived information to the *event data*.

Whenever a module requests event-independent information from the detector description, the request is relayed to a registry of managers capable of retrieving information from data stored in various sources and formats (see Fig. 3.2 for a schematic representation). In this way, a single interface may be used to access data stored in multiple locations, and users are sheltered from the underlying complexity.

With its new electronics, small PMT, and scintillator, the AugerPrime upgrade demands an enhanced flexibility in the surface detector classes, managers, and storage components of the *detector description*.

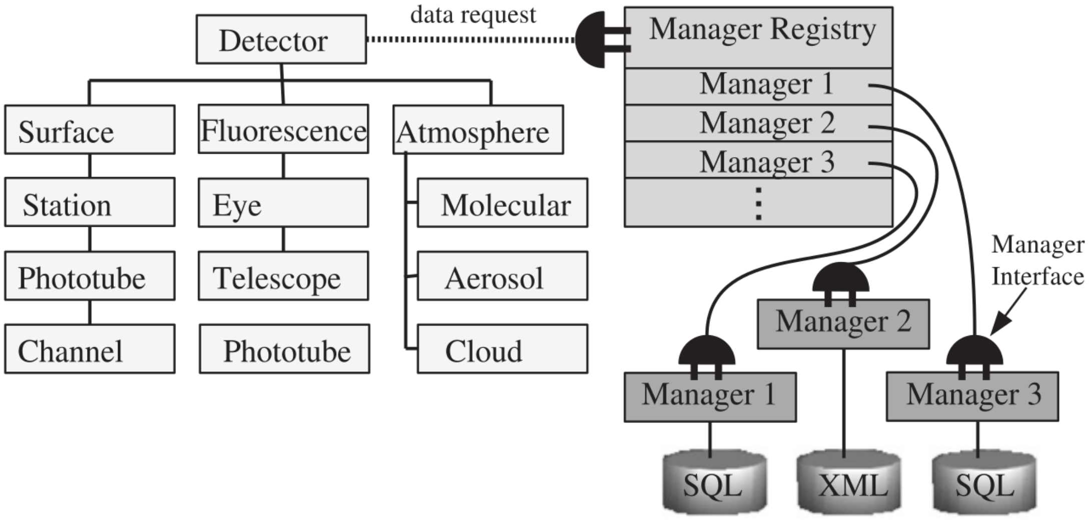


Figure 3.2: Schematic depiction of detector description machinery in $\overline{\text{Off}}\underline{\text{line}}$ (from [36]) including logically motivated classes (left) capable of sending requests for data to the manager registry (right). This machinery has been updated for greater flexibility in the grouping of hardware components and the time evolution of hardware on the level of individual surface detector stations.

Detector description classes corresponding to the new hardware were written. Perhaps most salient were those of the scintillator, which provide interfaces to acquire both mechanical and measured properties of the detector in addition to a number of geometric utilities useful for simulation and reconstruction algorithms. A number of existing detector classes also required modification as many properties of station electronics and PMTs are no longer static or must now be explicitly specified for differentiation purposes. Needless to say, detector update processes required enhancement including methods to construct and destruct various hardware components when appropriate.

Within the detector configuration, it was necessary to dissolve the description of existing station hardware components into separate entities, introduce descriptions of new hardware components, and allow for multiple differing instances of the same hardware type. Additionally, a time variance in the properties of always-existing hardware components in addition to the temporally dependent existence of some components had to be managed. As it is not guaranteed that all surface detector stations will be equipped with each new component of the AugerPrime hardware, it was also necessary to introduce the grouping of an arbitrary number of different types of hardware unique to individual stations at different points in time. Many hardware-model-specific properties were also migrated from specification in *module* configuration files to the detector configuration. This permits, for example, simple iteration over PMTs with differing properties such that simulation and reconstruction algorithms may be generically applied without the need for copious if/else statements and/or redundant code or even *modules*.

The *event data* class structures and constructors were also augmented to mirror that of the detector at the given event time and outfitted with new members and methods to aid modules in their handling of multiple sub-detectors and PMTs of different types.

All of these aforementioned demands have been accommodated in such a way that only the hardware version to which each station is upgraded need be specified in detector configuration files in addition to a time-stamp corresponding to the time at which the upgrade occurred. When accessing data for a given station at a given event time, the managers will pass to the detector classes the parameter values specified for the hardware of the requested version, which are specified in other configuration files corresponding to hardware model descriptions. Simple configuration examples may be found in Appendix B.

The enhancements to the detector configuration, construction, and update processes are invisible to the casual user. As long as detector configuration files are properly updated during deployment, the appropriate hardware classes will automatically be created for a given station within the detector description and the appropriate parameters will be returned through the manager registry in response to user requests.

3.2 Simulation Software and Tuning

A sequence of *modules* exists within the $\overline{\text{Off}}\underline{\text{line}}$ software framework to simulate the response of Auger's SD to the particle flux from extensive air showers arriving at ground. The SSD response is simulated in detail and includes, among other aspects,

- particle position and momenta distributions re-sampled from EAS simulation output
- step-wise simulation of particle interactions inside the SSD
- the total energy deposited by incident particles in the polystyrene bars
- decays and photon attenuation occurring in and along the bars and fibers
- photo-electron production time distributions at the PMT photocathode
- the response of the PMT and the current time distribution at its base
- the response of the station electronics

The details of each of these steps and others are covered in the following sections. A number of physical quantities and distributions either used in the simulations or used to tune simulation parameters were derived from measurements of a prototype SSD detector performed with a muon telescope. A description of this setup is briefly described.

3.2.1 Muon Telescope

In order to investigate the properties of the scintillator bars, optical fibers, and PMT of the SSD, measurements of a prototype SSD detector were performed using a muon telescope. The telescope is one of the 16 modules that made up the muon tracking

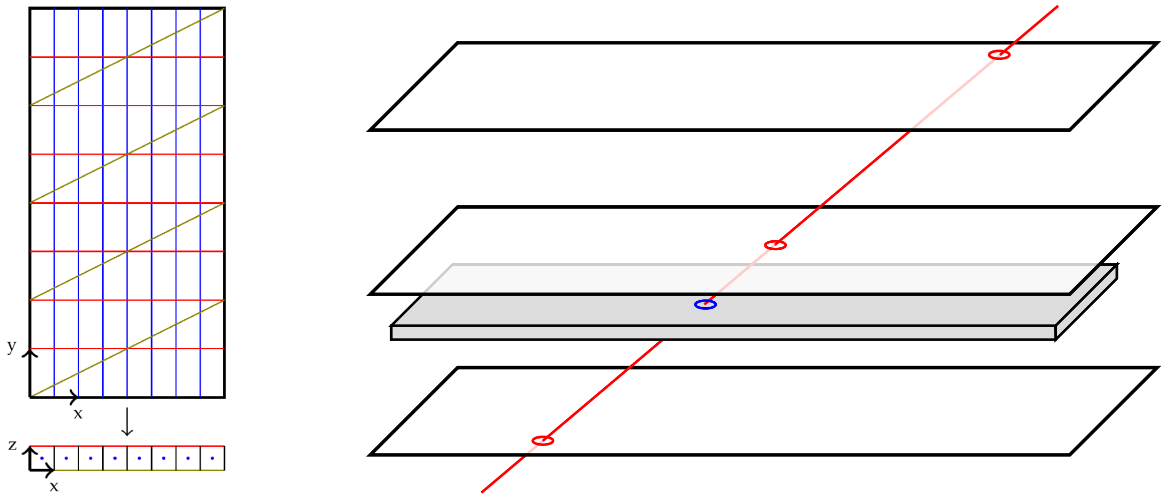
Figure 3.3: Muon telescope of the KASCADE muon tracker re-purposed for characterization of SSDs during the EA and production phases of AugerPrime.

detector (MTD) of the KASCADE experiment [37], which was operational from 1996 until 2009 at the Forschungszentrum Karlsruhe (now the northern campus of the Karlsruhe Institute of Technology). The modified telescope, pictured in Fig. 3.3, consists of three horizontal 4 m × 2 m planes of limited streamer tube (LST) chambers separated by a vertical distance of 82 cm and filled with pure CO_2. Charged particles passing through an LST chamber ionize the gas molecules, and electric fields between the negative pectinate cathode profiles and the positive anode wires inside the chambers accelerate electrons towards the wires, resulting in signals on the order of 100 mV. Perpendicular and diagonal pick-up strips, which by means of electrostatic induction measure the drift of positive ions, lie on top of the planes of LST chambers. Together, the LST and pick-up stripes can localize the crossing point of a charged particle. When all three planes provide a reconstructed position in coincidence, an axis may be fit with positional resolution of 1 cm.

For the analyses used to obtain parameters and distributions for use in simulations or the tuning thereof, one detector was placed between the lower two planes of the telescope and left to collect data. When the muon telescope reconstructed the axis of a crossing particle, readout of the time traces of the SSD was triggered and the reconstructed position of the particle crossing was recorded. For a schematic depiction of this process, see Fig. 3.4b.

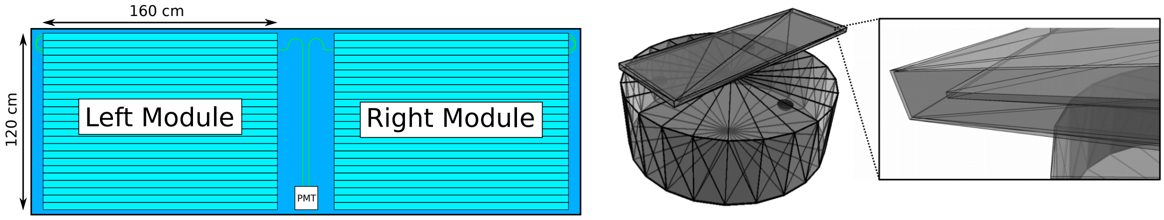
3.2.2 Detector Construction and Particle Injection

The basis of the upgraded SD simulation application is implemented using Version 4.10 of the GEANT4 simulation software package, and the latest geometric and material properties of the EA SSD design described in Section 2.2.1 are used to specify detector volumes. The scintillator bars are constructed of polystyrene and are situated within a larger extruded-polystyrene-foam casing housed inside an aluminum frame. Well considered specification of even non-sensitive volumes is important as they act as shielding thereby influencing measured signal by, for example, changing the the rates of electron absorption and gamma conversion above and below the active plane of the scintillator bars. The positioning of the SSD relative to the



(a) (b)

Figure 3.4: (a) Top-down and side depictions of one plane of the muon telescope. The blue lines and dots represent the anode wires inside the LST chambers. The red and brown lines represent the pick-up strips. (b) Schematic depiction of the three layers of the muon telescope used to reconstruct the tracks of incident particles and the point of crossing on the SSD. Both schematics taken from [34].



(a) (b)

Figure 3.5: (a) Schematic of a SSD, which is comprised of two symmetric modules consisting of 24 polystyrene scintillator bars each for a combined area of 3.84 m^2. (b) Visualization of SSD and WCD volumes as implemented in Geant4 within the Offline framework.

WCD in the Geant4 world volume is also of vital importance as many of the same particles traverse and deposit energy in both detectors, which manifests in correlations in signal, the strength of which should be well reproduced (see Section 2.1.2). A visualization of the SSD volumes above the WCD volumes may be observed in Fig. 3.5b.

In simulations of the detector response to EASs, particles are re-sampled from ground particle files (produced with simulation software CORSIKA [38]) according to the method described in [39]. These particles are then injected into a virtual cylinder that houses both the WCD and SSD that comprise an SD station. In the injection procedure, the momenta of re-sampled particles are preserved and the entry position of each particle is randomized in such a way that reflects the true probability distribution of intersecting the virtual cylinder at different points on its surface. These calculations involve the projection of the top and side planes of

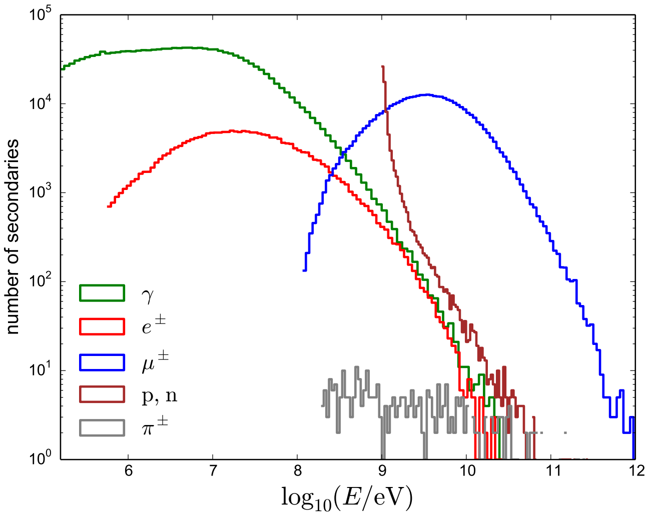


Figure 3.6: Spectrum of background flux of cosmic rays arriving at ground from low energy showers. A set of one hundred thousand 10 GeV muons was used for first-order tuning as this energy is close to the peak of the energy spectrum of atmospheric muons, which compose the vast majority of the data sample of the muon telescope. The vertical muons were simulated uniformly over the surface of the SSD with injection positions a few centimeters above the surface. Figure from [40].

the virtual cylinder into the plane perpendicular to the momentum of individual particles to form an ellipsoid. Particle injection positions are randomized on this ellipsoid and then transformed back onto the true surfaces of the virtual cylinder.

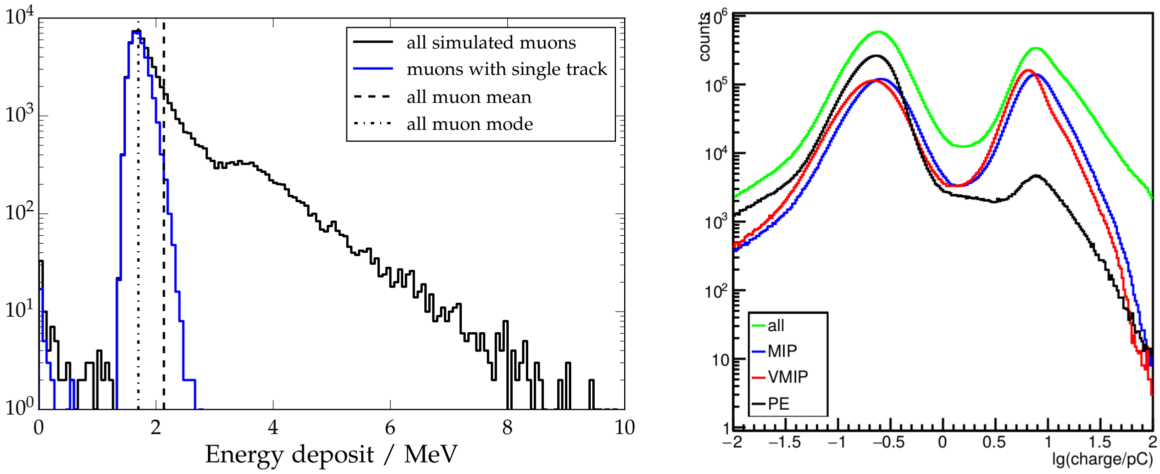
3.2.3 Energy Deposit and Photoelectrons

In the simulations, whenever a particle track crosses the scintillator bar volumes in GEANT4, the amount of energy deposited in the scintillator is extracted. The expected number of photoelectrons $\bar{N}(\mathbf{x})$ produced at the PMT's' photocathode is then calculated via

$$\bar{N}(\mathbf{x}) = N_{\mathrm{ref}} \frac{E_{\mathrm{dep}}}{E_{\mathrm{ref}}} f_{\mathrm{att}}(\mathbf{x}) \tag{3.1}$$

in which E_{dep} is the energy deposit, E_{ref} and N_{ref} are the respective reference energy and photoelectron numbers for the simulations, $f_{\mathrm{att}}(\mathbf{x})$ is a function which describes position-dependent signal attenuation, and $\mathbf{x}$ is the position of the crossing particle. This direct conversion from deposited energy to a photoelectrons number for simulations was pioneered within Auger for the AMIGA scintillators [41, 42].

The reference energy E_{ref} was obtained by simulating a sample of one hundred thousand 10 GeV vertical muons with randomized positions across the surface of the scintillator bars. The peak of the distribution of deposited energy, shown in Fig. 3.7a, is at 1.72 MeV. This peak energy deposit approximately corresponds to the peak in the distribution of vertical-equivalent photoelectron number as determined by an analysis of measurements of SSD modules performed using the muon telescope discussed in Section 3.2.1. In this analysis, the charges of all quality telescope events



(a) **(b)**

Figure 3.7: (a) Distribution of energy deposited by one hundred thousand simulated, vertical 10 GeV muons. The peak of the distribution is $E_{\mathrm{ref}} = 1.72$ MeV. (b) Distribution of charge measured by an SSD triggered using a muon telescope from the KASCADE experiment [37]. Peaks corresponding to single photoelectrons and minimum ionizing particles are visible.

were normalized by their unique track lengths to obtain a vertical- equivalent charge, namely

$$q_{\mathrm{v.e.}} = q/\sec\theta \tag{3.2}$$

where the zenith angle θ of a particle is defined relative to the vertical. The peak of this vertical-equivalent charge distribution was then divided by the peak of the charge distribution of single photoelectrons, resulting in an estimate of the "peak" number of photoelectrons per MIP for vertical events. The two distributions are shown in Fig. 3.7b. As acceptance of the muon tower is position-dependent, this "peak" multiplicity does not exactly correspond to a uniform distribution of intersection points on the SSD's active surface; however, any difference is assumed to lie within the detector-to-detector variance, which is supported by recent analysis of SSD detectors produced for the AugerPrime production.

The position-dependent signal attenuation function $f_{\mathrm{att}}(\mathbf{x})$ was obtained from an analysis of the same muon telescope measurement setup, which provided resolution of approximately a centimeter for positions of particles crossing the SSD. This analysis was performed in the joint context of this work and that of [43]. Theoretically, the dependence of measured charge on the the position of the crossing particle may be be described by

$$\bar{f}_{\mathrm{att}}(\mathbf{x}) = A(\mathbf{x})L(\mathbf{x})\,, \tag{3.3}$$

where $A(\mathbf{x})$ corresponds to the attenuation of photons along the fiber and is defined as

$$A(\mathbf{x}) = e^{-\ell_{\mathrm{c}}(\mathbf{x})/\lambda_{\mathrm{f}}} + e^{-\ell_{\mathrm{f}}(\mathbf{x})/\lambda_{\mathrm{f}}}\,. \tag{3.4}$$

The two summed exponential terms correspond to the fact that photons may travel in either direction along an optical fiber and still reach the PMT given the looped fiber layout schematically depicted in Fig. 3.8 and described in Section 2.2.1. The variables $\ell_{\mathrm{c}}(\mathbf{x})$ and $\ell_{\mathrm{f}}(\mathbf{x})$ correspond to the two different possible path lengths photons

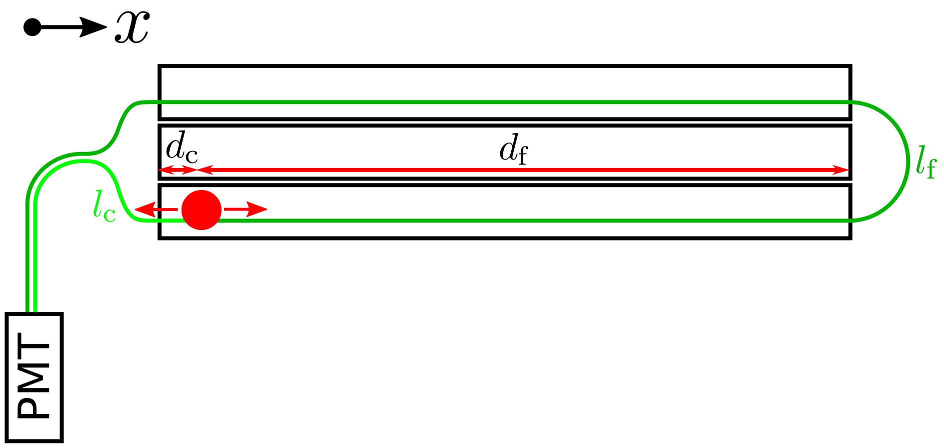


Figure 3.8: Schematic depiction of single particle scenario. Photons produced within a scintillator bar may travel to the PMT in either direction along the fiber. Depending on the direction, photons are attenuated to different degrees. Additional losses are observed for particles crossing near the edges of the scintillator bars.

travel in the optical fibers from the point of particle intersection to the PMT. Naturally, the photon distributions arriving at the PMT from these two different paths will, on average, exhibit different degrees of attenuation for any possible points where a particle enters the scintillator planes. λ_f corresponds to the attenuation length of the fiber. The term $L(\mathbf{x})$ in Eq. (3.3) is a factor accounting for the decreased yield observed for particle tracks close to the edge of the scintillator bars, namely

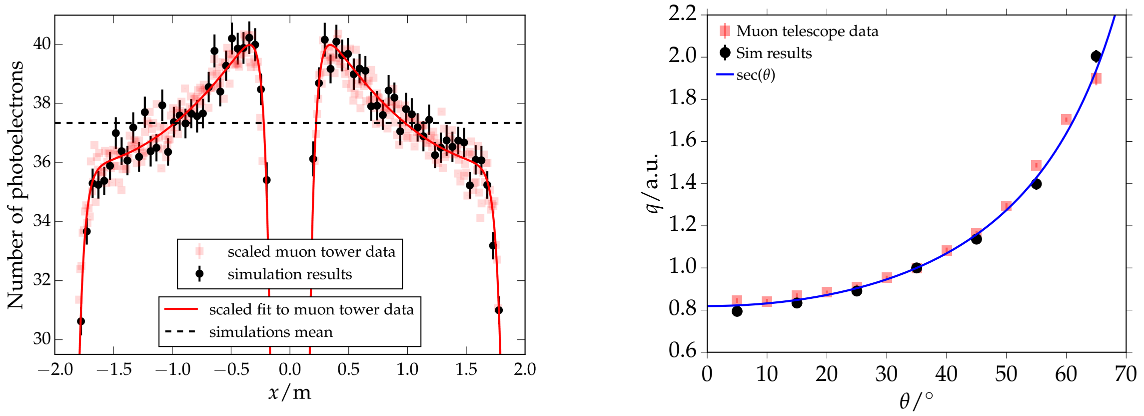
$$L(\mathbf{x}) = (1 - \alpha e^{-d_\mathrm{c}(\mathbf{x})/\lambda_\mathrm{b}})(1 - \alpha e^{-d_\mathrm{f}(\mathbf{x})/\lambda_\mathrm{b}}) \, , \tag{3.5}$$

where α corresponds to the maximum loss, λ_b corresponds to the effective attenuation length of this boundary effect, and $d_\mathrm{c}(\mathbf{x})$ and $d_\mathrm{f}(\mathbf{x})$ respectively correspond to the distance between the particle crossing point and the close and far ends of the scintillator bars. The exact cause of this edge effect has yet to be determined. The possibility of increased photon losses at the ends of scintillator bars was ruled out by performing a measurement in which the ends of the scintillator bars were covered with reflective materials. In this test, the edge effect remained unchanged. Nonetheless, the effect can be well modeled. The reference number of photoelectrons N_ref from Eq. (3.1) is acquired by mandating that the peak in the distribution of the number of photoelectrons produced by vertical minimum ionizing particles across the surface of the scintillator equals the estimate of 30 obtained from muon telescope analysis and denoted as $N_\mathrm{PE}/\mathrm{VMIP}$ in Eq. (3.7) [30]. This mandate is fulfilled by integrating f_att only along the path length of the fiber between the ends of the scintillator bars, namely

$$\bar{f}_\mathrm{att} = \frac{1}{p_\mathrm{f} - p_\mathrm{c}} \int_{p_\mathrm{c}}^{p_\mathrm{f}} f_\mathrm{att}(\ell) \, \mathrm{d}\ell, \tag{3.6}$$

where p_c and p_f respectively correspond to ends of the bars with shorter and longer path lengths along the fiber to the PMT. The quoted photoelectron count is then divided by the result to obtain N_ref,

$$N_\mathrm{ref} = \frac{N_\mathrm{PE}/\mathrm{VMIP}}{\bar{f}_\mathrm{att}} \, , \tag{3.7}$$



(a) **(b)**

Figure 3.9: (a) Simulated number of photoelectrons (black markers) alongside scaled muon telescope measurements (red markers). The dark red line represents the fit of the attenuation model $\bar{f}_{\mathrm{att}}(\mathbf{x})$ of Eq. (3.3) to muon telescope data. (b) Signal as a function of the angle of incident particles for simulations (black markers) and muon telescope measurements (red markers).

which is a quantity independent of the lengths of the fibers. In the actual simulation code, the expected number of photoelectrons is calculated for each individual leg of the signal instead of as one quantity as each leg exhibits a different start time in the photoelectron production time distribution. Nonetheless, the formulation separating the two legs is equivalent to the formulation of Eq. (3.3) when summed. A Poisson randomization is then performed on the expected number of photoelectrons for each leg of the signal to obtain the true number of photoelectrons produced at the PMT photocathode. The results of a set simulations of vertical muons may be observed in Fig. 3.9a alongside the fit of $f_{\mathrm{att}}(\mathbf{x})$ to muon telescope measurements. The mean number of photoelectrons observed here is greater than the peak value used for tuning due to the asymmetry in the distribution of energy deposit depicted in Fig. 3.7a, which may also be observed in the charge distribution from muon telescope measurements shown in Fig. 3.7b. By dividing by the ratio of the mean and peak of the energy deposit distribution, the peak number of photoelectrons of 30 is recovered.

Once the number of photoelectrons has been determined for a given particle, the production times at the photocathode must be simulated. The distribution of arrival times for photons having traveled nearly the same length of fiber before reaching the PMT was examined using muon telescope measurements. Although the true decay schema is likely considerably more complex, a model of two sequential exponential decays was found to provide a reasonable description of the data as shown in Fig. 3.10. The moderate differences between the model and the data could be explained by the re-absorption of photons produced in the fiber one or more times, which is possible given the overlap in the absorption and emission spectra of the optical fibers shown in Fig. 2.12a. A visual comparison between the time distributions at various positions measured with the muon telescope and the results of the parameterization are given in Fig. 3.11.

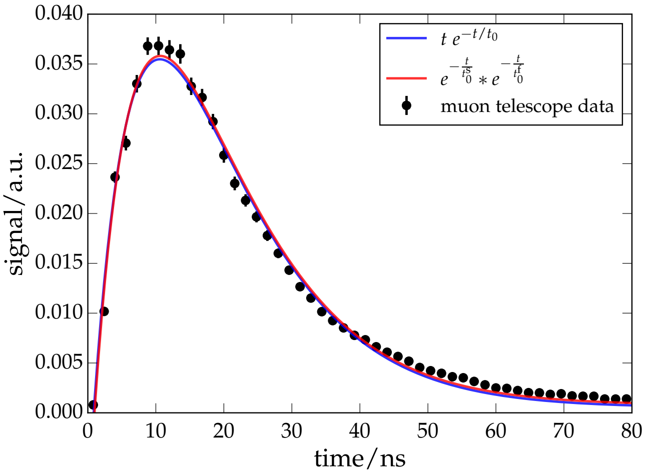


Figure 3.10: Fit of a model of two subsequent exponential decays to signals of the SSD from particles crossing at its end furthest from the PMT, where the distance traveled by the photons corresponding to the two legs of the signal is equal. This two parameter model may be simplified to a function with one parameter since the two decay constants are very similar in magnitude.

The decay constants from the fit are used to define two exponential decay time distributions, which are randomly sampled for each photoelectron. Each photoelectron's arrival time is then determined by summing the "global" time at which the particle crossed the scintillator bar, the time photons needed to travel within the fiber from the crossing point to the PMT, and the two decay times. The time delay due to travel along the fiber is calculated using an effective index of refraction also obtained from a template-based fitting procedure applied to muon telescope measurements. The time distribution from events incident on the active surface of the SSD at the ends of the scintillator bars furthest from the PMT were taken as a template for the shape of one leg of signal for the same reasons as those for the selection made for the fit of the decay constants. However, instead of performing a fit of this trace, a linear interpolation between time bins was used as a template. Two such templates shifted in time by an amount defined by the free parameter referred to as the "effective index of refraction" n_{eff} were fit to events with tracks crossing any point of the active surface of the SSD. Different scaling factors to account for differences in attenuation between the two signal legs were also applied to each of the templates. These scaling factors were both left as free parameters and fixed according to the quantity for the attenuation length of the fiber with little difference in results. The resulting index of refraction is referred to as "effective" because it assumes photons travel in a straight line along the fiber instead of repeatedly reflecting off of one of the multiple surfaces of the different layers of the multi-clad fiber. Nonetheless, it well reproduces the travel times of photons within the fibers.

3.2.4 PMT

Once the cumulative trace of photoelectron production times resulting from all incident particles is obtained, the PMT is simulated as follows. For each photoelectron, the measured single photoelectron pulse shape is scaled by a random draw from

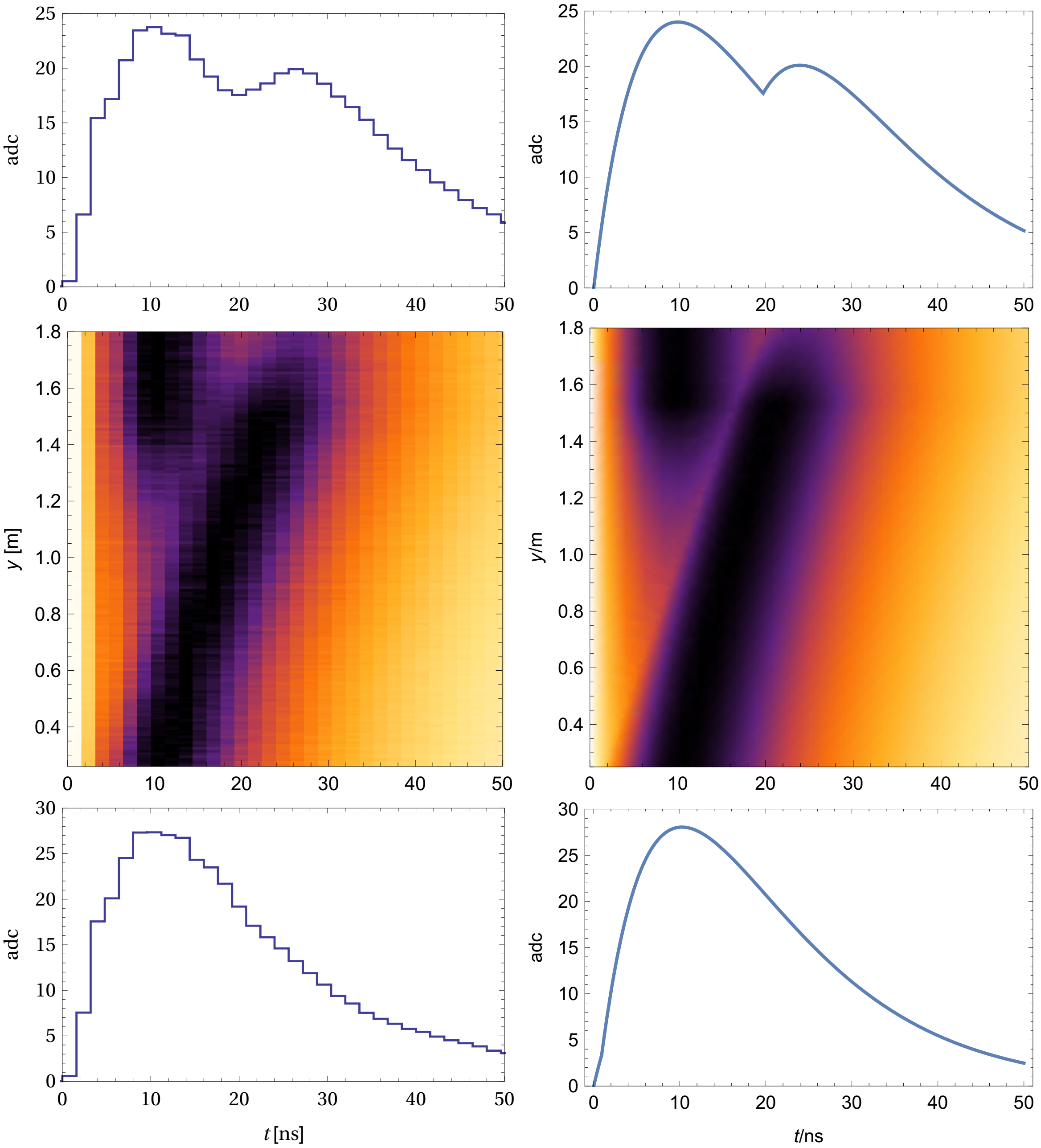


Figure 3.11: Depiction of the time evolution of the signal in one module of the SSD. Muon telescope measurements are shown in the plots on the left and the parameterization of these measurements are shown in those on the right. *Middle:* Density plot of the signal evolution over one module of the SSD (slices in the horizontal axis have been normalized to the same maximum). *Top and bottom:* Signal shapes at the close and far ends of the *middle* plot.

the measured charge distribution (see Fig. 3.13a). The resulting scaled pulses are linearly added to a base current time distribution. This procedure makes two assumptions. The first is that the shape of the pulse at the base of the PMT does not depend on charge. The second is that the response of the PMT is linear. Post-PMT signal amplification in the upgraded electronics [44] has been set such that the limit of the dynamic range of the electronics is reached prior to PMT departure from

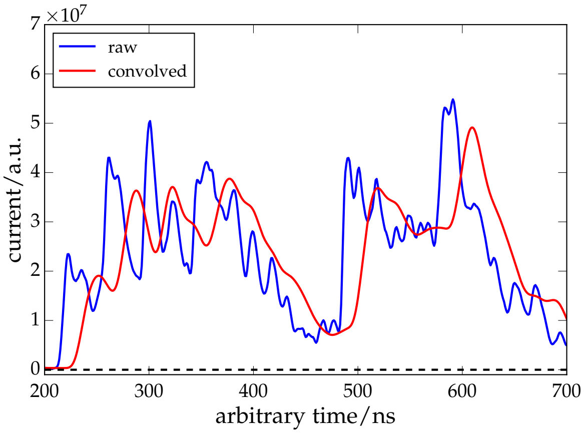
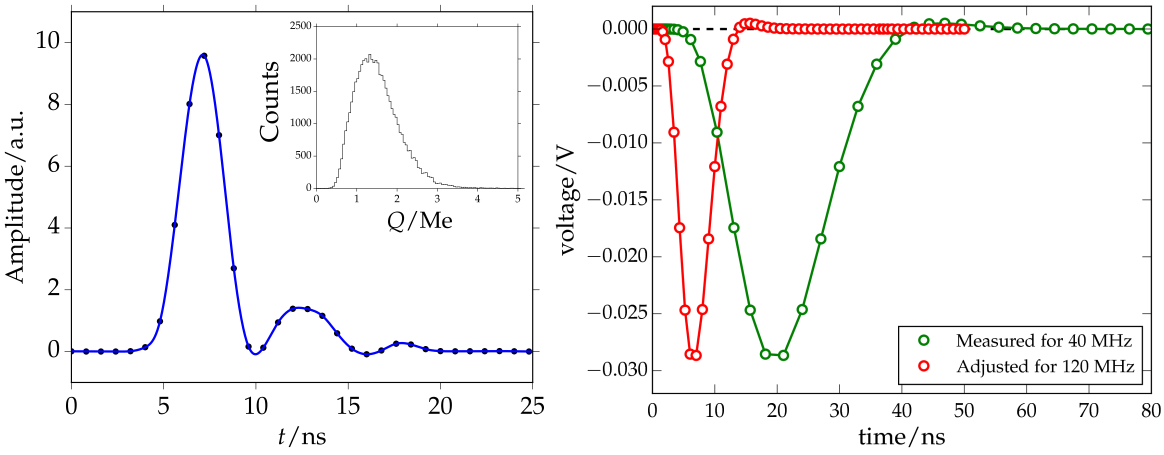


Figure 3.12: Sample simulated current at the base of the PMT and convolution with the transfer function of the station electronics.



(a)

(b)

Figure 3.13: (a) Pulse shape and charge distribution (inset) for properly amplified photoelectrons as measured using the the muon telescope setup. The small second peak may be due to unresolved issues in the determining the start times of these pulses or elsewhere in pulse-finding algorithms applied to the telescope data. (b) Transfer function measured for Auger's 40 MHz electronics along with scaled version used for the 120 MHz simulations up to the present.

non-linearity. The charge distribution and pulse shape were obtained from single photoelectron traces measured by the muon telescope. Future measurements in a dark box may provide higher quality estimations of these distributions.

3.2.5 Electronics

The procedure for simulating the post-PMT electronics is as follows. The current at the base of the PMT is convolved with the transfer function measured for the 40 MHz unified board [10] scaled to accommodate for the higher frequency of the 120 MHz upgraded unified board (see Fig. 3.13b). Given the similar designs of the standard and upgraded front ends, it is assumed that this scaling yields a reasonable estimate for the transfer function, the form of which should fall within the variance expected between boards. Nonetheless, a direct measurement of the transfer function will be considered once the exact production design of the upgraded unified board is confirmed. Once this convolution is performed, the result is scaled with a parameter analogous to gain and sampled at 120 MHz to obtain ADC traces. The parameter analogous to gain is tuned such that the mean charge (in ADC) of a vertical muon passing through any point of the SSD surface equals the charge measured by deployed EA prototypes as determined via a vertical equivalent muon calibration detailed in [14]. A distribution of charge obtained from different calibration periods for an SSD in the AugerPrime EA and used in this tuning was shown in Fig. 2.15b. A pedestal to each ADC trace is also added with mean values and bin-to-bin variances matching measurements of the EA, an example distribution of which was shown in Fig. 2.15a.

3.2.6 Calibration

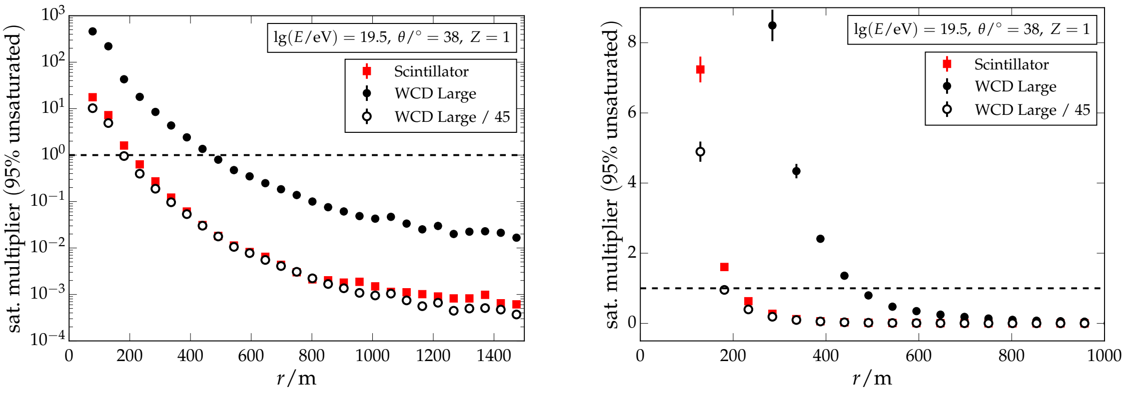
For validation and redundancy, both the simulated WCD and SSD signals are recalibrated with ensembles of vertical 1 GeV muons. In the case of the WCD, these muons are injected along the central axis of the tank. In the case of the SSD, they are injected evenly across its entire active surface. The charge resulting from this calibration is expected to match the charge to which the gain parameter of the simulations was tuned.

3.2.7 Dynamic Range

A cross check making use of the saturation multiplier defined in Eq. (2.15) was performed for both the SSD and WCD with simulated upgraded 120 MHz electronics. The objective was to set the gains of the the two channels of the SSD to match the dynamic range of the small PMT planned for the WCD, where both detectors should saturate at approximately 200 m from the shower axis at the highest energies. For the tested energy of $10^{19.5}$ eV, the saturation multiplier reaches a value of 1 at around 200 m.

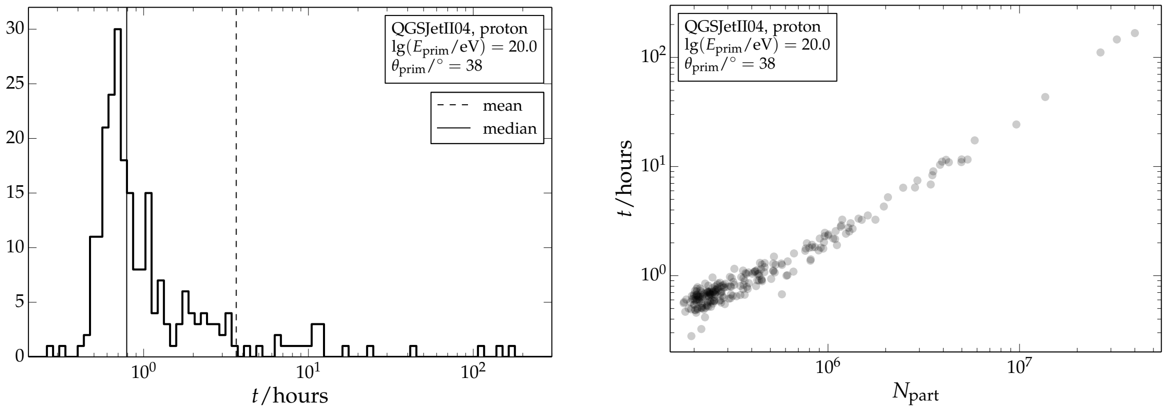
3.3 Optimization for Grid

When randomly throwing showers on an array during detector-response simulations, it is not so uncommon for the shower core to land near an SD station. In such cases, a massive number of particles intersect this station, and the resulting time required for its simulation, and consequently that of the event, is very high. In Fig. 3.15a, the distribution of simulation run-times t for a typical set of 240 Monte-Carlo SD simulations performed on a computing cluster at the Karlsruhe Institute of Technology for primaries of the same energy, zenith angle, and composition is shown. Run time ranges from a fraction of an hour to several dozen hours and depends on the number of particles injected in all stations. In the case of large numbers, the computation time is dominated by the mostly linear dependence of the `G4TankSimulator` module on the particle number N_{part}, as depicted in Fig. 3.15b.



(a) **(b)**

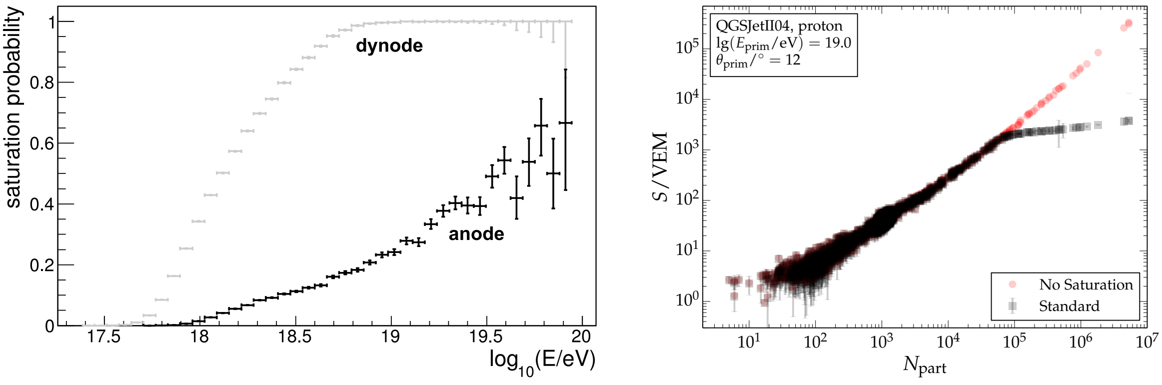
Figure 3.14: Saturation multiplier to guarantee less than 5% saturated stations as a function of distance in (a) logarithmic and (b) linear scales. The linear scale makes evident the range in distance where the slope of the LDF becomes very steep and demands on the upper limit of the dynamic range rapidly become unreasonable. Both the SSD and the planned small PMT in the WCD saturate at approximately the same distance, which is approximately 200 m for the energy and zenith angle depicted here.



(a) **(b)**

Figure 3.15: (a) Distribution of simulation run-times t for 240 jobs. There are a number of outliers at higher times due to shower cores landing near SD stations. (b) Dependence of the simulation time t on total number of particles N_{part} injected in all stations. The dependence is linear for large N_{part}.

With increasing primary energy, the probability of saturating both the dynode and anode of a WCD's PMT increases to the point of occurring in more than half of events at the highest energies (see Fig. 3.16a). Fig. 3.16b shows that the FADC signal S (black) of a station saturates above at a certain number of particles, at which point the dependency of S on the number of injected particles diverges from that of the unsaturated signal (red). It is clear that in stations with more than $\sim 10^5$ particles, an increasing amount of computational power is spent in simulating signals that will almost certainly saturate and be truncated. This is a gross waste of computing resources and should be avoided if possible.



(a) **(b)**

Figure 3.16: (a) Saturation probability as a function of the primary energy (in Auger data [45]). Signal estimation switches from using the dynode to the low-gain anode trace as soon as the former is saturated. Anode saturation is only partially recoverable with some empirical methods. (b) Dependence of station signal S on the number of injected particles N_{part}. Around a particular particle number, the dependency changes due to anode saturation.

3.3.1 Station Weighting Algorithm

To avoid this computational waste, a thinning procedure on station particle lists and a corresponding station-based weighting system is introduced. As the ordering of particles reaching the ground in shower simulations can have physics repercussions, it is crucial that such a thinning procedure ensures no biases are introduced with respect to arrival time, particle type, or particle energy. Each particle considered for injection must have an equal probability of being accepted.

These requirements are satisfied with the following algorithm. The SD simulation is initially permitted to proceed normally by indiscriminately adding particles resampled from a Cosmic Ray Simulations for Kascade (CORSIKA) ground particle file to the list of particles slated for injection. When the particle count reaches a predefined limit, N_{max}, a new, shorter list is randomly assembled from the existing one. To keep the procedure simple, the probability of keeping any one particle is mandated to be 50% and therefore the list is shrunk to approximately half of its size. At the same time the original station weight w of 1 is increased by factor 2. From this point on, new particles are added to the list with probability $1/w$. A depiction of this process may be observed in Fig. 3.17.

If the particle number limit is reached a second time later in the simulation, the list is randomly halved again and the weight is again multiplied by 2. New particles are then added to the list with a probability $1/w$, where w is the new weight of 4. This procedure is repeated until all particles have been processed. Most stations in an event do not reach the limit even once and, therefore, have a station weight of 1. In some events, one station will reach the limit and will have a decimated list and a station weight greater than 1. In rare cases of showers with extremely high energies, two stations will reach the limit.

In this way all particle injection lists are guaranteed to be shorter than N_{max} and for saturated stations, the number of injected particles will be somewhere between

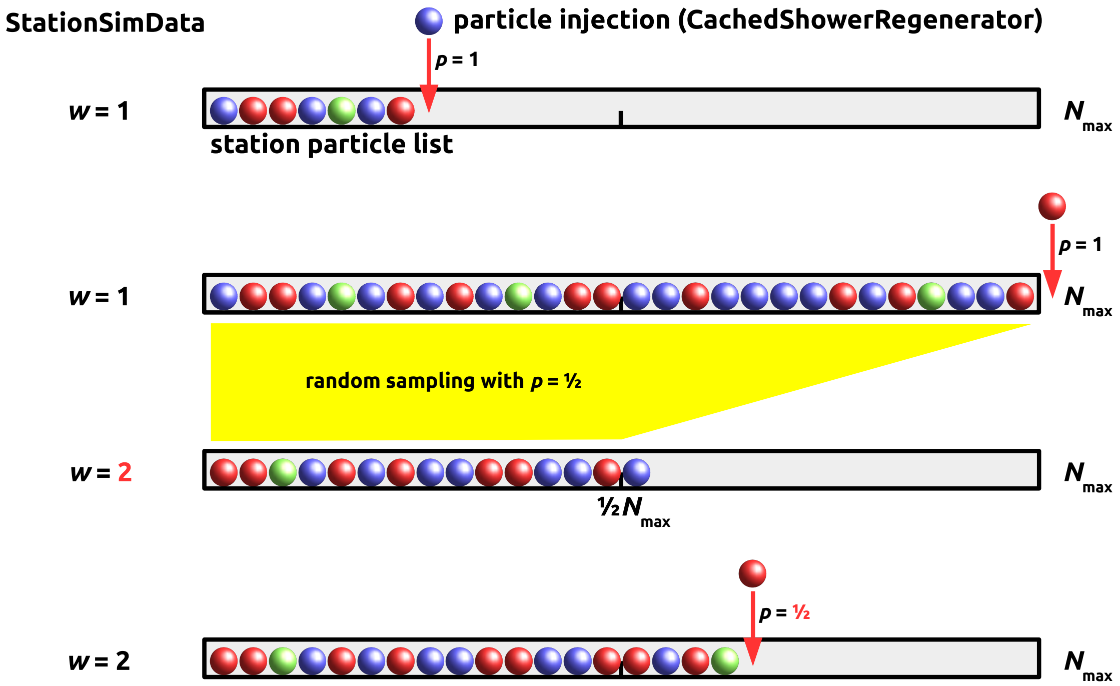


Figure 3.17: Schematic depiction of particle thinning algorithm.

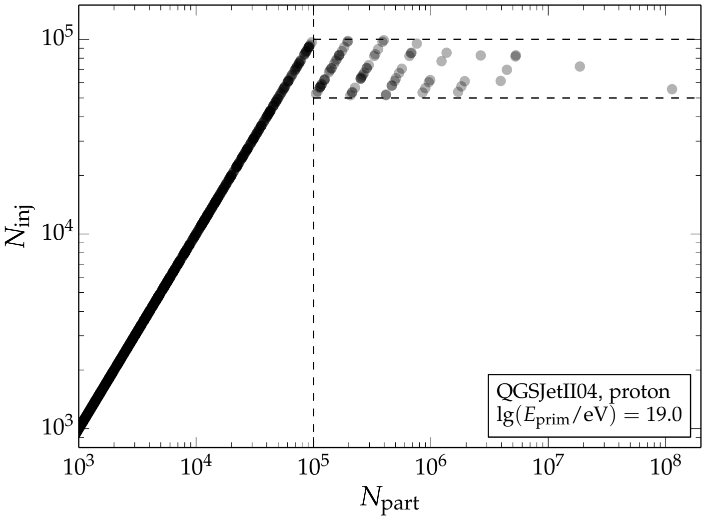
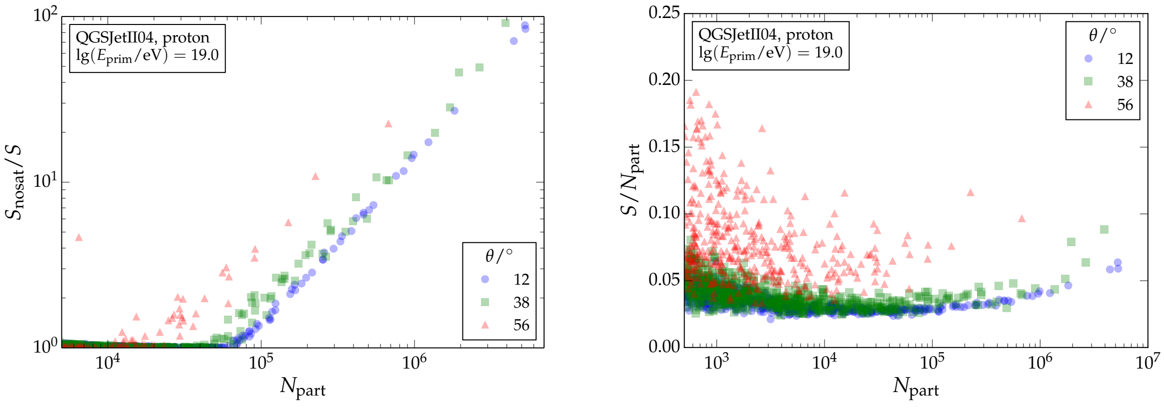


Figure 3.18: Number of particles injected during simulation as a function of total number of candidate particles re-sampled from the CORSIKA ground particle file for single stations. If the particle limit N_{max} is never reached for for a station, all re-sampled particles are injected and the ratio of injected to all particles is 1. Some particle lists undergo the decimation and refilling process one or more times as evidenced by the diagonal bands.

$N_{\mathrm{max}}/2$ and N_{max}. Put simply, this procedure ensures that when such thinning is applied for a certain station, the probability for the injection of a particle is independent of its type, momentum, arrival time, and its position in the shower ground file, and that this probability amounts to exactly $1/w$ for all particle candidates. Depictions of the thinning algorithm may be observed in Fig. 3.18.



(a) **(b)**

Figure 3.19: (a) Ratio of signal without and with simulation of electronics saturation as a function of station particle number. The particle number threshold at which saturation occurs depends on zenith angle. (b) Signal per particle as a function of station particle number. See Fig. B.2 in Appendix B for the corresponding muon fractions.

Once a final particle injection list is obtained, GEANT4 modules inject the particles into the WCD volumes and simulate photon production and propagation in order to obtain distributions of photo- electron release times at each PMT. The PMT simulation module `SdPMTSimulator` then converts these distributions into a current output at the bases of the PMTs. At this point, the base current is scaled by the ratio of all original particle candidates to injected particles, which is equal to $1/w$ on average. In doing so, the assumption is made that the simulation of all particles (i.e. no thinning) would produce a scaled version of the signal acquired using the station weighting algorithm.

3.3.2 Determination of N_{max}

The particle number threshold N_{max} at which the station weighting algorithm is employed was determined from the thresholds at which signals saturate for a variety of energies and zenith angles. As shown in Fig. 3.19a, these thresholds range between 10^4 and 8×10^4 particles, and depend on zenith angle. Part of the reason for this zenith angle dependence is the larger muon fraction in more inclined showers (see Fig. B.2 in Appendix B). Since muons deposit significantly more energy in a WCD than particles belonging to showers' electromagnetic component, the signal per particle is greater on average for showers of larger zenith angle as depicted in Fig. 3.19b. Other reasons may be a larger degree of time compression in particle arrival times for more inclined showers and differences in track lengths inside the WCD.

For testing purposes, the threshold N_{max} was chosen to be 10^5. All plots in the following sections made use of the algorithm with N_{max} set at this value.

3.3.3 Signal Verification

Analyses were performed to verify that no significant biases or changes in signal resolution result from the station-weights algorithm. SD response was simulated four times for each shower with the core fixed to the same location in each instance. The station-weights algorithm was enabled in two of the four instances and, in each of these two cases, one set used a different set of random numbers than the other.

Bias

To determine whether or not the station-weights algorithm introduced a bias in simulated signals, the difference between signals simulated with and without the algorithm were examined with the statistic of interest being

$$\frac{S_{\mathrm{sw}} - S_{\mathrm{orig}}}{\sqrt{\delta_{\mathrm{sw}}^2 + \delta_{\mathrm{orig}}^2}} \tag{3.8}$$

where S_{sw} is the signal with the algorithm enabled, S_{orig} is the signal with the algorithm disabled, and δ_{sw} and δ_{orig} are the respective uncertainties in the total signals from the uncertainty model implemented in $\overline{\text{Off}}\underline{\text{line}}$. The rationale of normalizing each signal difference by the associated uncertainty is to ensure that signals of differing magnitudes contribute approximately equally in calculating the parameters of the resulting distribution. Ideally, one would use an uncertainty model encompassing detector response but not fluctuations arising from the shower itself; however, such a model has yet to be obtained.

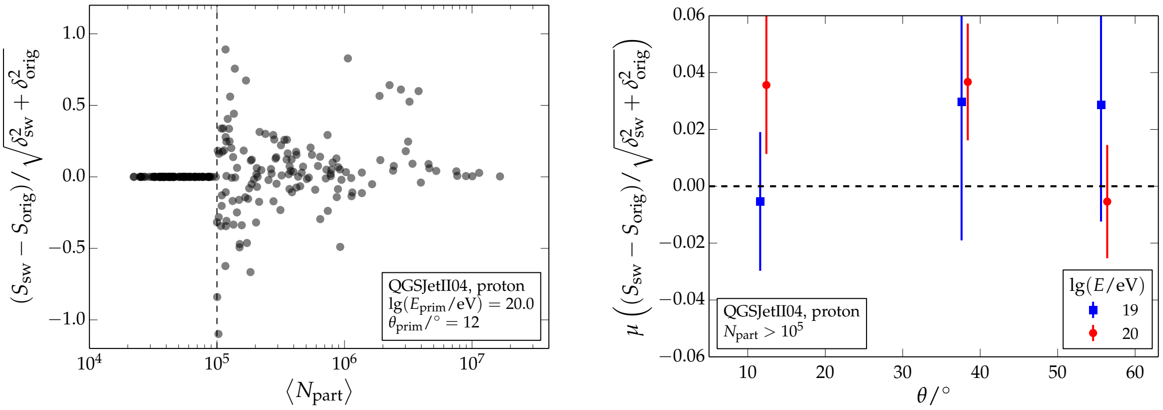
If the station-weights algorithm does not introduce a bias in simulated signals, the mean of chosen statistic's distribution (see Fig. 3.20a) should be zero. Fig. 3.20b shows that this is indeed the case within one or two standard errors for all tested energies and zenith angles.

Resolution

To determine whether or not the station-weights algorithm significantly alters signal resolution, the magnitude of signal Monte-Carlo fluctuations in instances where the algorithm was enabled and where it was not were examined. Estimations of the fluctuations in each case were obtained by performing simulations with different sets of random numbers. A distribution was constructed from the pairs of signals using the statistic

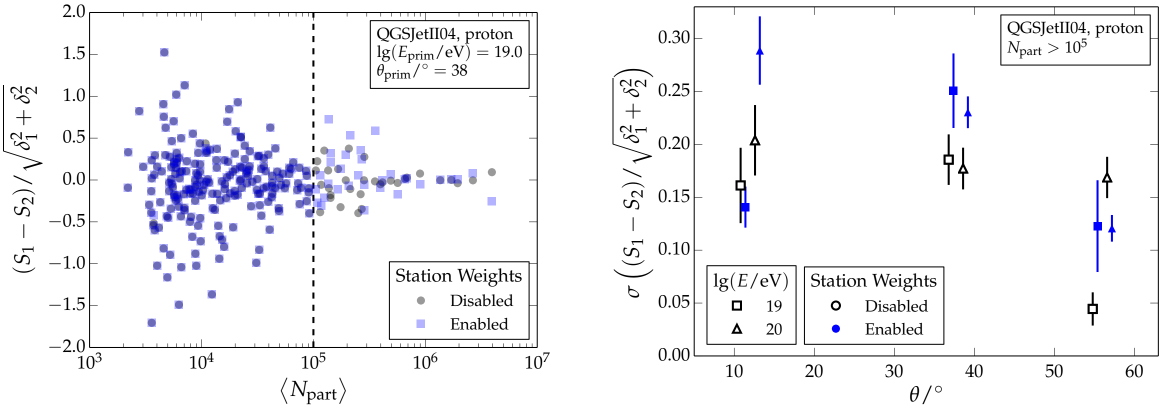
$$\frac{S_1 - S_2}{\sqrt{\delta_1^2 + \delta_2^2}} \tag{3.9}$$

where S_1 and S_2 correspond to the signals resulting from the use of different sets of random numbers, and δ_1 and δ_2 are the respective uncertainties in the total signals from the uncertainty model implemented in $\overline{\text{Off}}\underline{\text{line}}$. The rationale behind this normalization is identical to that described for the statistic of Eq. (3.8). The width of this statistic's distribution is representative of the Monte-Carlo fluctuations, and if the station-weights algorithm does not alter signal resolution, it should be the same in the case of algorithm's use and disuse. Fig. 3.21 shows that the fluctuations increased in some tested cases and decreased in others but were mostly within a couple of standard errors of one another.



(a) **(b)**

Figure 3.20: (a) Spread of normalized signal differences (see Eq. (3.8)) as a function of particle number for cases where the station weights algorithm was enabled and disabled. The dashed vertical line is the threshold at which the algorithm takes effect. For stations below this threshold, the algorithm is not applied, which is reflected in the fact that signals are identical for simulations where the use of the algorithm was enabled and disabled. (b) Mean normalized difference in signal for stations above the algorithm's threshold. In each case, the difference is compatible with zero within a couple standard errors.

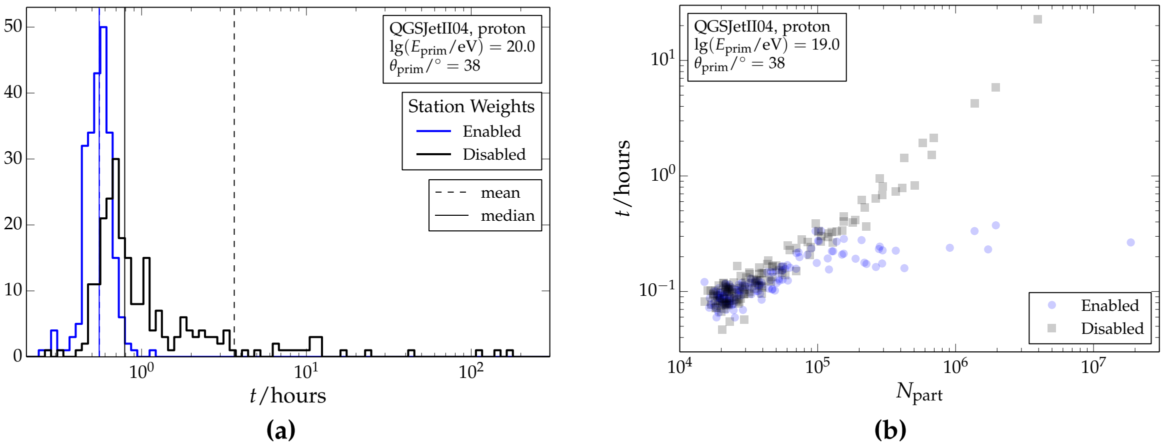


(a) **(b)**

Figure 3.21: (a) Spread of normalized signal differences (see Eq. (3.9)) as a function of particle number in the case of different sets of random numbers used during simulation. The dashed vertical line is the threshold at which algorithm takes effect. For stations below this threshold, the algorithm is not applied, which is reflected in the fact that signals are identical for both simulations where the use of the algorithm was enabled and disabled. (b) Estimators of the Monte-Carlo resolution for stations above the algorithm's threshold. The resolution estimator is larger without use of the algorithm in some cases and larger with use of the algorithm in others, but the estimates are, in most cases, within a couple of standard errors of one another.

3.3.4 Results

In Fig. 3.22, the simulation run-times with and without the station-weights algorithm enabled using the same set of showers are shown. Use of the algorithm removes



(a) **(b)**

Figure 3.22: (a) Distribution of simulation times t for identical simulations with the exception of the enabling of the station-weights algorithm. Use of the algorithm removes outliers and reduces the mean run time. (b) Dependence of the simulation time t on total number of particles N_{part} injected in all stations with and without station-weights algorithm enabled.

outliers and lowers both mean and median run-times. Additionally, all simulations for all energies and zenith angles completed successfully with the algorithm enabled whereas without the algorithm, a handful crashed or had to be killed due to memory and time limitations.

3.3.5 Conclusion

Use of the station-weights algorithm will benefit large-scale, detector- simulation productions of hundreds of events or more. When using the algorithm, average run-times are decreased by dozens of hours, and the occasional instances that would have taken hundreds of hours before crashing or being killed are eliminated. Additionally, geometries in which a station is very close to the shower axis, previously excluded from productions (explicitly or implicitly due to computational limitations), are now included. Users running only a few simulations but including stations with very large signals may also benefit, depending on their purposes.

Although a station weighting factor of 2 and a particle number threshold N_{max} of 10^5 were demonstratively used for the validation of the method in these sections, these variables can be set more conservatively or optimized for increased simulation speed. Both variables may be set in the XML configuration of the appropriate module in the upgraded $\overline{\text{Off}}\underline{\text{line}}$ software described in Section 3.1.

Many simulation algorithms involve signal parameterizations which are inadequate in their descriptions of the numerous geometric nuances, infrequent interactions, and other rarities which naturally arise and which only a full detector simulation can capture. Such inadequacies are often especially apparent in the simulation of low signals induced by only few particles, where oddities do not average out. The station algorithm addresses these issues by, in a manner of speaking, taking time from particle-rich stations, where averaging algorithms work well, and giving time to particle- poor stations, which require it.

3.4 Conclusions

Simulations of the AugerPrime SSD have been implemented within Auger's $\overline{\text{Off}}\underline{\text{line}}$ software framework. At the core of these simulations lies the use of models and parameterizations derived from real prototype detector measurements where possible and the use of the latest field-standard software, Geant 4.10, where measurements are impossible or unavailable. The simulations have been tuned to the latest data from the AugerPrime EA and promise to aid in both a detailed understanding of the upgraded SD as well as interpretation of future measurements of the observatory as the upgrade moves into the deployment phase. The $\overline{\text{Off}}\underline{\text{line}}$ software has also been modified and augmented with the classes and machinery necessary to accommodate the data of the new hardware with the enhanced flexibility required by a detector description for which both the existence and properties of hardware components vary between stations and in time.

CHAPTER 4

Reconstruction of Mass Composition

Estimation of the mass of individual UHECRs is the principle task in achieving the physics goals of the AugerPrime upgrade. With a comprehensive SD simulation application situated in the upgraded $\overline{\text{Off}}\underline{\text{line}}$ software framework, reconstruction algorithms may be developed and validated. This chapter presents the initial formulation and assessment of one such algorithm including the preliminary application to the selection of proton-enriched event samples from the background of two astrophysical scenarios in the region of the flux suppression. As will be demonstrated, a strength of this particular algorithm is its minimal dependence on hadronic interaction models.

4.1 Formalism

In this section, a simple algorithm used to obtain a parameter correlated with the number of muons reaching the ground in an EAS is introduced. This muon content estimator is, in turn, correlated with the mass of the primary that induced the shower and can therefore be used for mass discrimination purposes. As this algorithm is formulated as a matrix, it will often be referred to as the "matrix formalism." It is an adaptation of the formalism presented in [46] for a layered surface detector system.

The first step in formulation is the division of the flux of secondary particles arriving at ground into two components:

1. the "electromagnetic" component ($e^{\pm}, \gamma$), which is comprised of electrons, positrons, and photons
2. the "muonic" component ($\mu^{\pm}$), which is comprised of muons and anti-muons

Naturally, the flux could be split into additional components driven by the evolution of EAS; however, a two component solution, paired with the two observables that are the SSD and WCD signals (S_{ssd} and S_{wcd}, respectively) allows for an analytical formulation. The total energy deposited by these components in both sub-detectors is notated with $E^{\gamma,e^{\pm}}$ and $E^{\mu^{\pm}}$, which refer to the electromagnetic and muonic shower components, respectively. Expressing each of these as the sum of the energy deposited in the SSD and WCD gives

$$E^{\gamma,e^{\pm}} = E^{\gamma,\mathrm{e}^{\pm}}_{\mathrm{ssd}} + E^{\gamma,\mathrm{e}^{\pm}}_{\mathrm{wcd}} \tag{4.1}$$

$$E^{\mu^{\pm}} = E^{\mu^{\pm}}_{\mathrm{ssd}} + E^{\mu^{\pm}}_{\mathrm{wcd}} \tag{4.2}$$

where $E^{\gamma,\mathrm{e}^{\pm}}_{\mathrm{ssd}}$ represents the portion of the total energy deposited in the SSD coming from particles of the electromagnetic component, $E^{\mu^{\pm}}_{\mathrm{wcd}}$ represents the portion of the total energy deposited in the WCD by particles of the muonic component, and so on. Two additional variables may be defined, namely the fraction of energy deposited in the SSD out of the total energy deposited in both both detectors, namely

$$m = \frac{E^{\gamma,\mathrm{e}^{\pm}}_{\mathrm{ssd}}}{E^{\gamma,\mathrm{e}^{\pm}}_{\mathrm{ssd}} + E^{\gamma,\mathrm{e}^{\pm}}_{\mathrm{wcd}}} \tag{4.3}$$

for the electromagnetic component and

$$n = \frac{E^{\mu^{\pm}}_{\mathrm{ssd}}}{E^{\mu^{\pm}}_{\mathrm{ssd}} + E^{\mu^{\pm}}_{\mathrm{wcd}}} \tag{4.4}$$

for the muonic component. Eqs. (4.1) to (4.4) allow for the expression of the two observables (SSD and WCD signals) and two desired quantities in reconstruction ($E^{\gamma,e^{\pm}}$ and $E^{\mu^{\pm}}$) in a matrix, namely

$$\begin{pmatrix} E_{\mathrm{ssd}} \\ E_{\mathrm{wcd}} \end{pmatrix} = \begin{pmatrix} m & n \\ 1-m & 1-n \end{pmatrix} \begin{pmatrix} E^{\gamma,e^{\pm}} \\ E^{\mu^{\pm}} \end{pmatrix} \tag{4.5}$$

Through matrix inversion, one can obtain

$$E^{\gamma,\mathrm{e}^{\pm}}_{\mathrm{wcd}} = \frac{1-m}{m-n}\left[(1-n)E_{\mathrm{ssd}} - nE_{\mathrm{wcd}}\right] \tag{4.6}$$

and

$$E^{\mu^{\pm}}_{\mathrm{wcd}} = \frac{1-n}{m-n}\left[(m-1)E_{\mathrm{ssd}} + mE_{\mathrm{wcd}}\right] \tag{4.7}$$

This implies that with estimates of the energy deposited in the WCD and SSD and the parameters m and n, one can calculate the energy deposited in either detector by particles of either shower component. Of course, the energy deposit is not what is actually reported by the two sub-detectors. What is measured is a calibrated "signal," which is assumed to be proportional to the energy deposited, although the constant in the proportionality does differ between the two detectors. These proportionalities may be written as

$$S_{\mathrm{wcd}} = C_{\mathrm{wcd}}E_{\mathrm{wcd}} \tag{4.8}$$

and

$$S_{\mathrm{ssd}} = C_{\mathrm{ssd}}E_{\mathrm{ssd}} \tag{4.9}$$

Substituting Eq. (4.8) and Eq. (4.9) into Eq. (4.5), the formalism can be reformulated in terms of the real observables of signal, resulting in

$$\begin{pmatrix} S_{\mathrm{ssd}} \\ S_{\mathrm{wcd}} \end{pmatrix} = \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix} \begin{pmatrix} S^{\gamma,e^{\pm}} \\ S^{\mu^{\pm}} \end{pmatrix} \tag{4.10}$$

after simplification where the parameters a and b are defined as

$$a = \frac{S_{\mathrm{ssd}}^{\gamma,\mathrm{e}^{\pm}}}{S_{\mathrm{ssd}}^{\gamma,\mathrm{e}^{\pm}} + S_{\mathrm{wcd}}^{\gamma,\mathrm{e}^{\pm}}} \tag{4.11}$$

and

$$b = \frac{S_{\mathrm{ssd}}^{\mu^{\pm}}}{S_{\mathrm{ssd}}^{\mu^{\pm}} + S_{\mathrm{wcd}}^{\mu^{\pm}}} \tag{4.12}$$

and where once more through matrix inversion, one can obtain

$$S_{\mathrm{wcd}}^{\gamma,\mathrm{e}^{\pm}} = \frac{1-a}{a-b}\left[(1-b)S_{\mathrm{ssd}} - bS_{\mathrm{wcd}}\right] \tag{4.13}$$

and

$$S_{\mathrm{wcd}}^{\mu^{\pm}} = \frac{1-b}{a-b}\left[(a-1)S_{\mathrm{ssd}} + aS_{\mathrm{wcd}}\right]. \tag{4.14}$$

The electromagnetic and muonic signals $S_{\mathrm{wcd}}^{\gamma,\mathrm{e}^{\pm}}$ and $S_{\mathrm{wcd}}^{\mu^{\pm}}$ can thus each be written as a linear combination of the real observables S_{ssd} and S_{wcd}. Notice that there is no explicit dependency on the scaling parameters C_{ssd} and C_{wcd}. However, through knowledge of these parameters, or rather the mean energies associated with the units of MIP (approximately 2 MeV) and VEM (approximately 240 MeV) respectively, it is also possible to calculate the energy deposited by particles of either component in either detector.

Unfortunately, by reformulating in terms of signals instead of energy, the physical interpretability of the parameters a and b is considerably reduced, and the question remains as to how to obtain them. For the present moment, these must be acquired by simulating the WCD and SSD responses to the particles of simulated EASs reaching ground; however, their estimation from real shower measurements, especially with other complimentary detectors, could take place in the future.

It is of immense importance to note that in order for the formalism to provide unbiased estimates of the electromagnetic and muonic signal in either detector, the matrix parameters a and b must be similar for primaries of different mass. Additionally, in order for the formalism to have minimal model dependency, these parameters must be the same for different hadronic interaction models. These requirements are analyzed in Section 4.3.

4.2 Simulation Libraries

A fixed library of EASs simulated with CORSIKA were used for the derivation of the parameters a and b and the assessment of biases and resolutions resulting from

Table 4.1: Fixed library of CORSIKA simulations used for derivation of matrix parameters and assessments of biases and resolutions. For each combination of parameters, there are 10 unique showers.

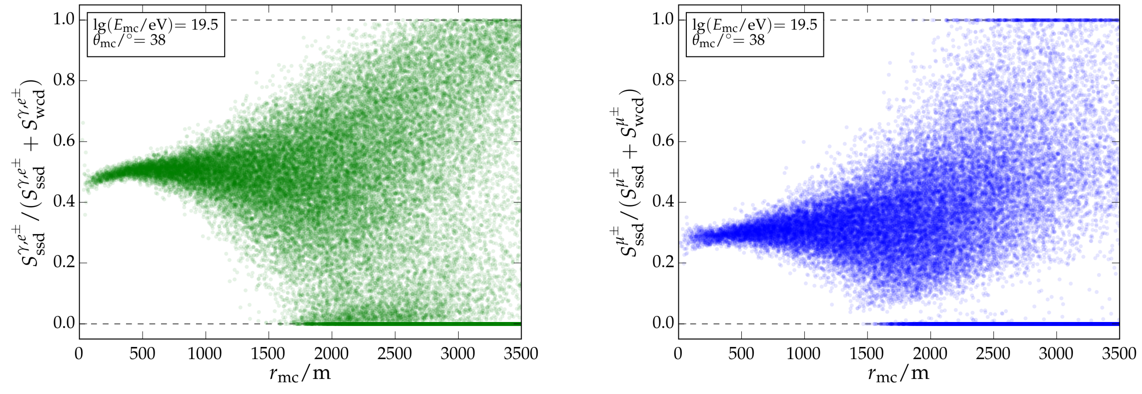
Primaries	proton, iron
Had. Int. Models	QGSJetII-04, EPOS-LHC
lg(E/eV)	18.0, 18.5, 19.0, 19.5, 20.0
$\theta/^\circ$	0, 12, 22, 32, 38, 48, 56, 65
$\phi/^\circ$	0-360, uniformly distributed
Atm. Models	1-12, uniformly distributed

the formalism's use. 10 unique air showers were simulated for each primary, energy, zenith angle, monthly atmosphere, and hadronic-interaction model. This equates to a total of 19200 unique air showers. A tabular summary of this library is provided in Table 4.1. For each of these showers, dense rings (described in Section 2.1.2) of 24 stations were simulated at 800 m and 1000 m from the shower axis in the SDP.

4.3 Matrix Parameters

The parameters a and b from equations Eq. (4.11) and Eq. (4.12) were calculated for each simulated station, and the distributions of these parameters are shown as a functions of distance in Fig. 4.1. At distances close to the shower axis, the distributions are more or less symmetric whereas at distances far from the shower axis, asymmetries with particular features may be observed. The matrix parameter a, which corresponds to the energy deposit in the detectors from particles of the electromagnetic component of EASs, exhibits zero values for some stations starting at the relatively large distance of approximately 1700 m from the shower axis for the energy of $10^{19.5}$ eV and zenith angle of 38° shown. Such values correspond to zero or negligible energy deposit in the scintillator whereas the WCD records a signal. This can largely be attributed to the significantly smaller area of the SSD and its only partial coverage of the WCD when the areas of the two detectors are projected into the SDP. In the case of small numbers of particles, there will inevitably be instances where no particles intersect the SSD while some do intersect the WCD. This effect can also be observed at larger distances in the distributions of the muonic b parameter. Although much more rare due to the positioning of the SSD relative to the WCD, there are also cases where particle trajectories only intersect the SSD and not the WCD, which can be observed in the bands where the respective parameters equal 1, which are observed at even larger distances where particle ground densities are very low.

The asymmetry in both parameters' distributions with a higher density of values above the mode can be attributed to the fact that particles which pass through the SSD are almost guaranteed to also intersect the WCD, whereas the opposite is not true the majority of the time. The population towards the bottom of Fig. 4.1a can perhaps be attributed to photons, which often deposit only a very small signal in the



(a) **(b)**

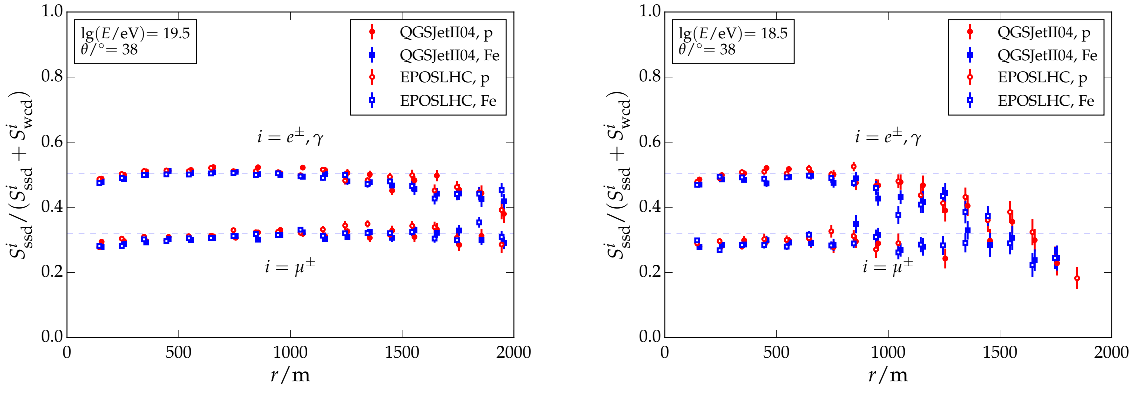
Figure 4.1: Values of the (a) electromagnetic and (b) muonic matrix formalism parameters as calculated for individual SSD and WCD pairs as a function of distance from the shower axis in the SDP. An ensemble of iron primaries with energies of $10^{19.5}$ eV and zenith angles of 38° was used to generate these values. The distribution of both parameters around their mean values begins to lose its symmetry at a given distance. At this point, the signals of the SSD and WCD no longer represent a response averaged over many particles, but instead exhibit fluctuations partly due to fluctuations in particle arrival positions.

SSD (or none at all), due to their lack of charge and need to interact via Compton scattering, pair production, or the photoelectric effect in order to produce charged secondaries. Whereas a significant rate of such interactions is not guaranteed in the SSD, the depth of the WCD all but guarantees deposition of all of a photon's energy within a few dozen centimeters.

Stated simply, the matrix parameters a and b represent the relationship between the respective sub-detectors averaged over many particles closer to the shower axis and are thus symmetrically distributed, whereas at larger distances various asymmetric geometric, positional, and interactive effects become influential. With that said, on the level of individual stations, the use of the average matrix parameters may only be applied in a simple manner up to a certain distance from the shower axis.

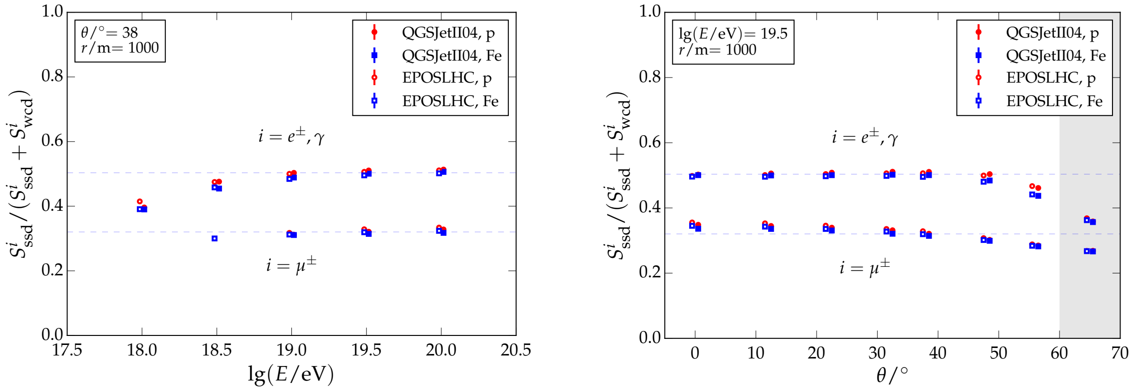
The distance dependency of the average matrix parameters may be observed in Fig. 4.2 for ensembles of showers with two different primary energies. It can be seen that the parameters are relatively constant up to a certain distance from the shower axis where the aforementioned asymmetries begin to be observed. This distance clearly increases with shower size. A slight decrease in the electromagnetic parameter a may be observed very close to the shower axis. This could be explained by an increase in the average energy of electrons or photons resulting in a larger relative energy deposit in the WCD.

The constancy of the matrix parameters may also be observed as functions of primary energy (Fig. 4.3a) and zenith angle (Fig. 4.3b) for the two different primaries of proton and iron and for two hadronic interaction models, QGSJetII-04 and EPOS-LHC. As discussed in Section 4.1, the necessity that the parameters be the same for different primaries is vital in avoiding excessive composition-dependent biases in



(a) (b)

Figure 4.2: Evolution of the matrix parameters as a function of distance from the shower axis in the SDP for primary energies of (a) $10^{19.5}$ eV and (b) $10^{18.5}$ eV. Although the parameters are fairly constant closer to the shower axis at both energies, the distance at which this ceases to be true is less far from the shower axis for lower energy primaries, where ground particle densities are lower.



(a) (b)

Figure 4.3: Evolution of the matrix parameters at 1000 m from the shower axis as functions of (a) energy and (b) zenith angle. The horizontal dashed lines represent the mean parameter values at an energy of $10^{19.5}$ eV and zenith angle of 38° as averaged over all simulated models and primaries. These purpose of these lines is exclusively to highlight deviations of the parameters from constancy. The deviations from constancy at low energies and high zenith angles are due to threshold effects as signals of both detectors become very low and no longer represent average responses of the detectors to many particles.

the results of the matrix formalism. The fact that the parameters are very similar for both interaction models allows one to interpret the results of the formalism as having minimal dependencies on hadronic interaction model, which is one of the formalism's salient benefits.

For the initial assessment of the matrix formalism reconstruction algorithm, constants were chosen for the values of a and b. The simulated average values of these two parameters at an energy of $10^{19.5}$ and zenith angle of 38° for the two simulated primaries and interaction models were used.

4.4 Station-level Evaluation

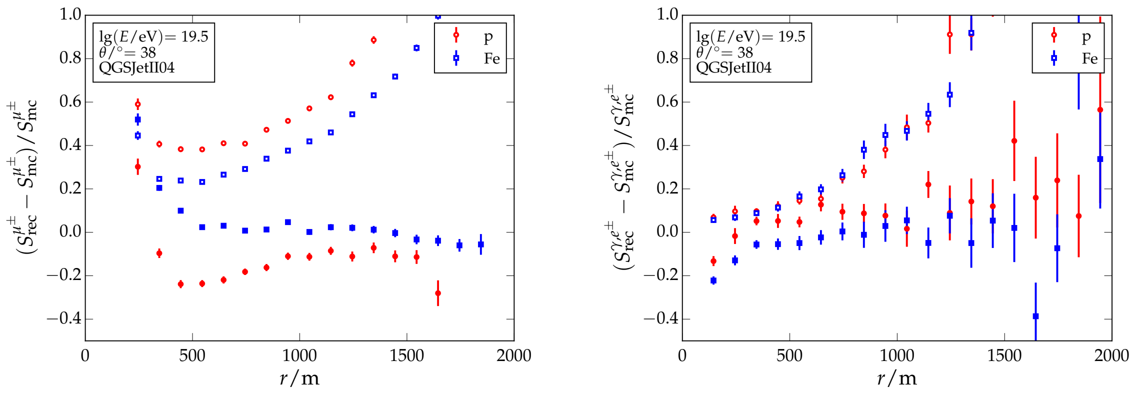
The matrix formalism may be evaluated for paired SSD and WCD measurements, or on the signals predicted by, for example, a fit of the lateral distribution of signals at a given distance from the shower axis. In this section, the performance of the formalism in its application to individual SD stations, comprised of an SSD and WCD, is evaluated.

The bias and resolution of the reconstructed muonic signal in the WCD are shown in Fig. 4.4a as a function of the stations' distances from the shower axis in the SDP. The same is shown for the reconstructed electromagnetic signals in Fig. 4.4b.

For the energy $10^{19.5}$ eV and zenith angle 38° shown, resolution in the reconstructed muonic signal improves when moving further away from the shower axis up to approximately 20% (40%) for iron (proton) at a distance of approximately 500 m. Naively, this almost appears counter intuitive as the relative uncertainty in the measured SSD and WCD signals improves closer to the shower axis; however, with increasing proximity to the shower axis, the electromagnetic component of the shower becomes increasingly dominant. At very close distances, muons and anti-muons comprise only a tiny fraction of the total measured signal in either sub-detector. Beyond approximately 500 m, resolutions in the reconstructed muonic signal become steadily worse once more. This can likely be attributed to the lower resolutions in both sub-detector signals. Biases in the reconstructed muonic signal are relatively constant and close to zero for iron primaries beyond approximately 500 m, whereas an underestimation of the muonic signal occurs for proton which gradually improves from approximately 25% at 500 m to 10% or less beyond 1000 m. The causes for this bias between proton and iron remains to be clarified, although it is likely due to differences in the energy spectra and particle densities of the two shower components between the two different primaries. Perhaps fortunately, the bias works in favor of mass discrimination as it underestimates the muonic content of proton primaries, which already produce less muons than heavier nuclei.

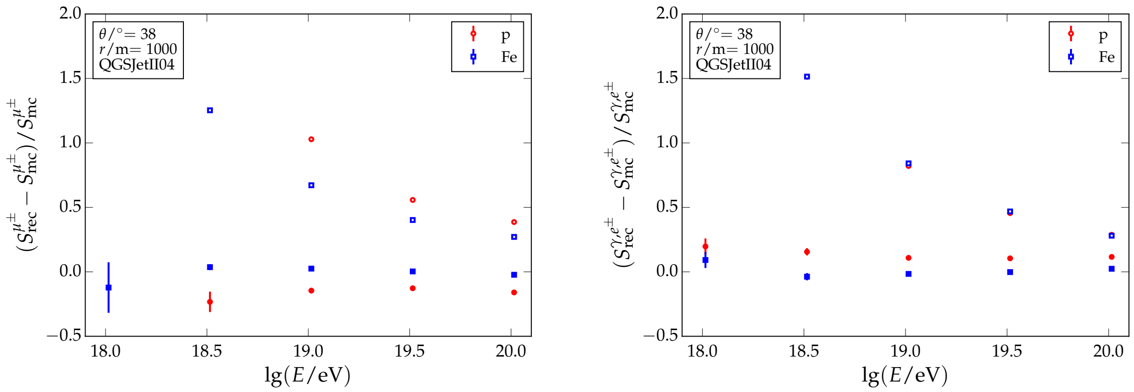
Regarding the electromagnetic component, resolutions are at around 5% close to the shower axis for the given energy and zenith angle and gradually become worse, with a resolution of approximately 40% at 1000 m from the shower axis. Below approximately 300 m from the shower axis, a bias presents itself, which can perhaps be attributed to the deviation from a constant in the a parameter of the matrix formalism, as distinctly visible in Fig. 4.1a.

As shown in Fig. 4.5 for a given reference distance, the resolutions in the reconstructed muonic and electromagnetic signals considerably improve with increasing primary energy. With increasing primary energy, there are more particles at a given distance and thus sampling uncertainties are reduced. Iron achieves a resolution in the reconstructed muonic signal at 1000 m on the order of 30% at the highest energies of 10^{20} eV, whereas proton, with its lower muon content, reaches a resolution of approximately 40%. The analogous resolution for the electromagnetic component is approximately 30% for both iron and proton. Zenith dependent biases and resolutions



(a) **(b)**

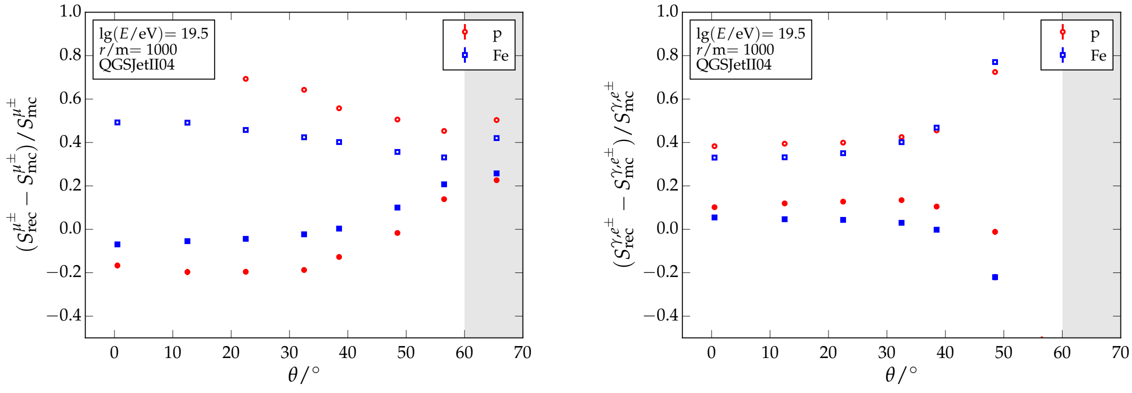
Figure 4.4: Bias (closed markers) and resolution (open markers) of the reconstructed (a) muonic and (b) electromagnetic signals in the WCD as functions of distance. It may be observed that in both cases, the optimal distance range for reduction of relative biases between proton and iron does not correspond to the optimal region of resolution.



(a) **(b)**

Figure 4.5: Bias (closed markers) and resolution (open markers) of the reconstructed (a) muonic and (b) electromagnetic signals in the WCD at 1000 m from the shower axis as functions of energy. For this distance, muonic resolutions improve to $\approx 20\%$ at the highest energies, where the difference in bias between iron and proton primaries is around 10-15%. The difference in biases between primaries is noticeably smaller in reconstructed electromagnetic signals.

may be observed in Fig. 4.6. More extensive attenuation of showers' electromagnetic component at higher zenith angles results in improved resolution in reconstruction of the muonic signal and worse resolution in that of the electromagnetic component.



(a) **(b)**

Figure 4.6: Bias and resolution of the reconstructed (a) muonic and (b) electromagnetic signals in the WCD at 1000 m from the shower axis as functions of zenith angle. The resolution of the muonic signal improves with zenith angle whereas the reconstructed resolution of the more attenuated electromagnetic component decreases alongside an increasing bias.

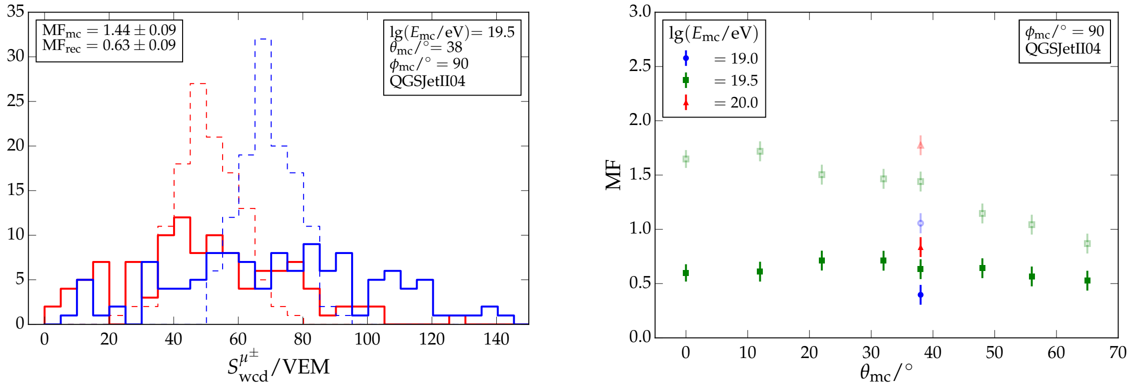
The Monte-carlo and reconstructed distributions of $S^{\mu^\pm}_{\rm wcd}$ are shown for a given energy and zenith angle in Fig. 4.7a. The ability to discriminate between proton and iron primaries can be quantified using a so-called “merit factor,” which is defined as

$$\mathrm{MF} = \frac{|\langle S_\mathrm{p}\rangle - \langle S_\mathrm{Fe}\rangle|}{\sqrt{\sigma^2_{S_\mathrm{p}} + \sigma^2_{S_\mathrm{Fe}}}} \tag{4.15}$$

where S_p and S_Fe refer to the muonic signals produced by proton and iron, respectively. In comparing the Monte-carlo distributions in the muonic signal to those reconstructed with the matrix formalism, the effects of the resolution of the matrix formalism can be observed. Whereas station-level discrimination between primaries is possible in principle, it is hampered by the poor station-level resolution in practice. The evolution of the power to discriminate between primaries, both for Monte-carlo and reconstructed values, can be observed in Fig. 4.7b as both a function of energy and zenith angle.

4.5 Event Reconstruction

In order to improve upon the single-station resolution in the reconstructed muonic signal as obtained with the matrix formalism, an event reconstruction was formulated. In the first iteration, the standard SD reconstruction described in Section 1.4 is used to obtain estimates for the shower geometry (axis and core position) and primary energy. The composition reconstruction is applied post factum. In future methods, the information from the SSD will be incorporated in the energy and geometry reconstruction procedure to reduce composition and asymmetry related biases and improve resolutions.



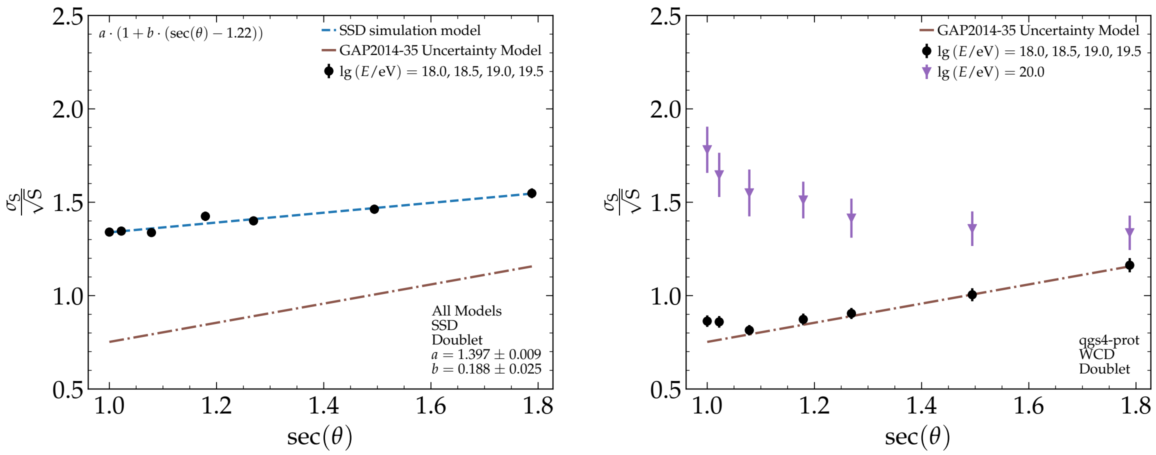
(a) **(b)**

Figure 4.7: (a) Merit factor for the discrimination of proton and iron primaries based on individual station resolution at a distance of 1000 m, energy of $10^{19.5}$ eV, and zenith angle of 38°. The distributions plotted with solid lines represent the reconstructed muonic signals in the WCD whereas the dashed lines represent the true Monte-carlo muonic signals for the same stations. Merit factors are calculated for both cases. (b) Merit factor as a function of zenith angle. The semi transparent markers represent merit factors calculated based on Monte-carlo muonic signals in WCDs, whereas the solid markers represent the analogous reconstructed quantities. No strong zenith dependence is observed in the discrimination power of reconstructed muonic signals of the WCD; however, an improvement in discrimination power with increasing energy is evident.

With that said, in this initial assessment, the lateral distributions of the SSDs are fit using the reconstructed SD geometry. Analogous to that of the WCD, a maximum likelihood method is used in which large signals are treated with a Gaussian probability density function (p.d.f.) and small signals, converted to effective particle numbers, with a Poisson p.d.f.. See Section 1.4.3 for a basic description of this procedure. Initial parameterizations for the SSD LDF shape parameters and the signal uncertainty model for SSD signals were derived in the joint context of this work and that of [47] using the SSD simulation application detailed in Chapter 3 and the EAS simulations library described in Section 4.2.

4.5.1 Signal Uncertainty Model

The signal uncertainty model for the SSD was derived using the same methods applied to doublet measurements of the WCD as described in Section 1.4.3; however, doublets were simulated as sufficient data from the EA multiplets has yet to be acquired. The simplest method for simulating doublets would be to place two sets of GEANT4 volumes for the SSD 10 m apart in the ground plane of simulations. However, due to the use of use of a thinning procedure in the library of EAS simulations used, such an arrangement would involve re-sampling the same set of weighted particles from CORSIKA ground particle output files for injection in the two different stations. This could result in a reduction in the fluctuations expected between the independent samples of particles incident on two doublet stations in reality. Initial investigatory analysis revealed indications for such a reduction, and



(a)

(b)

Figure 4.8: *secθ* dependence of the signal uncertainty of the (a) SSD and (b) WCD as derived from simulated doublet stations (black points) and compared with data-derived uncertainty model of the WCD (dashed line). The signal uncertainty of the SSD is up to twice as large as that of the WCD, which can be explained by the fact that less particles are required to produce the same signal magnitude. At the highest energies, the signal uncertainties of the simulated WCDs were shown to significantly deviate from the data-derived model, which only made use of low energy data. Figures are from [47].

therefore another method of doublet simulation was developed involving the dense ring arrangements discussed in Section 2.1.2.

The doublets located at azimuthal angles of 90° and 270° in the SDP were deemed *pseudo doublets* as their geometries relative to the shower axis are perfect reflections and the atmospheric depth traversed by the shower at these points is identical. Effects of east-west deflections of positive and negative charges due to the magnetic field of the Earth were investigated and found to be insignificant for the zenith angles involved in this study. Naturally, such *pseudo doublets* would exhibit fluctuations between their signals due to differences in shower development higher in the atmosphere, which are mostly driven by the initial interactions of the particle cascade. Such fluctuations are not present in real doublet measurements, but should, in fact, be included in the signal uncertainty model as the stations of the SD array also sample the lateral distribution of real showers at points separated in the SDP by large distances of many hundreds of meters.

For each set of *pseudo doublet* measurements, an estimator for the signal uncertainty was obtained via

$$\sigma_S = \frac{\sqrt{\pi}}{2}\,|\Delta S| \tag{4.16}$$

and the different estimates for each unique primary, zenith angle, and energy were then combined in a global fit of the function

$$\sigma_S(\theta, S) = a\,(1 + b\,(\sec\theta - \sec(35^\circ)))\,\sqrt{S} \tag{4.17}$$

Results of this fit for both the WCD and SSD may be observed in Fig. 4.8.

The parameterization derived for simulated WCD *pseudo doublets* is quite similar to the data-derived model of the WCD, which lends confidence to the detector and EAS simulations and provides a validation of the method described here for obtaining uncertainties. The model derived for the SSD shows uncertainties approximately 70% higher than those of the WCD. For both detectors, deviations in the behavior of the uncertainties were observed at very high energies and low zenith angles, where real doublet measurements were never performed for the WCDs. This suggests some physical causes, which could be related to shower age at ground, changing secondary energy spectra, or differing relative contributions of shower components. The cause was not investigated further here and the highest energy primaries were excluded from the model derivation.

In order to obtain the Poisson p.d.f.s for small signals, the concept of the Poisson Factor discussed in Section 1.4.3 was also employed for the SSD. Two options exist in writing the conversion from signal to the effective number of particles n_eff that produce the observed signal uncertainties under a Poisson assumption, namely

$$n_\text{eff} = pS\,, \tag{4.18}$$

which is the same formulation used for the WCD, or a formulation where the expected tracklength of particles is explicitly taken into consideration, namely

$$n_\text{eff} = p(S/\left\langle\text{tracklength}\right\rangle) = p(S/\sec\theta)\,. \tag{4.19}$$

Use of either formulation ends up equivalent as the tracklength dependency is accounted for implicitly in the Poisson factor as written in Eq. (4.18) whereas it is explicitly accounted for in Eq. (4.19). Nonetheless, explicitly accounting for the tracklength allows for interpretation of the Poisson factor as being the amount by which the uncertainty of the SSD deviates from the assumption that incident particles deposit, on average, a signal of $\sec\theta/\text{MIP}$.

The Poisson factor derived under both formulations is shown in Fig. 4.9 along with the pre-existing parameterization for the WCD alongside an additional parameterization for the WCD explicitly taking tracklength into consideration. It is immediately striking that while the Poisson factor the WCD is significantly greater than 1 over the majority of the range of zenith angle, that of the SSD is significantly less than one. For the WCD, this can be understood in the fact that many electrons (on the order of 20 or more) are required to produce a signal of 1 VEM as they deposit all of their energy within just a few centimeters of water. Electrons in the scintillator, however, will deposit a signal in MIP of around 1 over their tracklength if they have sufficient energy to pass through the SSD. As such, a Poisson factor of significantly less than 1 implies that each incident particle deposits a signal significantly greater than the naive expectation. The reason for this can be attributed, in part, to the energy deposit distribution of charged particles resembling the Landau distribution as shown in Fig. 3.7a. SSD signals are calibrated to the charge corresponding to the peak of this distribution, which due to its asymmetric long tail toward higher energies is considerably less than its mean.

4.5.2 Lateral Distribution Function Parameterization

With a model for signal uncertainties across the complete dynamic range of the SSDs in hand, it was possible to parameterize a model of the lateral distribution of signals

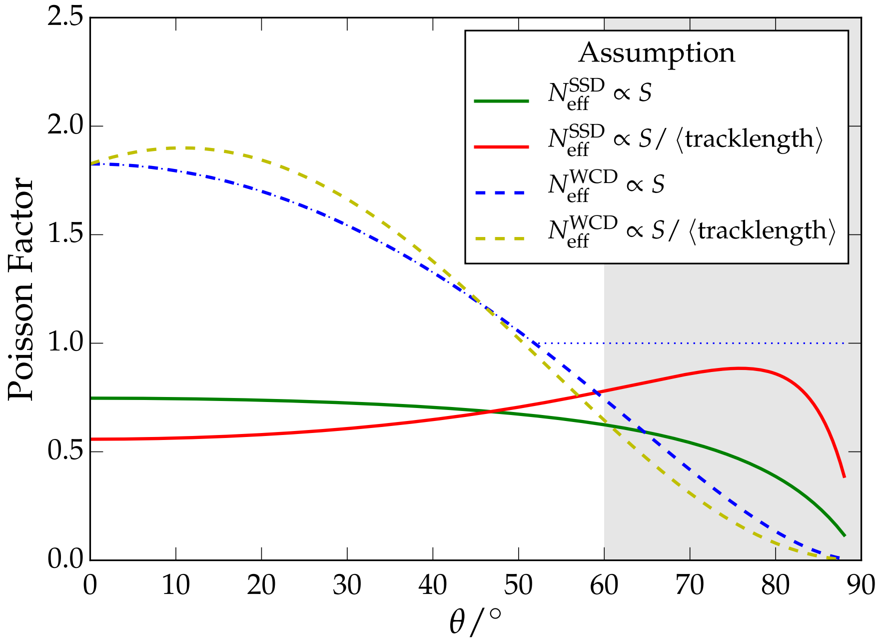


Figure 4.9: Poisson factor used to convert signals to "effective" particle numbers which reproduce the results of the signal uncertainty model under a Poisson assumption.

for the SSD. A modified NKG-like function identical in functional form to that used for the WCD given by Eq. (1.5) was used in this first effort. Given that the SSD array geometry is identical to that of the WCD, the values of r_{opt} of 1000 m and r_{s} of 700 m were also adopted from the WCD LDF.

The shape parameters, β and γ, of the SSD LDF were parameterized making use of the library of showers simulated with CORSIKA detailed in Table 4.1. The Monte-carlo geometry was used throughout this process. In order to obtain initial guesses for fitting the LDF shape parameters of individual events, global fits to the lateral distributions of all showers with a given energy and zenith angle were first performed. This involved normalizing the signals of each SSD in a given event by the signal at the optimum distance of 1000 m, which was obtained by taking a mean of the 24 dense station SSDs at this distance. The normalized signals of each of the events with the given energy and zenith angle were then fit together. The first step involved fixing γ to the value provided by the parameterizations for the WCD LDF and leaving β as a free parameter, with the guess given to the minimizer also provided by the WCD LDF parameterizations. In the next step, β was fixed to the result of the previous fit, and γ was left free. This process of iteratively fitting β and γ was continued until both values converged. The resulting values were then used as initial guesses when fitting individual events, which will be described shortly. In performing these global fits, a unique cut on distance was performed for each energy and zenith angle bin to exclude signals for distances greater than that at which 5% of signals were below 1 MIP.

The next step of the parameterization involved fitting individual events. The same simulated showers were thrown on the array multiple times in order to gain enough stations to fit the LDF shape parameters. For every event in the library, the

Table 4.2: Results of the parameterizations of the SSD LDF shape parameters.

i	a_i	b_i	c_i	d_i	e_i	f_i
β	4.2 ± 0.2	-0.6 ± 0.1	-10.0 ± 0.2	0.5 ± 0.2	4.07 ± 0.07	0.18 ± 0.06
γ	-19.6 ± 0.4	3.2 ± 0.3	25.0 ± 0.6	-4.2 ± 0.4	-9.2 ± 0.2	1.5 ± 0.2

normalization, S_{1000} was fixed to the average of the 24 dense station SSDs of the first throw at 1000 m. For each event, γ was fixed to the initial guess provided by the global fits for the energy and zenith angle in question. β was then fit as the only free parameter using the result of the global fit for the given event's energy and zenith angle as the initial guess. Once a value for β was obtained for each shower in the library, a parameterization of β in shower size (linear) and zenith angle (quadratic) was performed with the functional form

$$(a_\beta + b_\beta\, s) + (c_\beta + d_\beta\, s)\sec\theta + (e_\beta + f_\beta\, s)\sec^2\theta \tag{4.20}$$

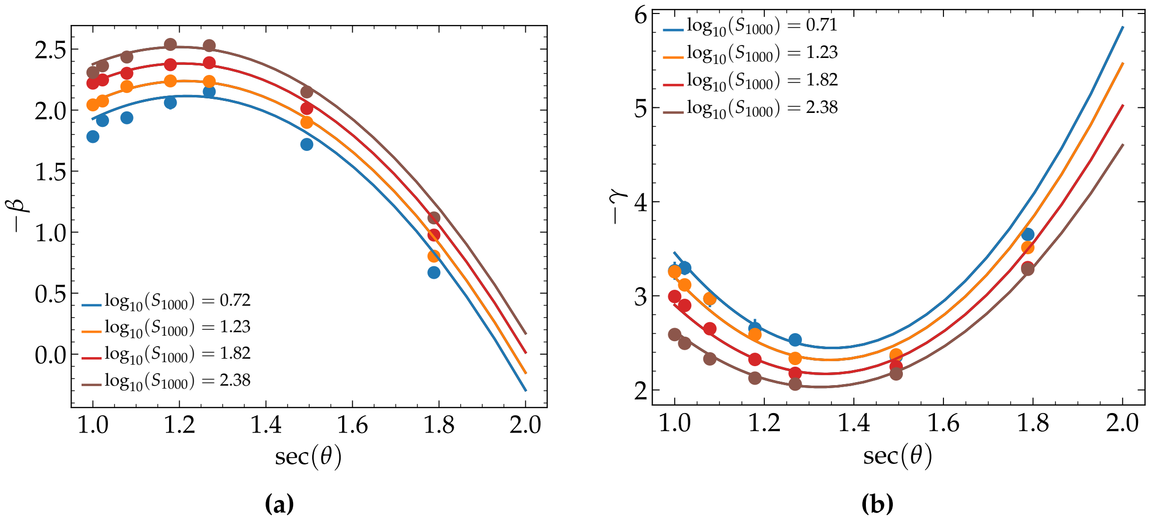
where $s = \lg S_{1000}$. In the next step, β was fixed to the result of its parameterization, and γ was fit using the results of the global fits as the initial guess. Following this step, a parameterization of γ was performed in shower size and zenith angle with a functional form analogous to that of β, namely

$$(a_\gamma + b_\gamma\, s) + (c_\gamma + d_\gamma\, s)\sec\theta + (e_\gamma + f_\gamma\, s)\sec^2\theta\ . \tag{4.21}$$

Thereafter, β and γ were alternately fit on a single event basis, using their last parameterizations as guesses or values to be fixed. Between each sequence of individual event fits, the appropriate parameterization of the recently fit shape parameter was re-performed. This was important in avoiding iteration-to-iteration jumps in β and γ values due to the strong correlation between the two. The results of these β and γ parameterizations are given in Table 4.2. The $\sec\theta$ dependencies of the two parameters are also plotted alongside the parameterizations in Fig. 4.10.

4.6 Event-level Evaluation

In order to evaluate the performance of the matrix formalism on the event level, LDFs were fit to both the WCD and SSD signals. The likelihoods for these fits were constructed as described in Section 1.4.3 making effective use of the Poisson Factor described in Section 1.4.3 and derived for the SSD in Section 4.5.1 to obtain the p.d.f.s for low signal stations. Stations with predicted signals corresponding to a c.d.f. greater than 0.05 at a signal of 1 MIP were excluded from the likelihood as trigger and use criteria for very low SSD signals has yet to be properly investigated. Only the LDF scaling parameter S_{1000} was fit for both sub-detectors. Values for β and γ were obtained via their parameterizations at each fitting iteration. Once the fit converged for both the SSD and WCD, the matrix formalism was applied at 1000 m from the shower axis to obtain an estimate of the muonic signal in the WCD. A depiction of this fitting procedure for a sample event is given in Fig. 4.11. The distance of



(a) **(b)**

Figure 4.10: (a) β and (b) γ parameterizations of the SSD lateral distribution function as a function of secθ at 1000 m for different bins of shower size. Figures are from [47].

1000 m was chosen arbitrarily out of convenience and strong indications exist that distances closer to the shower axis exhibit considerably enhanced resolutions in the reconstructed muonic signal in the WCD via the matrix formalism. This optimal distance will depend on the biases and resolutions of the respective sub-detectors' LDF predictions, as well as the difference in the expected number of muons that arrive at ground at a given distance from the shower axis between primaries of different mass. A statistic which could be optimized to determine the preferred distance is the merit factor defined in Eq. (4.15). Subtleties in the reconstruction process, such as decisions of whether to leave the slope parameters of the LDFs free, may have a sizable impact on this factor.

In addition to the resolution of the reconstructed muonic signal, there also exists a resolution on the bias of the reconstructed value. In order to estimate the magnitude of this bias resolution, 120 proton and iron showers with an energies of $10^{19.5}$ eV and zenith angle of 38° were each thrown randomly on a simulated detector array 30 times. An estimate for the reconstructed muon signal in the WCD was obtained for each of the 30 throws via the procedure described above, and the bias was calculated. The distribution of the average bias for 120 proton and 120 iron showers is shown in Fig. 4.12a. The widths of these distributions correspond to the "bias resolution," of the reconstructed muonic signal, which should be summed in quadrature with the uncertainty of the shower size estimate in order to obtain the complete measurement resolution. For the given energy and zenith angle, this bias resolution is relatively small at 5% and 6% for iron and proton, respectively. The average bias in the reconstructed muonic signal of 20% for proton and 15% for iron is more concerning; however, such a bias can be corrected for on average, leaving only a residual composition-dependent bias of 5% between proton and iron primaries.

The merit factor for discrimination between proton and iron primaries, again calculated using Eq. (4.15), is shown for the energy of $10^{19.5}$ eV and a zenith angle of 38° for both Monte-carlo and reconstructed values of $S^{\mu^{\pm}}_{\text{wcd}}$ in Fig. 4.12b. A merit factor of

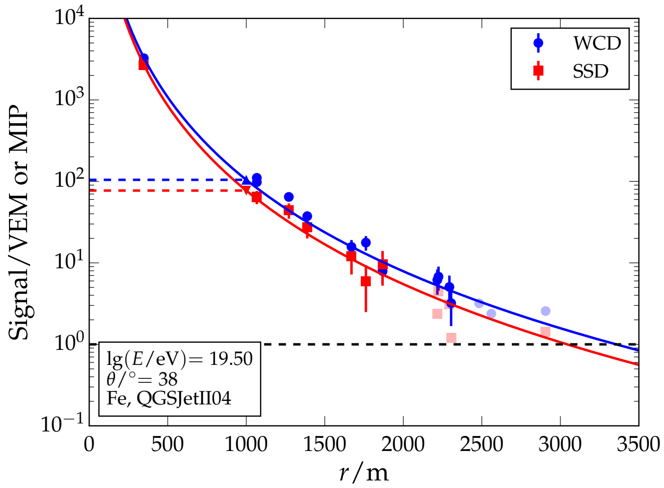
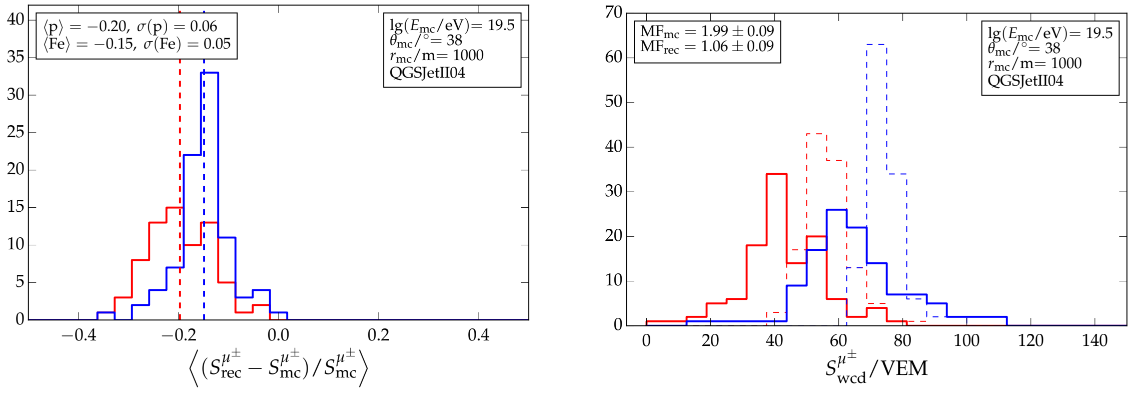


Figure 4.11: Fits to SSD and WCD lateral distributions for a sample event. The dashed lines and triangles depict the signal predicted by the LDF fits at a distance of 1000 m from the shower axis. The semi-transparent markers at lower signals were excluded from the likelihoods used in the fits as they did not satisfy statistical criterion intended to avoid biases due to threshold-level signals, which have yet to be studied in detail for the SSD.



(a) **(b)**

Figure 4.12: (a) Biases in reconstructed muonic signal in the WCD for ensembles of proton and iron primaries with energies of $10^{19.5}$ eV and zenith angle of 38°. The standard deviation of these distributions is taken to be the bias resolution, which must be added in quadrature with the resolution of the reconstructed signal to obtain the total uncertainty. (b) Distributions of reconstructed (solid lines) and Monte-carlo (dashed lines) muonic signals in the WCD for ensemble of proton and iron primaries of the same Monte-carlo energy and zenith angle.

nearly 2 in the Monte-carlo values is very good and very little overlap between the distributions of proton and iron is observed. The affects of reconstruction resolutions on the ability to discriminate is observed in the merit factor of the reconstructed values, which is closer to 1, which is still sufficient to perform discriminatory analysis.

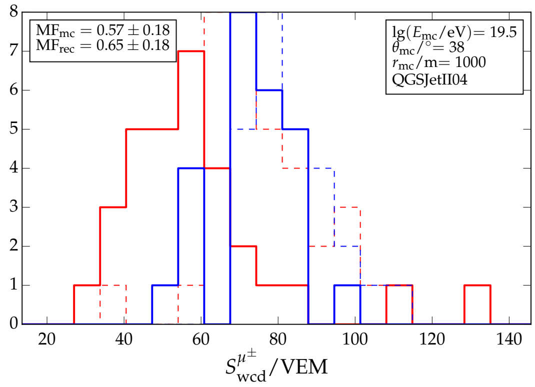


Figure 4.13: Distributions of reconstructed (solid lines) and Monte-carlo (dashed lines) muonic signals in the WCD for proton and iron primaries. Selection of the respective data samples was performed on the reconstructed energy and zenith angle such that the biases in these reconstructed variables are taken into account when examining the total discrimination power of the reconstruction algorithm.

Nonetheless, these merit factors are calculated for a fixed energy, which does not consider the selection effects of resolution and composition-dependent biases in the energy of the primaries reconstructed by the standard SD reconstruction. Selecting on reconstructed energy for a continuous library of CORSIKA showers results in the merit factors depicted in Fig. 4.13. The effect of the resolution and composition-dependent bias in the reconstructed energy is clear, although the power to discriminate is still visible.

4.6.1 Composition Sensitive Parameter

Whereas the muonic signal in the WCD correlates with the muon content of an air shower, which in turn correlates to its composition, this parameter is not directly comparable between showers of different energies and zenith angles. In order to obtain an energy and zenith composition sensitive parameter, a parameterization of the muonic signal at a distance of 1000 m from the shower axis was performed. Named R^{μ}_{38}, this estimator was defined by

$$R^{\mu}_{38} = S^{\mu^\pm}_{1000}/f(\theta, E) - 1 \tag{4.22}$$

where

$$f(\theta, E) = g(E)h(\theta) \tag{4.23}$$

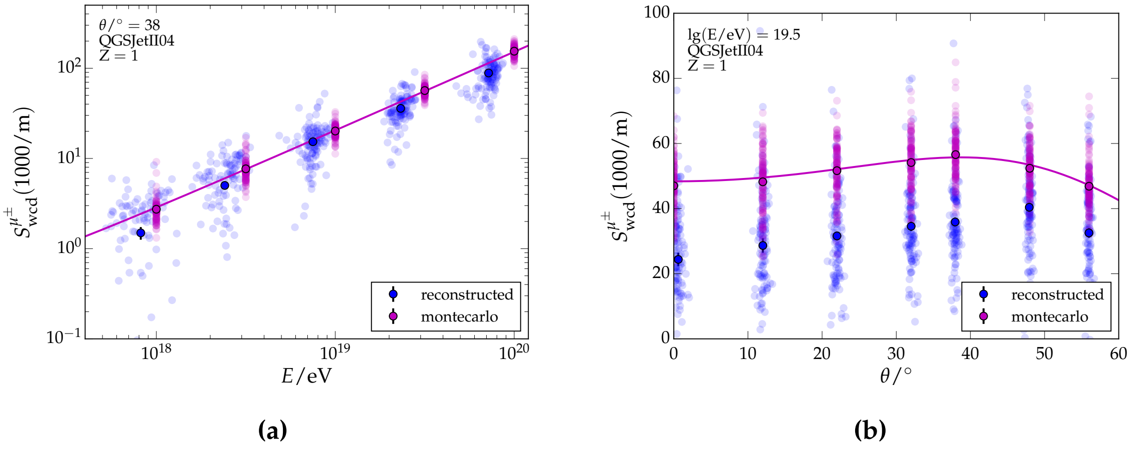
in which

$$g(E) = A + (E/\mathrm{EeV})^{B} \tag{4.24}$$

and

$$h(\theta) = C + Dx + Ex^2 + Fx^3 \tag{4.25}$$

where $x = \cos^2(\theta) - \cos^2(38/°)$. A global fit was performed over all showers in the library described in Table 4.1 using Monte-carlo values to obtain the parameters A through F, after which R^{μ}_{38} could be calculated for reconstructed and Monte-carlo muonic signals for each event using the parameterization. The energy and zenith dependencies in the parameterization may be observed in Fig. 4.14.



(a) **(b)**

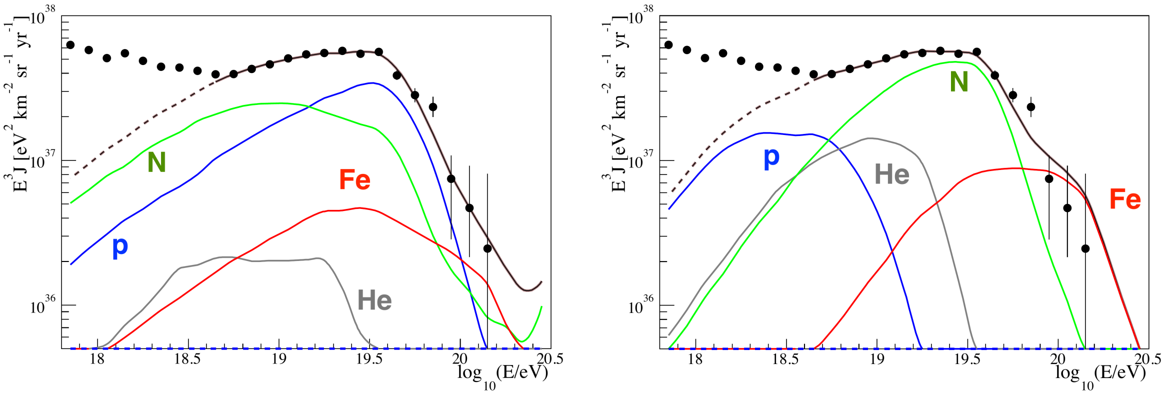
Figure 4.14: Monte-carlo and reconstructed muonic signals in the WCD, $S^{\mu^\pm}_{\rm wcd}$, as a function of (a) energy and (b) zenith. Also plotted by magenta lines is the parameterization to Monte-carlo data to obtain the zenith and energy independent estimator muon content estimator R^{μ}_{38}.

4.7 Astrophysical Scenarios

Given that identification of the sources of UHECRs remains an objective for Auger, it is prudent to analyze the potential of AugerPrime in enhancing this pursuit. In such an effort it is necessary to consider the impact of propagation effects (e.g. photo-disintegration and photo-pion production) and astrophysical sources reaching their maximum acceleration potential on the flux of different primaries measured at Earth. Depending on the magnitude of the relative contributions of these effects, the elemental composition of the flux of UHECRs could be very different at and beyond the energy spectrum's suppression region. Given that higher charge primaries suffer from larger deflections in galactic and extra-galactic magnetic fields, selection of light primaries is warranted for further attempts at performing charged-particle astronomy. Such selection can either be complicated or aided depending on the elemental composition of the total flux.

In this section, the the impact of the two extremes of a photo-disintegration dominated flux scenario and a maximum rigidity dominated flux scenario are examined on attempts at selecting a proton-enriched sample of primaries above 40 EeV. The differential fluxes used as inputs for this analysis are shown in Fig. 4.15 and are taken from [48].

The photo-disintegration scenario assumes that sources accelerate nuclei to an energy greater than the threshold for photo-disintegration via interactions with CMB photons. As such, lighter elements arriving at Earth could be products of heavier nuclei produced during propagation. In this scenario, the suppression at the end of the all particle flux would be the natural result of energy loss processes such as photo-disintegration and the lightest elements would arrive with energies shifted by the the ratio of the daughter to parent mass numbers. In this way, the highest energy particles would be related to those with energies around the spectrum's ankle.



(a) **(b)**

Figure 4.15: (a) Scenario in which distributions of lighter elements resulting from photo-disintegration of heavier elements (from [48]). (b) Scenario of maximum rigidity, where UHECR sources progressively reach their maximum acceleration potential for increasing elemental mass (from [48]).

The maximum-rigidity scenario makes the assumption that sources accelerate nuclei up to a maximum energy proportional to their charge. Thus protons with an energy of approximately $10^{18.5}$ would be related to iron nuclei at an energy 26 times higher, which is roughly coincident with the flux suppression. The suppression in this scenario would be explained by extra-galactic sources reaching their maximum acceleration potential and would not heavily rely on propagation effects.

In order to obtain the elemental spectra of Fig. 4.15, numerical fits were performed with the all-particle flux and distributions of the depth of shower maximum as inputs in different energy intervals. A homogeneous distribution of sources injecting identical power-law spectra with mass composition independent of energy was also assumed. Neither of the two idealized scenarios provided a satisfactory description of both the flux and composition results of Auger; however, a hard source spectra of E^{-1} and a low mass cutoff at 10^{18} eV provided the best results. An improved description was obtained when adding an additional light component in a limited energy range at high energies.

With the two astrophysical scenarios in hand, a method for simulating the different elemental fluxes that comprise the total cosmic ray energy spectrum was implemented. The latest reported flux of Auger [49] was integrated for the region above 40 eV and scaled to obtain an expected event count for 10 additional years of data acquisition, which equates to approximately 330 additional events in this energy range. A Poisson randomization was then applied to this expected number of events, and each event was randomly assigned an energy with a probability according to the same energy spectrum. The species of the primary particle of each event was then assigned to be proton, helium, nitrogen, or iron with a probability given by energy-dependent primary fractions taken from the simulation of different astrophysical scenarios described in more detail shortly. A random zenith angle was chosen under the assumption of a uniform flux in solid angle. With an energy, species, and shower geometry in hand, a set of continuous CORSIKA libraries including proton, helium, nitrogen, and iron primaries, was scanned for the closest matching event. A full SD response to this event was simulated, and the energy and muon content estimator, R^{μ}_{38}, of the event were reconstructed using the matrix formalism described in this chapter. These reconstructed values and the corresponding Monte-carlo values were then selected for analysis.

This entire process was iterated 100 times for each astrophysical scenario to also allow for the analysis of fluctuations in event rate and elemental abundance distributions expected in a 10 year period of additional data acquisition to, in short, answer the question of how much AugerPrime will rely on nature in its pursuit of charged particle astronomy and how much it must rely on its own technical prowess in light of different astrophysical scenarios.

Some parameters of interest in evaluating the ability of the SD to select proton-enriched event samples are purity and efficiency. For the subsequent analyses, a cut is made on R^{μ}_{38}, where primaries below the cut value are included in a selected sample. Purity is defined as the fraction of the selected sample that is comprised of protons, namely

$$\text{purity} = \frac{n_{\text{p}}}{n_{\text{tot}}} \tag{4.26}$$

where n_{p} is the number of protons in the selected sample and n_{tot} is the total number of selected primaries. Efficiency is defined as the fraction of all protons in the flux that are included in the sample, namely

$$\text{efficiency} = \frac{n_{\text{p}}}{N_{\text{p}}} \tag{4.27}$$

where N_{p} is the total number of protons in the flux, both selected and unselected.

4.7.1 Photo-disintegration Scenario

The distributions of R^{μ}_{38} from one of the 100 10-year simulated fluxes are shown in Fig. 4.16a for the Monte-carlo values and in Fig. 4.16b for the reconstructed values. In Fig. 4.16c, the Monte-carlo and reconstructed distributions are shown together, where the populations have been binned into signal (proton) and background (all other nuclei). The effects of reconstruction resolution are evident. As may be observed in Fig. 4.16d, an efficiency of 50% can be achieved in the Monte-carlo R^{μ}_{38} before dropping below 100% purity, and excellent compromises in purity ($\approx 90\%$) and efficiency ($\approx 90\%$) can be struck. For the reconstructed values, a sample with 100% purity is not observed for efficiencies above a few percent. Nonetheless, a sample with a compromise of approximately 80% purity and 80% efficiency can be obtained. Some of this is due to the mass discrimination power of the detector and reconstruction algorithm; however, the fact that a large fraction of the flux above 40 EeV is comprised of protons plays a significant role.

In Fig. 4.16e and Fig. 4.16f, the purity and efficiency in selected samples for two cuts on R^{μ}_{38} are shown for the 100 possible flux scenarios in 10 years of additional data. The spread in purity and efficiency for the Monte-carlo values is on the order of a few percent and are only slightly larger for the reconstructed values. This implies that in a photo-disintegration scenario, luck with nature would not play a significant role in selecting a proton-enriched event sample.

4.7.2 Maximum Rigidity Scenario

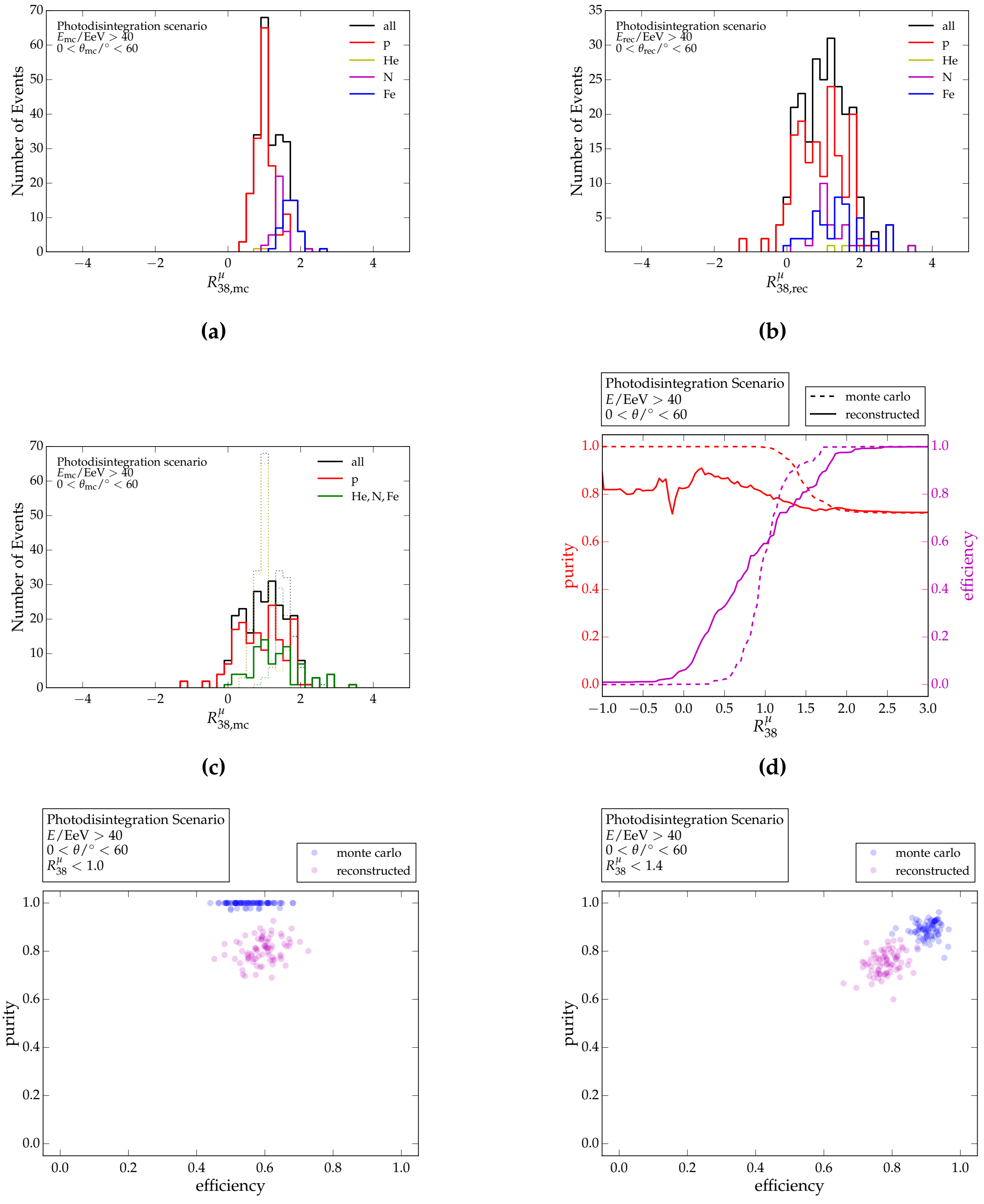
Whereas the photo-disintegration scenario exhibits a significant proton fraction at the highest energies, the maximum rigidity scenario, by nature of its premise, results in hardly any light primaries at the spectrum's end. Nonetheless, adding a small fraction of light composition improves the quality of Auger's combined fits to its spectrum and depth of shower maximum. Additionally, above the suppression energy, the event count of the FD is insufficient to obtain an estimate of the flux composition. This leaves the possibility open for the existence of a nearby proton source with significant contributions at the very highest energies. The ability to obtain a proton enriched event sample is examined both phenomenologically for a small proton fraction and the hypothetical case of a nearby high energy proton source.

10% Proton Fraction

For exploratory purposes, the flux for the maximum rigidity scenario depicted in Fig. 4.15b was supplemented with a 10% proton fraction constant in energy. 100 10-year flux samples above an energy of 40 EeV were then simulated with this addition proton component. The distributions of R^{μ}_{38} for one 10-year flux are shown in Fig. 4.17a for the Monte-carlo values and in Fig. 4.17b for the reconstructed values. In Fig. 4.17c, the Monte-carlo and reconstructed distributions are plotted together, where the populations have been binned into signal (proton) and background (all other nuclei) once more. The effects of a considerably smaller proton fraction in the total flux above 40 EeV are evident. Viewing the same cuts in R^{μ}_{38} as for the photo-disintegration scenario (see Fig. 4.17e and Fig. 4.17f), it is clear that different 10-year periods can result in very different purities and efficiencies both for the Monte-carlo and reconstructed values. It is highly improbable to obtain a purity above 25% when selecting on the reconstructed values for R^{μ}_{38} and efficiencies span almost the entire range of possibility. The lower proton fraction results is a lower signal to noise ratio and larger fluctuations for different 10-year flux scenarios.

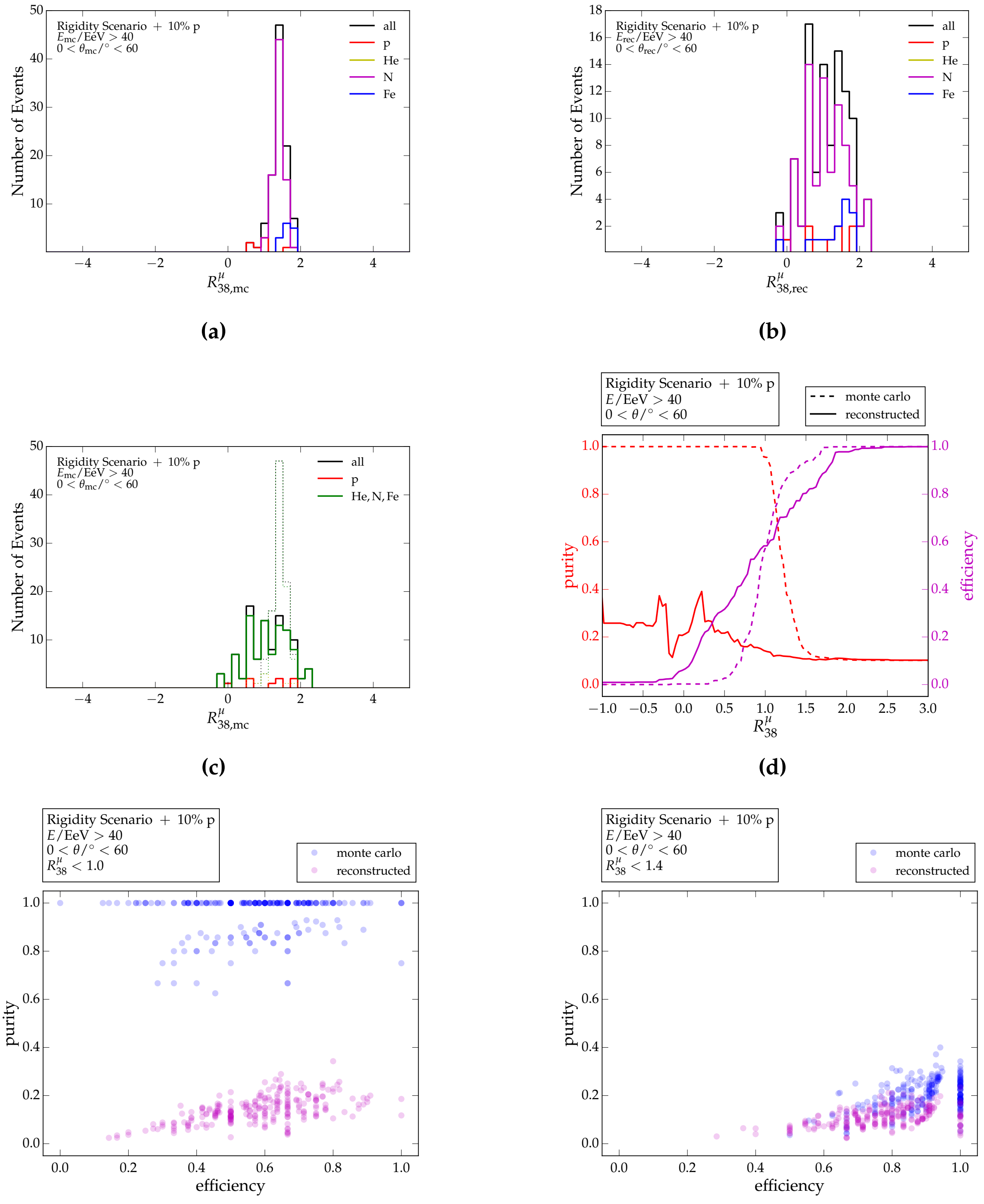
Nearby Proton Source

To explore a maximum rigidity scenario in which contributions from a nearby proton source significantly contributing to the total flux appears at the highest energies, an additional protonic flux contribution was added to the the maximum rigidity scenario, using the spectral shape of iron, with an intensity equal to that of iron. In this exploratory effort, abundances of other elements were not, however, adjusted to account for the reduction of the iron flux. The analogous set of plots to those in Fig. 4.16 and Fig. 4.17 are given in Fig. 4.18. A smaller spread in the purity and efficiency than the case of a constant 10% proton fraction is visible, likely due to the larger proton fraction. The situation is otherwise very similar.



(a) (b) (c) (d) (e) (f)

Figure 4.16: Results of 100 simulated 10-year photo-disintegration scenario fluxes. (a) Monte-carlo energy and zenith independent muon content estimator R^{μ}_{38} for sample 10-year flux. (b) Reconstructed R^{μ}_{38} for the same flux. (c) Monte-carlo (dashed lines) and reconstructed (solid lines) distributions of R^{μ}_{38} for the signal (proton) and background (all other nuclei). (d) Purity and efficiency as a function of $R^{\mu}_{38} < R^{\mu}_{38,\rm cut}$ for Monte-carlo (dashed lines) and reconstructed (solid lines) quantities. (e) Distribution of possible efficiency and purity values for a cut on R^{μ}_{38} of 1.0. Values calculated using Monte-carlo values are shown in magenta while those calculated with reconstructed values are shown in blue. Over the next 10 years, nature would only provide us with one sample from these distributions. (e) The same for a cut of R^{μ}_{38} of 1.4.



(a)

(b)

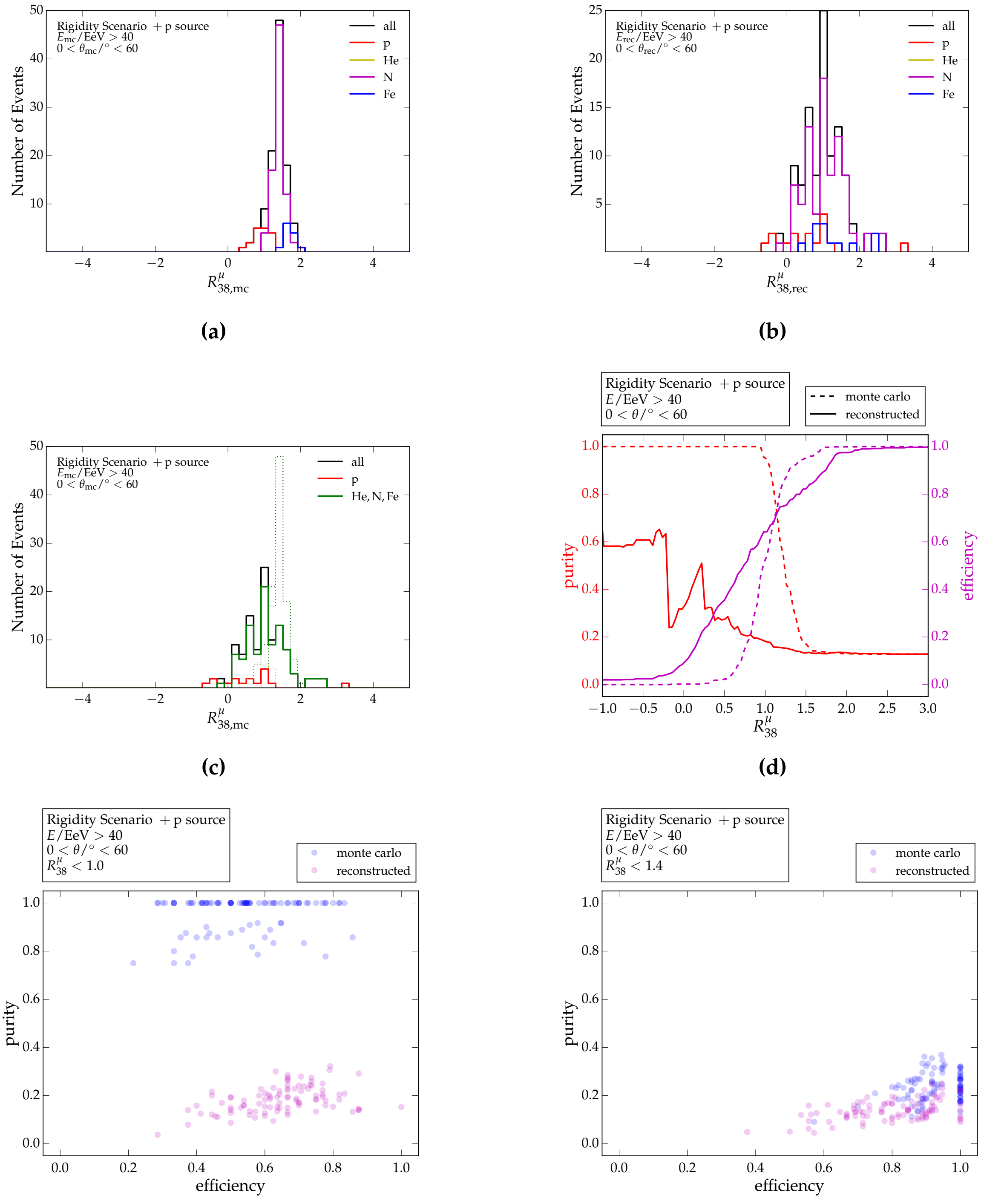
(c)

(d)

(e)

(f)

Figure 4.17: Results of 100 simulated 10-year maximum rigidity scenario fluxes with addition 10% proton fraction constant in energy. (a) Monte-carlo energy and zenith independent muon content estimator R^{μ}_{38} for sample 10-year flux. (b) Reconstructed R^{μ}_{38} for the same flux. (c) Monte-carlo (dashed lines) and reconstructed (solid lines) distributions of R^{μ}_{38} for the signal (proton) and background (all other nuclei). (d) Purity and efficiency as a function of $R^{\mu}_{38} < R^{\mu}_{38,\mathrm{cut}}$ for Monte-carlo (dashed lines) and reconstructed (solid lines) quantities. (e) Distribution of possible efficiency and purity values for a cut on R^{μ}_{38} of 1.0. Values calculated using Monte-carlo values are shown in magenta while those calculated with reconstructed values are shown in blue. Over the next 10 years, nature would only provide us with one sample from these distributions. (f) The same for a cut of R^{μ}_{38} of 1.4.



(a) **(b)** **(c)** **(d)** **(e)** **(f)**

Figure 4.18: Results of 100 simulated 10-year maximum rigidity scenario fluxes with a hypothetical nearby proton source at the highest energies accounting for 50% of the flux. (a) Monte-carlo energy and zenith independent muon content estimator R^{μ}_{38} for sample 10-year flux. (b) Reconstructed R^{μ}_{38} for the same flux. (c) Monte-carlo (dashed lines) and reconstructed (solid lines) distributions of R^{μ}_{38} for the signal (proton) and background (all other nuclei). (d) Purity and efficiency as a function of $R^{\mu}_{38} < R^{\mu}_{38,\rm cut}$ for Monte-carlo (dashed lines) and reconstructed (solid lines) quantities. (e) Distribution of possible efficiency and purity values for a cut on R^{μ}_{38} of 1.0. Values calculated using Monte-carlo values are shown in magenta while those calculated with reconstructed values are shown in blue. Over the next 10 years, nature would only provide us with one sample from these distributions. (f) The same for a cut of R^{μ}_{38} of 1.4.

CHAPTER 5

Summary and Conclusions

A number of enigmas remain in the field of UHECR physics. Some, such as the origins of these incredibly energetic particles, are long-standing questions. Others, such as the cause of the flux suppression at the end of the energy spectrum, have only recently been called into question due to the surprising trend towards heavier elements at higher energies, as measured by the Pierre Auger Observatory. A common thread throughout these quandaries is the salience of information on the mass of these cosmic rays. The work presented in this dissertation enters the picture here in the scope of the Pierre Auger Observatory's detector upgrade (AugerPrime), which aims to provide the information necessary to estimate the mass of cosmic rays on a per event basis. The work described in this dissertation contributed to this effort both technically and in terms of physics-related analysis. The principle contributions are summarized below.

Informing AugerPrime design decisions

Estimating the number of muons produced in an EAS and reducing mass-dependent biases in reconstructed energy were specified as the objectives of Auger's detector upgrade. A number of proposals were put forth including a layered surface detector and buried scintillators. A surface scintillation detector was determined as the the least invasive, most cost effective proposal capable of fulfilling the physics-based demands. An active detection area of $4\,\mathrm{m}^2$ was found to be warranted due to 20-100% distance-dependent gains in the resolution of the reconstructed muonic signal at the single-station level. Correlations between the scintillator and WCD signals were found to have a moderate dependency on the relative positioning of the two subdetectors, reaching correlation coefficients of up to 0.4 in the examined parameter spaces, although only small effects on the resolutions of reconstructed muonic signals were observed. A maximum difference in resolution of approximately 25% was observed between the two extremes of no area overlap between the detectors in the shower detector plane and complete overlap. Measurement at a distance of 200 m from the shower axis at the highest energies was found to be the approximate limit which could be feasibly achieved with the upper end of the scintillator dynamic range, as the slope of the lateral distributions of particles rapidly increases closer to the shower axis.

Upgrade of the $\overline{\text{Off}}\underline{\text{line}}$ cosmic ray detector simulation and reconstruction software

Extensive modifications and augmentations to the $\overline{\text{Off}}\underline{\text{line}}$ cosmic ray detector simulation and reconstruction software framework were required in order to handle the needs of a large-scale detector upgrade including new electronics for existing hardware, new hardware within an existing detector, and an entirely new sub-detector. These needs included storage of and access to vital data and reconstructed variables as well as methods for data processing during reconstruction. Changes and upgrades were performed with simplicity for the user and physics developer in mind.

Comprehensive Scintillator Surface Detector simulations

Simulations of the AugerPrime Scintillator Surface Detector were conceptualized and implemented in the $\overline{\text{Off}}\underline{\text{line}}$ software framework. This required the extraction of a number of distributions and variables from measurements of prototype detectors performed using a muon telescope. These simulations also included electronics and calibration properties measured with the same prototype detectors once deployed in the AugerPrime engineering array on site at the observatory. These simulations were validated with real measurements performed by a centimeter precision muon telescope at the Karlsruhe Institute of Technology. The number of photoelectrons, positional and angular dependencies of charge, and signal timing properties of data were shown to be well produced by the simulations. Initial impressions of the lateral distribution functions and signal uncertainties measured with the AugerPrime engineering array are also promising in their resemblance to simulation predictions. These simulations will be useful in parameterizing and validating future reconstruction algorithms as well as interpreting real measurements.

In addition to the development of the simulations, a station-level particle thinning algorithm was developed and implemented. This algorithm reduces simulation times and eliminates exceptionally long jobs or jobs failing due to memory or processing limitations. No significant biases or artificial fluctuations are observed in normal station signals when using this algorithm and as such, it promises to be an effective tool for geometrically complete, large-scale simulation productions, such as those performed for the GRID.

Matrix-based reconstruction of UHECR mass

The basic concept behind the combined use of scintillation and water-Cherenkov detectors is to capitalize on their differing responses to particles of the electromagnetic and muonic components of EAS. The total signals of the two detectors and the two desired quantities (electromagnetic and muonic signal in the WCD) may be formulated in a matrix. Through matrix inversion, one can analytically solve for each of the desired variables in terms of only the two observables and an additional two parameters, which can be obtained through simulations. This simple formulation was tested both the level of the event and for individual stations over two decades of primary energies. In order to perform an event-level reconstruction of the muonic and electromagnetic signals, the signal uncertainties and lateral distributions of the new scintillator detectors were also parameterized. At an energy of $10^{19.5}$ eV and zenith angle of 38°, a resolution in the reconstructed electromagnetic (muonic) signal of approximately 10% (25%) with a bias on the order of 5% (10%) is observed at the station level at a distance of 1000 m from the shower axis. At the event level, the resolution in the muonic signal improves to approximately 15% with a relative bias between proton and iron primaries on the order of 5%. It should be noted that the biases in the reconstructed muon signal are such that the muonic signal is overestimated for iron primaries and underestimated for proton primaries, which actually aids in discrimination between the two.

Potential for proton astronomy

The ability of the matrix-based reconstruction in selecting protons from scenarios with a background spectrum resembling idealized expectations for (1) photo-disintegration and (2) maximum-rigidity scenarios was examined. This analysis included fluctuations in the arriving flux for different elements over an additional 10 years of observatory operation as well as full detector simulations and event reconstruction. For the photo-disintegration scenario, a reasonable balance between efficiency (80%) and purity (80%) could be achieved with the dependence in each of these values on the possible fluctuations in the flux on the order of only a few percent. In the maximum-rigidity scenario, however, selection purity above 25% was improbable with a large spread in the possible selection efficiency which could be observed. This was the case both when a 10% proton component was added uniformly in energy and when a proton component identical in size and energy-dependency that of the maximum-rigidity iron component was added at the highest energies.

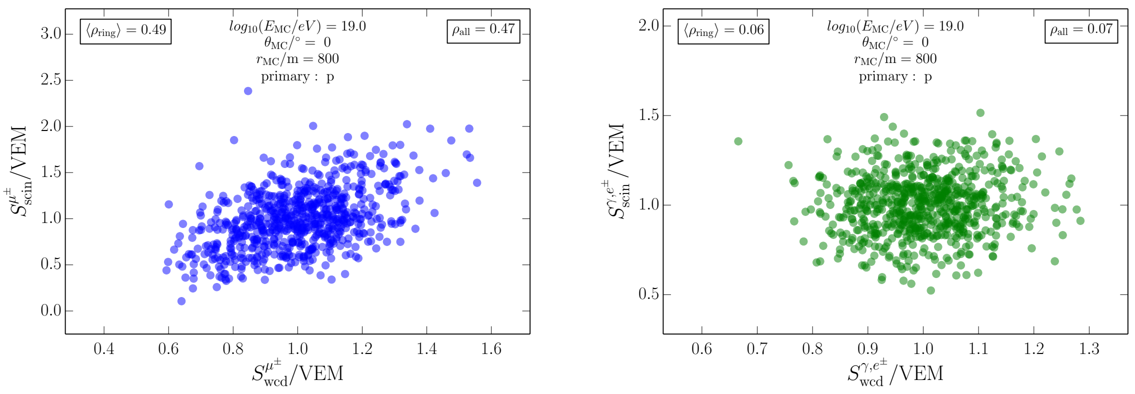
Acronyms

This is a list of acronyms used within this work sorted alphabetically.

AMIGA	Auger Muon Detectors for the Infill Ground Array	8
Auger	the Pierre Auger Observatory	4
CDAS	the central data acquisition system	10
CLF	Central Laser Facility	7
c.d.f.	cumulative distribution function	30
CMB	cosmic microwave background radiation	2
CORSIKA	Cosmic Ray Simulations for Kascade	52
EA	engineering array	32
EAS	extensive air shower	3
XLF	Extreme Laser Facility	7
FD	Fluorescence Detector	4
FNAL	Fermi National Accelerator Laboratory	32
GDAS	Global Data Assimilation System	7
ITeDA	the Instituto de Tecnologías en Detección y Astropartículas	19
LDF	lateral distribution function	12
LST	limited streamer tube	41
MIP	minimum ionizing particle	30
MTD	muon tracking detector	40
NKG	Nishimura-Kamata-Greisen	12
Auger	the Pierre Auger Observatory	4
p.d.f.	probability density function	68
PMT	photo-multiplier tube	5
RPC	resistive plate chamber	17
SD	Surface Detector	4
SDP	shower detector plane	12
SNR	supernova remnant	1
SPMT	Small Photo-multiplier Tube	18
SSD	Scintillator Surface Detector	32
UB	unified board	18
UMD	Underground Muon Detector	18
UUB	upgraded unified board	18
UHECR	ultra-high energy cosmic ray	3
VEM	vertical-equivalent muon	19
WCD	water-Cherenkov detector	4

APPENDIX A

AugerPrime Design



(a) **(b)**

Figure A.1: Correlations in the (a) muonic signal and (b) electromagnetic signal for the same set of stations. The muonic signal is the dominant contribution to correlations in the total signal.

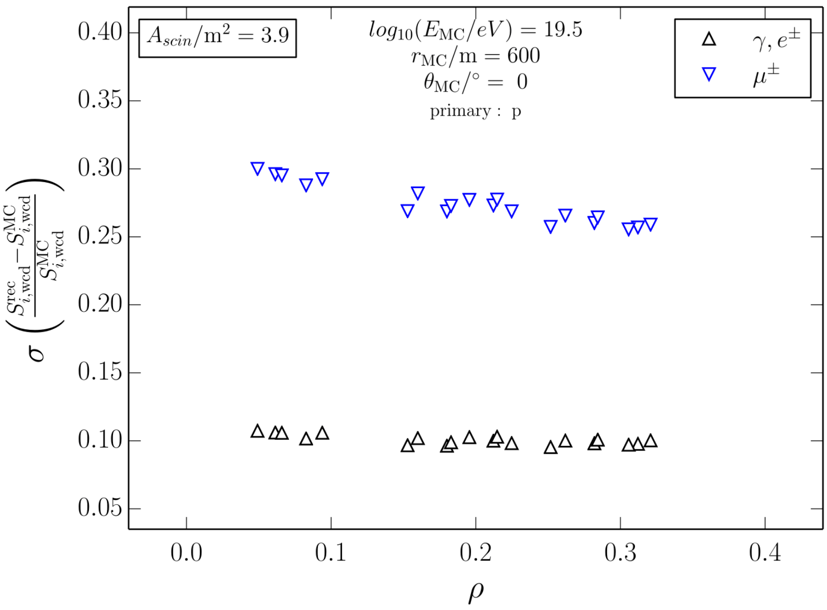
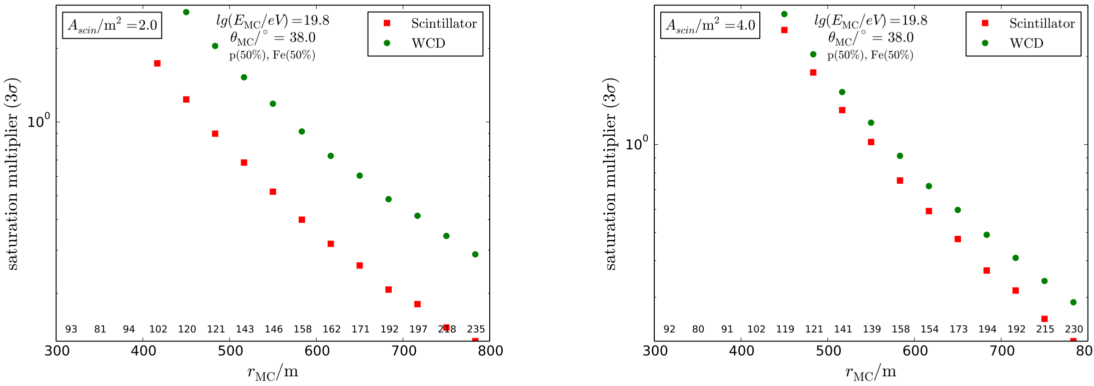


Figure A.2: Resolution of muonic signal reconstructed via the matrix formalsim as a function of correlation strength.



(a) **(b)**

Figure A.3: Saturation multipliers for (a) $2\,\mathrm{m}^2$ and (b) $4\,\mathrm{m}^2$ scintillators. The effect of active scintillation area on saturation distances is evident.

APPENDIX B

Simulations

B.1 Geometric Effect

For the majority of simulations where the shower core is randomly thrown onto the array, the distance between the shower axis and the closest station is on the same order of magnitude as the distance between the axis and the second closest station. While the closest station does receive more particles than any other station in such cases, it does not receive an overwhelming majority of all particles simulated in all stations. In the few cases, however, where the shower core lands very close to a station, this station receives a massive number of particles and and does dominate the all-station particle count. As a result, simulation times are extraordinarily large, which may be observed in Fig. B.1. This phenomenon accounts for the outliers in Figure 3.15a.

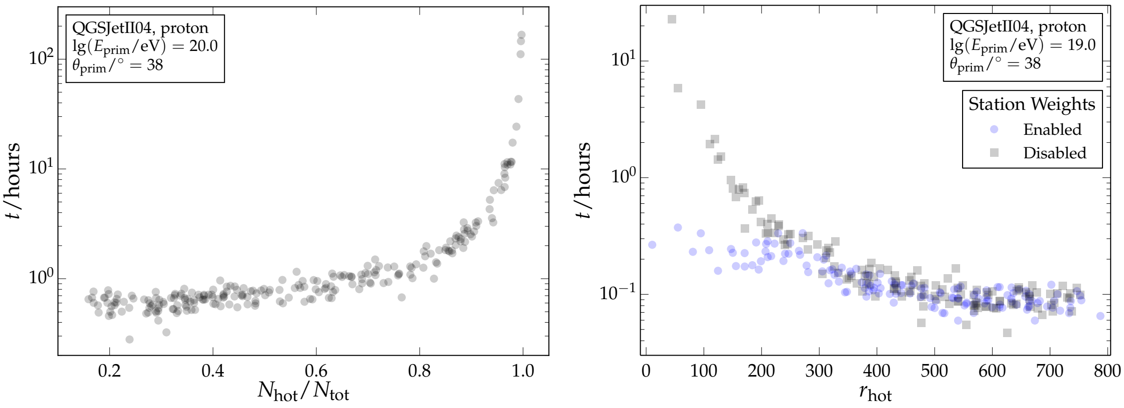


Figure B.1: *Left:* Event simulation time as a function of the fraction of particles injected in the hottest station prior to implementation of the station weights algorithm. As this fraction approaches one, event simulation times become very large. *Right:* Simulation time as a function of the distance between the shower axis and the hottest station in the shower plane. Whereas simulations with a station very close to the axis take massive amounts of time with the stations weights algorithm disabled, enabling the algorithm permits simulation of such stations without significant costs in time.

B.2 Muon Fraction

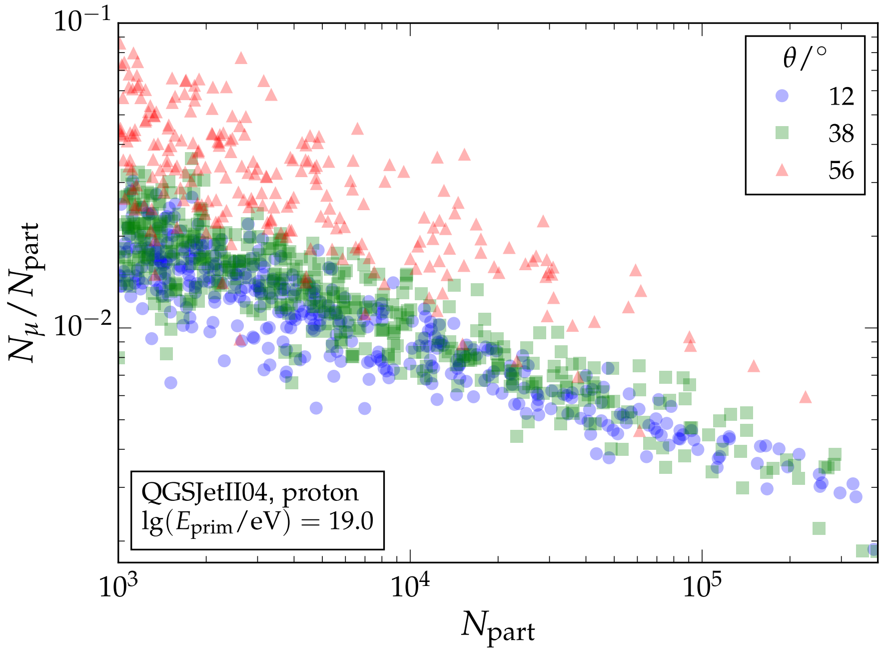


Figure B.2: Muon fraction as a function of particle number. More inclined showers have a larger muon fraction due to a larger attenuation of the electromagnetic component.

B.3 Offline Configuration Samples

Station descriptions

```
<station id="1687">
  <northing unit="meter"> 6115220.51 </northing>
  <easting unit="meter"> 504624.7 </easting>
  <altitude unit="meter"> 1348.8 </altitude>
  <name> Kopal </name>
  <commission> 2008-02-26T09:26:37Z </commission>
  <decommission> 2029-12-31T23:59:59Z </decommission>
  <hardwareUpgrade> 2000-01-01T00:00:00Z </hardwareUpgrade>
  <hardwareVersion> 1 </hardwareVersion>
  <inGrid> 1 </inGrid>
  <groupId> 0 </groupId>
  <axes> 131 115 </axes>
  &UTMdata;
</station>
```

Hardware grouping

```
<hardware hardwareVersion="1">
```

```
  <tankCasing id="rotoplas"/>
  <tankPMTConfig id="standardPlusSmall"/>
  <water id="standard"/>
  <liner id="1073B-tyvek"/>
  <electronics id="120MHz"/>
  <scintillator id="ssdPrototype"/>

  <!-- Three large PMTs (Tank) -->
  <PMT PMTId="1" id="photonis-XP1805"/>
  <PMT PMTId="2" id="photonis-XP1805"/>
  <PMT PMTId="3" id="photonis-XP1805"/>
  <!-- One small PMT (Tank) -->
  <PMT PMTId="4" id="hamamatsu-R6095-100"/>
  <!-- One scintillator PMT -->
  <PMT PMTId="5" id="ssdPlaceholder"/>

  <!-- PMT-specific electronics properties -->
  <pmtElectronics PMTId="1" id="wcdLarge120MHz"/>
  <pmtElectronics PMTId="2" id="wcdLarge120MHz"/>
  <pmtElectronics PMTId="3" id="wcdLarge120MHz"/>
  <pmtElectronics PMTId="4" id="wcdSmall120MHz"/>
  <pmtElectronics PMTId="5" id="scin120MHz"/>

</hardware>
```

Hardware model properties

```
  <electronics id="120MHz">
    <frequency unit="MHz"> 120 </frequency>
    <traceBins> 2304 </traceBins>
    <bits> 12 </bits>
    <impulseResponse>
      <x unit="ns">
        ...  3.33567e-01  3.33667e-01  3.33689e-01  ...
      </x>
      <y unit="V">
        ... -1.29953e-32 -7.55255e-32 -1.08883e-31 ...
      </y>
    </impulseResponse>
  </electronics>
```

APPENDIX C

Acknowledgments

I would like to thank the two reviewers of this dissertation, Professors Blümer and Etchegoyen, for facilitating this work and for laying the foundations of this joint German and Argentine program through which I have now passed. It has truly been a life-changing experience on many levels. To my advisor, Markus Roth, I am appreciative for his support as well as the trust and freedom I was given. Were it not for him, I would very likely not have come to Karlsruhe to pursue a doctoral degree. Sabine Bucher and Marie-Christine Kaufmann deserve my sincerest gratitude for their outstanding logistical support.

Immense appreciation goes to my family; to my mother, Monika Schmidt, and father, Karl Schmidt, for their steadfast nature; to my brother, Kevin Schmidt, who is my oldest friend; to my wife, Karina Hoyos, who stood by my side with profound acceptance during my most joyful moments, but also throughout my darkest, most hellish struggles. Here, I would also like to thank Ariel Bridgeman and Darko Veberic, who I also consider family, like a sister and sometimes exasperating uncle. To my relatives in Germany: Helene, Karlheinz, Niklas, and Lena Brückner and Monika, Robert, and Markus Staudter, for their welcome company. To my relatives in Argentina: Nelly Rueja and Ciro, Vanina, Jimena, Florencia, and Pablo Hoyos, for ther welcome company.

I am very grateful to have had the pleasure to spend the last years in the presence of many wonderful people. To Federico Suarez, who essentially took me in during my time in Buenos Aires. To Nicolas Gonzalez and Ana Botti, who helped me through a particular time of need. To Peter Secor and Erica Getz, who always welcomed me with open arms in Colorado. To Kimberly Cara and Toty Perea for their welcome company. To Alexander Streich and Alexander Herve, who have always been readily willing to calmly lend their ears and hands. To Alexander Schulz and Daniela Mockler, as my first office mates, who hold a special place in my initial memories at KIT in particular. To Felix Riehn for many a drink and non-judgemental discussion. To Stefan Grimm, for the many discussions on non-standard subject matter. To Tanguy Pierog and Detflef Maurel, for the evening squash. To Aaron Mull and Esteban Chavez, for the precious moments. To Alvaro Taboada and Johannes Hulsman, for their welcome company. To Raul Prado, Ana Laura Müller, Isabel Goos, Martin Schimassek, William Painter, Ewa Holt, Sarah Müller, Mariela Josebachuili, Maximilian Reininghaus, Monica Cifuentes, María Ituarte, Romina Ollaze, and Alan Coleman, for the good times and company. To the students I had the pleasure to work with, Manuel Frietsch, Matias Kessler, and Tobias Schulz, who likely taught me more than I taught them. To the many other colleagues and friends with whom I had meaningful contact, but are not mentioned explicitly here.

I am truly fortunate.

I would also like to cite Matplotlib [50], which was used to generate many of the figures in this work.

Bibliography

[1] R. Engel, Private communication concerning cosmic ray energy spectrum (2017).

[2] T. Antoni *et al.* (KASCADE Collab.), KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems, Astropart. Phys. 24 (1-2) (2005) 1 – 25. `doi:10.1016/j.astropartphys.2005.04.001`.

[3] W. D. Apel *et al.* (KASCADE-Grande Collab.), Kneelike Structure in the Spectrum of the Heavy Component of Cosmic Rays Observed with KASCADE-Grande, Phys. Rev. Lett. 107 (2011) 171104. `doi:10.1103/PhysRevLett.107.171104`.

[4] A. M. Hillas, Topical Review: Can diffusive shock acceleration in supernova remnants account for high-energy galactic cosmic rays?, J. Phys. G 31 (2005) R95–R131. `doi:10.1088/0954-3899/31/5/R02`.

[5] V. Berezinsky, A. Gazizov, S. Grigorieva, On astrophysical solution to ultrahigh energy cosmic rays, Phys. Rev. D 74 (2006) 043005. `doi:10.1103/PhysRevD.74.043005`.

[6] T. Gaisser, R. Engel, and E. Resconi, Cosmic Rays and Particle Physics, Cambridge University Press, 2016.

[7] R. A. Batista *et al.*, Effects of uncertainties in simulations of extragalactic UHECR propagation, using CRPropa and SimProp (2015). `arXiv:arXiv:1508.01824`.

[8] A. M. Hillas, Cosmic Rays: Recent Progress and Some Current Questions, in: Conf. on Cosmology, Galaxy Formation and Astro-Particle Physics on the Pathway to the SKA, Oxford, England, 2006. `arXiv:astro-ph/0607109`.

[9] A. Bridgeman, Determining the Mass Composition of Ultra-high Energy Cosmic Rays Using Air Shower Universality, Dissertation, Karlsruhe Institute of Technology (2018).

[10] A. Aab *et al.* (Pierre Auger Collab.), The Pierre Auger Cosmic Ray Observatory, Nucl. Instrum. Methods Phys. Res. A798 (2015) 172–213. `doi:10.1016/j.nima.2015.06.058`.

[11] I. A. *et al.*, The surface detector system of the pierre auger observatory, Nucl. Instrum. Meth. A586 (2008) 409 – 420. `doi:10.1016/j.nima.2007.12.016`.

[12] J. A. *et al.*, The fluorescence detector of the pierre auger observatory, Nucl. Instrum. Meth. A620 (2010) 227 – 251. `doi:10.1016/j.nima.2010.04.023`.

[13] D. Veberic, Private communication (2018).

[14] X. Bertou *et al.* (Pierre Auger Collab.), Calibration of the surface array of the Pierre Auger Observatory, Nucl. Instrum. Methods Phys. Res. A568 (2006) 839–846. `doi:10.1016/j.nima.2006.07.066`.

[15] A. Aab *et al.* (Pierre Auger Collab.), Prototype muon detectors for the AMIGA component of the Pierre Auger Observatory, J. of Instrum. 11 (02) (2016) P02012.

[16] J. Abraham *et. al.* (Pierre Auger Collab.), Measurement of the energy spectrum of cosmic rays above 10^{18} eV using the Pierre Auger Observatory, Phys. Lett. B 685 (2010) 239–246. `arXiv:1002.1975`, `doi:10.1016/j.physletb.2010.02.013`.

[17] M. Settimo *et al.* (Pierre Auger Collab.), An update on a search for ultra-high energy photons using the Pierre Auger Observatory, in: Proc. 32nd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2011.

[18] P. Abreu *et al.* (Pierre Auger Collab.), Ultrahigh Energy Neutrinos at the Pierre Auger Observatory, Adv. High Energy Phys. 708680 (2013).

[19] A. Aab *et al.* (Pierre Auger Collab.), Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8×10^{18} eV, Science 357 (2017).

[20] J. Abraham *et al.* (Pierre Auger Collab., Measurement of the Depth of Maximum of Extensive Air Showers above 10^{18} eV, Phys. Rev. Lett. 104 (2010) 091101. `doi:10.1103/PhysRevLett.104.091101`.

[21] P. Abreu *et. al.* (Pierre Auger Collab.), Measurement of the Proton-Air Cross Section at $\sqrt{s}$=57 TeV with the Pierre Auger Observatory, Phys. Rev. Lett. 109 (2012) 062002. `doi:10.1103/PhysRevLett.109.062002`.

[22] A. Aab *et al.* (Pierre Auger Collab.), Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events, Phys. Rev. Lett. 91 (3) (2015) 032003. `arXiv:1408.1421`, `doi:10.1103/PhysRevD.91.032003`.

[23] K. Greisen, Progress in elementary particle and cosmic ray physics, Vol. 3, North-Holland Publishing, Amsterdam, 1956.

[24] K. Kamata, J. Nishimura, The Lateral and the Angular Structure Functions of Electron Showers, Prog. Theoret. Phys. Suppl. 6 (1958) 93–155. `doi:10.1143/PTPS.6.93`.

[25] D. Newton, J. Knapp, A. A. Watson, The optimum distance at which to determine the size of a giant air shower, Astropart. Phys. 26 (2007) 414–419. `arXiv:astro-ph/0608118`, `doi:10.1016/j.astropartphys.2006.08.003`.

[26] D. Veberic *et al.* (Pierre Auger Collab.), Estimation of Signal in Saturated Stations of Pierre Auger Surface Detector, in: Proc. 33th Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[27] M. Ave, P. Bauleo, A. Castellina, A. Chou, J. L. Harton, R. Knapik, and G. Navarra (Pierre Auger Collab.), The accuracy of signal measurement with the water Cherenkov detectors of the Pierre Auger Observatory, Nucl. Instrum. Meth. A578 (2007) 180–184. `doi:10.1016/j.nima.2007.05.150`.

[28] J. H. *et al.*, Observations of Extensive Air Showers near the Maximum of Their Longitudinal Development, Phys. Rev. Lett. 6 (1961) 22–23. `doi:10.1103/PhysRevLett.6.22`.

[29] A. Castellina *et al.* (Pierre Auger Collab.), The dynamic range of the AugerPrime Surface Detector: technical solution and physics reach, in: Proc. 35th Int. Cosmic Ray Conf., Busan, South Korea, 2017.

[30] R. Smida *et al.* (Pierre Auger Collab.), Scintillator detectors of AugerPrime, in: Proc. 35th Int. Cosmic Ray Conf., Busan, South Korea, 2017.

[31] A. Dyshkant *et al.*, About NICADD extruded scintillating strips, FERMILAB-PUB-05-010-E (2005).

[32] Wavelength shifting fibers, `http://kuraraypsf.jp/psf/ws.html`.

[33] NOvA Collaboration, The NOvA Technical Design Report, FERMILAB-DESIGN-2007-01 (2007).

[34] A. Streich, Scintillator Surface Detector for the Upgrade of the Pierre Auger Observatory, Master's thesis, Karlsruhe Institute of Technology (2017).

[35] S. Agostinelli *et al.* (Geant4 Collab.), GEANT4: A Simulation toolkit, Nucl. Instrum. Meth. A506 (2003) 250. `doi:10.1016/S0168-9002(03)01368-8`.

[36] S. Argirò, S. Barroso, J. Gonzalez, L. Nellen, T. Paul, T. Porter, L. P. Jr., M. Roth, R. Ulrich, D. Veberič, The Offline software framework of the Pierre Auger Observatory, Nucl. Instrum. Meth. A580 (2007) 1485–1496. `doi:http://dx.doi.org/10.1016/j.nima.2007.07.010`.

[37] P. Doll, *et al.*, Muon tracking detector for the air shower experiment KASCADE, Nucl. Instrum. Methods Phys. A488 (2002) 517–535. `doi:10.1016/S0168-9002(02)00560-0`.

[38] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. Capdevielle, CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers (1998).

[39] P. Billoir, A sampling procedure to regenerate particles in a ground detector from a “thinned” air shower simulation output, Astropart. Phys. 30 (2008) 270–285. `doi:10.1016/j.astropartphys.2008.10.002`.

[40] A. Taboada, Private communication (2016).

[41] B. Wundheiler, Rayos còsmicos con energìas entre 10^{17} y 10^{19} eV, Dissertation, Universidad de Buenos Aires (2013).

[42] B. Wundheiler *et al.* (Pierre Auger Collab.), The AMIGA muon counters of the Pierre Auger Observatory: performance and first data, in: Proc. 32th Int. Cosmic Ray Conf., Beijing, China, 2011.

[43] M. Kessler, Position-sensitive time-series analysis of SSD signals, Bachelor Thesis, Karlsruhe Institute of Technology (2016).

[44] T. Suomijärvi *et al.* (Pierre Auger Collab.), New electronics for the surface detectors of the Pierre Auger Observatory, in: Proc. 35th Int. Cosmic Ray Conf., Busan, South Korea, 2017.

[45] D. Veberic *et al.* (Pierre Auger Collab.), Estimation of Signal in Saturated Stations of Pierre Auger Surface Detector, in: Proc. 33th Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[46] A. L.-S. *et al.*, Layered water Cherenkov detector for the study of ultra high energy cosmic rays, Nucl. Instrum. Meth. A767 (2014) 41 – 49. `doi:10.1016/j.nima.2014.08.029`.

[47] T. Schulz, Enhanced UHECR Event Reconstruction by means of Sampling Lateral Distributions with Multiple Surface Sub-Detectors, Bachelor Thesis, Karlsruhe Institute of Technology (2018).

[48] Pierre Auger Collaboration, The Pierre Auger Observatory Upgrade - Preliminary Design Report (2016). `arXiv:arXiv:1604.03637[astro-ph.IM]`.

[49] F. Fenu *et al.* (Pierre Auger Collab.), The cosmic ray energy spectrum measured using the Pierre Auger Observatory, in: Proc. 35th Int. Cosmic Ray Conf., Busan, South Korea, 2017.

[50] J. D. Hunter, Matplotlib: A 2d graphics environment, Computing In Science & Engineering 9 (3) (2007) 90–95. `doi:10.1109/MCSE.2007.55`.

Erklärung

Karlsruhe, den 01.06.2018

Erklärung der selbständigen Anfertigung meiner Dissertationsschrift

Hiermit versichere ich, dass ich die Dissertationsschrift mit dem Titel

Sensitivity of AugerPrime to the masses of ultra-high-energy cosmic rays

selbständig und ohne unerlaubte fremde Hilfe verfasst habe. Dabei habe ich keine anderen, als die von mir angegebenen Hilfsmittel benutzt.

Your Name