2015

Search for TeV Gamma-Ray Sources in the Galactic Plane with the HAWC Observatory

Hao Zhou
Michigan Technological University, hzhou1@mtu.edu

SEARCH FOR TEV GAMMA-RAY SOURCES IN THE GALACTIC PLANE WITH THE HAWC OBSERVATORY

By

Hao Zhou

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Physics

MICHIGAN TECHNOLOGICAL UNIVERSITY

2015

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Physics.

Department of Physics

Dissertation Advisor:	*Dr. Petra Huentemeyer*
Committee Member:	*Dr. Brian Fick*
Committee Member:	*Dr. Robert Nemiroff*
Committee Member:	*Dr. Robert Pastel*
Department Chair:	*Dr. Ravindra Pandey*

Contents

List of Figures

List of Tables

Preface

This thesis is based on the data taken with the HAWC Gamma-Ray Observatory, the subject of an intentional collaboration between more than 100 scientists from institutions in the US and Mexico. The timing calibration in Chapter 5, the likelihood analysis in Chapter 7, and the results in Chapter 8 are my original work, with the assistance of many other collaborators.

Part of the contents presented in Chapter 7 and 8 have been accepted by the Astrophysical Journal as "Search for TeV Gamma-Ray Emission from Point-Like Sources in the Inner Galactic Plane with a Partial Configuration of the HAWC Observatory" [1].

Acknowledgments

First, I would like to express my deepest gratitude to my advisor Petra Huentemeyer, for your guidance and continuous support of my Ph.D. study. You are the best advisor and mentor I have ever worked with, from accepting me as your first Ph.D. student to providing me an excellent atmosphere and freedom for doing research, from proofreading my thesis that was written at the last minute to your creative style of writing comments on margins. You make it possible to achieve my childhood dream of being an astrophysicist.

To Michelle Hui and Hugo Ayala, I am exceedingly glad that we worked together during most of my Ph.D. study. I enjoyed our long group meetings, from finding the toughest bugs in my code to brainstorming ideas in data analyses. We made the great Michigan Tech HAWC group in snowy Upper Peninsula of Michigan.

Thanks to the HAWC collaboration. I would like to thank Brenda Dingus, John Matthews, Robert Lauer, Francisco Salesa Greus for working together on the calibration of HAWC. Thank you, Jim Braun, John Pretz, Andy J. Smith, and Gus Sinnis, for helping me understand the reconstruction algorithms of HAWC data. Special thanks go to HAWC likelihood team for developing the likelihood framework and for being a constant source of statistical knowledge, especially Robert Lauer, Giacomo

Vianello, Pat Harding, and Patrick Younk.

I would like to thank all my committee members, Dr. Brian Fick, Dr. Robert Nemiroff and Dr. Robert Pastel, for their insightful questions and comments on my thesis.

Last but not least, thanks to my family. Mom and Dad: you provide me the longest support and inspire me to be who I am today. My wife Wei Zhang: you encourage me to chase my childhood dream and always believe in me through good and bad time. As this acknowledgement was written, our first child Carl was just born. Thanks Wei!

Abstract

Cosmic rays, with an energy density of $\sim 1\,\mathrm{eV\,cm^{-3}}$, play an important role in the evolution of our Galaxy. Very high energy (TeV) gamma rays provide unique information about the acceleration sites of Galactic cosmic rays. The High Altitude Water Cherenkov (HAWC) Gamma-Ray Observatory is an all-sky surveying instrument sensitive to gamma rays from 100 GeV to 100 TeV with a 2 steradian instantaneous field of view and a duty cycle of $> 95\%$. The array is located in Sierra Negra, Mexico at an elevation of 4,100 m and was inaugurated in March 2015. Thanks to its modular design, science operation began in Summer 2013 with one third of the array. Using this data, a survey of the inner Galaxy region of Galactic longitude $l \in [+15^\circ, +50^\circ]$ and latitude $b \in [-4^\circ, +4^\circ]$ is performed. To address the ambiguities arising from unresolved sources in the data, a maximum likelihood technique is used to identify point source candidates. Ten sources and candidate sources are identified in this analysis. Eight of these are associated with known TeV sources but not all have differential fluxes compatible with previous measurements. Three sources are detected with significances $> 5\,\sigma$ after accounting for statistical trials, and are associated with known TeV sources. With data taken with the full array and improved reconstruction algorithms, the significance on the Crab nebula increases from $3.1\,\sigma/\sqrt{\mathrm{day}}$ to $5.5\,\sigma/\sqrt{\mathrm{day}}$, which allows more sensitive sky surveys and more precise spectral and morphological analyses on individual sources.

Chapter 1

Introduction to Gamma-Ray Astrophysics

1.1 The "Violent" Universe

Our knowledge of the universe beyond the solar system comes from observations of photons over an enormous range of energies. The majority of photons propagating in the universe and observed on the Earth are thermal radiation, which is produced by hot objects such as stars and follows a black-body spectrum. Black-body radiation is a type of electromagnetic radiation that a black body, which absorbs all electromagnetic radiation, emits to stay in thermal equilibrium. The spectrum of black-body radiation

only depends on the temperature of the source. Thermal radiation can reach into the kilo electron volt (keV) energy range and beyond under extreme conditions. However, there are other photons, for example the cosmic gamma rays, which span at least 14 energy decades, from 10^6 to $\geqslant 10^{20}$ eV [2]. This is well beyond the energy range thermal emission can reach. These photons are produced by "violent" non-thermal collective mechanisms, focusing the energy outflow on a relatively small number of particles. Instead of a black-body spectrum, photons or particles produced by non-thermal processes follow a power-law spectrum:

$$\frac{dN}{dE} \propto E^{-\Gamma}, \tag{1.1}$$

where Γ is the spectral index.

Cosmic rays are high-energy radiation that are produced outside the solar system. They are composed primarily of protons and heliums, as well as heavier nuclei, electrons and positrons, and gamma rays. Fig. 1.1 shows the energy spectrum of cosmic rays measured on the Earth. Cosmic rays follow a power law, and there are two features, which are caused by the change in slope at the "knee" (4×10^{15} eV) and at the "ankle" (5×10^{18} eV). The cosmic rays up to at least the knee are thought to be accelerated in our milky way Galaxy, and are called Galactic cosmic ray (GCR).

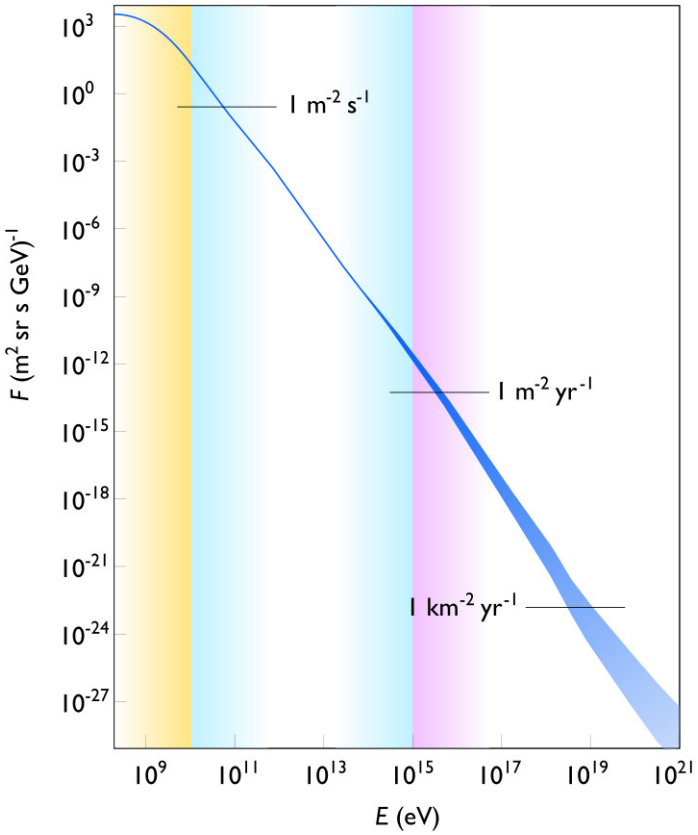

Figure 1.1 The cosmic ray spectrum above 10^9 eV measured at the top of the atmosphere. It follows a power law, and there are two features, the change in slope at the "knee" (4×10^{15} eV) and at the "ankle" (5×10^{18} eV). Image credit: S. Lafebre (see Appendix A).

Cosmic rays span a very wide energy range. The cosmic rays with the highest energies, beyond energies that any laboratory on the Earth can reach, provide unique information about the most energetic astrophysical phenomenas at most "violent" sites, which cannot be studied elsewhere. They are abundant with energy density of $\sim 1\,\mathrm{eV\,cm^{-3}}$, comparable to that contained in the cosmic microwave background (CMB), and thus play an important role in the evolution of our Galaxy. However, after more than a century since cosmic rays were first discovered by Victor Hess in 1912, the origin of cosmic rays remains one of the biggest questions in the field. There

is no unambiguous evidence on where cosmic rays are accelerated. The studies on the acceleration site, acceleration populations, and acceleration mechanisms are very important. However, cosmic rays are charged particles that are deflected and diffused when propagating in interstellar magnetic fields. The observation of cosmic rays on the Earth can therefore not provide information on where they come from. On the other hand, gamma rays that are produced by interactions between cosmic rays and other particles or fields can be used to probe cosmic rays since gamma rays point back to their production sites.

1.2 Galactic Gamma-Ray Sources

Gamma rays are photons with the highest energies in the electromagnetic spectrum. Traditionally, based on the energy per photon, gamma rays are defined as low (MeV), high (GeV), very high (TeV), ultra high (PeV) and extremely high (EeV) energy gamma rays [2]. There are three main astrophysical processes that contribute to the gamma-ray emission:

1. synchrotron emission, emitted by high energy electrons when gyrating in a magnetic field near a compact object;

2. inverse Compton scattering of low energy photons (CMB, star light, or IR from dusts) to high energies by ultra-relativistic electrons;

3. decay of π^0 that are produced by the interaction of cosmic rays with inter stellar medium (ISM).

The first two processes are referred to as leptonic processes and the π^0 decay is referred to as a hadronic process.

The Galactic gamma-ray emission consists of contributions from pulsars, pulsar wind nebulae, supernova remnants, compact object binaries, and the Galactic diffuse emission. The discrete gamma-ray sources produce gamma rays near the acceleration sites of cosmic rays. The diffuse emission is produced when the cosmic rays propagate in ISM.

1.2.1 Pulsar Wind Nebulae

Pulsars are rapidly rotating magnetized neutron stars formed at supernova explosions. A strong co-rotating magnetic field induces an electric field that accelerates electrons and positrons to ultra-relativistic energies. Pulsars steadily lose their rotational energy through formation of these relativistic particle winds called pulsar winds. Pulsed emission is observed from pulsars when the beams of emission are pointing toward the Earth. The existence of TeV pulsed emission from pulsars is still in doubt because of the heavy absorption of TeV gamma rays in strong magnetic fields of pulsars.

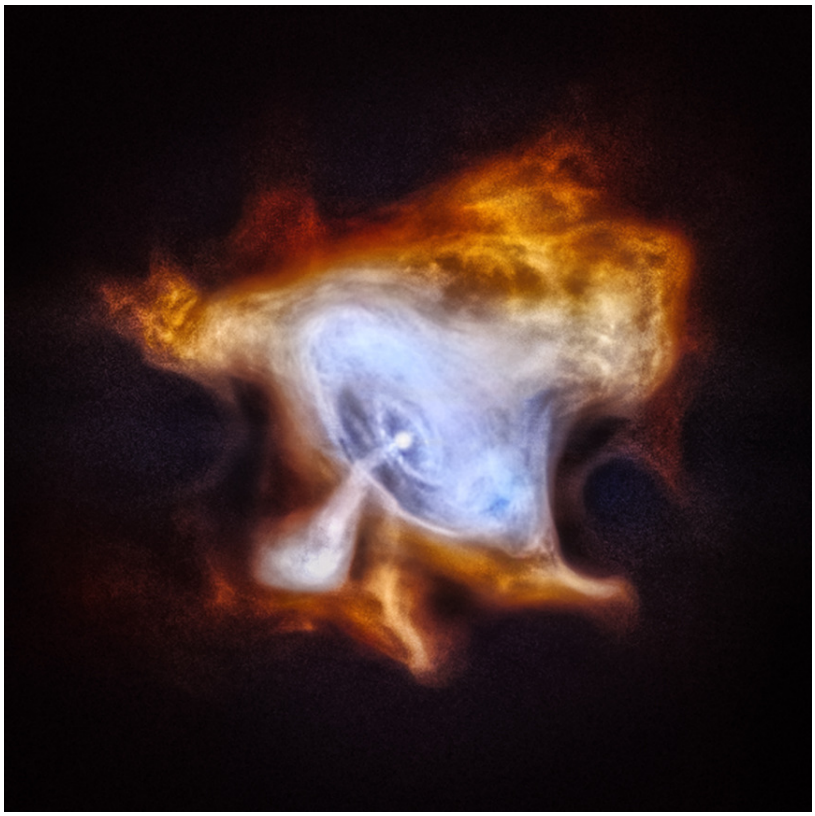

Figure 1.2 The Crab nebula in X-ray by Chandra X-ray telescope. Lower-energy X-rays are red, medium energy X-rays are green, and highest-energy X-rays are blue. Image credit: NASA/CXC/SAO (see Appendix A).

The confinement of pulsar winds by the ISM causes the formation of a so-called pulsar wind nebula (PWN). The interactions between pulsar winds and surrounding ISM generate shocks, which accelerate electrons and positrons to ultra-relativistic energies through the process of Fermi acceleration [3]. PWNe have been observed from the radio band to TeV gamma rays. The photons are produced mainly through two important mechanisms. At TeV energies, inverse Compton scattering dominates. Synchrotron radiation contributes to a very broad energy range from radio emission to GeV gamma rays. Fig. 1.2 shows an image of the Crab nebula in X-ray.

1.2.2 Supernova Remnants

Supernova remnants (SNRs) are the remains of massive star explosions. Fig. 1.3 shows an X-ray image of the Tycho SNR. The shock wave launches into the ISM after the star explosions and particles are accelerated by the shock wave through the process of Fermi acceleration [3]. Gamma rays are produced from the decay of π^0 when protons that are accelerated by the shock wave interact with the atoms in ISM. Models predict that the particles are accelerated with an efficiency of 10% (i.e. 10% of the kinetic energy of the shock wave transfers to the particles) and up to at least 10^{15} eV (PeV). In this regard, SNRs have been postulated to be the dominant cosmic ray acceleration sites since they can provide sufficient power to explain the cosmic ray flux as observed from the Earth. TeV gamma-ray observations are important in order to proof that SNRs can accelerate protons up to PeV energies, in these so-called Pevatrons,.

1.3 Galactic Gamma-Ray Survey

Besides the types of gamma-ray sources that are listed above, many of the gamma-ray sources in the Galactic plane are unidentified (UID) due to the fact that the measurements can envelope multiple sources identified at other wavelengths. Morphological and spectral studies are crucial for making associations with observations at other

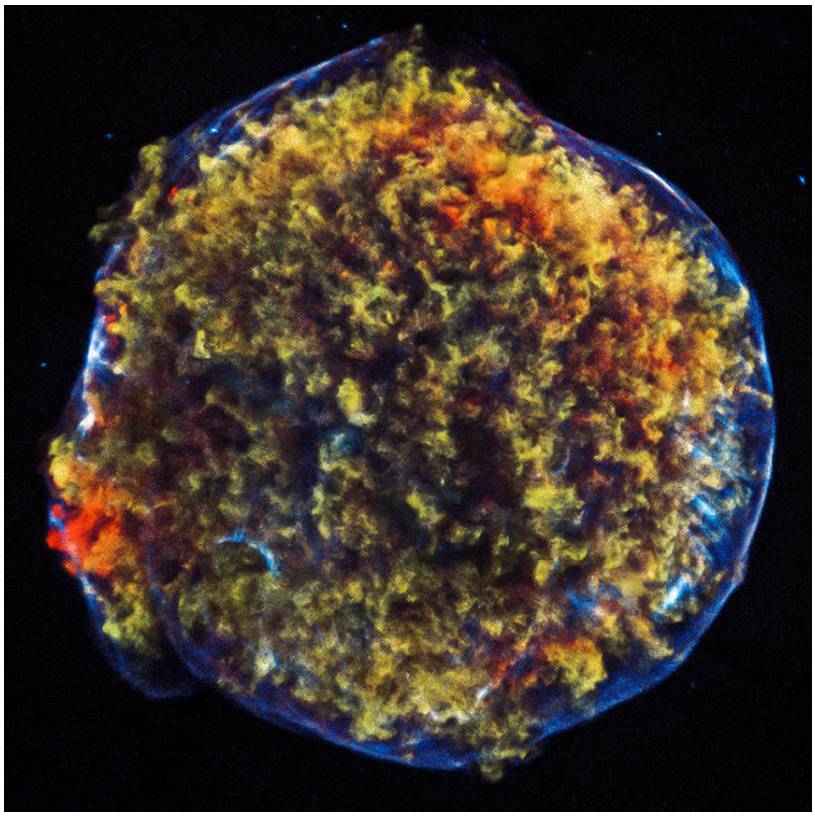

Figure 1.3 The Tycho SNR in X-ray by Chandra X-ray telescope. Image credit: NASA/CXC/SAO (see Appendix A).

wavelengths and for distinguishing leptonic and hadronic gamma-ray production processes which will aid in source identification.

On the other hand, survey observations are also important in order to study the acceleration populations of cosmic rays. Galactic plane surveys and observations have been performed by the H.E.S.S. imaging atmospheric Cherenkov telescopes (IACTs) (for example, [4, 5, 6, 7, 8]). Over 15 sources were discovered by H.E.S.S. (see Fig. 1.4) within the inner Galaxy region presented in this thesis (Galactic longitude $l \in [+15^\circ, +50^\circ]$ and latitude $b \in [-4^\circ, +4^\circ]$). Similar surveys within this region have

been performed by Milagro [9] and ARGO [10], along with targeted observations by the IACTs VERITAS and MAGIC (for example, [11, 12, 13]). At lower energies, the *Fermi* Large Area Telescope (*Fermi*-LAT) has published its third source catalog in the 100 MeV-300 GeV energy range (3FGL) based on the first four years of science operation [14] (see Fig. 1.5). The region surveyed with the partial HAWC array includes the locations of 73 sources from this catalog, 47 of which are without known astronomical associations. The first catalog with sources > 10 GeV (1FHL) has been published based on the first three years of *Fermi*-LAT data [15]. The region surveyed in this thesis contains twelve sources from this catalog, four of which are not in the 3FGL catalog.

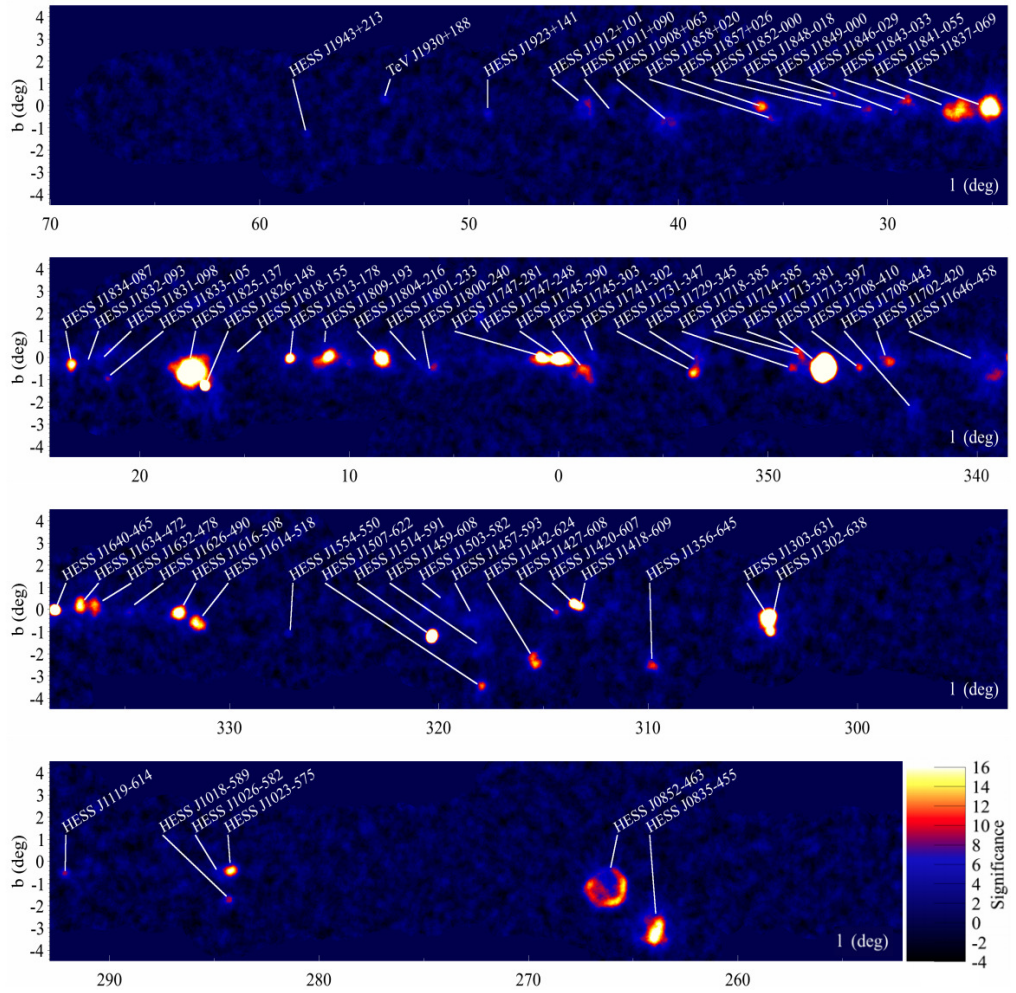

Figure 1.4 Galactic plane survey performed with H.E.S.S. Retrieved from [4]. Image credit: The H.E.S.S. collaboration (see Appendix A).

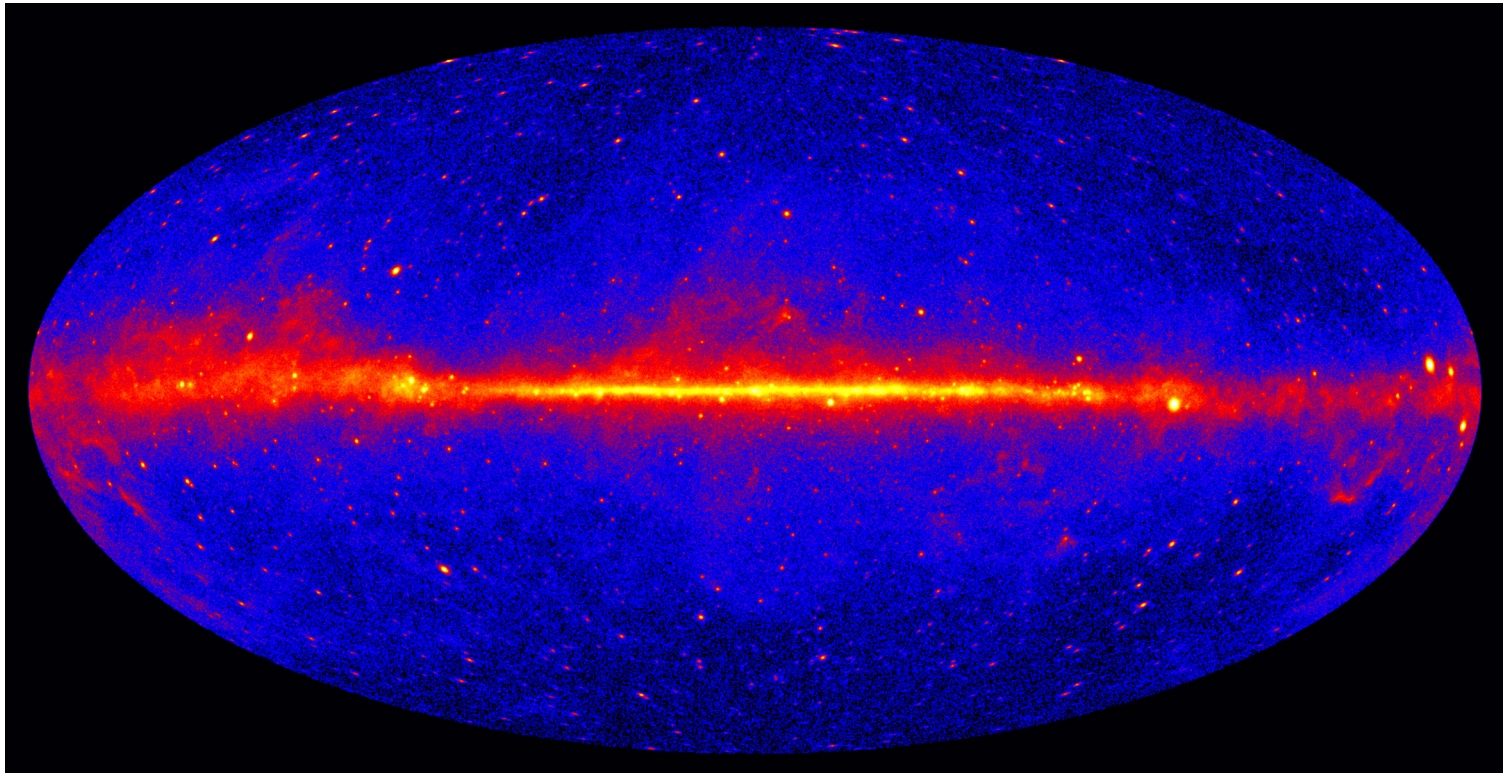

Figure 1.5 Counts Map $> 1\,\mathrm{GeV}$ with five years of *Fermi*-LAT data. Image credit: NASA (see Appendix A)

This thesis is structured as follows. Chapter 2 discusses the detection methods of gamma rays and Chapter 3 introduces the HAWC observatory. Chapter 4-6 describes the event reconstruction, calibration, and map making for HAWC. Chapter 7 and 8 describes the maximum likelihood method and presents the results and systematic studies. Chapter 9 concludes this thesis.

Chapter 2

Detection Techniques of Gamma Rays

Earth's atmosphere is opaque to gamma rays, which makes direct detection on the ground impossible. With the energy of gamma rays spreads over six orders of magnitude, it is not surprising that various types of gamma ray instruments are developed and operated. Different approaches with focus on different energy ranges are described in this chapter.

2.1 Space-Based Observatories

One straight forward solution is to launch detectors above all or most of the atmosphere, on-board the balloons or satellites. The effort of observing gamma rays with space-based telescopes began with several balloon missions and the first gamma-ray telescope, on-board the Explorer 11 satellite [16], which was launched in 1961. Followed by other satellite-based telescopes, these experiments revealed cosmic gamma-ray emission for the first time. Localized gamma-ray sources outside of our solar system were not discovered until the launch of the second Small Astronomy Satellite [17] in 1972, which surveyed the sky in the energy range between 30 MeV and 1 GeV and detected a handful of gamma-ray sources including the Crab, Vela, and Geminga pulsar. Almost a decade later in 1981, a catalog of 25 gamma-ray sources was published using the data taken with the Cos-B satellite [18]. The Energetic Gamma Ray Experiment Telescope (EGRET) on board of the Compton Gamma Ray Observatory was operated from 1991 to 2000, and sensitive to gamma rays in an energy range between 20 MeV and 30 GeV. The third EGRET catalog contains 271 new gamma-ray sources above 100 MeV [19]. The most sensitivity satellite gamma-ray experiment to date is the *Fermi* Gamma-Ray Space Telescope, which started operating in 2008 [14].

Unlike optical photons, gamma rays cannot be refracted by a lens or reflected by a

mirror, thus cannot be focused. The detection of gamma rays with space-based observatories is usually based on electron-positron pair production. A detector consists of three basic components:

1. a thin layer of plastic anti-coincidence detector, where charged cosmic rays create a flash of light, allowing the detector to reject the background particles;
2. tracking chambers that convert gamma rays into electron-positron pairs and record the trajectories of these particles, and ultimately reconstruct the direction of the primary gamma ray;
3. calorimeters that measure the total energy deposited.

The *Fermi* Gamma-Ray Space Telescope has two instruments on-board. The main instrument, the Large Area Telescope (*Fermi*-LAT), monitors the gamma-ray sky from $\sim$ 20 MeV to more than 300 GeV with excellent angular resolution ($< 0.15^\circ$ for 10 GeV), $< 10\%$ energy resolution, and high duty cycle. The large FOV of 2 sr allows a whole sky coverage every three hours. The third *Fermi*-LAT source catalog (3FGL) [14] was released in 2015 based on four years of data. The 3FGL catalog includes 3033 sources, increasing the number of gamma-ray sources by an order of magnitude compared to previous instruments. Fig 2.1 shows the positions of the sources in Galactic coordinates, with Galactic sources distributed near the Galactic plane while the AGNs uniformly distributed in the sky.

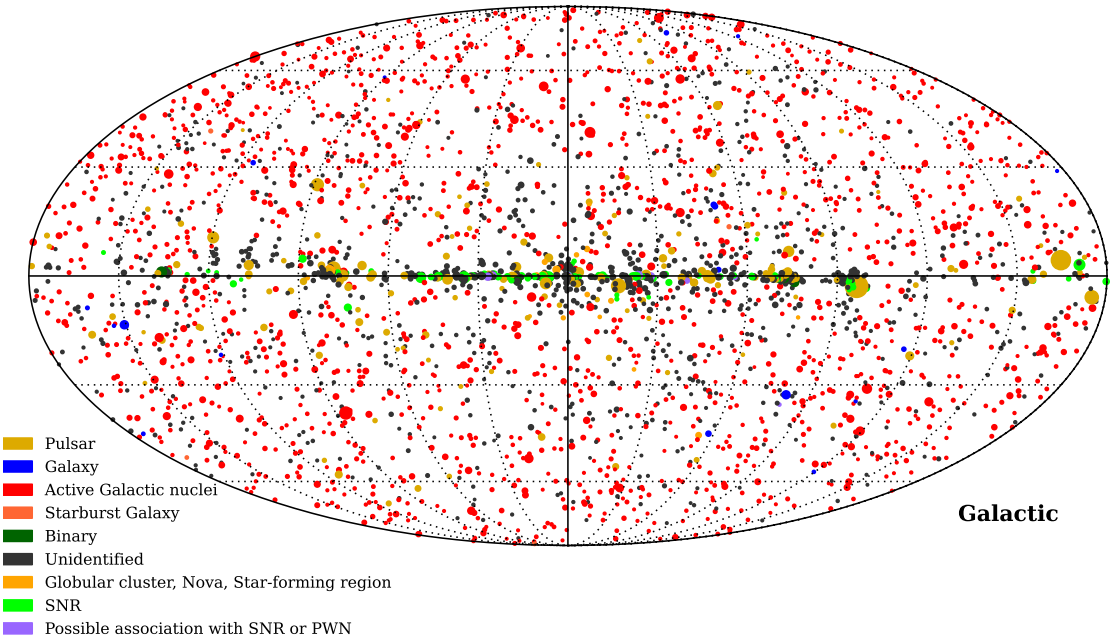

Figure 2.1 3FGL sources by source class in Galactic coordinate [14].

Space-based gamma ray telescopes have a large FOV, excellent energy resolution and angular resolution, and high duty cycle. But the size of the spacecraft limits the effective area to $\leq\sim 1\mathrm{m}^2$. Balloon borne experiments have a larger effective area of a tens of m^2 but they have very limited live time. As the gamma-ray sources follow a power-law spectrum with a spectral index of typically between 2.0 and 3.0. The number of gamma rays emitted from a source rapidly decreases with energy. This limits the energy range of space-based gamma-ray observatories from several MeV to several hundreds of GeV. To detect gamma rays with higher energies, instruments with larger effective area are needed.

2.2 Ground-Based Observatories

Given the difficulty of launching larger spacecrafts, instruments that are able to collect TeV gamma rays have to be built on the ground. As the gamma rays cannot penetrate the Earth's atmosphere, these ground-based instruments employ the Earth's atmosphere as an intrinsic part of the detection technique. It is thus important to understand how gamma rays interact with the molecules in the atmosphere.

2.2.1 Extensive Air Shower

The dominating processes in an air shower development are electron-positron pair production and bremsstrahlung. Earth's atmosphere is a very deep calorimeter with $\sim$ 30 radiation lengths above the sea level. A radiation length is defined as the distance over which an electron loses all but 1/e of its energy by bremsstrahlung, which equals to 7/9 of the mean free path of pair production [2].

A gamma-ray photon enters the atmosphere and collides with a nucleus, producing a electron-positron pair, so called the first interaction. The altitude of the first interaction varies but is approximately after the gamma ray travels a radiation length in the atmosphere, i.e. at an altitude of $\sim$ 20 km. The secondary electron and

positron share the energy of the primary gamma ray and produce a new generation of gamma rays by bremsstrahlung, which initiate more pair production. Approximately in each radiation length a particle produces two more secondary particles that share the energy of the primary gamma ray. These iterative processes create a cascade of secondary particles, so call an extensive air shower (EAS). The number of secondary particles increases with the depth the particles travel through the atmosphere, until the average energy of the secondary electrons and positrons drops to the critical energy ($\sim 80\,\mathrm{MeV}$) that the cross section for ionization losses starts to exceed that for bremsstrahlung. This is referred to as the shower max, typically at an altitude of 5-10 km for TeV gamma ray showers. As the air shower continues to propagate after the shower max, the number of secondary particles in the air shower rapidly drops since the dominating process, the ionization losses, does not produce additional air shower particles. Fig. 2.2 illustrates the simulation of the development of an air shower induced by a 1 TeV gamma-ray photon.

In a gamma-ray air shower, the secondary particles are ultra relativistic and the dominating processes are sharply peaked forward. Consequently, the air shower particles form a thin front, referred to as the shower front, only a few meters thick, while Coulomb scattering causes the lateral spread of a shower front of the order of one hundred meters. The shape of the shower front is often described as a pancake.

As the main background for gamma-ray observations, charged cosmic rays induce

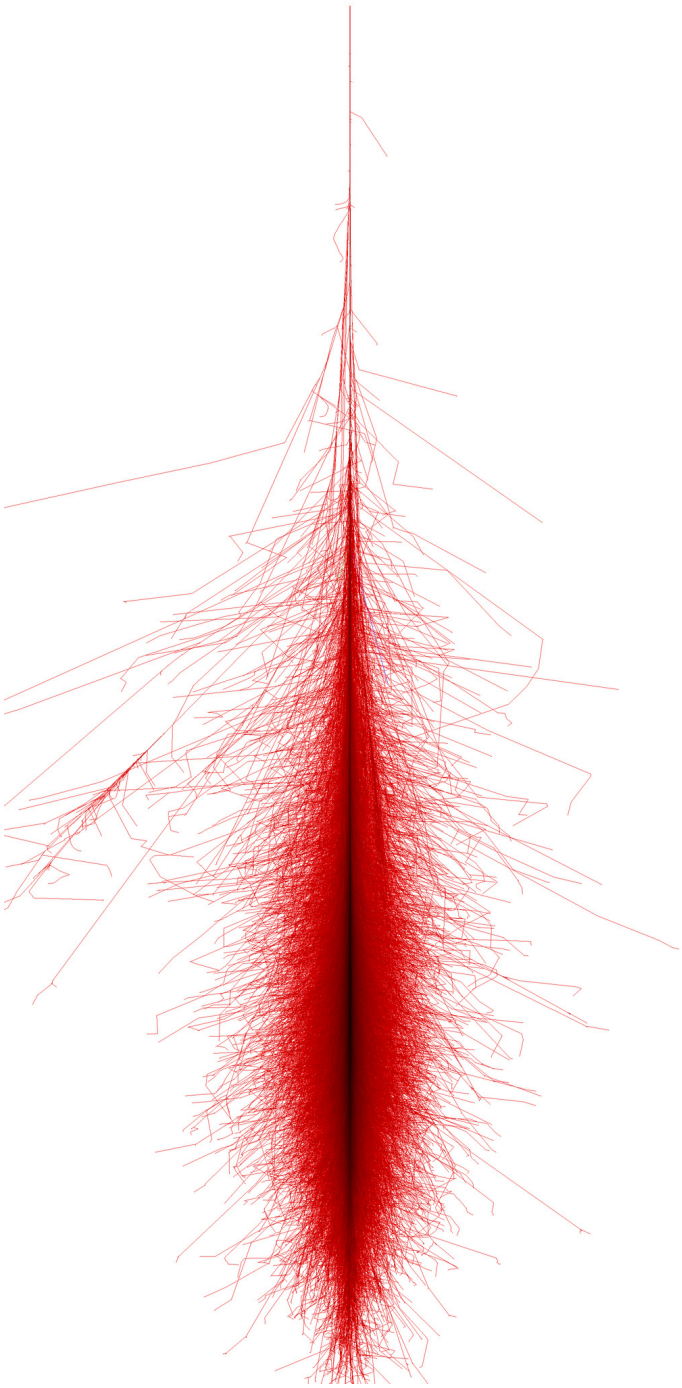

Figure 2.2 The simulation of the development of an air shower induced by a 1 TeV gamma-ray photon. Red lines represent the tracks of electromagnetic component (electrons, positrons, and gamma rays) in the air shower. Image credit: F. Schmidt and J. Knapp (see Appendix A).

similar but more complicated air showers, except electrons[1]. A proton or heavier nucleus collides with the nuclei in the atmosphere, producing pions and other nuclei. Pions have short life time and the π^0 decay into gamma rays while the $\pi^{\pm}$ decays into $\mu^{\pm}$ and neutrinos. The produced gamma rays create sub-showers of electromagnetic components, while muons often reach the ground due to their relatively longer life time. Fig. 2.3 illustrates the simulation of the development of an air shower induced by a 1 TeV proton.

Despite the similarity between gamma ray and hadronic air showers, the existence of large number of muons in hadronic air showers provides ground-based gamma ray observatories a tool to distinguish between these two types of primary particles.

Unlike the space-based instruments, ground-based instruments cannot directly detect the primary gamma rays. Instead, they directly or indirectly observe the air shower particles induced by gamma ray showers and reconstruct the features of the primary gamma rays using observables. These are two main types of ground-based gamma ray observatories: Imaging Atmospheric Cherenkov Telescopes (IACTs) and extensive air shower or particle-sampling arrays. Both types of instruments have large effective area of the order of $10^5\,\mathrm{m}^2$, which is sufficient to detect TeV gamma rays.

[1]Electrons produce similar electromagnetic air showers as gamma rays do.

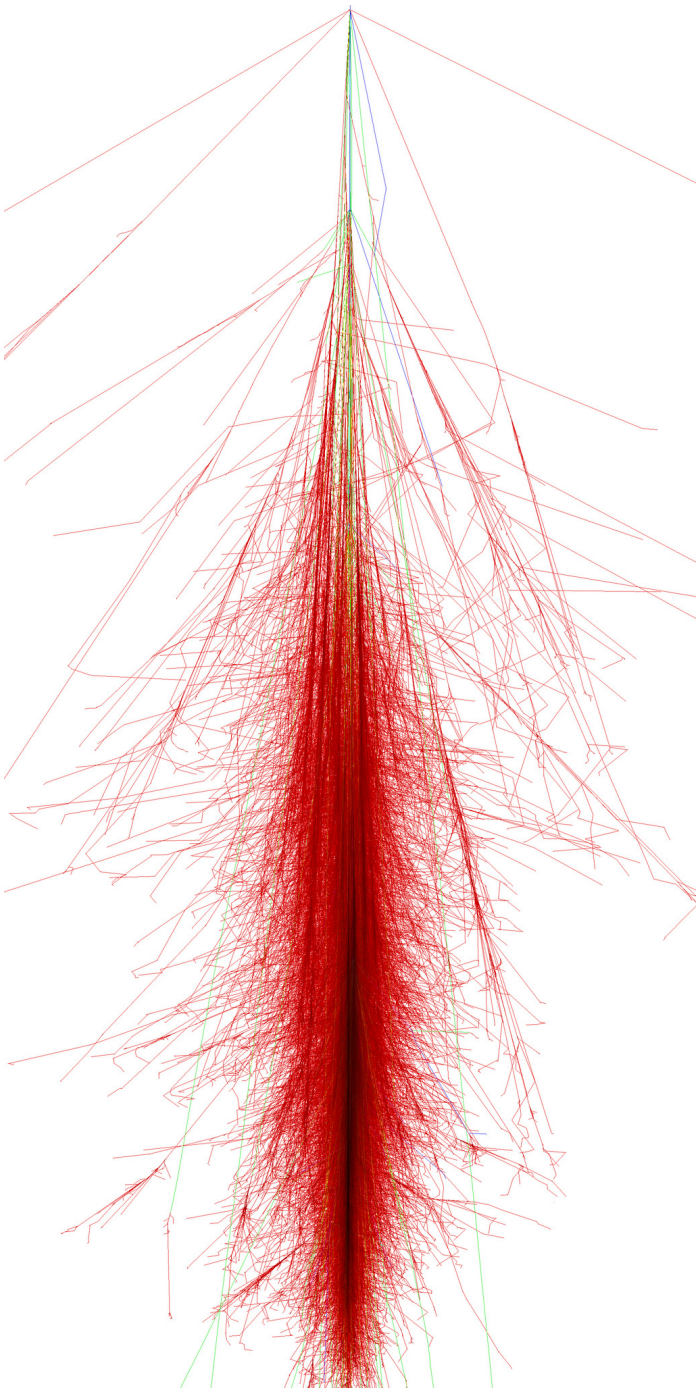

Figure 2.3 The simulation of the development of an air shower induced by a 1 TeV proton. Red, blue, and green lines represent the tracks of the electromagnetic component, hadronic component, and muons. Image credit: F. Schmidt and J. Knapp (see Appendix A).

2.2.2 Imaging Atmospheric Cherenkov Telescopes

IACTs detect the Cherenkov radiation produced by EAS particles. Cherenkov radiation is produced when a charged particle travels faster than the speed of light in the medium. The charged particle disrupts the local electromagnetic field when it travels. As the wave in the field travels slower than the particle itself, a shock wave forms as shown in Fig. 2.4. The angle between the emitted Cherenkov radiation with respect to the track of the particle θ, referred to as Cherenkov angle, is determined by the index of refraction of the medium n and the ratio of the particle speed to the speed of light in vacuum, $\beta = v/c$:

$$cos\,\theta = \frac{1}{n\beta} \tag{2.1}$$

As the charged EAS particles are ultra relativistic, the speed of these particles is close to the speed of light in vacuum, thus $\beta \approx 1$. The refractive index in the atmosphere is 1.00029 at the sea level and decreases with altitude. The Cherenkov angle in air is $\sim 1^{\circ}$. Given the shower max at a few km in altitude, the Cherenkov light cone on the ground is typically of order of one hundred meters in radius.

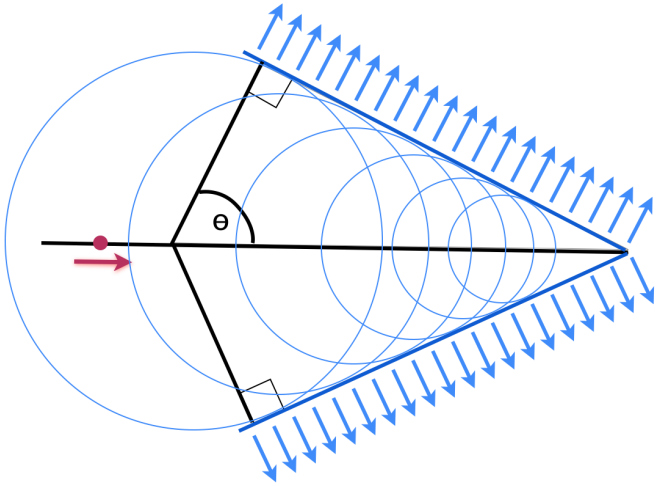

Figure 2.4 A diagram illustrating Cherenkov radiation.

Sampling atmospheric Cherenkov telescopes are the earliest attempts using this technique. They often were based on solar farms covering a large area with each mirror viewed by one photomultiplier tubes (PMT) [20, 21]. Since the first solid detection of the Crab Nebula at TeV energies by Whipple [22] in 1991, modern IACTs, including VERITAS [23], H.E.S.S. [24], and MAGIC [25] have discovered more than one hundred TeV gamma ray sources. Fig. 2.5 shows a picture of the H.E.S.S. gamma-ray observatory located in Namibia. Modern IACTs employ an array of telescopes with $\geqslant 10\,\mathrm{m}$ in diameter. An array of PMTs on the local plane of each telescope is used to record the image of the Cherenkov light produced by EAS particles. This stereoscopic approach provides excellent background rejection as well as angular and energy resolution.

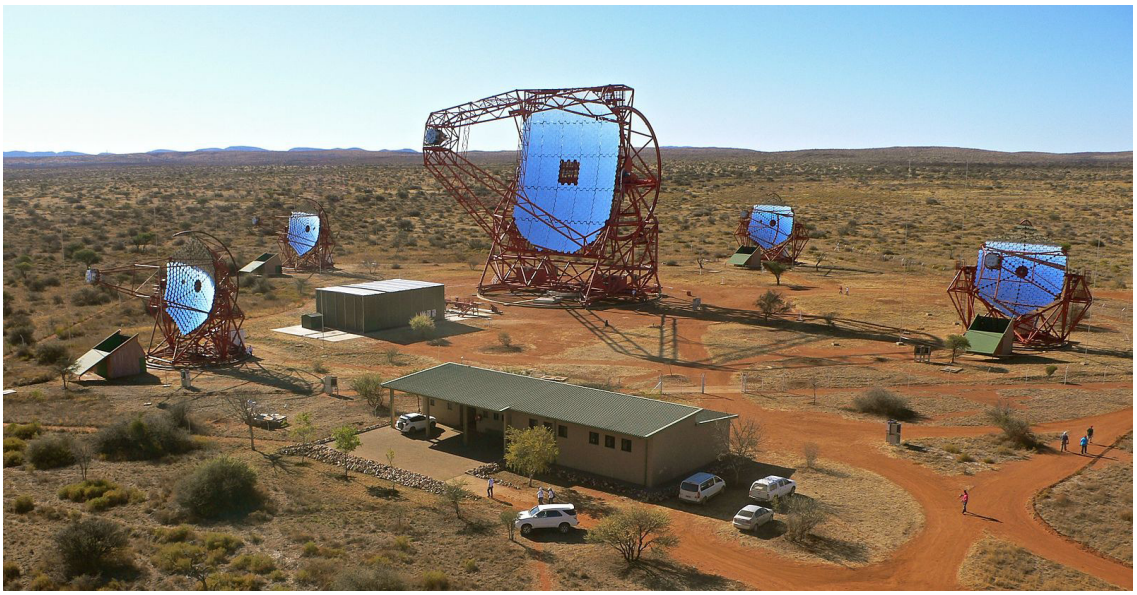

Figure 2.5 The H.E.S.S. gamma-ray observatory located in Namibia. Image credit: S. Klepser (see Appendix A).

IACTs cover the energy range from $< 100\,\mathrm{GeV}$ to a few tens TeV with great sensitivity. They have a low energy threshold and good energy resolution down to $< 20\%$ because they can detect the Cherenkov radiation that is produced in the earlier stage of an air shower thus can provide a complete picture of the longitudinal shower development. As pointing instruments, the angular resolution is better than 0.1° while the field of view (FOV), typically a few degrees in diameter, is relatively small. In addition, the observations can be made only during clear and mostly moon-less nights, which limits the duty cycle to $\sim 10\%$.

2.2.3 EAS or Particle-Sampling Arrays

Different from IACTs, EAS or particle-sampling arrays only observe the footprint of an EAS at the detector level. EAS arrays, e.g. ARGO-YBJ [10] and Tibet AS-gamma [26], use scintillation counters or resistive plate counters to detect charged EAS particles that make to the ground level. As the secondary gamma rays outnumber electrons and positrons by a factor of ~ 5, layers of lead are often used to convert gamma rays in an EAS to electrons and positions that can be detected by particle counters. An array with large area is required to cover the footprint of an EAS. However, EAS arrays are designed as sampling instruments with a relatively small fraction ($\sim 1\%$) of area covered by particle counters. As a result, the sensitivity to lower energy gamma rays is limited. Fig. 2.6 shows a picture of the Tibet AS-gamma experiment. The Tibet AS-gamma experiment is located in Tibet, China at an altitude of 4,300 m above sea level. High altitudes are preferred for EAS arrays since the detector level is closer to the shower max, where more EAS particles are detectable, which led to a better sensitivity and a lower energy threshold.

The timing and energy of EAS particles detected by each particle counter are recorded. The direction of the primary gamma ray is reconstructed using the timing information as the shower plane triggers each detector at a different time. The measured energy of EAS particles are used to reconstruct the energy of the primary gamma ray. The

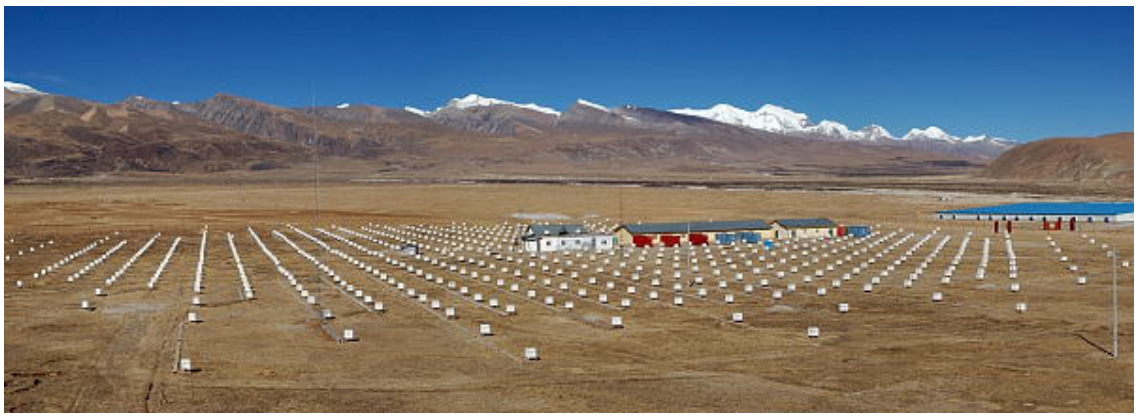

Figure 2.6 The Tibet AS-gamma experiment located in Tibet, China. Courtesy of Institute for Cosmic Ray Research, the University of Tokyo, Japan (see Appendix A).

fluctuations in the EAS development limit the angular and energy resolution of EAS arrays.

As survey-type instruments, EAS arrays monitor the entire overhead sky at the same time, which led to a very large FOV of $\sim 2\,\mathrm{sr}$. And they can operate during any weather condition and any time of day, having a nearly 100% duty cycle. EAS arrays are good complements of IACTs with large FOV, high duty cycle, and highest energy reach.

2.2.4 Water Cherenkov Experiments

Another type of EAS particle-sampling instruments are water Cherenkov detectors. Water Cherenkov experiments also benefit from high altitude. The Milagro gamma-ray detector near Los Alamos, New Mexico, is sensitive to gamma rays between

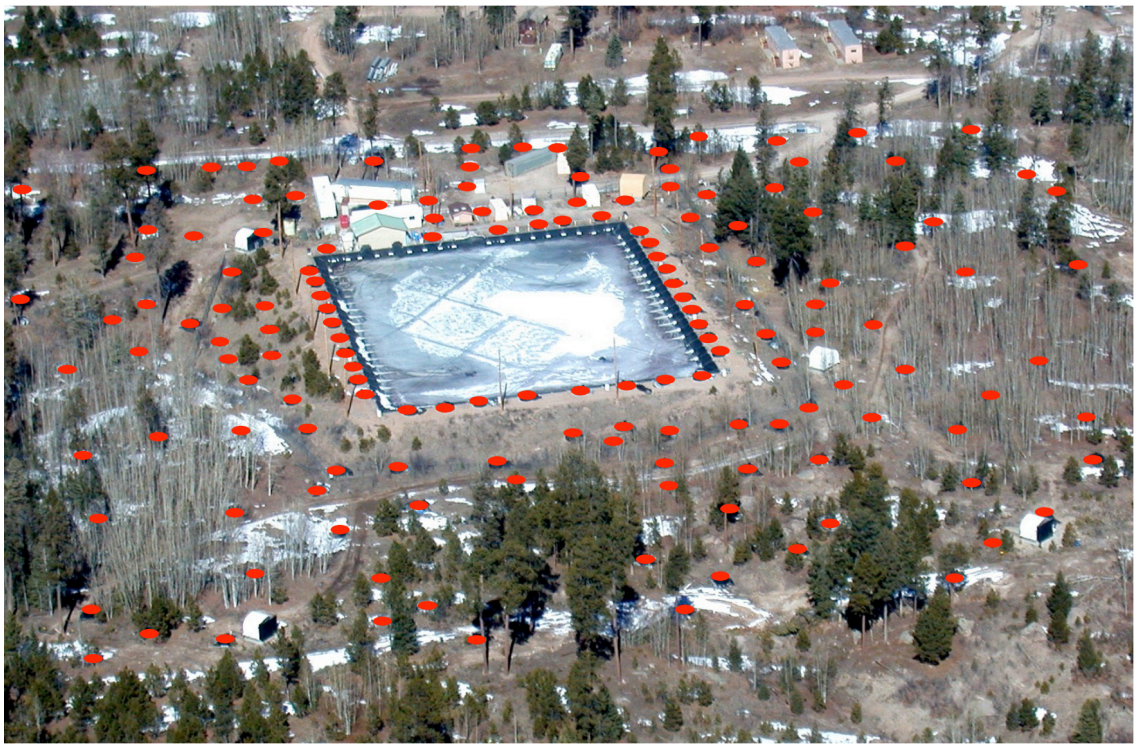

Figure 2.7 The Milagro detector with the central pond and surrounding outriggers marked with red. Image credit: The Milagro collaboration.

100 GeV and 100 TeV. Operated between 2000-2008, it demonstrated the sensitivity of this technique with the first detection of the galactic diffuse emission at TeV energies [27] and of several new TeV gamma-ray sources [9]. Fig. 2.7 shows a picture of the Milagro detector, which consists of a central pond and outrigger tanks. Two layers of PMTs are deployed in the central pond. The top layer of 450 PMTs is used for event triggering and reconstruction, while the bottom layer of 273 PMTs is used for background rejection and energy estimation. 175 outrigger tanks with one PMT in each increase the physical area of the detector by a factor of 8.

The Cherenkov angle in water is 41°. The large Cherenkov cone ensures that every

charged EAS particle that enters the water in the pond should be observed by at least one PMT. The secondary gamma rays are detectable as well since they are converted to electrons and positrons as they enter the water. The high detection efficiency of EAS particles increases the sensitivity especially on lower energy (sub-TeV) gamma rays.

The HAWC gamma-ray observatory is a second-generation water Cherenkov detector that is build on the experience of Milagro. More detailed descriptions about the HAWC detector and its performance will be given in the next Chapter.

Chapter 3

The HAWC Gamma-Ray Observatory

The High Altitude Water Cherenkov observatory, or short HAWC, is a ground-based gamma-ray detector that utilizes the water Cherenkov technique. It is located at Sierra Negra, Mexico (18°59'41" N 97°18'30.6"W) at 4,100 m above sea level, and is sensitive to gamma rays and cosmic rays between 100 GeV and 100 TeV [28]. The array consists of 300 water Cherenkov detectors (WCDs) over an area of 20,000 m^2 as shown in Fig. 3.1. HAWC has a large instantaneous field of view of 2 sr and a high duty cycle of $> 95\%$. With this large field of view, HAWC monitors two-thirds of the sky each day, covering the sky between -26° and 64° in declination. The science operation began in August 2013 with one-third of the array and the complete array

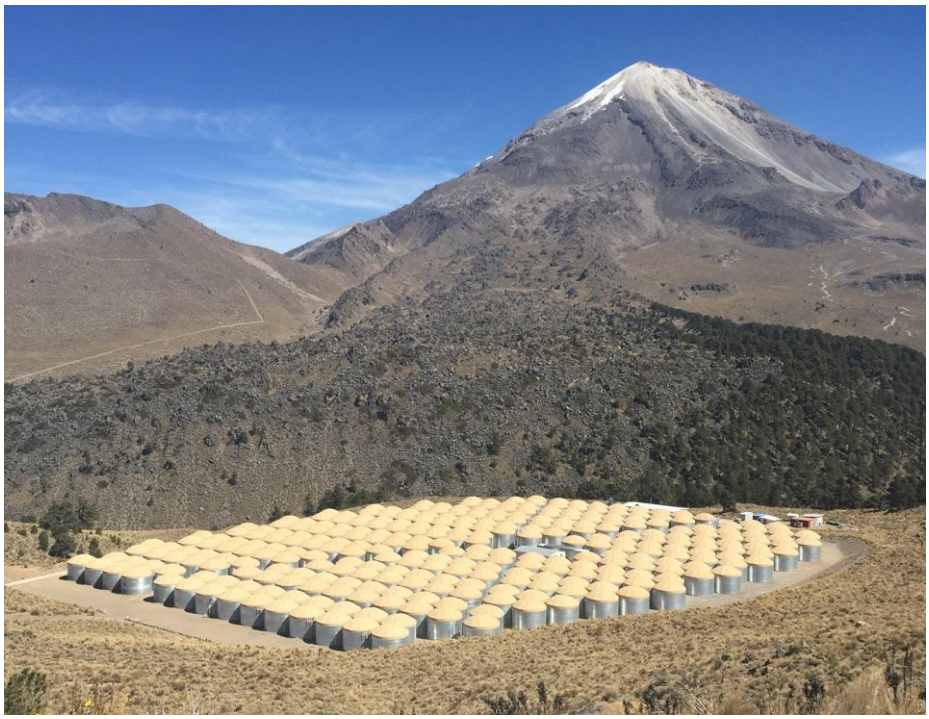

Figure 3.1 The complete HAWC array with Pico de Orizaba in the background. Image credit: A. Carramiñana (see Appendix A).

of HAWC was inaugurated in March 2015.

This chapter discusses the technical details and expected performance of HAWC, and describes the data sets and simulations that are used in this thesis.

3.1 Location

The HAWC site is located on a relatively flat piece of land in the saddle region between the Sierra Negra (4580 m above sea level) and the Pico de Orizaba (5745 m above sea level). The altitude of the site corresponds to an atmospheric depth of $630\,\mathrm{g\,cm^{-2}}$, or

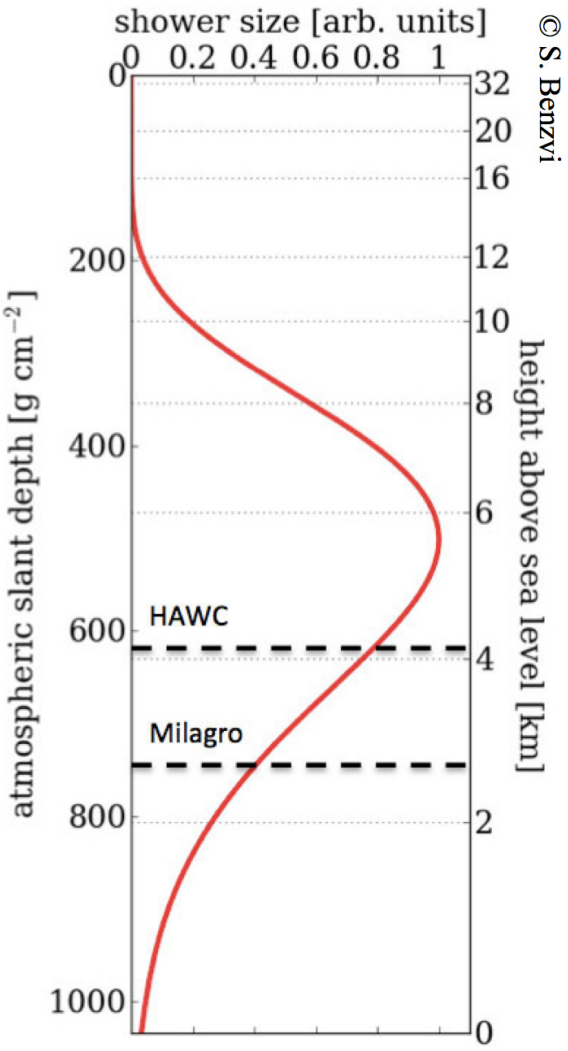

Figure 3.2 The shower size of a typical TeV gamma-ray air shower as a function of height above sea level. Dashed lines indicate the altitude of HAWC and Milagro. Image credit: S. BenZvi (see Appendix A).

approximately 17 radiation lengths [29] (see Fig. 3.2). Compared to its predecessor, the Milagro gamma-ray observatory, which was located at 2630 m above sea level [9], HAWC collects air shower particles closer to its shower max (5-10 km above sea level for a typical TeV gamma-ray air shower).

Located inside the Parque Nacional Pico de Orizaba, a Mexican national park, HAWC

is not the only scientific facility in this area. The Large Millimeter Telescope (LMT) is located on the summit of Sierra Negra and has a single dish with a diameter of 50 m. The LMT was inaugurated in 2006. HAWC has benefited from the presence of the LMT by using the infrastructure that was put in place during the construction of LMT, such as roads, transmission lines, and network lines.

3.2 WCDs and Photomultiplier Tubes

Fig. 3.3 shows a graphic and a picture of an installed WCD. Each WCD is a corrugated steel tank of 7.3 m in diameter and 5 m in height. A water-tight plastic bladder inside of each tank contains approximately 200,000 L of purified water, corresponding to a water level of 4.5 m. WCDs detect secondary air shower particles through Cherenkov radiation that is produced in a WCD as the relativistic secondary particles travel through the water. Both secondary gamma rays and charged particles in an air shower can be detected by WCDs. The latter directly emits Cherenkov radiation while gamma rays produce electron and positron pairs once they enter the water and emit Cherenkov radiation as described in the previous chapter.

Four photomultiplier tubes (PMTs) are attached to the bottom of each WCD to detect the Cherenkov radiation inside the water: one high-quantum efficiency 10-inch Hamamatsu R7081-MOD PMT at the center and three 8-inch Hamamatsu R5912

Figure 3.3 Left: A sketch of a HAWC WCD. Right: A picture of an installed WCD. PMTs [30], spaced 120° apart at 1.8 m from the center. The 900 8-inch Hamamatsu PMTs are from the Milagro experiment and refurbished for use in HAWC.

3.3 Electronics and Data Acquisition

More than 600 feet of RG59 cable connects each PMT through a spark gap to the data acquisition system (DAQ) in the counting house located at the center of the array. The cables deliver high voltage to the PMTs and carry PMT signals back to the DAQ. The PMT signals are processed by custom-made front end boards. The signals are shaped and discriminated by two voltage threshold at 20 mV and 50 mV. The times at which a pulse crosses each threshold, so call edges, are recorded by CAEN VX1190A time-to-digital-converter (TDC) modules with a precision of 100 pico seconds (ps). The signal size and ultimately the charge, i.e. the number of photoelectrons (PEs),

Figure 3.4 The electronics racks inside the counting house.

can be inferred from the time-over-threshold (ToT), as will be described in Chapter 5 on the detector calibration. Fig 3.4 shows the electronic racks inside of the counting house.

HAWC utilizes two different DAQs. The main DAQ, utilizing TDCs, reads out PMT signals within a $2\,\mu$s time window once a trigger condition is met. A simple multiplicity trigger is applied that counts the number of PMT signals in a sliding trigger window. The trigger threshold varies for different size of the detector during different construction stages to maintain a data rate of $\sim 20\,\mathrm{MB\,s^{-1}}$. For the full HAWC array, the trigger threshold is set to 28 PMTs within a sliding window of 100 ns. The digitized PMT signals are recorded for reconstruction of air shower events. The second DAQ, the scaler system, counts the number of signals in each PMT within a 10 ms

window. It is sensitive to transient events, such as gamma ray bursts, which can be detected by a statistically significant excess over the noise level [31]. The analysis presented in this thesis only uses data that are collected with the first, the TDC DAQ system.

3.4 The Performance of HAWC

The performance of the HAWC observatory is determined by the angular resolution, effective area, and the ability to reject cosmic ray background [2]. These characteristics improve with the size of the air shower on the ground, as larger air showers, which contain more secondary air shower particles, are better measured and more precisely reconstructed. HAWC is located at a higher altitude than Milagro, where more secondary air shower particles are detectable at the detector level. This section presents the expected performance of the HAWC detector derived from simulations.

Timing information from more secondary air shower particles, along with a larger area of sensitive detector compared to the central pond of Milagro, provides an angular resolution unprecedented for a survey-type air shower array. Fig. 3.5 shows the angular resolution of HAWC compared to Milagro. The angular resolution of HAWC (40% containment) is reaching $\sim 0.1^{\circ}$ at above 10 TeV.

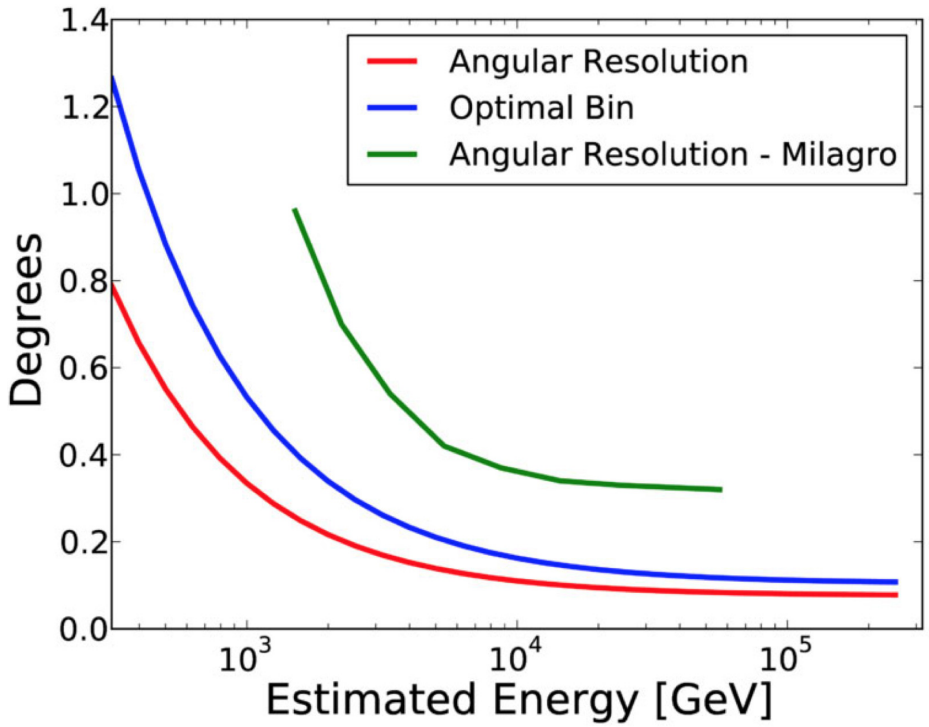

Figure 3.5 The angular resolution (40% containment, red) and the optimal bin (70% containment, blue) of HAWC as a function of primary gamma-ray energy, compared to the angular resolution (40% containment, green) of Milagro [28].

More detectable secondary air shower particles also leads to a larger effective area, in particular for low energy air showers, which have too few particles survived at the detector level of Milagro and yet trigger the detector. Fig. 3.6 shows that HAWC has a similar effective area at higher energies but much a larger effective area in lower energies than Milagro. The dashed lines indicate the effective area after background rejection, which slightly reduces the effective area of gamma rays but greatly improves the signal-to-background ratio.

As mentioned above, the hadronic cosmic rays are the main background for the analysis of gamma-ray emissions. The number of muons, which are a signature of hadronic

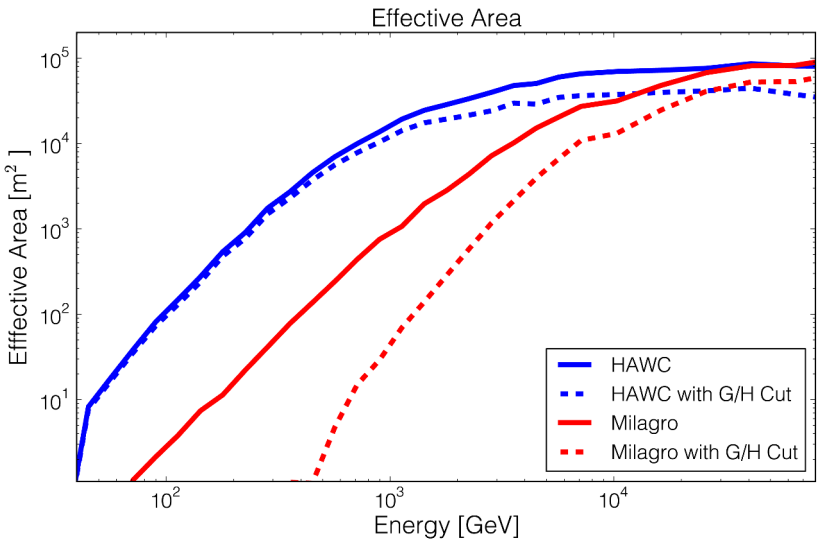

Figure 3.6 The effective area of HAWC and Milagro without (solid lines) and with (dashed lines) background rejection as a function of primary gamma-ray energy [28].

cosmic rays, increases with the primary cosmic ray energy, and so the ability to reject cosmic ray background improves with energy. HAWC, with a larger muon detection area, has a better background rejection ability than Milagro, as shown in Fig. 3.7. A ten-to-one signal-to-background ratio is anticipated at above 10 TeV for a source with a Crab-like spectrum.

3.5 HAWC Sensitivity to Steady Sources

Fig 3.8 shows the sensitivity of HAWC to known TeV sources derived from detector simulations with the known intensities measured by other instruments. The sensitivity of HAWC is a function of source declination and the grey shaded bands of different intensity indicate the sensitivity of HAWC with 1-year of data [28]. The

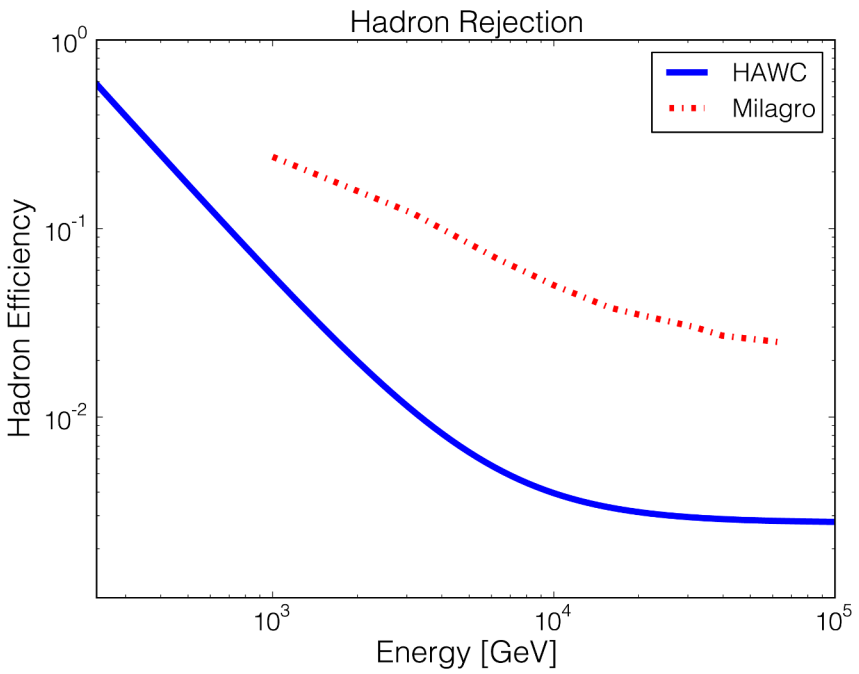

Figure 3.7 The background efficiency of HAWC and Milagro as a function of primary particle energy.

best sensitivity is achieved for the sources that transit through the zenith, which is at 19° N for the HAWC site, and decreases for those at larger zenith angle as the gamma rays travel through thicker atmosphere.

3.6 Science Goals

Supernova remnants have been postulated as the cosmic ray acceleration sites since they can provide sufficient power to explain the cosmic ray flux as observed from the Earth. HAWC has unprecedented sensitivity above 10 TeV as shown in Fig. 3.9. The observations in this energy range is the key to distinguish the gamma ray that are

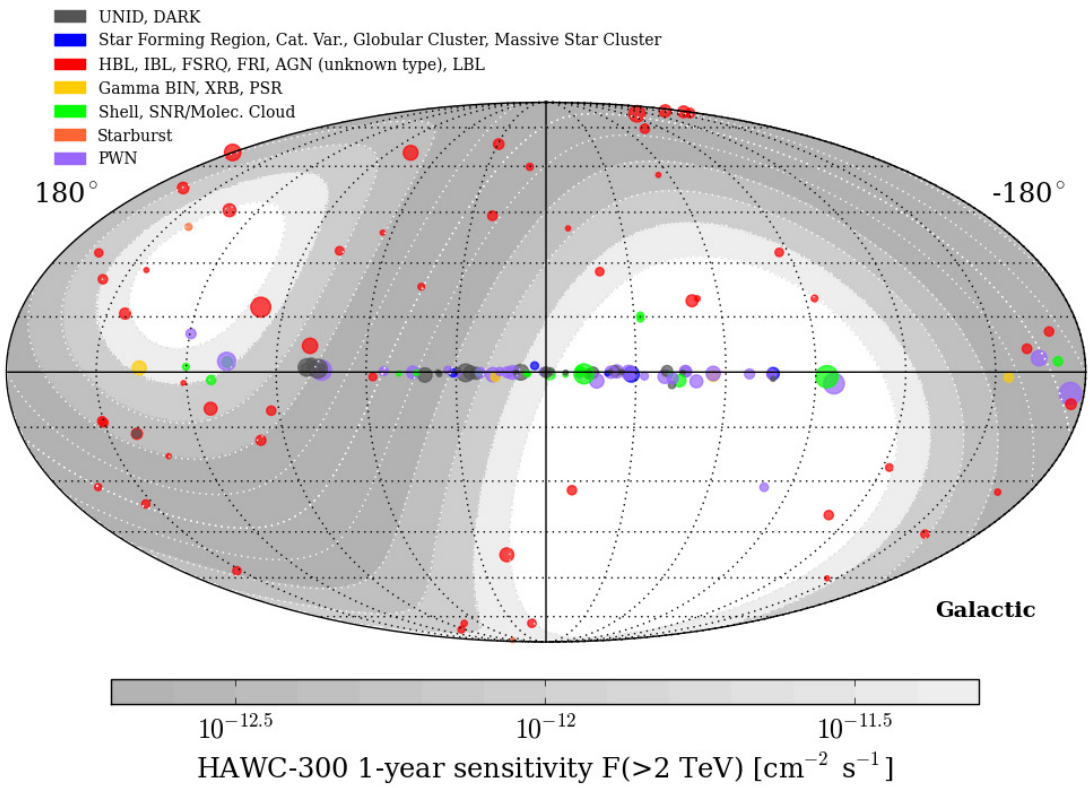

Figure 3.8 The sensitivity map of HAWC in Galactic coordinates [28]. Dots mark the known TeV sources from the TeV Catalog [32]. The color and size represent the type and strength of a source.

produced by electrons from those are produced by hadrons. The spectrum in this energy range provides unique information on the acceleration sites and the highest energies that particles are accelerated to.

With its this large FOV, HAWC monitors two-thirds of the sky each day, which makes it a perfect instrument to perform an unbiased survey of the TeV gamma-ray sky and to look for new gamma-ray sources. The survey of the inner Galaxy region will be described in detail in this thesis.

Consisting of light-tight WCDs, HAWC can operate during both daytime and night

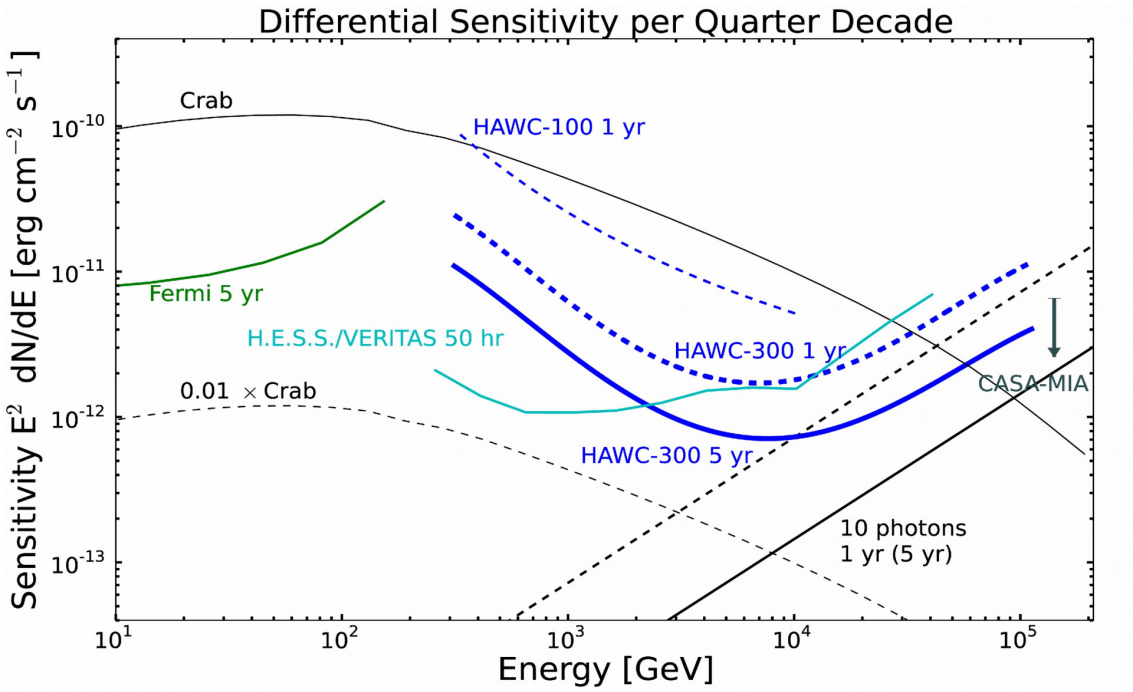

Figure 3.9 The differential sensitivity of HAWC [28].

time. With its high duty cycle, HAWC continuously monitors the sky for transient sources, including active Galaxies, gamma-ray bursts, and Galactic transients. HAWC can alert IACTs of TeV outbursts, allowing follow-up observations performed by IACTs for these time-sensitive objects. The search for TeV gamma-rays from the bright GRB 130427A has been recently published by HAWC [31].

Previous observations reveal a number of extended structures in the gamma-ray sky, including the extended sources from the Cygnus region [33], the emission near the Geminga pulsar [34], and the Fermi bubbles [35]. IACTs have difficulties observing the sources when their sizes are comparable or larger than the FOV of IACTs. Thanks to its large FOV, HAWC has good sensitivity on these extended structures.

HAWC is not only a gamma-ray observatory, but also observing the hadronic cosmic-rays at the same time. HAWC perform searches for small-scale and large-scale cosmic-ray anisotropy in the northern sky. The results in early HAWC data has been summarized in [36].

This thesis focuses on a search for point-like emission from the Galactic plane, that could come from objects such as PWNe, SNRs, or binary objects.

3.7 Detector Construction

The construction of HAWC began with extending the road, power lines, and network fiber from the present LMT infrastructure in 2010. Between 2010 and 2011, a prototype array called Verification And Monitoring Of Systems (VAMOS) was constructed in a few tens meter northwest from the HAWC site and was operational for eight months between October 2011 and May 2012 [37]. In the mean time, the HAWC platform was being prepared and ready for WCD installation in December 2011. The first WCD was installed in March 2012 and by September 2012 an array of 30 WCDs (HAWC-30), which is the northeast corner of the HAWC array, was deployed and verified. The last WCD was installed in December 2014 and the complete array was inaugurated in March 2015. Fig 3.10 shows the HAWC site at different time in the past few years.

Figure 3.10 Google Earth images of the HAWC site during 2006-2014. The first image shows the HAWC site before construction, while the last image shows the site after construction finished. Map data: Google, DigitalGlobe (see Appendix A).

3.8 Data Sets

Thanks to its modular design, HAWC was taking data during the construction phase. Data taking started with HAWC-30, with more and more WCDs integrated into this growing array. Science operation started in August 2013 with about one-third of

the array (HAWC-111). By November 2014, about 250 WCDs started taking data, with new WCDs channels added into the data stream until March 2015, when the full array started operating. Different sizes of operational detector lead to different sensitivity. The update of electronics and use of different calibrations can also change the property of the detector. To take into these variations, the data are split into several data sets:

1. HAWC-111 data, taken with about 1/3 of the full array between August 2, 2013 and July 9, 2014. The live time of this data set is $\sim$ 283 days. During this period, data taking was occasionally interrupted for detector construction work and maintenance, and the duty cycle is $\sim$ 84%. After taking into account the geographical effect when a region of the sky was blocked by Earth, the exposure time for the inner Galaxy region is 275 ± 1 days. The number of operational WCDs increased from 108 to 134 during this period. This HAWC-111 data set is sub-divided into three epochs based on different operational tanks and use of calibrations. This is the most well understood data set at the time this thesis is written. Detailed systematic studies were performed and are discussed in Chapter 7. The analysis and results presented in this thesis are based on HAWC-111 data unless otherwise noted.

2. HAWC-250 data, taken with the almost-full array with 247 to 293 WCDs between November 26, 2014 and May 6, 2015 with 150 days of live time. This

is a preliminary data set. In Chapter 8, the analysis results are presented as a systematic check of the results derived with HAWC-111 data. A brief discussion of the spectral analysis with more sensitive HAWC-250 data with improved reconstruction is also given in Chapter 8.

3.9 Monte Carlo Simulations

To convert the gamma ray count measured by HAWC to a source flux, the response of the HAWC detector needs to be fully understood. Different sets of simulations are generated to properly model the detector for different data sets. Monte Carlo simulations are generated with the following two steps.

3.9.1 Simulations of Extensive Air Showers

The air shower simulation program CORSIKA (COsmic-Ray Simulations for KAskade) [38] is used to simulate the development of air showers in the atmosphere. It simulates the interactions between particles and tracks the productions. Air shower simulation is computationally intensive and the computation time is approximately proportional to the energy of the primary particle. For the secondary particles under

a certain energy, a statistical sampling method is used instead of tracking each particle. This reduces computation time and makes simulation of high energy air showers practically feasible.

Air showers induced by different types of primary particles are simulated to the detector level of HAWC, including gammas, protons, and heavier nucleus with hundreds of millions of air showers for each species. The particles are thrown with 0° to 75° from zenith. The gammas and protons are simulated in an energy range of 5 GeV to 500 TeV and 5 GeV to 2 PeV, respectively, while other heavier nuclei have higher starting energy. The particles are thrown with a power law spectrum with a spectral index of 2. A spectrum harder than typical astrophysical spectra is used to have enough statistic on high energy particles. The particles are then properly weighted to follow different spectra that different analyses require.

3.9.2 Simulations of the Detector

The detector simulation package GEANT4 (GEometry ANd Tracking) [39] is used to simulate the detector response to the air shower particles that are simulated by CORSIKA. The core positions, i.e. where the extrapolated trajectory of the primary particle intersects the detector level, of the simulated air showers are shifted from 0 to 1000 m with respect to the center of the HAWC array. GEANT4 simulates

the interaction of the air showers particles in the HAWC WCDs, the production of Cherenkov radiation, and finally the time and charge detected by each PMT. The simulations have the same format as the data so the same reconstruction software can be used.

Chapter 4

Air Shower Event Reconstruction

Once an air shower triggers the detector, the data, as a series of edges are recorded by the TDC. The data are passed on to the reconstruction chain, which includes a so-called edge finder, core reconstruction, and angular reconstruction, to determine the core position and direction of the primary particles. There are two types of reconstructions employing similar algorithms. The online reconstruction is performed with the computers at the HAWC site using preliminary calibration values. This reconstruction provides instant and first results for time-critical analyses, for example to alert other telescopes so that they can perform follow-up observations. The offline reconstruction is performed with finalized calibration values and other improvements on two computer clusters located at University of Maryland and National Autonomous University of Mexico. HAWC analyses intended for journal publication, like the one

presented in this thesis, are based on results from the offline reconstruction.

In the following the different reconstruction stages are described in detail.

4.1 Edge Finding and Hit Selection

A PMT hit is the base unit in the reconstruction of an air shower. It contains the information of the position, charge, and time of a signal in a PMT that detects Cherenkov radiation from an air shower. The determination of the charge and time of a hit is described in detail in Chapter 5 about the calibration of the observatory. The analog PMT pulses are digitized by the custom-made front end boards (FEBs), and stored as continuous data stream of leading/trailing edges, i.e. the time stamp that is recorded when a PMT signal crosses a voltage threshold. A small pulse that only crosses the low threshold creates a "2-edge hit" (see Fig. 4.1), while a large pulse that crosses both low and high thresholds will create a "4-edge hit" (see Fig. 4.2). The purpose of the edge finding algorithm is to identify individual PMT pulses in this continuous data stream.

A "2-edge hit" occurs when a PMT pulse only crosses the low threshold. The FEB digitizes the pulse into a single square pulse (a leading edge followed by a trailing edge) (see Fig. 4.1) and the times of the edges t_0 and t_1 are stamped and recored

by the TDC. The width of the square pulse ($t_{01} = t_1 - t_0$) is the time over the low threshold (low ToT). The hardware of the FEB ensures that the low ToT recorded by the TDC is large or equal to ~ 53 ns. If a low ToT is recorded with less than ~ 53 ns, the trailing edge will be pushed further to ensure a minimum low ToT of ~ 53 ns.

There is probability that two small PMT pulses in the continuous data stream mimic a longer pulse (with a long low ToT). Thus the pair of small pulses may be misidentified as a large PMT pulse. An additional high threshold is introduced to avoid this ambiguity. A true large pulse crosses both the low and the high threshold and is digitized as two square pulses of opposite polarity (see Fig. 4.2). Four edges: leading t_0, trailing t_1, leading t_2, and trailing t_3 are recored by the TDC and $t_3 - t_0$ and $t_2 - t_1$ represent the time over threshold for the low and the high threshold (low ToT and high ToT).

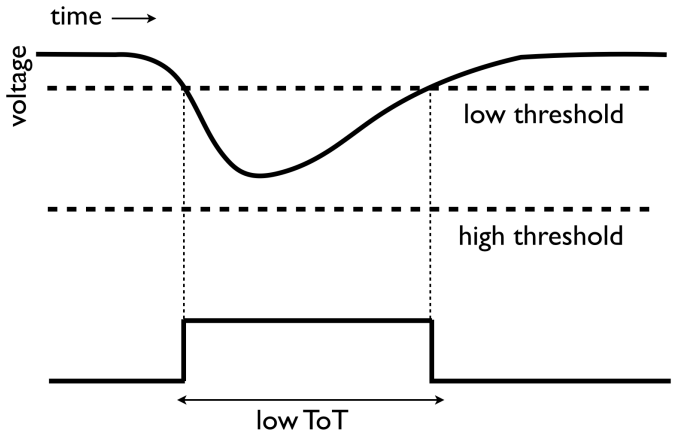

Figure 4.1 Schematic of a smaller analog pulse and corresponding edges recorded by TDCs as a 2-edge hit.

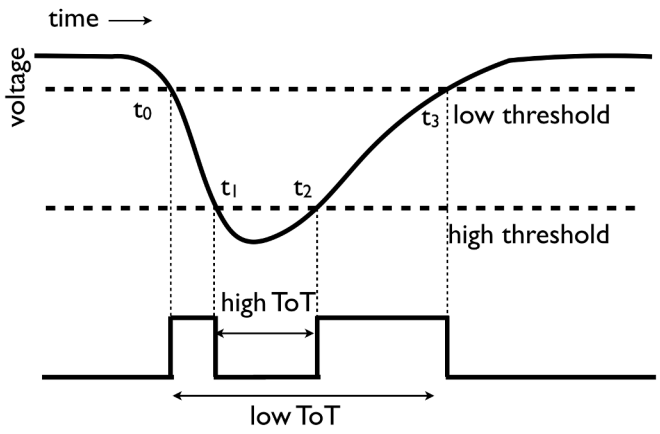

Figure 4.2 Schematic of a smaller analog pulse and corresponding edges recorded by TDCs as a 4-edge hit.

An edge-finding algorithm is developed to identify and to reconstruct 2-edge and 4-edge hits from the data stream of edges. The rise time of a 4-edge hit t_{01} decreases with a larger pulse size. For most 4-edge hits, $t_1 - t_0$ is less than 53 ns, which is the minimum value of low ToT for a 2-edge edge. We use t_{01} as the discriminator of 2-edge and 4-edge hits. If $t_1 - t_0$ is less than 53 ns, we look for the following leading-trailing edge pair and register these two square pulses as a 4-edge hit. Otherwise a 2-edge hit is registered.

Occasionally, very long 2-edge hits are created by the combination of two small pulses. These hits are marked with a "BAD" flag. In more rare cases, a prompt afterpulse[1] are followed by a large PMT pulse and the voltage level rises above the high threshold again before the original pulse drops back below the low threshold, producing a 6-edge hit as shown in Fig. 4.3. It is hard to accurately determine the charge from ToT for

[1]Afterpulses are spurious pulses that appear in the wake of true pulses.

6-edge hits. An "AMBIGUOUS" flag is set for this type of hits.

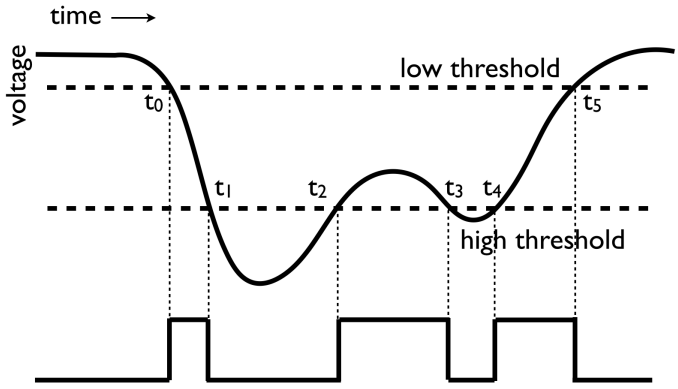

Figure 4.3 The analog pulse and edges as recorded by TDCs for a 6-edge hit.

The edges that are recorded by the TDCs are re-grouped as hits. These hits in an air shower event are filtered by more stricter timing and charge requirements (e.g. > 1 PE within 450 ns time window), referred to as standard selection cuts, and used for the reconstruction of the core position and cosmic and gamma-ray direction. Hits with "BAD" or "AMBIGUOUS" flag are removed.

4.2 Core Reconstruction

We use the charge detected by each PMT to reconstruct the core position of an air shower. Two core finding algorithms are applied. Firstly, a fast and simple Center-Of-Mass (COM) core finder provides a first guess of the core position. It calculates the centroid location of the detected charge of an air shower:

$$x_{COM},\, y_{COM} = \frac{\sum x_i M_{PEi}}{\sum M_{PEi}},\, \frac{\sum y_i M_{PEi}}{\sum M_{PEi}} \tag{4.1}$$

where M_{PEi} is the charge measured by the PMT at position x_i and y_i.

The results of the COM core finder are used as the initial seeds for a more accurate and time-consuming core finder that applies a χ^2 minimization to fit the lateral charge distribution with the Nishimura-Kamata-Greisen (NKG) function. The NKG function [40, 41] describes the lateral distribution of charged particles produced by an electromagnetic air shower.

$$\rho(r, s, N) = \frac{N}{2\pi R_{Mol}^2} \frac{\Gamma(4.5 - s)}{\Gamma(s)\Gamma(4.5 - 2s)} (\frac{r}{R_{Mol}})^{s-2} (1 + \frac{r}{R_{Mol}})^{s-4.5} \tag{4.2}$$

where $\rho(r, s, N)$ is the particle density as a function of r - the radial distance from the shower axis, s - the shower age, and N - the amplitude, i.e. the total number of secondary particles in the shower, which is a function of the energy of the primary gamma ray and the shower age. The age s describes the stage of the shower development in the atmosphere. The number of secondary particles in an air shower increases when $s < 1$, reaches its maximum when $s = 1$, and decreases when $s > 1$ (see Fig. 3.2). The Molière radius R_{mol} [42], determined by multiple Coulomb scattering, is the radius that contains 90% of the secondary particles. The value of R_{mol}

is calculated as

$$R_{Mol} = X_0 \frac{E_s}{E_c \rho} = (37.15 \mathrm{g\,cm^{-2}}) \frac{21\,\mathrm{MeV}}{84.4\,\mathrm{MeV}\,7.4 \times 10^{-4}\,\mathrm{g\,cm^{-3}}} = 124.21\,\mathrm{m} \tag{4.3}$$

where X_0 is the radiation length of the atmosphere, E_s is the scattering energy, and E_c is the critical energy where an electron loses equal amount of energy per radiation length through bremsstrahlung and through ionization. At the HAWC altitude, the Molière radius is $R_{Mol} = 124.21\,\mathrm{m}$.

Because two types of PMTs with different quantum efficiencies are used in HAWC WCDs, the charge detected by a 10-inch PMT is scaled by 0.4554, called effective charge. We assume that the effective charge detected by PMTs in an air shower (the "footprint" of the shower in the detector) follows the same NKG lateral distribution and in the fitting procedure minimize:

$$\chi^2 = \sum_{i=1}^{n} \frac{(M_{PEi} - E_{PEi})^2}{E_{PEi}} \tag{4.4}$$

where M_{PEi} and E_{PEi} are the measured effective charge and the effective charge expected with the NKG function at a given PMT location, and n is the number

of PMTs that participate in the air shower event after hit selections. During this minimization procedure, the most likely core position is determined as well as the amplitude and age of the shower. These parameter are used for energy reconstruction algorithms that are under development. Fig. 4.4 shows the distributions of effective charge as a function the distance from the shower core. Fig. 4.5 shows the modified reduced-χ^2 distribution for gamma rays and cosmic rays (divided by the number of available PMTs is used instead of degree of freedom in the current analysis). As the NKG function describes the electromagnetic components in an air shower, gamma ray showers on average have a smaller modified reduced-χ^2. Fig. 4.6 shows the core distribution in the detector plane. The χ^2 fit is rather time-consuming since it implies a variable in the exponent. Several other core fitting algorithms have been tested or are under development in order to speed up the fit and to increase the accuracy.

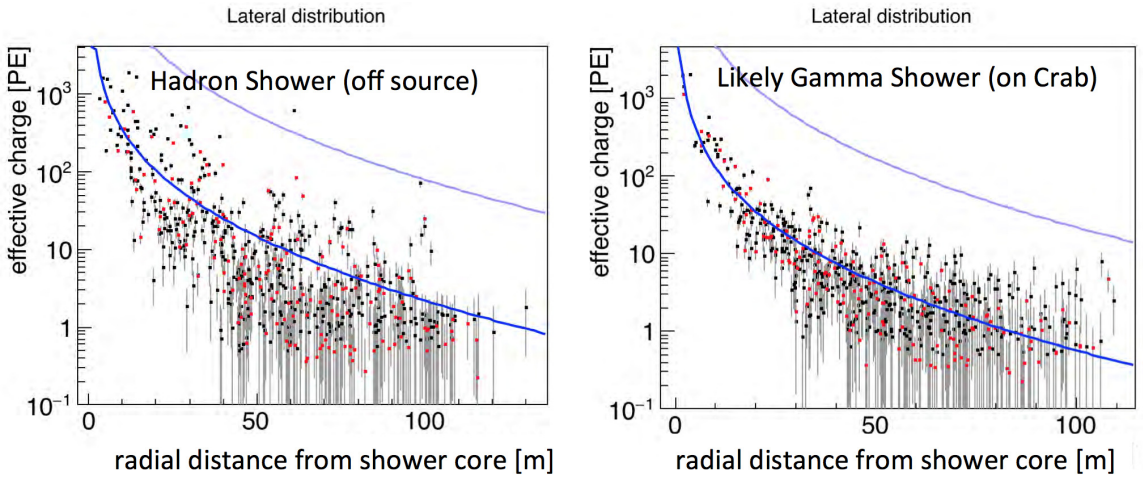

Figure 4.4 Effective charge as a function the distance from the shower core for a hadron-like (left) and gamma-like (right) shower in data. The red (black) dots represent 10-inch (8-inch) PMTs. The thick blue line is the NKG fit result.

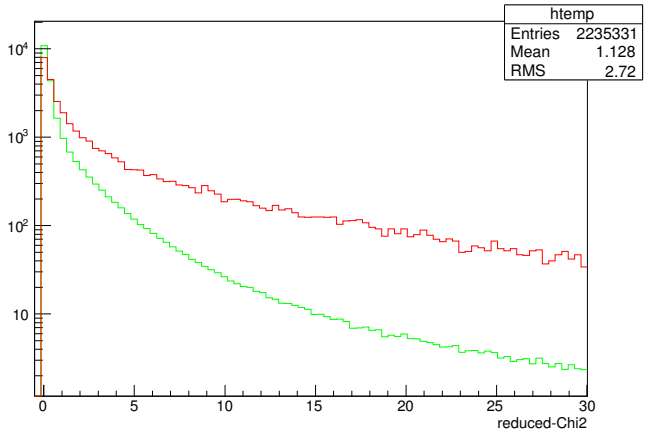

Figure 4.5 The modified reduced-χ^2 distribution of the core fit for gamma rays (green) and cosmic rays (red) from simulations.

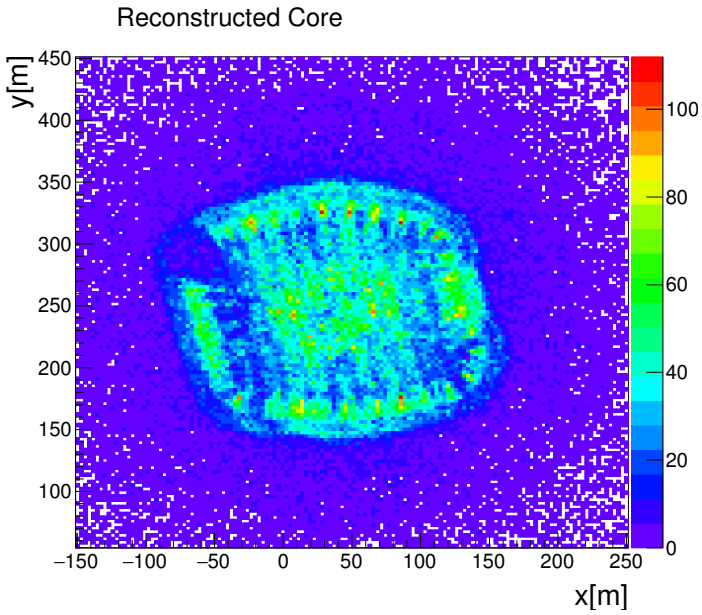

Figure 4.6 A core distribution with online reconstruction using data taken with the full array.

4.3 Angular Reconstruction

The angle fitter reconstructs the direction of the primary particle using the time and charge measurements by the PMTs detecting an air shower. The shower front as shown in Fig. 4.7 can be treated as a cone that is perpendicular to the direction of the primary particle.

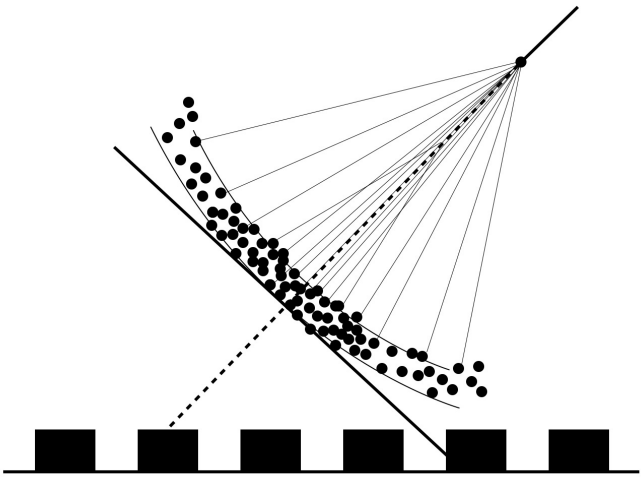

Figure 4.7 A sketch of the shower front.

However, fitting the arrival time of the shower plane to a cone is mathematically more complicated, thus more time-consuming than a planar fit. To be able to use a planar fit, there are two effects that need to be corrected for. The first is the shower curvature. The shower front (see Fig. 4.7) is approximately a cone because all the secondary particles in an air shower have approximately the same velocity, which is very close to the speed of light. The curvature of the shower front is corrected

depending on the distance from the shower core. The PMT hit times are shifted so that they land on a flat plane (corresponding to a projection of the shower front). The typical number of curvature correction is 0.07 nano seconds (ns) per meter. The second correction is the sampling correction. The shower front has a finite thickness of the order of a few ns that increases with the distance from the shower core. On average secondary shower particles that produce hits with larger number of PEs in the PMTs are detected earlier [43]. This results in a sampling bias of a few ns. To remove the bias a sampling correction is applied as a function of the number of PEs measured by a PMT. Both the curvature and sampling corrections are obtained using the simulations of a large number of air showers.

After these two timing corrections, the shower front approximately forms a flat plane. A weighted χ^2 fit to the corrected arrival times detected by the PMTs is performed to obtain the most likely shower front plane and, ultimately, the corresponding direction of the primary particle.

Fig. 4.8 shows the distribution of the shower direction in horizontal coordinate. The zenith angle is defined as 0° at the zenith and 90° at the horizon. The azimuth angle is 0° to the East and increases counter-clockwise.

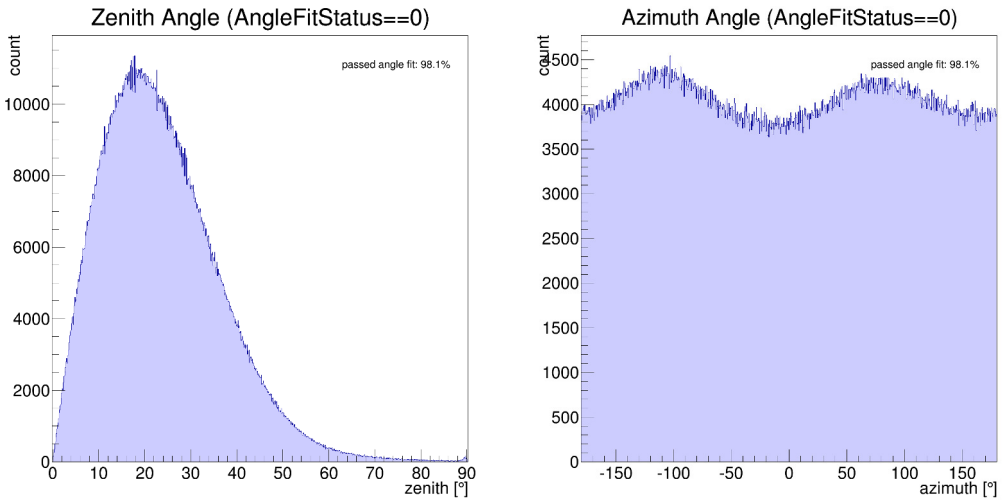

Figure 4.8 Zenith and azimuth angle distribution with online reconstruction using data taken with the full array.

4.4 Data Quality Cuts and Gamma Hadron Separation

The energy, angular resolution, and background rejection of the observed gamma rays are correlated with the shower size measured in the array. Therefore, the data are divided into 10 bins according to the fraction f of PMTs in the array that are triggered by an air shower event (see Table 4.1). To separate the gamma ray signal from the cosmic-ray background, standard selection requirements are applied to:

1. The ratio of the chi-square of the fit to the core location and the number of available PMTs in the array, $\frac{\chi^2_{core}}{N_{PMTavail}}$;

Table 4.1. Definition of Ten f bins.

Bin	0	1	2	3	4
f	0.04-0.07	0.07-0.11	0.11-0.16	0.16-0.25	0.25-0.37
Bin	5	6	7	8	9
f	0.37-0.51	0.51-0.66	0.66-0.78	0.78-0.88	0.88-1.00

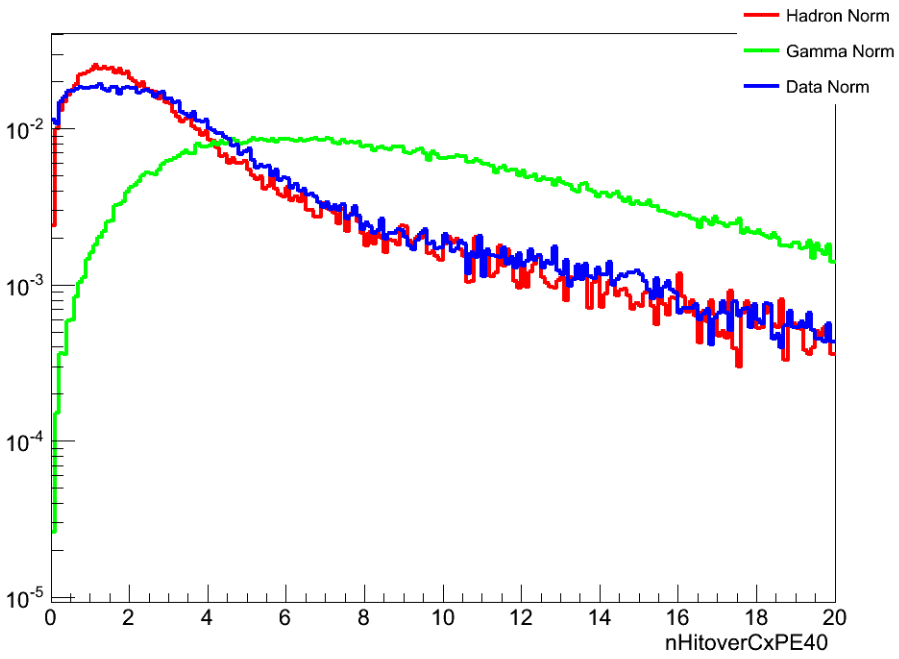

Figure 4.9 Distribution of "compactness" of data (blue), simulated gamma rays (green), and simulated cosmic rays (red) in f-bin 3.

2. The "compactness" of the charge distribution in the array, $\frac{N_{hit}}{\mathrm{CxPE}}$;

3. The reduced-χ^2 of the fit to the angle direction, $\frac{\chi^2_{angle}}{N_{\mathrm{dof}}}$.

Fig. 4.9 and Fig. 4.10 show the distributions of "compactness" of data, and simulated gamma rays, simulated cosmic rays. The cuts are explained in detail in [44].

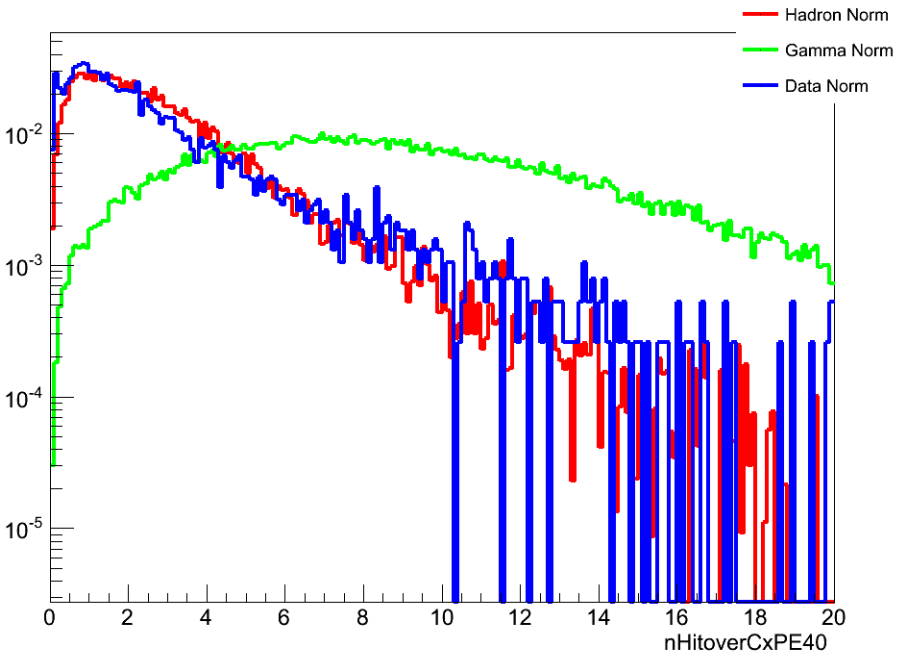

Figure 4.10 Distribution of "compactness" of data (blue), simulated gamma rays (green), and simulated cosmic rays (red) in f-bin 7.

The primary background of this analysis are cosmic ray air showers induced by hadrons. The shape of the lateral distribution of gamma-induced air showers differs from that of hadron-induced showers as previously seen in fig 4.4. The NKG function used for fitting the shower core describes the lateral distributions of gamma-ray showers. Consequently, the χ^2 values resulting from the core fit are on average smaller for gamma-ray showers than for cosmic-ray showers with the same number of triggered PMTs, and this parameter is used for background discrimination. Furthermore, hadronic showers produce pions with large transverse momenta that decay to gamma rays and muons. These sub-showers are likely to produce large signals in PMTs far from the shower core. Fig. 4.11 shows a typical simulated hadronic event

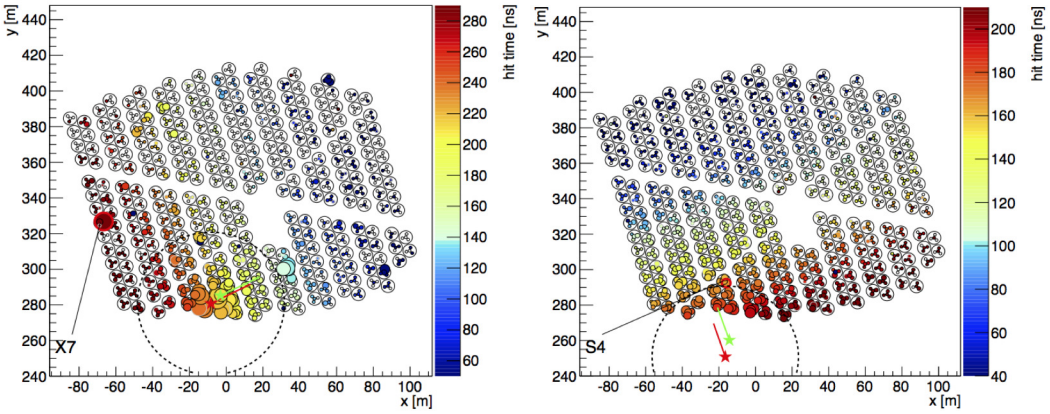

Figure 4.11 A hadron event (left) and a gamma event (right) from simulations. Each filled circle is a PMT signal, with the color indicating the hit time and the size indicating the charge. The reconstructed shower core is marked with a red star and a dashed circle indicates the 40 m radius around the reconstructed core. The highlighted red circle represents the location of the PMT with the maximum observed PE outside of the 40 m radius from the core.

and gamma-ray event in the simulation. Fig. 4.12 shows a typical hadron-like event and gamma-like event observed with HAWC. The ratio of the number of triggered PMTs to the number of PEs in the PMT that detects the strongest signal outside of a radius of 40 m from the shower core is found to result in good gamma/hadron separation performance. A third selection requirement applied to the reduced chi-square distribution of the shower angle fit removes poorly reconstructed air showers. In the lowest f bin, about 79% of gamma ray events and 38% of cosmic ray events pass the cuts, whereas for the highest f bin these numbers drop to about 13% and 0.03% respectively.

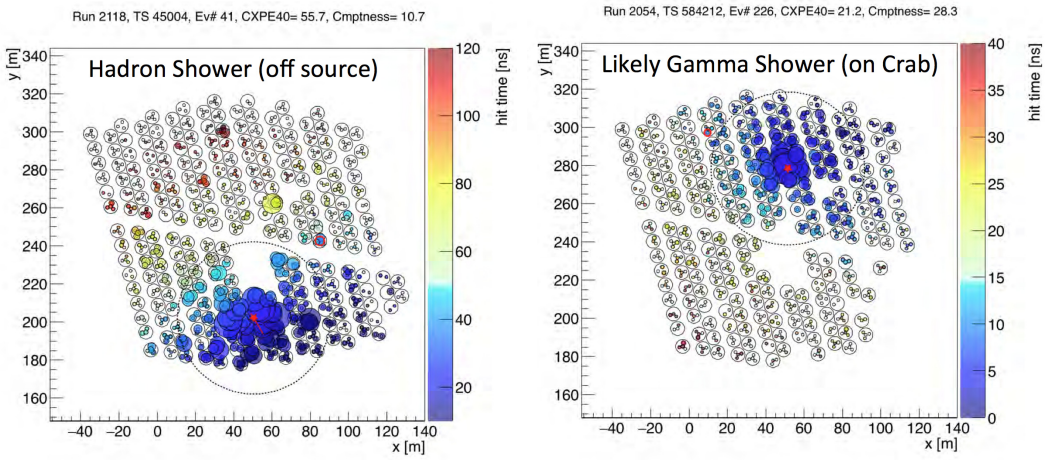

Figure 4.12 A hadron-like event (left) and a gamma-like event (right) in data taken with the full array.

4.5 Cuts Optimization on the Crab Nebula

All selection requirements are optimized in each f bin by maximizing the sensitivity to emission from the Crab Nebula, the brightest steady source at TeV energies and therefor the benchmark source of TeV gamma-ray analyses. Table 4.2 lists the cuts used for HAWC-111 and HAWC-250 data. Due to their similarity, Epoch 1 and 2 of HAWC-111 data are combined and use the same cuts, while Epoch 3 data uses a different set of cuts.

Table 4.2. Optimized Data Quality and Gamma/Hadron Cuts for HAWC-111 and HAWC-250

f	Core reduced-χ^{2}[a]			Compactness[b]			Angle reduced-χ^{2}[c]		
	Epoch 1+2	Epoch 3	HAWC-250	Epoch 1+2	Epoch 3	HAWC-250	Epoch 1+2	Epoch 3	HAWC-250
0	13	16	13	2.0	2.0	4.0	13	13	9
1	18	17	20	3.0	3.3	9.0	12	12	10
2	21	18	19	4.5	5.5	12.0	11	10	10
3	20	17	16	6.0	8.0	16.0	10	10	9
4	14	11	13	8.0	8.0	15.0	8	8	10
5	15	9	12	11.0	5.0	18.0	7	7	10
6	11	15	10	12.0	5.0	19.0	8	7	9
7	12	13	9	15.0	12.0	14.0	6	8	8
8	12	23	11	20.0	17.0	14.0	14	10	8
9	22	27	9	13.0	12.0	12.0	14	8	10

[a]Modified reduced-χ^2, defined as the ratio of the chi-square of the fit to the core location and the number of available PMTs in the array, $\frac{\chi^2_{core}}{N_{PMTavail}}$. Air shower events with smaller values are gamma-like and pass the cut.

[b]Air shower events with larger values are gamma-like and pass the cut.

[c]Air shower events with smaller values are gamma-like and pass the cut.

4.6 Measurement of the Point Spread Function on the Crab Nebula

Because the air shower events with the shower cores on the HAWC array have better core and angular reconstruction, the point spread function (PSF) of HAWC is described as a double-Gaussian function, i.e. a linear combination of two Gaussian functions:

$$PSF = \frac{A \cdot G(\sigma_1) + B \cdot G(\sigma_2)}{A + B} \tag{4.5}$$

where $G(\sigma)$ is a Gaussian function with width of σ. Fig. 4.13 shows the signal vs. the angle from the source center for a simulated source.

The Crab Nebula is used to measure the point spread function in the data since its extent is much smaller than the PSF of the HAWC detector. We can directly measure the PSF from the data around the Crab. However, in HAWC-111 data, the statistic in individual f bins near the Crab is not sufficient to distinguish a double-Gaussian function from a Gaussian function. As an approximation, a Gaussian function is fit to the data around Crab (see Fig. 4.14).

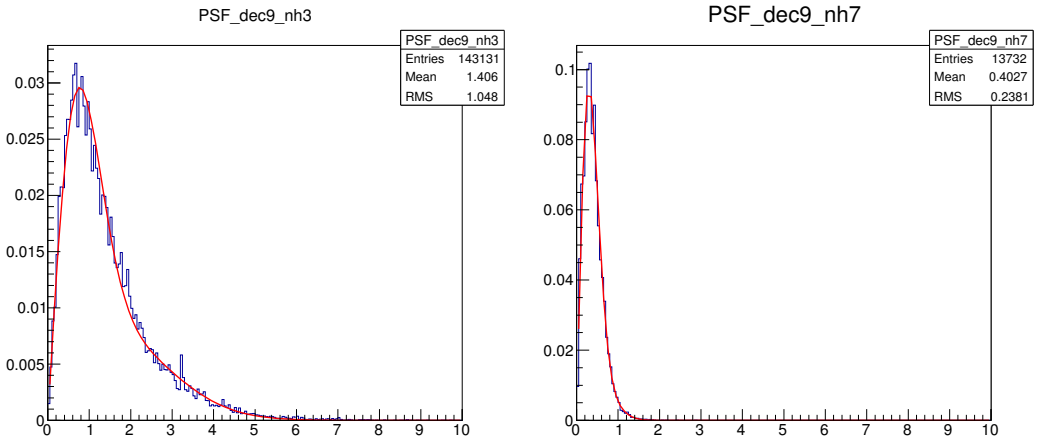

Figure 4.13 Signal vs. angle from the source center for a simulated source in f bin 3 and 7 simulation. The red lines represent double-Gaussian fits.

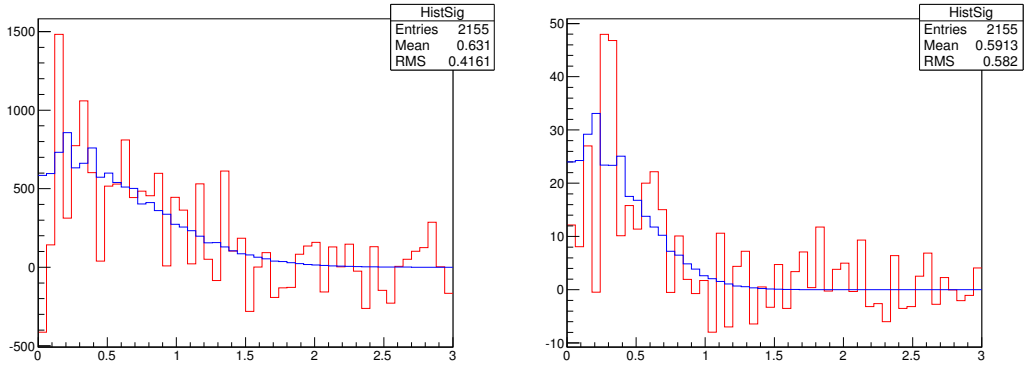

Figure 4.14 Signal vs. angle measured in data around Crab in f bin 3 and 7. The red lines represent the signal and blue lines are the fits.

With the current reconstruction of HAWC-111 data, fitting the emission from the Crab Nebula region in each f bin, the PSF (68% containment) is found to vary from 2.5° to 0.6°. It is on average $\sim$ 0.3° wider (in quadrature) than the expected PSF from simulation. In the current standard analysis presented in this thesis, the PSF derived from simulation with 0.3° added in quadrature is used (see Fig. 4.15).

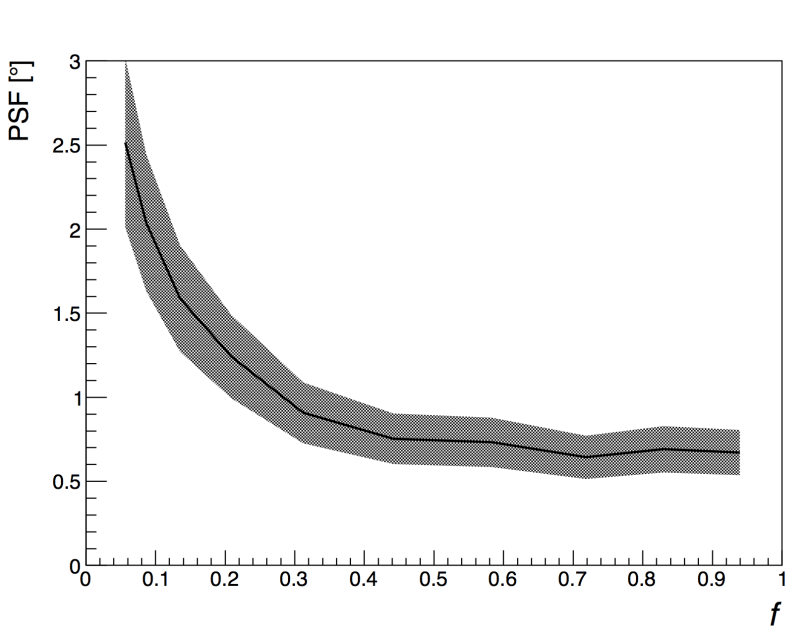

Figure 4.15 PSF (68% containment) used in the current standard HAWC-111 analysis as a function of the fraction f of PMTs used in event reconstruction. The systematic uncertainty of the measured PSF (gray) is $< 20\%$.

Chapter 5

Calibration of the Observatory

The proper reconstruction of the position of the shower core and of the shower plane, and therefor the reconstruction of ultimately the direction of a primary gamma ray in the sky requires a precise and accurate knowledge of the arrival time and charge measured by each PMT in the array.

To reach the designed angular resolution of down to 0.1°, a timing calibration with an accuracy better than 1 ns is necessary. In order to achieve this, there are two effects that need to be calibrated. The first is a slewing effect, i.e. the response time of a PMT and electronics as a function of pulse size. The time of a PMT pulse crossing the low threshold the first time is used as the time of this signal. Since the threshold is at a fixed voltage value, the time of the first edge also depends on the pulse size.

As the pulse size increases, the pulse crosses the thresholds earlier, yielding an earlier edge and a shorter response time as shown in Fig. 5.1. This pulse size dependent slewing time must be corrected for depending on the ToT values. The second effect is the relative offset among PMT channels. While the cables connecting the PMTs to the central DAQ are cut to be the same length, there might be still small difference in propagation time due to different characteristics of individual cables as well as the different characteristics of individual PMTs. This systematic timing offset between PMTs must be corrected for.

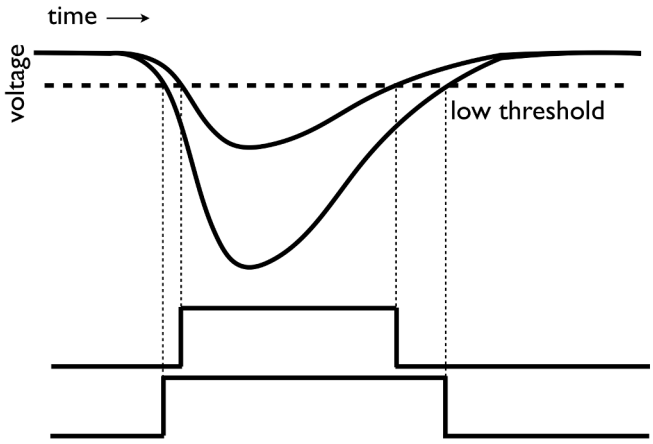

Figure 5.1 A sketch demonstrating slewing effect. A larger pulse produces an early signal recored by DAQ.

The following sections describes the setup of the calibration system, the charge and timing calibration, and methods using air shower data to correct relative timing and detector pointing.

5.1 Calibration System Setup

To calibrate the ToT to charge conversion and to correct for the two timing effects that are described above, a laser calibration system was designed and installed at the HAWC site [45]. Before shipment and installation at the ultimate high-altitude, the core system consisting of the laser, three filter wheels (FWs), the optical switch, and four radiometers was tested in the lab at Michigan Technological University. In parallel, a second laser calibration system akin to this main system intended for HAWC was tested with a DAQ system installed on a WCD prototype at Colorado State University.

Fig 5.2 shows an illustration of the calibration layout. A green laser with a wavelength of 532 nm is used. It is triggered with a square-wave that is generated by a pulse generator. The laser beam is split by an optical splitting cube. One of the sub-beams travels through a 1:19 splitter. One of the 19 output fibers is connected to a radiometer labeled RAD 1, to monitor the laser power. The other sub-beam goes through a set of three computer-controllable FWs that each holds six different neutral density filters, providing a wide range of intensities of laser pulses over six orders of magnitude. Table 5.1 lists the optical depth (OD $= -\log_{10}$(transmittance)) provided by each filter.

Table 5.1. OD Values of Installed Filters

FW position	FW 1	FW 2	FW 3
1	5.0	0.0	opaque
2	0.0	0.2	0.0
3	1.0	0.4	0.3
4	2.0	0.6	1.0
5	3.0	0.8	1.3
6	4.0	1.0	2.0

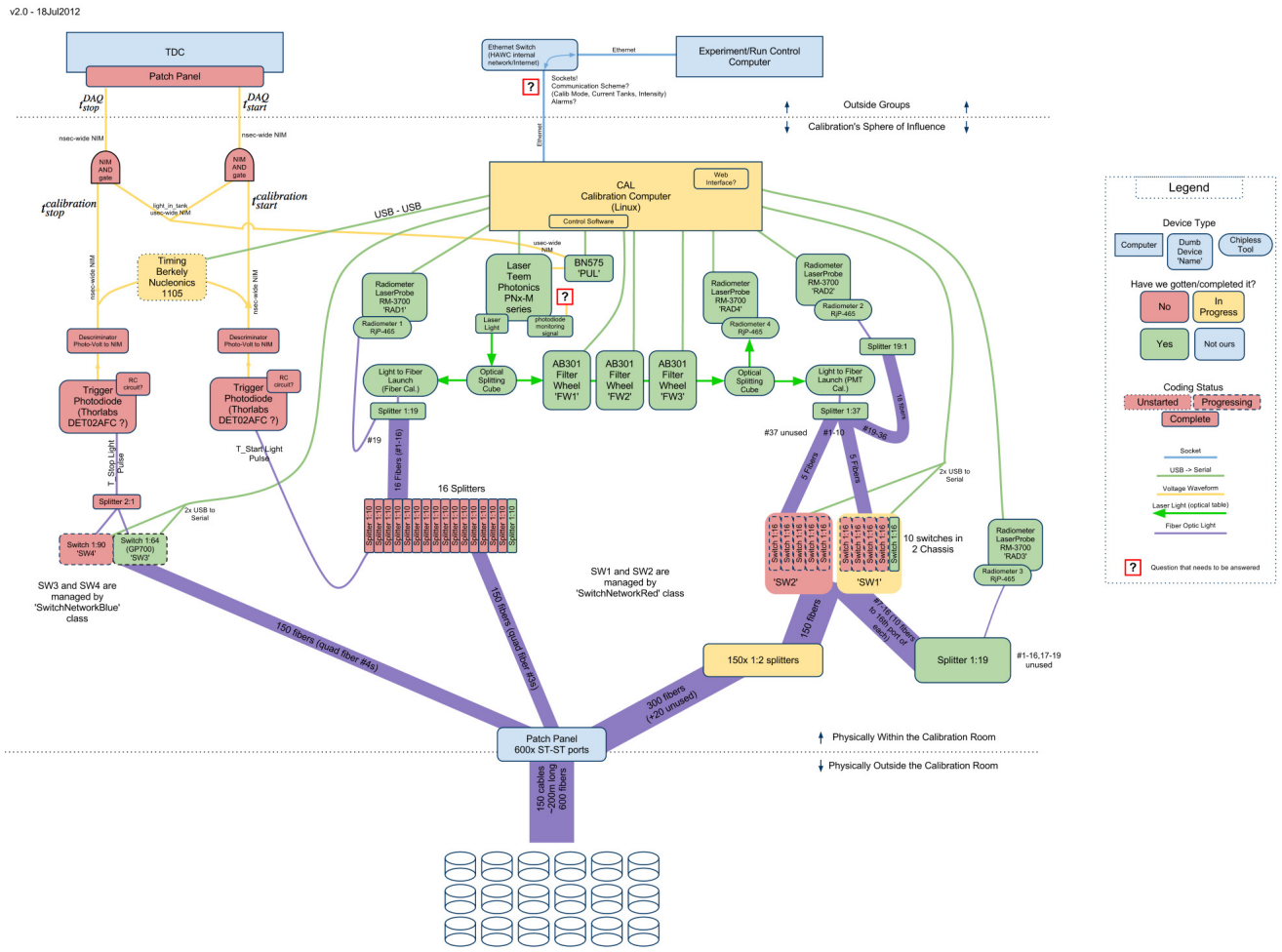

Figure 5.2 The layout of the calibration system.

The sub-beam after the filter wheels is split into two sub-beams again: one is monitored by another radiometer (RAD 4) of the same type as RAD 1 and the other sub-beam is launched into a 1:37 splitter. Eight of 37 output fibers are connected to eight 1:16 optical switch modules. Each output channel runs through a 1:2 or 1:4 splitter and approximately 600 feet of optical fiber into each of 300 WCDs. The connection of the 1:2 and 1:4 splitters are carefully selected so that the light intensity is similar in each WCD. Fig. 5.3 shows the fiber connections inside the counting house.

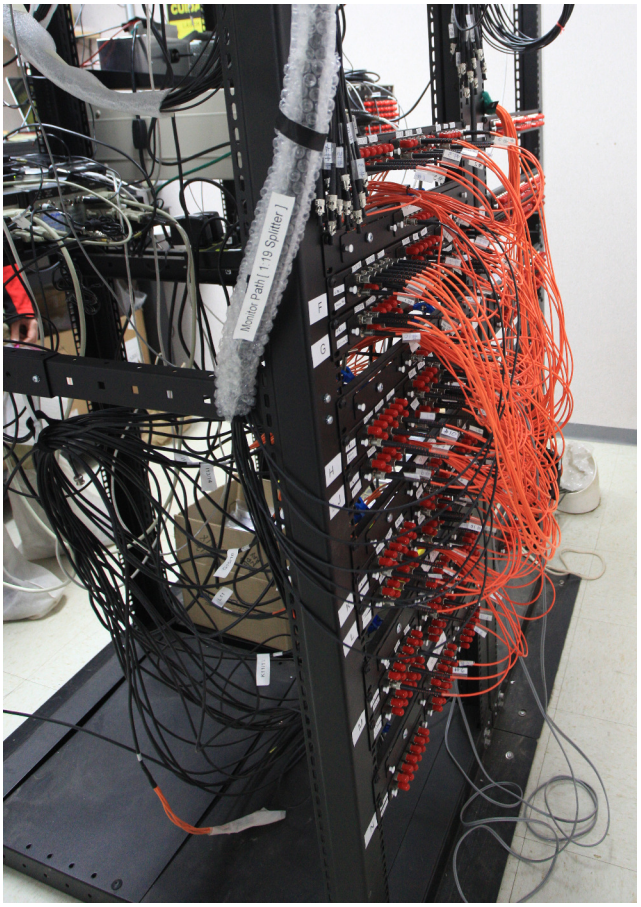

Figure 5.3 The calibration rack inside the counting house during installation of calibration devices.

In each WCD, a custom-made diffuser is installed at the end of the optical fiber and

attached to a float assembly. The assembly is attached to the central PMT via kevlar strings so that the diffuser is floating under the water about 3 m above the central PMT. The diffuser expands the opening angle of the output light cone to cover both central and outer PMTs. Once the PMTs detect the laser light, the signals are sent back and recorded by the DAQ system in the same way as signals due to air showers.

When there is laser light being sent into the WCDs, a "light-in-tank" flag is generated and sent to the DAQ system to ensure that the PMT signals due to laser light are not misidentified as air shower signals. Also being sent to the DAQ system are a T_{start} and a T_{stop} flag, where T_{start} indicates the time the laser pulse is sent as measured by a fast-response Thorlab photo sensor. The signals in a pre-defined time window after each T_{start} flag ($\sim 2\,\mu s$) are then selected for the analysis of charge and timing calibration.

There are 180 non-opaque FW combinations, but some of the filter combinations yield identical ODs. To reduce the time of a calibration run, currently 68 FW combinations are used providing transmittance from 1 to $10^{-6.5}$. For each FW combination 2000 laser pulses are sent. Thus the laser intensity can be regulated from < 1 PE to $> 1,000$ PE in the majority of the PMT channels.

5.2 Charge Calibration

Using RAD 4, the relative laser intensity that is sent to WCDs with different FW combinations can be measured. However, the charge detected by each PMT is different channel-to-channel since the laser light travels through different splitters and fibers that do not necessarily have the same transmittance. An occupancy method is used to determine the charge in number of PEs (N_{PE}) detected by each PMT [46].

In this method it is assumed that the N_{PE} produced at the PMT photocathode follows a Poisson distribution. The probability of producing n PEs given a certain number of PEs given a mean value of $< N_{PE} >$ is:

$$P(\mathrm{n}, < \mathrm{N_{PE}} >) = \frac{< \mathrm{N_{PE}} >^{\mathrm{n}}}{\mathrm{n!}} e^{-<\mathrm{N_{PE}}>} \tag{5.1}$$

The occupancy is defined as the ratio of the number of laser pulses detected by a PMT (N_{laser}) and the total number of laser pulses that are sent ($N_{trigger}$). The occupancy equals to the probability of observing at least one PE:

$$\eta = \frac{\mathrm{N_{laser}}}{\mathrm{N_{trigger}}} = P(n > 0, < \mathrm{N_{PE}} >) = 1 - P(n = 0, < \mathrm{N_{PE}} >) = 1 - e^{-<\mathrm{N_{PE}}>} \tag{5.2}$$

A time window is selected with a pre-defined time delay after each laser trigger ($\sim 2\,\mu s$), which is chosen to include all the PMT signals due to laser pulses. $\mathrm{N_{laser}}$ is estimated by counting the number of time windows that have at least one hit. However, noise caused by cosmic ray particles is also present in the time windows and cannot be distinguished from the laser pulses. Assuming the probability of detecting a laser signal and detecting a noise signal is independent, the joint probability of observing at least one hit is:

$$P(\text{any}) = P(\text{laser} + \text{noise}) = P(\text{laser}) + P(\text{noise}) - P(\text{laser})P(\text{noise}) \tag{5.3}$$

The occupancy η can the be derived from:

$$\eta = P(\text{laser}) = \frac{P(\text{any}) - P(\text{noise})}{1 - P(\text{noise})} = \frac{\mathrm{N_{any}} - \mathrm{N_{noise}}}{\mathrm{N_{trigger}} - \mathrm{N_{noise}}} \tag{5.4}$$

where $\mathrm{N_{any}}$ is the number of time windows that have at least one signal and $\mathrm{N_{noise}}$ can be estimated from the average PMT signal rate and the duration of the time window.

From Equ. 5.2 the average number of PE, $< \mathrm{N_{PE}} >$, can be calculated for each FW combination as:

$$< \mathrm{N_{PE}} >= -ln(1-\eta) \tag{5.5}$$

This method only works at low laser intensities, otherwise the error on $< \mathrm{N_{PE}} >$ grows quickly when the occupancy is close to 1. As $< \mathrm{N_{PE}} >$ is proportional to the laser light intensity measured by RAD 4:

$$< \mathrm{N_{PE}} >= A(I/I_{ref}) + B \tag{5.6}$$

where I is the RAD 4 measurement, I_{ref} is the reference value, and B is a constant offset, which should be ≈ 0. The occupancy can also be written as a function of laser light intensity:

$$\eta = C[1 - e^{-A(I/I_{ref})+B}] \tag{5.7}$$

Here C, which should be ≈ 1, takes into account the effect that the maximum occupancy may not reach 1 due to missing signals. Fig. 5.4 shows the occupancy as a function of laser light intensity. The function of equation equ:occfitis fit to the data in the occupancy range of 0.05-0.9, resulting in values for B of ~ 0 and C of ~ 1, as expected. Thus if the laser data is not dominated by noise or other systematic effect,

the fit result can be extrapolated and the number of PEs, $< N_{PE} >$, can be calculated from the occupancy, for each FW combination, i.e. for the light intensity.

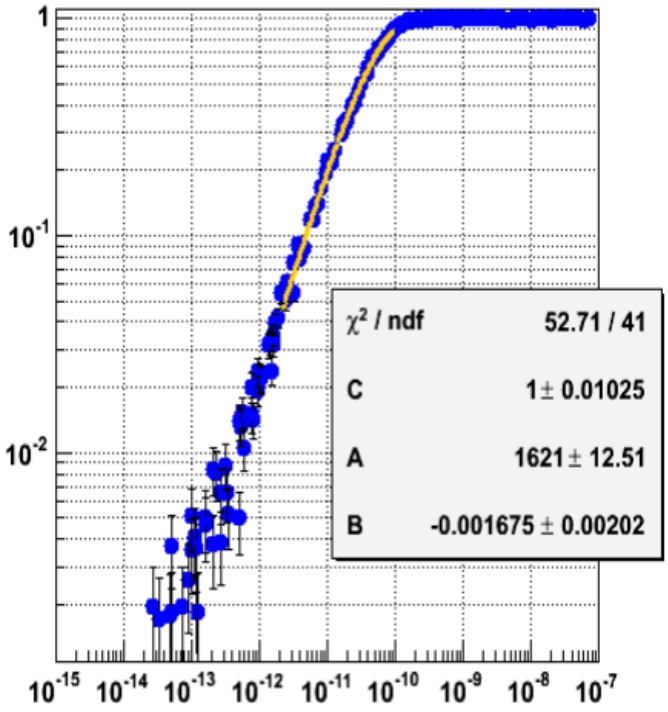

Figure 5.4 The occupancy as a function of laser energy as measurement by RAD 4 for one PMT. Image credit: F. Salesa (see Appendix A).

After the determination of $< N_{PE} >$ at a given laser intensity, distributions of N_{PE} values are simulated based on a Poisson distribution with a mean of $< N_{PE} >$. Then each discrete value of N_{PE} from the Poisson distribution is smeared by a factor of 35% using a Gaussian function, which represents the charge resolution of PMTs. The simulated distribution of N_{PE} is then compared to the data distribution of ToT for the same FW combination (laser light intensity). Both distributions are plotted in Fig. 5.5.

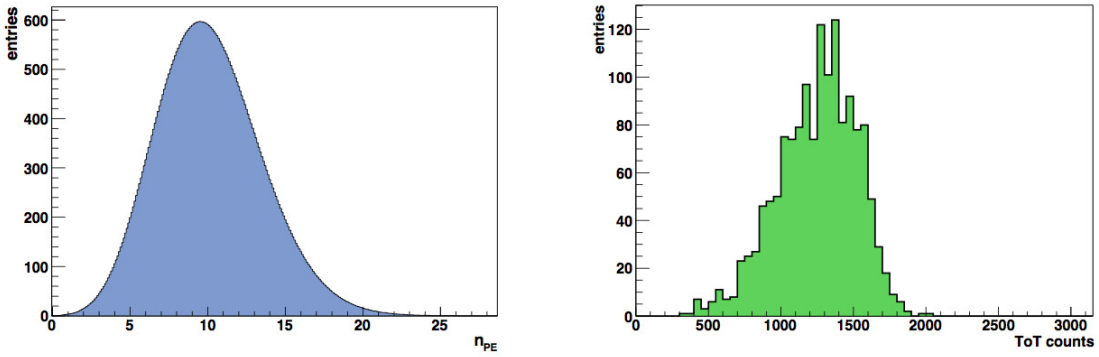

Figure 5.5 (Left) A simulation of N_{PE} distribution with $< N_{PE} >= 10$ PE. (Right) The ToT distribution measured by a PMT with the same FW combination, i.e. at the same light intensity. Image credit: F. Salesa (see Appendix A).

The percentiles of step of 10% between 10% and 90% for both distributions are computed and the values of N_{PE} and ToT are generated by the various FW combinations. This is done for both low ToT and high ToT and all laser light intensities. Fig. 5.6 shows the resulting $\log_{10}(N_{PE})$ as a function of ToT. A broken power-law function which was used in MILAGRO is adopted for fitting these curves, which is a broken power-law function:

$$\log_{10}(n_{PE}) = \begin{cases} p_1 + p_2 \cdot ToT & \text{for} \quad p_{\min} < ToT \leq p_4 \\ (p_1 + p_2 \cdot ToT) \cdot \left[\exp\left(\frac{ToT - p_4}{p_3}\right) - \frac{ToT - p_4}{p_3}\right] & \text{for} \quad p_4 < ToT \leq p_{\max} \end{cases} \tag{5.8}$$

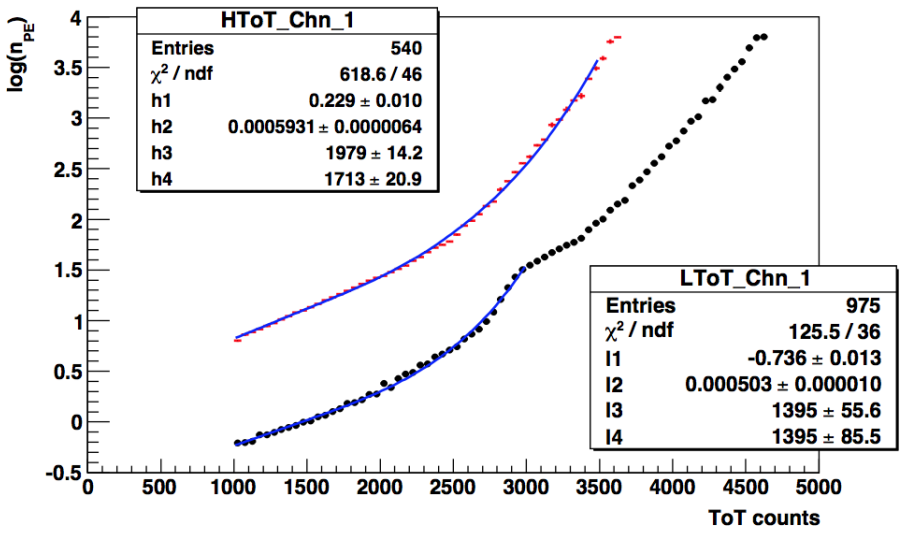

Figure 5.6 $\log_{10}(N_{PE})$ as a function of low ToT(lower curve)/high TOT(upper curve) in TDC counts (0.1 ns) for one PMT. Image credit: F. Salesa (see Appendix A).

5.3 Timing Calibration

The response time of each PMT channel is a function of ToT and varies from channel to channel. The response time can also be measured using the laser calibration system. The time from firing the laser (T_{start}) to a PMT pulse crossing the low/high threshold is measured as the raw slewing time. It includes not only the response time of the PMT and electronics, but also the time of the laser light travels in the splitters, fibers and water. The time of the laser light passing through splitters and fibers was measured at the site and is subtracted from the raw slewing time. In a WCD, the central PMT is closer to the diffuser than the peripheral PMTs (see Fig. 5.7). This difference in the light path corresponds to $\approx$ 2.2 ns in water and is added to the raw slewing time of the central PMTs to correct this effect. After applying these

corrections, the slewing time represents the response time of the PMT and electronics.

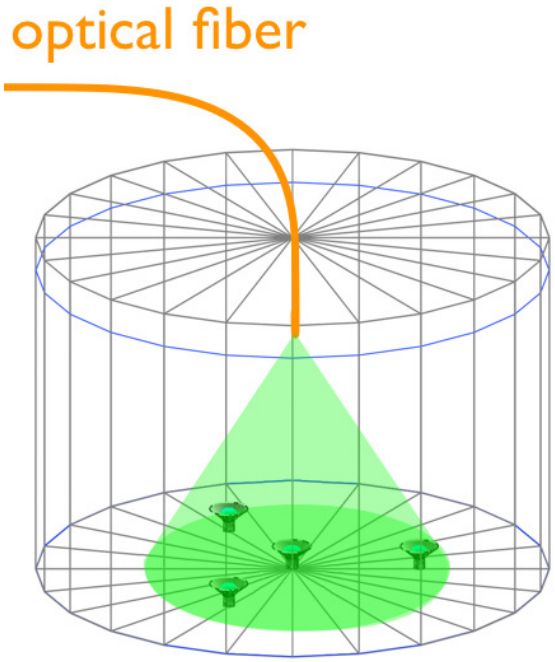

Figure 5.7 Sketch of a WCD with optical fiber and diffuser, showing that the central PMT is closer to the diffuser.

Fig. 5.8 (left) and Fig. 5.9 show the slewing time as a function of ToT for low and high thresholds. As expected, the slewing time for both thresholds is shorter when the ToT (i.e. the pulse size) is larger.

The same edge finder is used to process calibration data and air shower data. However, since the laser can be much brighter than the expected Cherenkov light produced by air showers, there are many "4-edge" pulses that are misidentified as pairs of "2-edge" pulses. These misidentified "2-edge" pulses have short low ToT values and appear as earlier pulses, respect to other true small "2-edge" pulses in the plot of Fig. 5.8.

To remove the misidentified "2-edge" pulses, firstly the laser data collected with FW transmittance of greater than $10^{-0.6}$ is not used in the timing calibration. Secondly, a parameter DT is defined as the time from the first edge of the current pulse to the first edge of the following pulse in the same PMT channel. If a "2-edge" pulse with a short ToT has a short DT, there is a large chance that this "2-edge" pulse rather represents the first two edges of a "4-edge" pulse. These pulses are removed from the data and not used for calibration purposes. The optimal cut is found to be low ToT $>$ 100 ns —— DT $>$ 1000 ns. Fig. 5.10 shows slewing time of the low threshold as a function of low ToT after the data quality cuts.

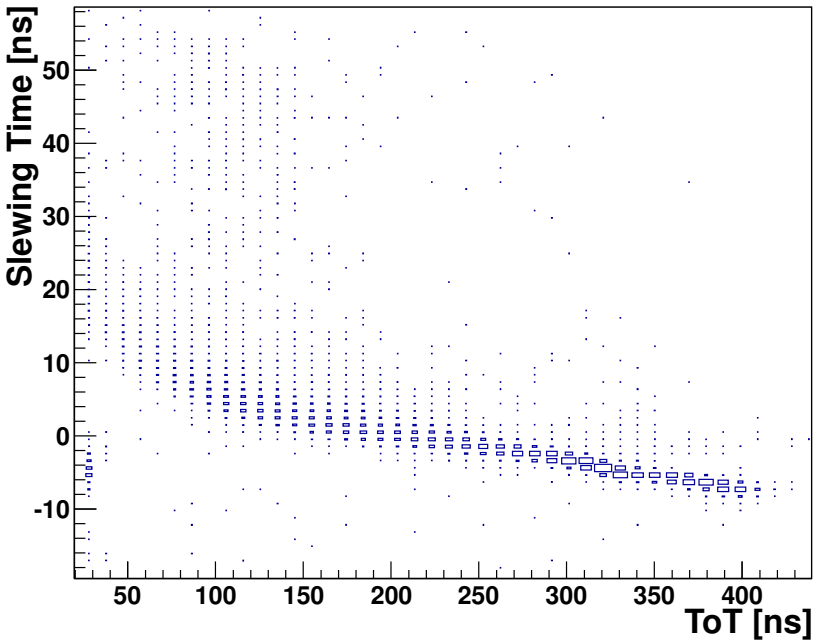

Figure 5.8 Low slewing time as a function of low ToT before the data quality cuts.

In order to define slewing curves, the histogram of slewing time vs. ToT is binned

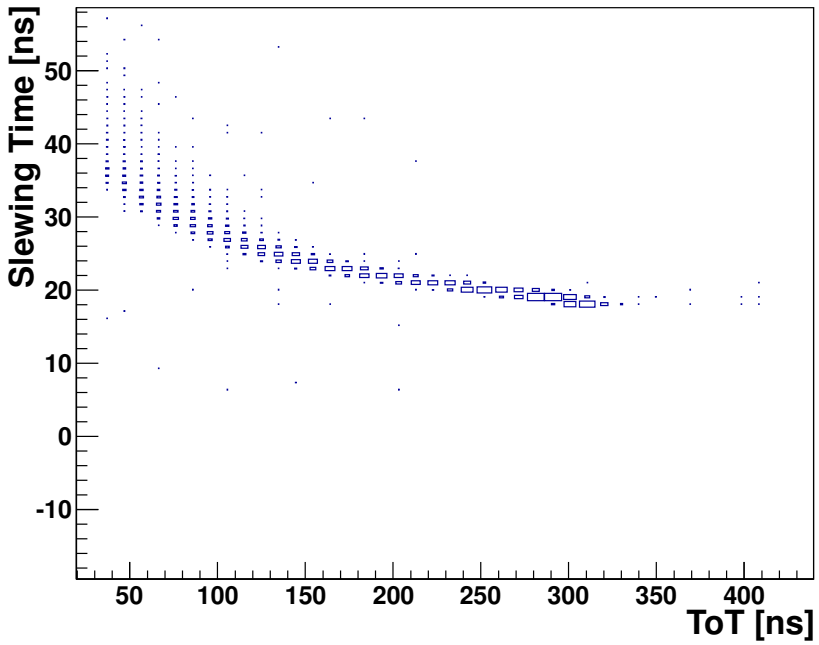

Figure 5.9 High slewing time as a function of high ToT.

by the ToT values. A bin size of 9.8 ns is chosen to achieve sufficient curve sampling as well as enough statistic in each ToT bin. Only ToT bins with more than 200 entries are used. In each ToT bin, three methods, a Gaussian fit to the slewing time, an asymmetrical Gaussian fit and a method based on 16%, 50%, and 84% quantiles of the distribution, were tested to find the mean slewing time. Because the angular reconstruction assumes that the distribution of signal time follows a Gaussian function, the Gaussian mean is the best representative of the slewing time in each ToT bin. Fig. 5.11 shows the slewing curves of four PMTs in one WCD with the Gaussian mean and sigma of the slewing time in each ToT bin.

Since the shape of slewing curves can be explained by the shape of PMT pulses,

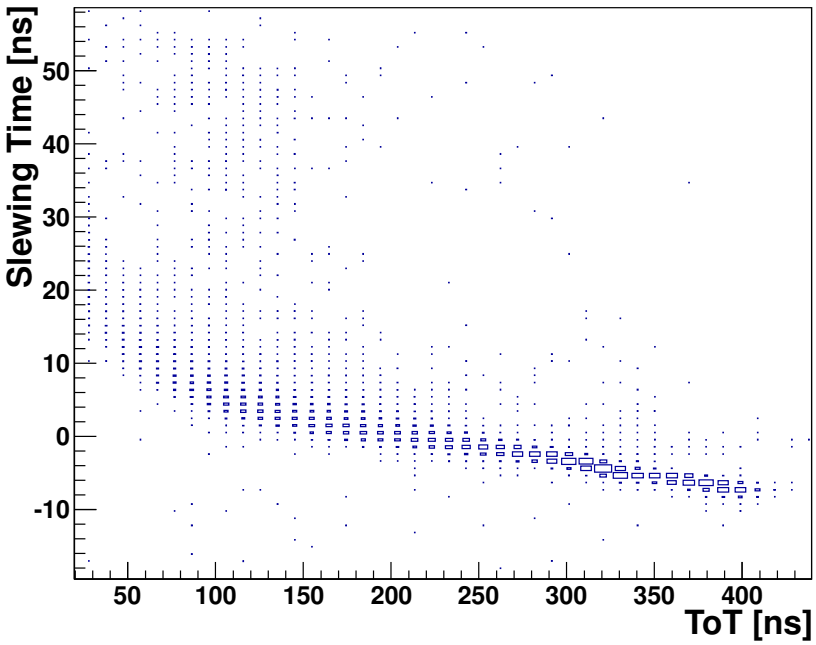

Figure 5.10 Low slewing time as a function of low ToT after data quality cuts.

a dedicated fitting function is derived representing the features of the curves. The rising part of a PMT pulse can be estimated as a cumulative Gaussian function. The curvature at the lower ToT end of both curves for the low and high thresholds is due to the threshold sampling near the top of the pulse. At the higher end of the low threshold curve, there is curvature expected too due to the threshold sampling near the bottom of the pulse. This feature does not show up in Fig. 5.11 because the laser intensity is not high enough to produce that big PMT pulse. The middle part of both curves is approximately linear.

Taking all these considerations into account, a function with six parameters is constructed as

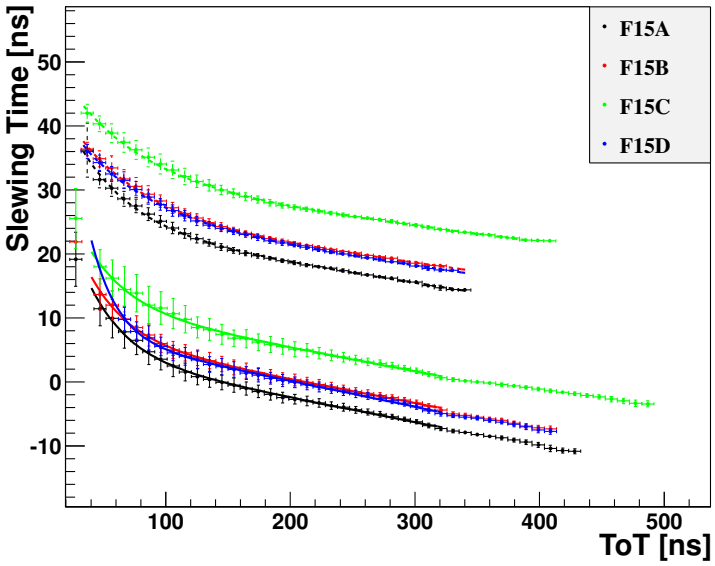

Figure 5.11 Slewing curves of the PMTs in one WCD fit with a 6-parameter function. The upper population of data points group is for the high threshold (high slewing time vs. high ToT) and the lower population for the low threshold (low slewing time vs. low ToT). "C" represents the central PMT, and "A" "B" and "D" represent the peripheral PMTs.

$$\text{Slewing Time (ToT)} = e^{\frac{-\text{ToT}-\text{p0}}{\text{p1}}} - e^{\frac{\text{ToT}-\text{p2}}{\text{p3}}} + \text{p4} - \text{p5} * \text{ToT} \tag{5.9}$$

where $p0$ and $p1$ represent the break point and scale of the exponential term at the lower ToT values, $p2$ and $p3$ represent the break point and scale of the exponential term at the higher ToT values, and $p4$ and $p5$ represent the intercept and slope in the linear part. In practice, since the curvature of the slewing curves at higher ToT values is trivial for both low and high threshold, $p2$ and $p3$ are fixed in order to reduce the number of parameters.

The values of the parameters are found by an iterative fitting procedure. First, the middle range (between 150 ns and 350 ns) of a curve is fit with a linear function. The resulting fit values for $p4$ and $p5$ are then used as initial seeds when the four-parameter fit is performed. The curves in Fig. 5.11 represents the fit results. As previously mentioned, each WCD is instrumented with three 8-inch PMTs and one central 10-inch PMT. The timing characteristics of the PMT models are different. This can also be seen in Fig. 5.11. The response time of the 10-inch PMT (green) is larger than that of the 8-inch PMT. The resulting fit functions for th PMT channels serve as input to the air shower event reconstruction code.

5.4 Relative Time Calibration Using Air Shower Data

The timing measurements of deployed fibers are subjected to varying temperature at the site, which can cause a different laser travel time as large as $\sim$ 1 ns. To achieve a better relative timing, a time residual method is applied to correct for inaccurate timing measurements of fibers and other unknown effects. Time residual is the difference between a PMT hit time and the expected time on the fit air shower plane. If a PMT is systematically off in time, the distribution of time residuals on

this PMT calculated from a large number of air showers will be peaked at a non-zero value. To calculate time residuals, 200,000 air shower events with $N_{hit} > 200$ and within 30° from zenith are used. The events are reconstructed with the slewing correction described above applied and the time residual is computed per PMT per event. A Gaussian function is fit to the time residual distribution of each PMT to find the mean offset of this PMT channel. This procedure repeats for three times with the time residuals obtained from the previous step being corrected in the next step. Fig. 5.12 shows the time residual distribution of a PMT before/after time residual correction. The mean time residual is much more closer to zero with time residual correction.

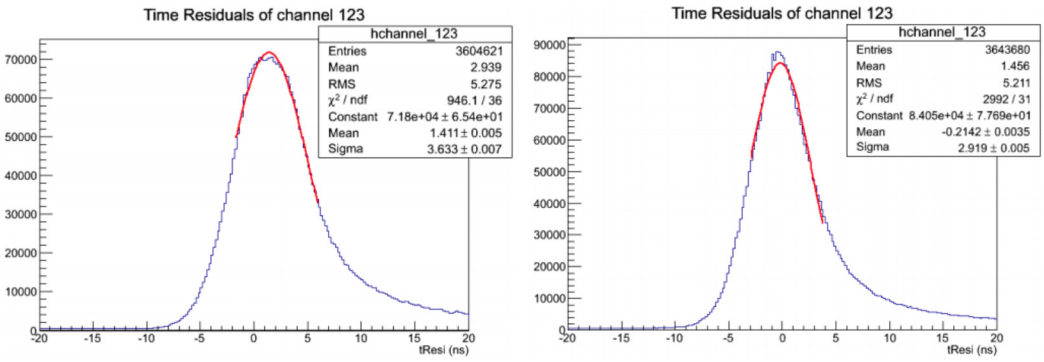

Figure 5.12 The time residual distribution of a PMT before (left)/after (right) time residual correction. The mean is off from zero without the correction.

After three iterations, the time residuals from three steps are added up and this additional time offset is combined with the slewing correction obtained above. This is the final timing calibration that is being applied to the event reconstruction. Fig. 5.13 shows after the timing calibration the relative timing among PMTs are better than

1 ns.

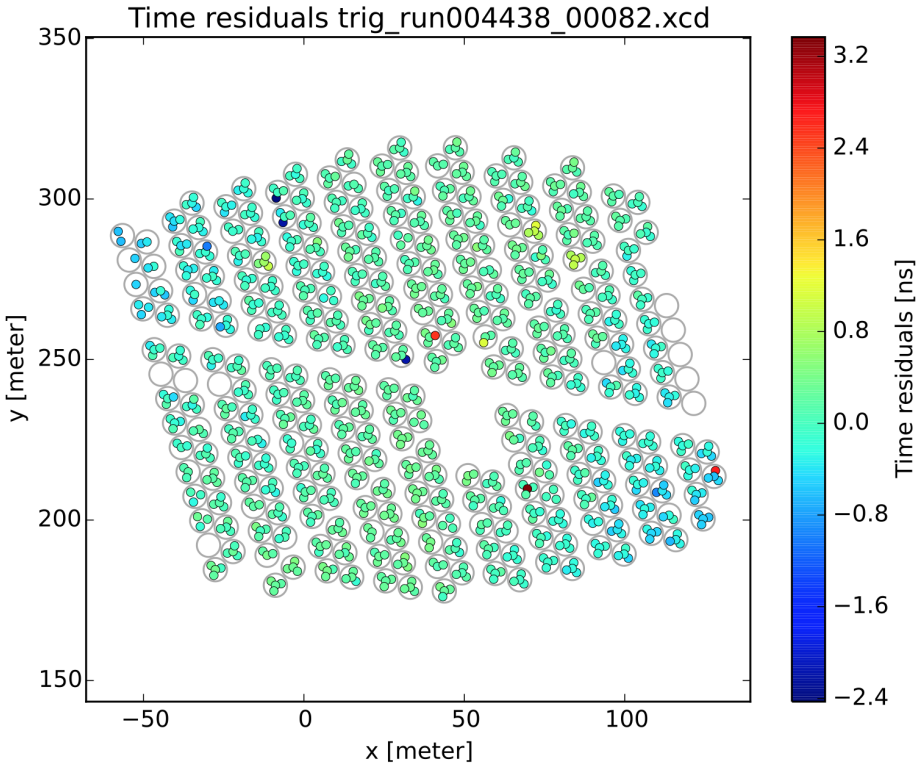

Figure 5.13 Time residual as a function of PMT position after timing calibration with online reconstructed data taken with the full array. The relative timing among PMTs are better than 1 ns, except a few PMT channels.

5.5 Zenith Alignment

The air showers that come from larger zenith angles travel through thicker atmosphere thus have fewer secondary particles that survive at the detector level. The arrival directions of the primary particles should have the maximum rate at zenith. This effect can be used to correct the pointing of the detector. Fig. 5.14 shows the

distributions of the air shower arrival direction along E-W and N-S axes with each distribution peaking near zero as expected. A Gaussian function is fit to each distribution to find the peak, i.e. the apparent zenith. The apparent zenith may change with different number of WCDs in the detector and with different charge/timing calibration. Fig. 5.15 shows how it changes with time during the construction stage of HAWC. To have sufficient statistics, each data point is the Gaussian mean using five days of data.

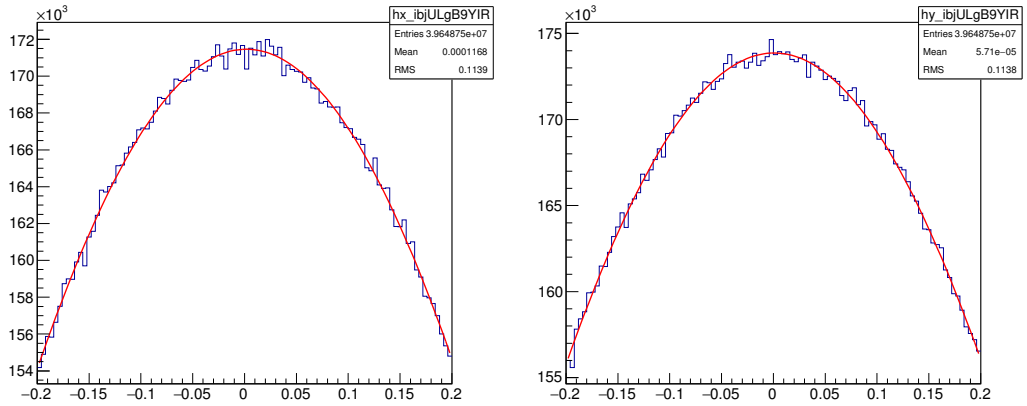

Figure 5.14 The distribution of air shower rate along East-West (left) and North-South (right) axes fit with a Gaussian function. The unit is radian.

An additional zenith correction is applied as a function of time after the angular reconstruction. It corrects the arrival direction of each primary particle, or equivalently, tilts the detector plane to make the apparent zenith the same as the true zenith.

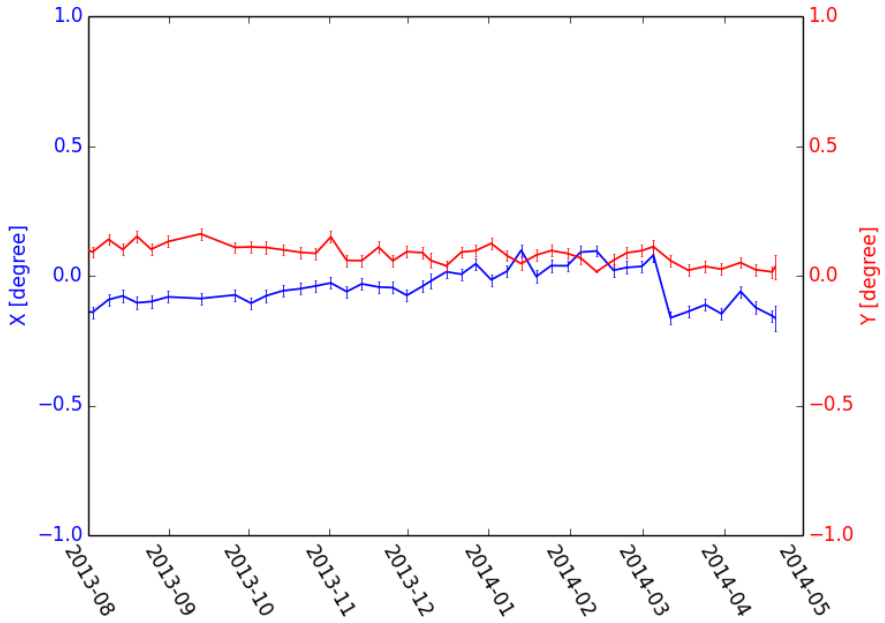

Figure 5.15 The apparent zenith, the direction that has the maximum air shower rate, along X (East-West) and Y (North-South) axes as a function of time during the construction stage of HAWC. The slope between December 2013 and March 2014 is due to more WCDs added to the detector and the jump at March 2014 represents the change of calibration.

Chapter 6

Mapping the Gamma-Ray Sky

The arrival directions of air shower events are reconstructed as previously described in Chapter 4. To search for gamma ray sources, a data map is generated and compared to a background map, which is also based on the observed data. As the HAWC detector is co-rotating with Earth, the reconstructed direction of events are converted to right ascension (RA) and declination (Dec) in equatorial coordinate [47] using the event time. The angular direction of events are corrected for the J2000 epoch [1]. For the events in each f bin, the RA and Dec values are then binned using HEALPIX [48], a pixelation scheme that divides the unit sphere into equal-area pixels. It splits the unit sphere into twelve equal-area tessellations. Each of the twelve tessellations are subdivided into an $N_{\text{side}} \times N_{\text{side}}$ grid, giving $12\,N_{\text{side}}^2$ pixels. In this analysis, N_{side}

[1]The celestial coordinate varies with time. J2000 epoch refers to the celestial coordinate at the instant of 12h on 1 January 2000.

is chosen to be 512, which provides an average pixel size of 0.11°. These HEALPIX maps of events are referred as raw data maps.

This chapter describes the background estimation, the map smoothing algorithm, the Li-Ma method that calculates the significance, and finally shows the all-sky significance maps using the two HAWC data sets.

6.1 Background Estimation

The hadronic cosmic rays that pass the gamma/hadron cut are the main background for the analysis of gamma-ray sources. They are produced by distant sources and diffused by Galactic magnetic fields. Therefore the flux of hadronic cosmic rays is approximately isotropic and stable in time. The main factor that affects the background rate is the variations in the atmosphere and the detector response. The direct integration method described below is used to estimate the background from data that reflects these variations [49].

The direct integration method assumes that the local angular distribution (zenith and azimuth) of events is independent of the cosmic ray rate. It is applied to data for each Dec band (with the width of a HEALPIX pixel). The length of integration period Δt is defined and the data are split into integration periods. The procedure

of direct integration is demonstrated in Fig. 6.1 [49]. In each integration period, the all-sky event rate as a function of sidereal time τ in this period is $R(\tau)$. The angular distribution of events in hour angle (h) and declination (δ) [47] is normalized and forms a efficiency map $\epsilon(h, \delta)$. The estimated background as a function of RA α and Dec δ is the convolution of the efficiency map with the all-sky rate:

$$B(\alpha, \delta) = \int \epsilon(h, \delta) R(\alpha - h) dh \tag{6.1}$$

The final background map is the sum of the background map derived with data in each integration period. Using the direct integration method, the variations in the atmosphere and the detector that affect the data map are measured in the background map. Therefore the effect is canceled when the two maps are compared to search gamma-ray sources. The direct integration method reliably produces background with systematic errors of $< 10^{-3}$[50]. It also assumes that the cosmic-ray background dominate over gamma-ray events. Its limitations are a non-uniformly distributed cosmic-ray background, as well as strong localized and extended gamma-ray emission in higher f bins, where the gamma-to-hadron ratio greatly increases due to the gamma-ray selection cuts. These effects can be partly removed by from the background estimation by excluding region of interests, where strong gamma-ray signals are expected. This is yet to be applied to the analysis presented in this thesis.

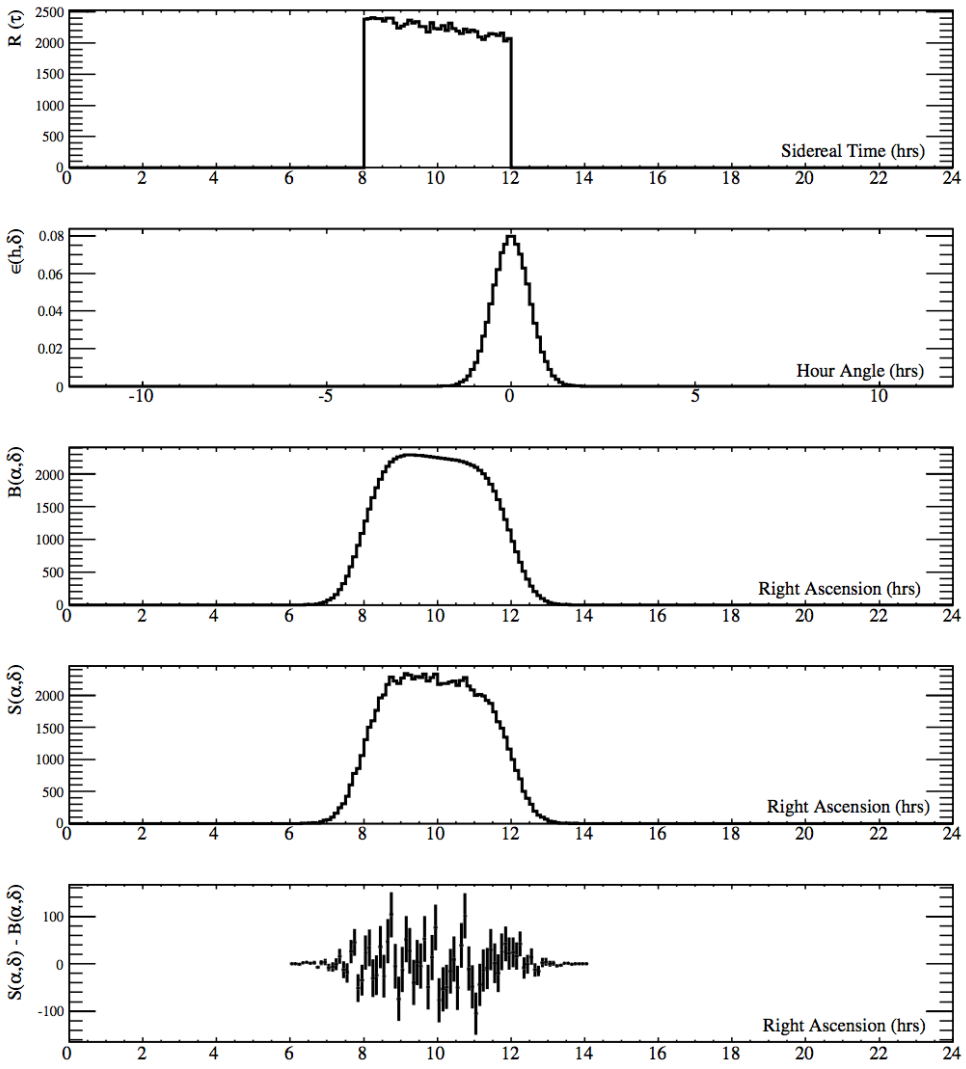

Figure 6.1 An example of direct integration in a fixed declination band. From top to bottom: the all-sky event rate $R(\tau)$ as a function of sidereal time in a 2 hr integration period; the normalized efficiency map $\epsilon(h, \delta)$ depending on hour angle; the estimated background $B(\alpha, \delta)$ as the convolution of $R(\tau)$ and $\epsilon(h, \delta)$; the raw data map $S(\alpha, \delta)$; the signal map $S(\alpha, \delta) - B(\alpha, \delta)$. [49]

Since the direct integration method is basically averaging the number of events in a RA range of $15^\circ\,\mathrm{hr}^{-1} \times \Delta t$ in a declination band, the method with integration period Δt (hr) is sensitive to the gamma-ray sources smaller than $15^\circ\,\mathrm{hr}^{-1} \times \Delta t$ while eliminating larger gamma-ray structures. For this analysis on point and extended

gamma-ray sources, a choice of $\Delta t = 2\,\mathrm{hr}$ is sufficient. Longer integration period may be used for the analysis on larger gamma-ray structures such as Fermi bubbles.

6.2 Map Smoothing

The HAWC detector has a finite angular resolution and the pixel size of raw data map is much small than the PSF. The map needs to be smoothed in order to enhance features with different sizes. Two smoothing functions are commonly used: Gaussian function, which often describes the PSF of a detector but is computational intensive; top-hat function, which simply integrates the events with in a certain radius. In this section, the method of performing a Gaussian smoothing is described.

For point source analysis, both raw data map and background map are convoluted with the PSF of the detector, referred to as map smoothing. The PSF of HAWC is described as a double-Gaussian function (see Fig. 4.13):

$$PSF = \frac{A \cdot G(\sigma_1) + B \cdot G(\sigma_2)}{A + B} \tag{6.2}$$

where $G(\sigma)$ is a normalized Gaussian function with width of σ, and A and B are the relative amplitude. On a two-dimensional sphere, the equivalent of a Gaussian

function is a Fisher-Von Mises function [51],

$$B(\theta) = \begin{cases} \frac{\kappa \exp(\kappa \cos\theta)}{4\pi \sinh(\kappa)} & \\ \frac{\kappa \exp(\kappa\theta^2/2)}{2\pi} & \text{for } \kappa \gg 1 \end{cases} \tag{6.3}$$

where $\kappa = 1/\sigma^2$. For small Gaussian width ($\kappa \gg 1$ or $\sigma < 0.1\,\text{radian}$), a Gaussian function is a good approximation of the Fisher-Von Mises function [51].

Since a Gaussian function (and the Fisher-Von Mises function) involves long tails that contribute to a large number of pixels in the map, the conventional convolution algorithm that sums the number of events from a large number of pixels is computational expansive. To reduce CPU time, a fast map smoothing algorithm in the sphere harmonic domain is performed. It only requires the calculation of the multipole moment up to the Nyquist frequency ($l_{max} \approx 2N_{\text{side}}$) [52].

A_{lm} is the spherical harmonic coefficients of an un-smoothed map. The spherical harmonic coefficients H_{lm} that corresponds to the smoothed map is calculated as:

$$H_{lm} = B_l A_{lm} \tag{6.4}$$

where the beam function (B_l) is obtained by integrating the smoothing kernel in

Equ. 6.3 over the sphere:

$$B_l = \int_0^{2\pi} \int_0^{\pi} B(\theta) P_l(\cos\theta) \sin\theta d\theta d\phi \tag{6.5}$$

where P_l is the Legendre polynomials. Given by [51], the beam function of the Fisher-Von Mises function is:

$$G_l = \exp(-l(l+1)/2\kappa) \quad \text{for } \kappa \gg 1 \tag{6.6}$$

and for large angles, a recursion relation is used to calculate the beam function term:

$$G_{l+1} = G_{l-1} - (2l+1)G_l/\kappa$$

$$G_0 = 1$$

$$G_1 = \coth(\kappa) - \frac{1}{\kappa}$$

As the PSF of HAWC follows a double-Gaussian function, the corresponding beam function is:

$$B_l = \frac{A \cdot G_l(\kappa_1) + B \cdot G_l(\kappa_2)}{A + B} \tag{6.7}$$

where $\kappa_1 = 1/\sigma_1^2$ and $\kappa_2 = 1/\sigma_2^2$ are related to two Gaussian components in Equ. 4.5. The spherical harmonic coefficients of the smoothed map H_{lm} are computed by substituting B_l in Equ. 6.4 by Equ. 6.7. The spherical harmonics coefficients are finally converted back to a smoothed map. This smoothing procedure is performed on the raw data map and background map in each f bin with the same PSF function corresponding to that bin.

6.3 Li-Ma Significance

The method of Li and Ma is used to calculate the statistical significance of the gamma-ray excess (or deficit) in each pixel of the sky [53]. The Li-Ma method is based on the likelihood ratio test and perfect for high-energy gamma-ray air shower detectors, with a large number of background events and a relatively small number of events from gamma-ray sources. The method calculates the statistical significance, S to be:

$$S = \sqrt{2N_{on}\ln(\frac{1+\alpha_{\text{Li-Ma}}}{\alpha_{\text{Li-Ma}}}\frac{N_{on}}{N_{on}+N_{off}}) + 2N_{off}\ln(\frac{(1+\alpha_{\text{Li-Ma}})N_{off}}{N_{on}+N_{off}})} \tag{6.8}$$

where N_{on} is the "on-source" counts, N_{off} is the "off-source" counts, and $\alpha_{\text{Li-Ma}}$ is the relative exposure of "on-source" over "off-source" observations. As the direct integration method uses the events from a large area of the sky to estimate the background, $\alpha_{\text{Li-Ma}}$ is usually $\ll 1$. The values of $\alpha_{\text{Li-Ma}}$ are calculated based on geometrical considerations and are given by:

$$\alpha_{\text{Li-Ma}} = \frac{S}{I - S} \tag{6.9}$$

where S is the source exposure and I is the total exposure of the background including the source region.

For a map that is smoothed with a Gaussian PSF of width σ, the source exposure is calculated as,

$$S = \pi(3.544\,\sigma)^2 \tag{6.10}$$

and the total exposure is calculated as,

$$I = \Delta t_{hrs} \times 15^{\circ}/hr \times \cos\delta \times 2(3.544\,\sigma) + \pi(3.544\,\sigma)^2 \tag{6.11}$$

where Δt is the integration period and $3.544\,\sigma$ is the effective bin radius of a Gaussian-smoothed bin compared to a top-hat smoothed bin. The PSF of HAWC is described by a double Gaussian function (see Equ. 4.5), and the corresponding effective Gaussian width σ' is

$$\sigma' = \frac{A+B}{A/\sigma_A + B/\sigma_B} \tag{6.12}$$

In practice, since the data are divided into ten f bins according to the shower size relative to the array size, these ten smoothed data maps (background maps) need to be combined before the calculation of Li-Ma significance. To combine the sky/background maps, a gamma-hadron weight, GHW, is applied to each map. The parameter GHW is proportional to the signal-to-background ratio at the position of the Crab nebula, to take into account different gamma-hadron separation power in each f bin. The maps from different data sets can be also combined with the proper value of GHW. As the PSF is different among f bins, the average $\alpha_{\text{Li-Ma}}$ is calculated from the average effective Gaussian width, given by:

$$< \sigma' >= \frac{\sum_{i}^{\text{f bins}} N_i \cdot GHW^2/\sigma'_i}{\sum_{i}^{\text{f bins}} N_i} \tag{6.13}$$

where N_i is the total number of events in the ith f bin.

In the absence of any signals, the Li-Ma significance should follow a normal distribution. This provides a useful check of the background estimation, map smoothing, and significance calculation. Fig. 6.2 shows the histogram of Li-Ma significances with HAWC-111 data. It follows a normal distribution. The long tail to the positive side shows the existence of sources.

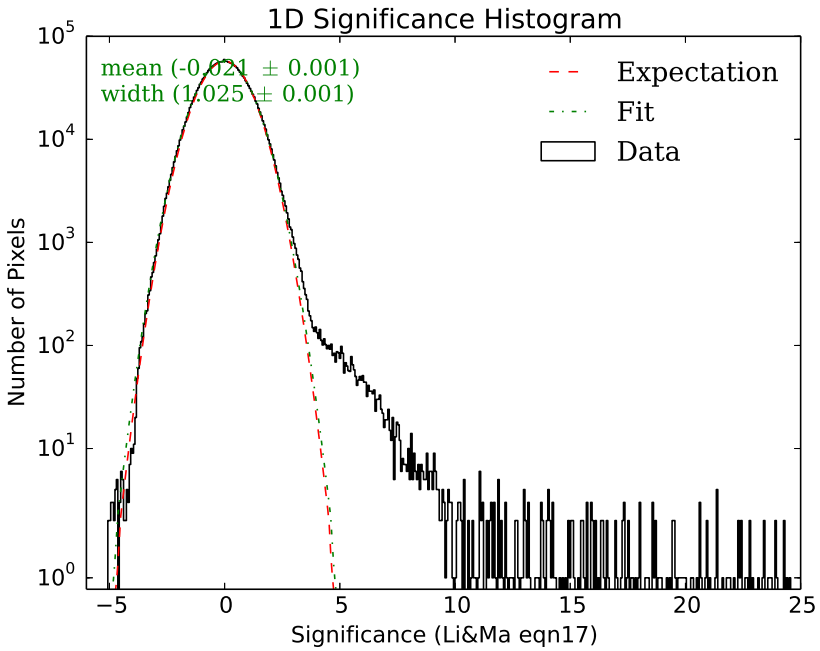

Figure 6.2 Histogram of Li-Ma significances with HAWC-111 data.

6.4 All-Sky Map

Using the Li-Ma method, the all-sky significance map is obtained. Fig. 6.3 and Fig. 6.4 show the significance map with HAWC-111 data in equatorial and Galactic coordinates [47]. The white regions in the maps are regions in the sky that HAWC is not sensitive on due to its location. Fig. 6.5 and Fig. 6.6 show the significance map with HAWC-250 data in equatorial and Galactic coordinates [47].

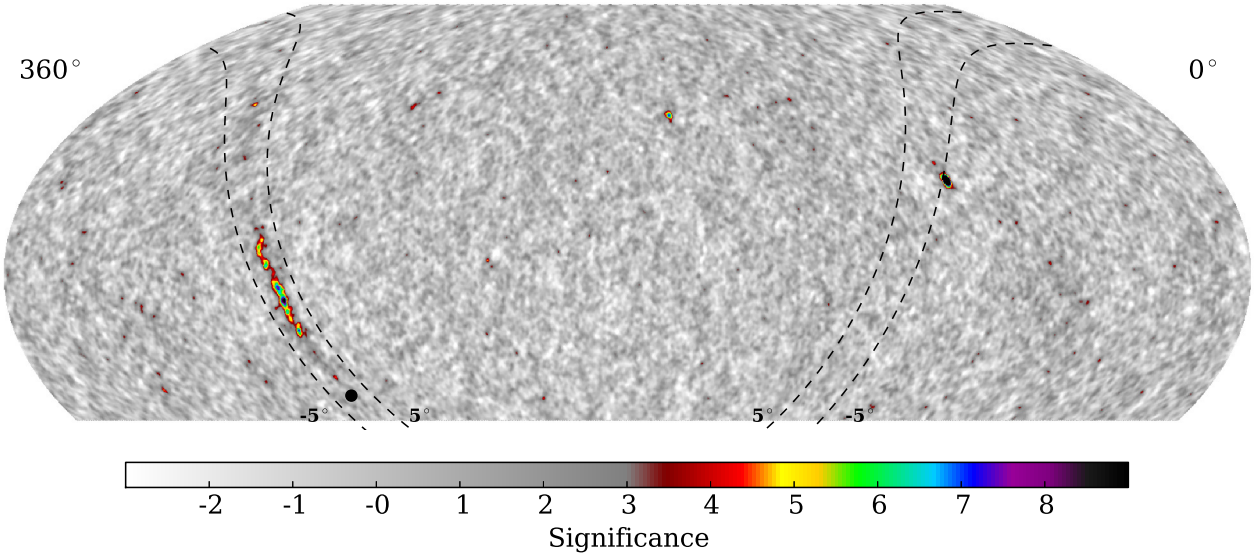

Figure 6.3 Significance map with HAWC-111 data in equatorial coordinates. The black dot marks the Galactic center and the dashed lines mark ±5° around the Galactic plane.

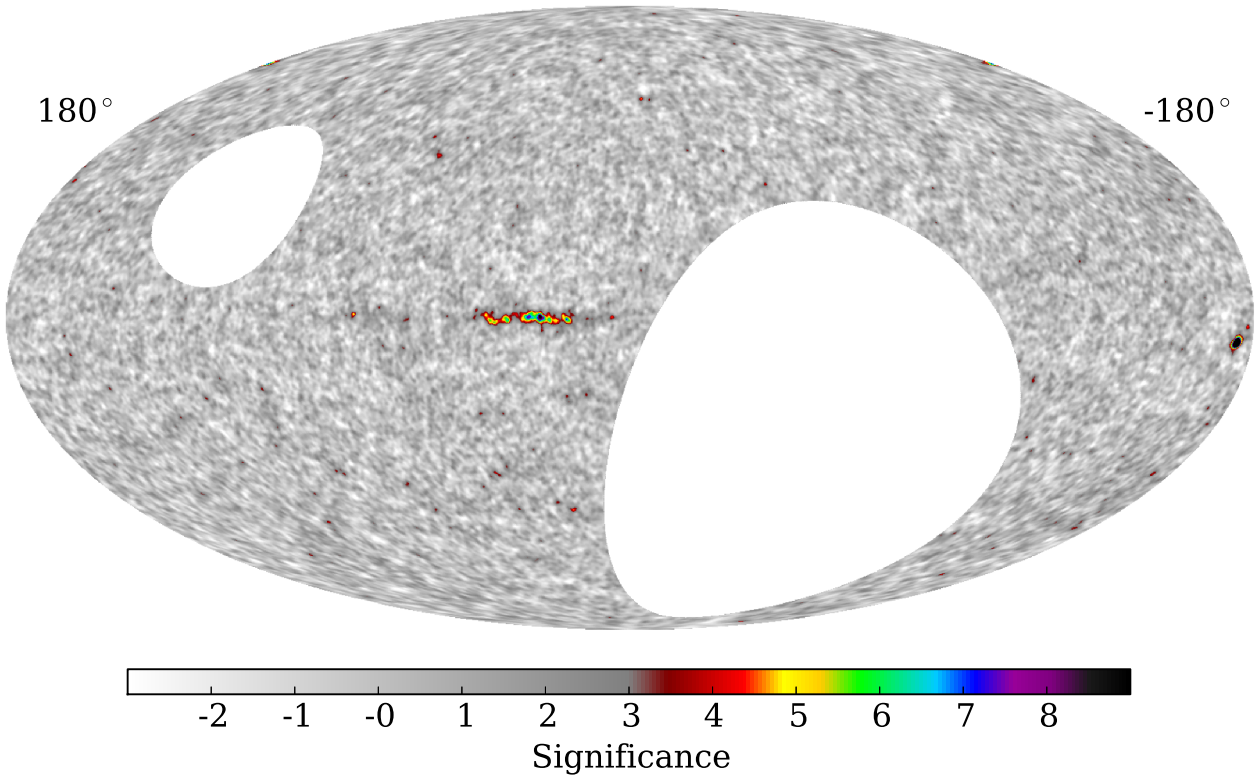

Figure 6.4 Significance map with HAWC-111 data in Galactic coordinates. The Galactic plane is in the middle of the map with the Galactic center at the center of the map.

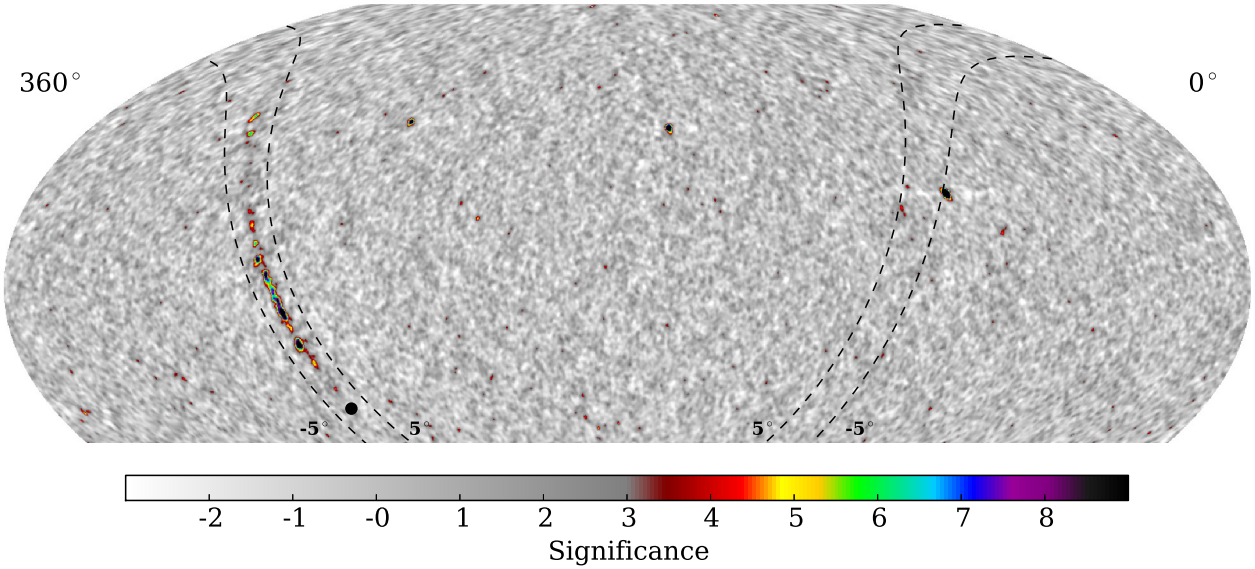

Figure 6.5 Significance map with HAWC-250 data in equatorial coordinates. The black dot marks the Galactic center and the dashed lines mark ±5° around the Galactic plane.

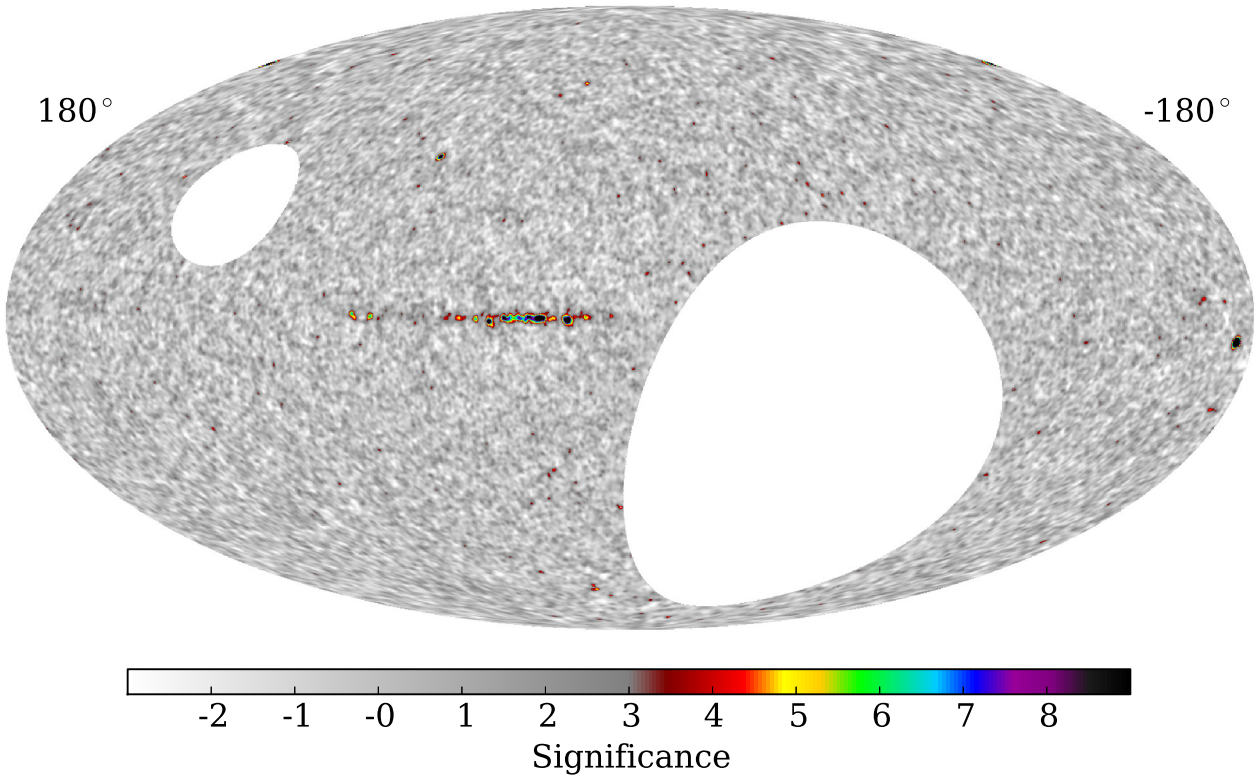

Figure 6.6 Significance map with HAWC-250 data in Galactic coordinates. The Galactic plane is in the middle of the map with the Galactic center at the center of the map.

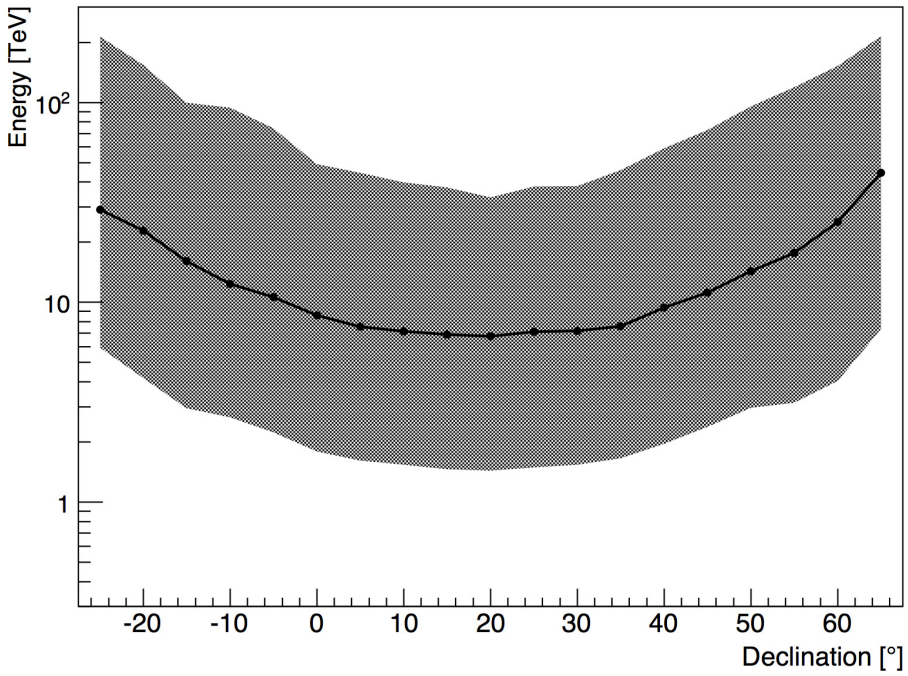

Figure 6.7 Median gamma-ray energy of the HAWC-111 data as a function of declination along with the 10% and 90% quantiles, derived from simulations and assuming a spectral index of -2.3.

The energy range in which the present HAWC-111 analysis is estimated using detector and air shower simulations. Since more inclined air showers travel through more atmosphere, the median gamma-ray energy of the HAWC-111 data is a function of declination. The declination dependence is shown in Fig. 6.7 for the spectral index assumption of 2.3, which represents the average indices for Galactic sources. The median energy increases from 7 TeV for a declination of $+19^\circ$ to $\sim$ 30 TeV for declinations of -26° and $+64^\circ$ assuming a spectral index of 2.3.

6.5 Limitation of the Li-Ma Method

The Li-Ma method is an ideal statistical test for data with large background and small signals, which are diverging from the background expectation. However, there are a few limitations of this method:

1. It only provides the significance assuming a point source. To test for source extent, maps need to be smoothed by different smoothing angles larger than the width of the PSF, which is time-consuming.

2. It calculates the significance of a putative point source. However, for source-rich regions, e.g. the Galactic plane, multiple sources may have overlapping contributions in the same pixel. The method cannot separate contributions from individual sources.

Chapter 7

The Maximum Likelihood Analysis

As there are certain limitations of the conventional Li-Ma method, an analysis based on the maximum likelihood method, which is similar the one used by *Fermi* [14], is developed and applied to HAWC data. In this chapter, the maximum likelihood method will be presented, which is an approach that is able to simultaneous fit the position, extent, and spectrum of multiple sources. In this chapter, a brief description of the maximum likelihood method is given, and the application of this method to HAWC-111 data will be discussed and tested.[1]

[1]Part of the contents presented in this chapter have been accepted the Astrophysical Journal [1].

7.1 The Likelihood Method

The maximum likelihood method is a well-known method of estimating the parameters of a model by applying it to data. For a fixed set of data and underlying statistical model, the method gives the set of parameters that maximizes the likelihood function. For example, suppose x_1, x_2, ..., x_n are n independent observations, and $\overrightarrow{\theta}$ is a set of parameters that characterizes the model. The likelihood is the probability of having such observations given the model with a certain set of parameters:

$$\mathcal{L}(\overrightarrow{\theta}\,|x_1, x_2, ..., x_n) = P(x_1, x_2, ..., x_n; \overrightarrow{\theta}) = \prod_{i=1}^{n} P(x_i; \overrightarrow{\theta}) \tag{7.1}$$

where $P(x_i; \overrightarrow{\theta})$ is the probability of having observation x_i given the model with parameters $\overrightarrow{\theta}$. As observation x_i often follows a certain distribution (a Poisson distribution in case of HAWC data), $P(x_i; \overrightarrow{\theta})$ can be calculated accordingly.

To apply the method to HAWC data, the first step is to build a source model characterized by the position and spectrum of the source. RA α and Dec δ are used to describe the source position. The source spectrum in the present analysis is assumed to follow a simple power law

$$\frac{dN}{dE} = I_0 \left(\frac{E}{E_0} \right)^{-\Gamma}, \tag{7.2}$$

where I_0 is the differential flux normalization, E_0 is the pivot energy, which is chosen at where the differential flux normalization is least dependent on the spectral index, and Γ is the spectral index. The study with HAWC-111 data assumes a fixed index of 2.3 due to the limited sensitivity in this data set. The index of 2.3 is representative of measured values for known Galactic objects [32]. The fit that allows spectral index floating is also tested with more recent HAWC-250 data with more statistics and improved reconstructions (see Section. 8.3).

The source model is convoluted with a detector response, established through Monte Carlo simulations, and converted to the expected gamma-ray signal in each pixel of each f bin. The maximum likelihood fit is performed by comparing the observed signal to the expected signal.

The region of interest (ROI) used for a likelihood fit needs to be larger than the angular resolution of the detector to include most photons from a given source. However, it is not always possible to find an ROI with photons from one single source due to a high potential of source confusion in the Galactic plane data taken with HAWC-111. Therefore, the source model may need to contain more than one source and in this case the expected event count becomes

$$\lambda_{ij} = b_{ij} + \sum_k \gamma_{ijk}, \tag{7.3}$$

where b_{ij} is the number of background events in the jth pixel of the ith f bin, and γ_{ijk} is the expected number of gamma rays from the kth source in the jth pixel of ith f bin. The event count is convolved with the detector response. As the observed event count in each pixel is distributed according to a Poisson distribution, the probability of observing N number of events given an expected count λ from the source model is

$$P(N;\lambda) = \frac{\lambda^N e^{-\lambda}}{N!}. \tag{7.4}$$

The likelihood given a parameter set $\overrightarrow{\theta} = (\alpha, \delta, I_0)$ in the source model is the product of the likelihood of each pixel in an ROI and in each f bin:

$$\mathcal{L}(\overrightarrow{\theta}\,|\overrightarrow{N}) = \prod_i^{\text{f bins}} \prod_j^{\text{ROI}} P(N_{ij};\lambda_{ij}). \tag{7.5}$$

where N_{ij} and λ_{ij} are the observed and expected event count in the jth pixel of the ith f bin, respectively. The logarithm of the likelihood is used for ease of computation:

$$\ln \mathcal{L}(\overrightarrow{\theta}|\overrightarrow{N}) = \sum_{i}^{\text{f bins}} \sum_{j}^{\text{ROI}} (N_{ij} \ln \lambda_{ij} - \lambda_{ij}), \tag{7.6}$$

The term $N_{ij}!$ is discarded from Eq. 7.6 since it is independent of the parameters in the source model. Multiple f bins from different data sets can be also used in the fit simultaneously and treated as independent analysis bins in the likelihood calculation. The log likelihood is maximized with respect to the parameter set $\overrightarrow{\theta}$ in the source model using the MINUIT package [54].

A likelihood ratio test is performed to decide how many sources are needed to properly model an ROI. To decide if the one-source model is preferred over the background-only model, the log likelihood of the background-only model $\ln \mathcal{L}_0$ is computed first. Then the log likelihood $\ln \mathcal{L}_1$ of the one-source model is computed. The test statistic (TS) defined by

$$TS = -2(\ln \mathcal{L}_0 - \ln \mathcal{L}_1) \tag{7.7}$$

is used to compare the goodness of one model over the other. The TS-value is converted to a p-value, which is the probability of the data being consistent with the background-only hypothesis. The same likelihood ratio test can be used to compare between two models with N and N+1 sources.

7.2 Source Searching Procedure

This section presents the procedure of searching for Galactic gamma-ray sources, which is applied to HAWC-111 data. The analysis focuses on the inner Galaxy region of Galactic longitude l$\in [15, 50]$ and Galactic latitude b$\in [-4, +4]$. This is a source-rich region with emission from point-like and extended emissions overlapping each other, as well as with diffuse emission from the Galactic plane. Source identification is a major challenge when analyzing the emission from this region. The maximum likelihood method is applied to the data to simultaneously fit the positions and spectra of multiple sources. With the limited statistics in HAWC-111 data, all sources are assumed to be point sources with a simple power law spectrum with a spectral index of 2.3.

7.2.1 Identification and Localization

Fig. 7.1 shows the significance map of the inner Galaxy region with HAWC-111 data. The map is made by moving a putative point source through each pixel, performing a maximum likelihood fit of differential flux normalization with the spectral index fixed at 2.3, and converting the TS value to a significance according to Wilks' theorem (With one free parameter in the model, TS follows a Chi square distribution with

1 degree of freedom). There are multiple $> 5\,\sigma$ hotspots in the surveyed region as shown in Fig. 7.1.

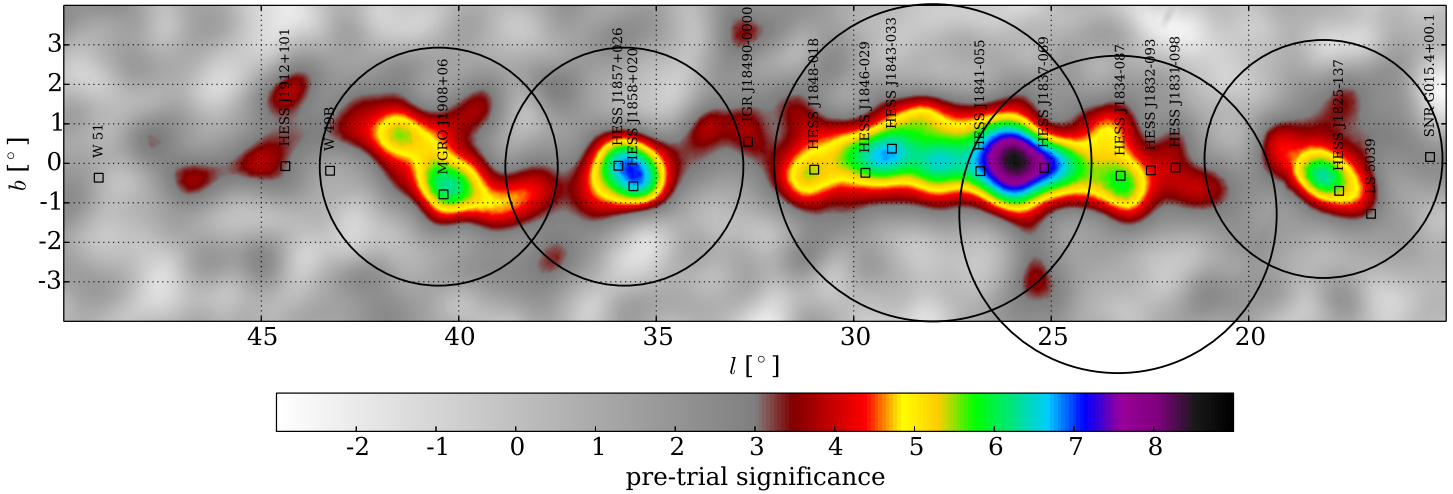

Figure 7.1 Significance map with five ROIs (large circles). The open squares mark TeV sources in the TeV Catalog [32].

Table 7.1. Definition of Five ROIs

Region	RA	Dec	Radius
1	286.4°	6.7°	3.0°
2	284.2°	2.5°	3.0°
3	280.7°	−4.1°	4.0°
4	279.6°	−9.2°	4.0°
5	275.9°	−13.1°	3.0°

The inner galaxy region is further divided into five ROIs to reduce the number of sources in each likelihood fit. This was done by visually checking the map to avoid a source cut by the edge of a ROI, although later study shows that the results of source searching do not heavily depend on the selection of ROIs. Table 7.1 shows the definition of five ROIs. In each of ROIs, a likelihood ratio test is performed to decide how many sources are needed to properly model the ROI. First, a putative point source is placed in each pixel to calculate the TS value over the background model. The pixel with the maximum TS is found and used as the initial position of the first source. Then a likelihood fit is performed to fit three free parameters, RA, Dec, and differential flux normalization, simultaneously. If $TS > 15$, the one-source model is preferred over the background-only model and the two-source model will be tested. This is an iterative process during which one more source is added into the model (three more free parameters) at each iteration step and the ΔTS of this model over the previous model is calculated. The procedure is repeated until $\Delta TS < 15$ with an additional source.

The ΔTS threshold of 15 was chosen *a priori*. ΔTS follows a Chi square distribution

with the degree of freedom (DoF) being equal to the difference between the number of free parameters in the two models, according to Wilks Theorem. But it is valid only under certain conditions. First of all, two models have to be nested, and either:

1. the additional parameters are not on the boundary of the parameter space in the null hypothesis;

   or

2. the additional parameters is restricted to one side of an (n-1) dimensional plane when the alternative (null) model has n (n-1) parameters.

Unfortunately, in this analysis the likelihood fit has three free parameters (RA, Dec, and differential flux normalization) and the differential flux normalization is at the its boundary (0) in the null hypothesis, so Wilks Theorem does not apply to this case. A toy Monte Carlo study is performed to check the ΔTS distribution. 2000 source-less pseudo event maps are generated based on the the measured background. The counts in each pixel of an event map is a random value from a Poisson distribution with the measured background counts as the mean. The one-source model with three parameters (same as the one used in the analysis presented before) is fit to each pseudo map in Region 1 and Region 5 as defined in Table 7.1 and the ΔTS value of the one-source model over the background-only model is computed. Fig. 7.2 shows the complementary cumulative distribution function (tail function) of ΔTS, which

is the probability of having a ΔTS value larger than a certain value. ΔTS of 15 corresponds to a p-value of 1.0% (0.4%) for Region 1 (Region 5). Conservatively, we conclude that the threshold of ΔTS at 15 has a $< 1.0\%$ p-value.

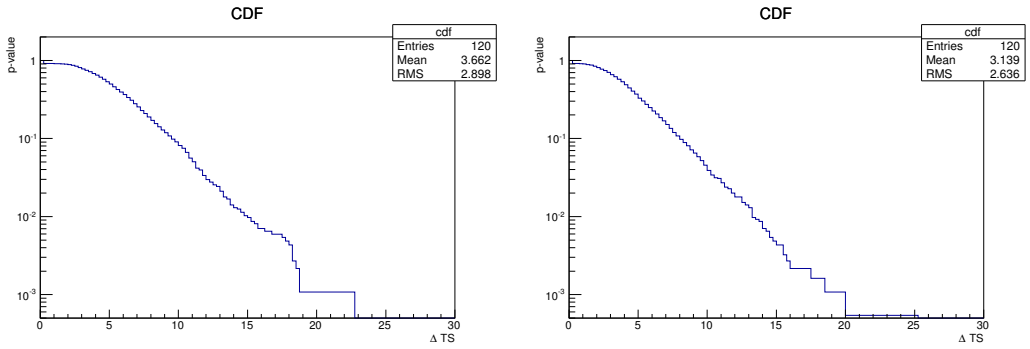

Figure 7.2 The complementary cumulative distribution function of ΔTS for Region 1 (left) and Region 5 (right).

7.2.2 Flux and Significance Determination

After the iteration in each ROI, a simultaneous fit of the differential source fluxes in all regions is performed while the source positions are fixed. This is done to take into account photons from a source with its centroid position just outside of an ROI but still contribute to the ROI due to the PSF of the detector. Finally, the differential flux and TS value of each source are obtained by fitting a single source while treating other sources as part of the background with fixed positions and differential fluxes.

7.3 Pivot Energy

In the analysis with HAWC-111 data, the spectral index is fixed at 2.3. The chosen of 2.3 represents the average spectral index for Galactic sources, however, the measured spectral indices of known sources spread approximately between 2.1 to 2.8. It is important to choose the pivot energy at where the differential flux normalization is least dependent on the spectral index. Studies based on the detector response are performed to find the best pivot energy. It first injects simulated sources with different spectral indices of 2.1 to 2.8 on source-less pseudo maps, then, using the likelihood method, fits are performed to find the flux of the simulated sources with the spectral assumption of 2.3. Finally the differential fluxes from the fit results are compared to the expected value of the simulated sources. Since a spectral assumption of 2.3 is used for all simulated sources, it is not possible to find one pivot energy at which the differential flux from the fit results are equal to the expected values. However, a best pivot energy can be found so that the systematic change in the fit differential flux for all simulated sources is minimized. As the energy response of the detector is a function of source declination, this study is repeated for different source declinations.

For the sources in the region of the Galactic plane analyzed here a spectral index of 2.3 is assumed and the best pivot energy varies from 4 TeV to 9 TeV depending on declination when a spectral index of 2.3 is assumed. In order to minimize the

variance in the flux normalization determination due to different spectral indices, different pivot energies are used for sources as a function of declination. According to simulations, this leads to a systematic change of -10% to $+20\%$ in the fitted differential flux for a source spectral index ranging between 2.1 and 2.8.

7.4 Number of Trials

For a source-searching analysis, it is important to understand the number of trials[55]. The significances at different locations in the inner Galaxy region are calculated. The distribution of significances in a source-less map follow a Gaussian function. If large number of tests are performed, it is expected to have high significance points due to statistical fluctuation in the background. Since the significance is calculated in each pixel of the map, if all pixels are independent, the number of trial is equal to the number of pixels in the maps. However, the PSF of the HAWC detector is larger than the pixel size of the maps, so the pixels are correlated. A similar toy Monte Carlo method that is described in section 7.2.1 is required to estimate the number of trials .

In this study, 1000 source-less pseudo maps are generated and the highest significance value in the inner Galaxy region from each map is found. For one trial, the distribution of significance follows a Gaussian function, so the probability of having a significance

value smaller than a given value Y is:

$$P(\sigma_i < Y) = \frac{1 + erf(Y/\sqrt{2})}{2} \tag{7.8}$$

where erf is the error function. For M independent trials, the probability to have all the values smaller than Y is:

$$\prod_{i=1}^{M}(P(\sigma_i < Y)) = (\frac{1 + erf(Y/\sqrt{2})}{2})^M \tag{7.9}$$

Then the probability to have at list one value greater than Y is:

$$1 - (\frac{1 + erf(Y/\sqrt{2})}{2})^M \tag{7.10}$$

The tail distribution of the highest significance values obtained in the region from pseudo maps is fit to Equ. 7.10. The resulting number of trials, M, is 424 ± 3. Fig. 7.3 shows the distribution of the highest significance values, the corresponding tail distribution, and the best fit results. The post-trial for each source is then calculated based on the new p-value, which is the p-value of the pre-trial significance multiplied by the number of trials.

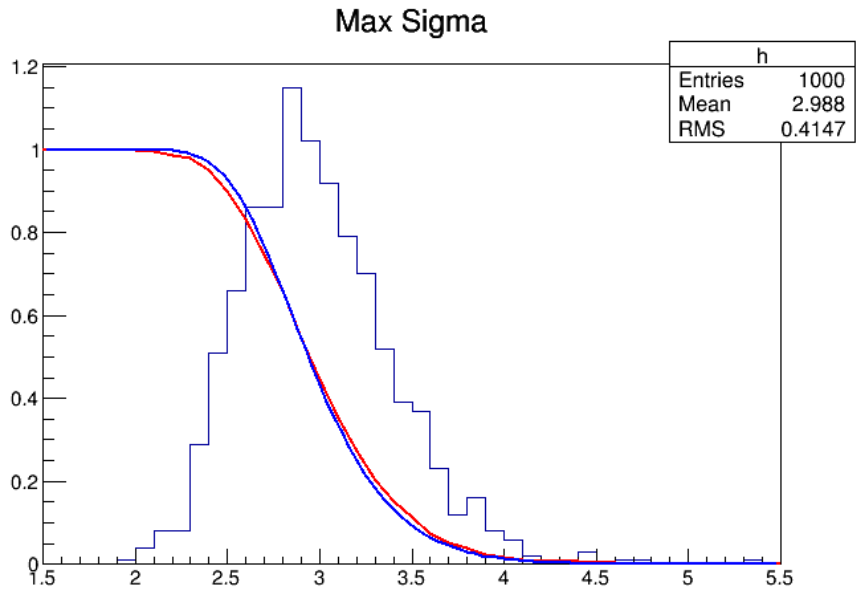

Figure 7.3 The distribution of the highest significance values (histogram), the corresponding tail distribution (red line), and the best fit results (blue line).

7.5 Testing the Method on Crab and Mrk421

The likelihood fit is tested using the data from the Crab nebula and the blazar Markarian (Mrk) 421. The pointing of the detector is also verified. Both source regions are relatively isolated in the sky and do not suffer overlap with additional sources. A likelihood fit with RA, Dec, and differential flux normalization is performed using a power law spectral assumption with the fixed spectral index of 2.6 for the Crab Nebula and 3.0 for Mrk 421. The results are summarized in Table 7.2. The Crab Nebula spectral index assumption is chosen based on previous IACT measurements, and for Mrk 421 a soft spectral index of 3.0 is chosen due to known spectral cutoff

Table 7.2. Pointing in J2000

Source	SIMBAD Database		HAWC-111			
	RA (°)	Dec (°)	RA[a](°)	Dec[a](°)	*TS*	significance
Crab	83.63	22.01	83.53 ± 0.06	22.06 ± 0.06	491.4	$22.2\,\sigma$
Mrk 421	166.11	38.21	166.22 ± 0.18	38.14 ± 0.18	69.0	$7.8\,\sigma$

that is not modeled here. The positions of both sources in the HAWC-111 data are consistent with measurements by IACTs (for example [56, 57]). Changing the spectral index used in the fit of the Crab Nebula between 2.0 and 3.0 shifts the best fit position by $< 0.07°$. The significance maps in the vicinity of the Crab Nebula and Mrk 421 are shown in Fig. 7.4. These maps are made by moving a putative point source through each pixel and performing a maximum likelihood fit of the differential flux normalization with spectral index fixed at 2.6 and 3.0, respectively. Then the TS value in each pixel is converted to significance according to Wilks' theorem.

The Crab Nebula is the most significant source in the HAWC-111 data. The spectrum of this source has been well measured by IACTs. The HAWC-111 data cover a similar gamma-ray energy range as the IACT data. The pivot energy for the Crab Nebula analysis is selected to be 4 TeV in order to minimize the dependence of the differential flux normalization on the spectral index in the fit. It differs from the pivot energy of 1 TeV used by Whipple, HEGRA, H.E.S.S., and VERITAS and 0.3 TeV used by MAGIC, so the IACT differential flux normalizations at 4 TeV are computed from the respective flux normalizations and spectral indices [22, 23, 24, 25, 56]. Table 7.5

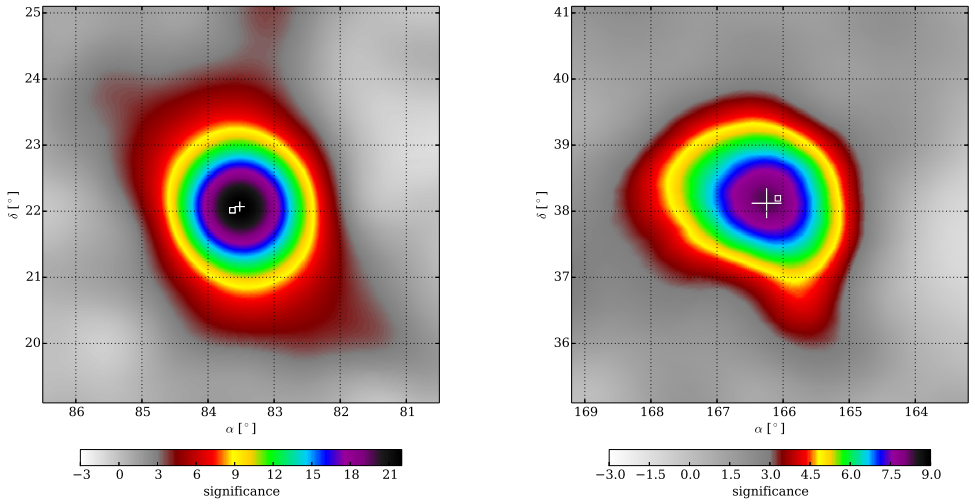

Figure 7.4 Significance maps of the Crab Nebula (left) and Mrk 421 (right). The white squares indicate the source positions from the TeV Catalog [32] and the white crosses are measured positions in the HAWC-111 data along with $1\,\sigma$ errors.

shows the differential flux normalization from the Crab Nebula with a spectral index assumption of 2.3, which is used for the analysis of the Inner Galaxy region, as well as the differential flux normalization with a spectral index assumption of 2.6, which is close to the index measured by IACTs. The differential flux of the Crab Nebula as measured in the HAWC-111 data is within 15% of the IACT measurements. The flux of Mrk 421 is beyond the scope of this analysis due to known variability of the blazar during the period covered by this dataset.

The determination of the differential flux normalization depends on the assumption of the spectral index, but the variance can be minimized by selecting the best pivot energy. For example, the derived Crab differential flux normalization increases by 5% if an index of 2.3 is chosen instead of 2.6 for a pivot energy of 4 TeV, while a pivot

Table 7.3. Differential Flux Normalization Comparison of the Crab Nebula Assuming a Simple Power Law

Instrument	Flux Normalization at 4 TeV (10^{-13} TeV^{-1}cm^{-2}s^{-1})	Reported Flux Normalization at 1 TeV (10^{-11} TeV^{-1}cm^{-2}s^{-1})	Spectral Index[a]
HAWC-111	8.57 ± 0.45	-	2.30[b]
HAWC-111	8.25 ± 0.40	-	2.60[b]
Whipple	10.1	$3.2 \pm 0.17 \pm 0.6$	$2.49 \pm 0.06 \pm 0.04$
HEGRA	7.49	$2.83 \pm 0.04 \pm 0.6$	$2.62 \pm 0.02 \pm 0.05$
H.E.S.S.	9.00	$3.45 \pm 0.05 \pm 0.7$	$2.63 \pm 0.01 \pm 0.09$
MAGIC	9.25	$57 \pm 2 \pm 6$[c]	$2.48 \pm 0.03 \pm 0.2$
VERITAS	8.83	$3.48 \pm 0.14 \pm 1.08$	$2.65 \pm 0.04 \pm 0.3$

[a]The first quoted uncertainty is statistical uncertainty and the second is systematic uncertainty.

[b]Assumed spectral index.

[c]At 300 GeV.

energy of 9 TeV results in a 30% increase in flux normalization.

There are three other major contributions to the systematic error of the flux normalization that have been studied using data and simulations:

1. Detector configuration variability. The number of active PMTs changed continuously in HAWC-111 dataset but only three configurations are modeled with the detector simulations. The event passing rate is found to vary $< 20\%$ among the three simulated configurations, which is equivalent to the resulting uncertainty on the flux estimate.

2. Angular resolution. The measured PSF on the Crab Nebula is used to compute the expected number of gamma rays in each pixel in a given model. The error in the measured PSF width from the Crab Nebula is $< 20\%$, which corresponds to 15-20% uncertainty in the flux estimate.

3. Charge scale. The core fitter and gamma/hadron cuts rely on the charge measurements by PMTs, which are based on the ToT-charge calibration. Up to 20% change in the flux estimate is observed due to the uncertainty in the charge scale estimated in studies of single muons with HAWC.

Each source of systematic uncertainties contributes $\sim 20\%$. In addition, a minor contribution of $\sim 8\%$ from the uncertainty of atmospheric modeling is taken into

account [31]. We add systematic uncertainties in quadrature for a total systematic uncertainty in the flux normalization of $\sim 40\%$.

Chapter 8

HAWC Source Catalog from a Survey of the Inner Galaxy Region

The results of the survey in the inner Galaxy region based on HAWC-111 data is presented in this chapter, followed by discussions for each source and their possible associations. The same method is also applied to HAWC-250 data as a systematic check. Finally a preliminary result on the spectrum of Crab is shown using the latest data with improved reconstructions. [1]

[1]Part of the contents presented in this chapter have been accepted by Astrophysical Journal [1].

8.1 Galactic Survey with HAWC-111 Data

Using the likelihood method that is described in Chapter 7, eleven seed sources are initially identified with $\Delta TS > 15$ criterion and are used in the source model. Fig. 8.1 and Fig. 8.2 show the model and residual map with the eleven seed sources. The residual map is derived by subtracting the model map (Fig. 8.1) from the data map (Fig. 7.1). After accounting for trials, ten source detections and candidates remain with $> 3\,\sigma$. "1HWC" is used as the identifier of this catalog, considering this is the first source catalog that is expected to be published with HAWC data. Table 8.1 lists the epoch J2000 positions, differential flux normalizations, TS, and the post-trials significances of the detections and candidates from this analysis. Table 8.2 lists the possible counterparts of each source and differential flux normalization comparison with known TeV sources. The discussion is separated into TeV source detections and source candidates below using a criterion of $5\,\sigma$ significance after trials.

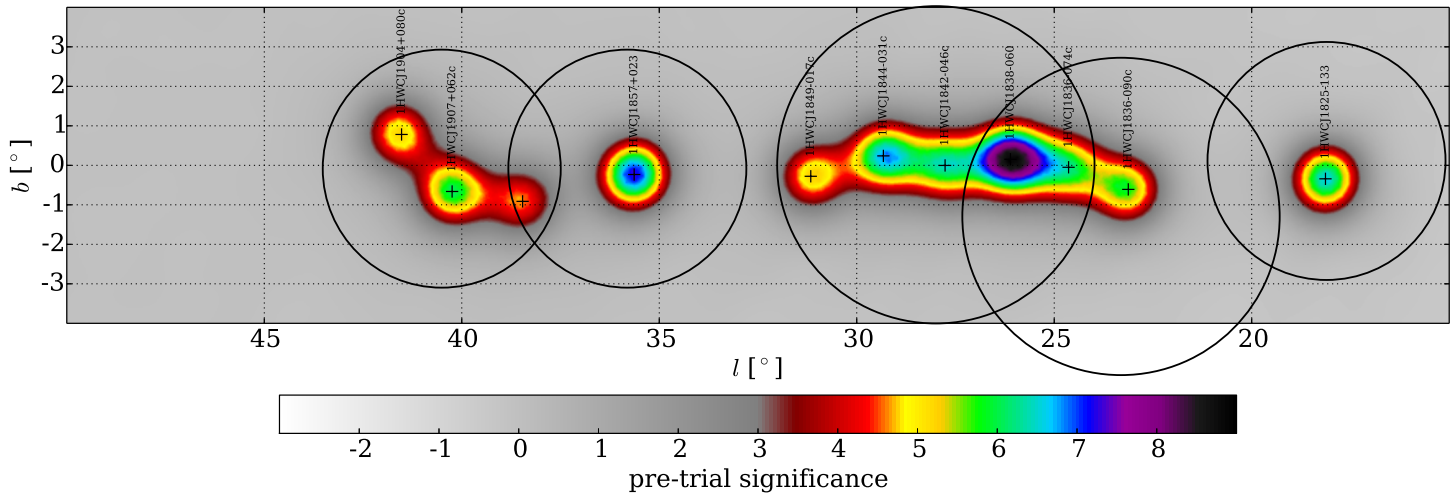

Figure 8.1 Model significance map with five ROIs (large circles). The crosses mark 1HWC sources.

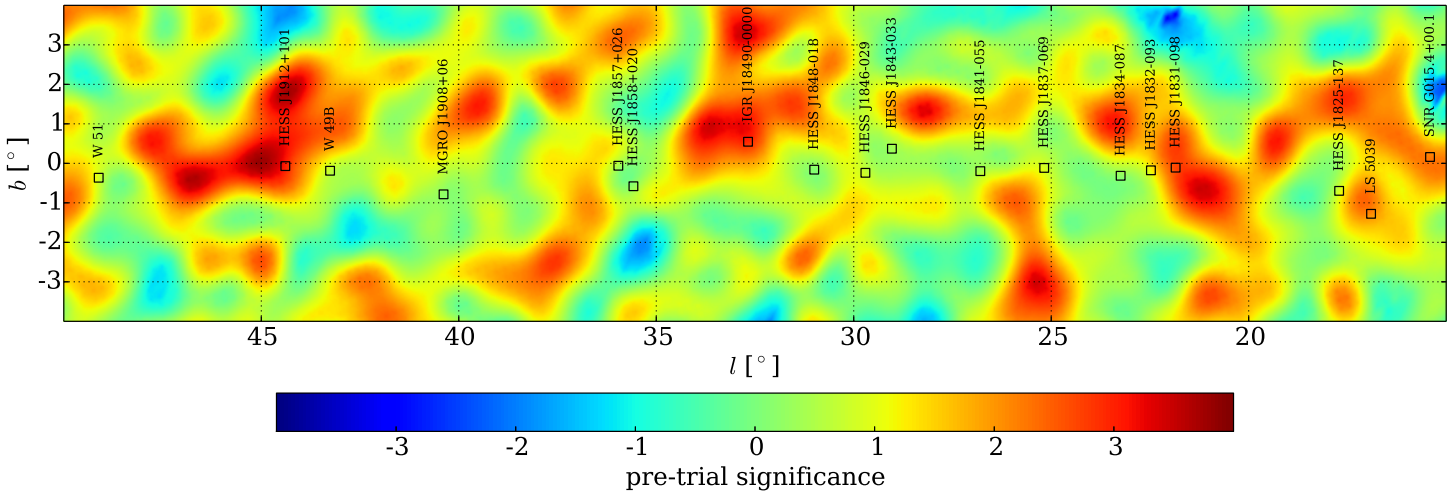

Figure 8.2 Residual significance map. The open squares mark TeV sources in TeV Catalog [32].

Table 8.1. Galactic Source Detections and Candidates with HAWC-111 Data

Region	Source[a]	ΔTS[b]	RA (°)[c]	Dec (°)[c]	l (°)[c]	b (°)	Differential Flux (Pivot Energy) (10^{-14}TeV^{-1}cm^{-2}s^{-1})	TS[c]	Post-trials Significance
1	1HWC J1907+062c	40.9	286.8±0.2	6.2±0.2	40.2±0.2	-0.7±0.2	22.0±4.6 (4 TeV)	32.8	4.6 σ
	1HWC J1904+080c	26.8	286.1±0.2	8.0±0.2	41.5±0.2	0.8±0.2	19.0±4.4 (4 TeV)	26.5	3.9 σ
	–	16.2	286.2±0.4	4.5±0.3	38.5±0.4	-0.9±0.4	N/A	17.2	2.5 σ
2	1HWC J1857+023	52.1	284.3±0.2	2.3±0.2	35.6±0.2	-0.2±0.2	18.0±3.0 (5 TeV)	50.2	6.2 σ
3	1HWC J1838-060	74.7	279.6±0.3	-6.0±0.2	26.1±0.3	0.2±0.3	11.3±1.2 (7 TeV)	48.9	6.1 σ
	1HWC J1844-031c	47.4	281.0±0.2	-3.1±0.2	29.3±0.2	0.2±0.2	11.8±2.4 (6 TeV)	33.7	4.7 σ
	1HWC J1849-017c	25.2	282.3±0.3	-1.7±0.2	31.2±0.3	-0.3±0.3	9.1±2.2 (6 TeV)	24.9	3.7 σ
	1HWC J1842-046c	23.7	280.5±0.3	-4.6±0.3	27.8±0.3	0.0±0.3	7.0±1.6 (7 TeV)	23.2	3.4 σ
4	–	70.7	279.7±0.2	-6.1±0.3	26.1±0.3	0.0±0.3	11.3±1.2 (7 TeV)	48.9	same source as J1838-060
	1HWC J1836-090c	33.6	278.9±0.3	-9.0±0.2	23.1±0.3	-0.6±0.3	5.8±1.3 (8 TeV)	26.6	3.9 σ
	1HWC J1836-074c	18.4	279.1±0.3	-7.4±0.3	24.6±0.3	0.0±0.3	6.9±1.4 (7 TeV)	22.0	3.2 σ
5	1HWC J1825-133	40.8	276.3±0.1	-13.3±0.2	18.1±0.2	-0.3±0.2	7.3±1.4 (9 TeV)	40.6	5.4 σ

[a] List of Galactic source detections and candidates. The positions reported here are in epoch J2000, and the differential flux normalization assumes a spectral index of 2.3. Only statistical uncertainties are quoted in this table.

[b] ΔTS of a model with one more source (ΔDoF=3) over the previous model.

[c] TS of a source over the background model while treating other sources as part of the background.

Table 8.2. Possible TeV Gamma-Ray Source Counterparts

Source	Possible Counterpart	Counterpart Classification	Distance to Counterpart (°)	Published Angular Extent (°)	Extrapolated Published Flux	Flux Normalization[a]	Pivot Energy (TeV)
1HWC J1907+062c	MGRO J1908+06	UID	0.38	< 2.6	36	22.0±4.6	4
	HESS J1908+063	UID	0.19	$0.34^{+0.04}_{-0.03}$	22.5		
	MGRO J1908+06 (ARGO)	UID	0.29	0.49 ± 0.22	61		
	MGRO J1908+06 (VERITAS)	UID	0.04	0.44 ± 0.02	20.0		
1HWC J1857+023	HESS J1857+026	UID	0.37	$(0.11 \pm 0.08) \times (0.08 \pm 0.03)$	13.0	18.0±3.0	5
	MAGIC J1857.2+0263	PWN	0.33	$(0.17 \pm 0.03) \times (0.06 \pm 0.03)$	16.6		
	HESS J1858+020	UID	0.35	$(0.08 \pm 0.02) \times (0.02 \pm 0.04)$	1.8		
1HWC J1838-060	HESS J1841-055 (ARGO)	UID	0.16	$0.40^{+0.32}_{-0.22}$	41	11.3±1.2	7
	HESS J1841-055	UID	0.77	$(0.41 \pm 0.04) \times (0.25 \pm 0.02)$	11.7		
	HESS J1837-069	PWN	0.97	$(0.12 \pm 0.02) \times (0.05 \pm 0.02)$	6.1		
1HWC J1844-031c	HESS J1843-033	UID	0.32	extended	N/A	11.8±2.4	6
	HESS J1846-029	PWN	0.61	point-like	1.1		
	ARGO J1841-0332	UID	0.87	point-like	N/A		
1HWC J1849-017c	HESS J1848-018	MSC[b]	0.20	0.32 ± 0.02	2.5	9.1±2.2	6
1HWC J1836-090c	HESS J1834-087	UID	0.31	point-like+(0.17 ± 0.01)	1.0	5.8±1.3	8
	HESS J1834-087 (MAGIC)	UID	0.41	0.14 ± 0.04	2.0		
1HWC J1836-074c	HESS J1837-069	PWN	0.55	$(0.12 \pm 0.02) \times (0.05 \pm 0.02)$	6.1	6.9±1.4	7
1HWC J1825-133	HESS J1825-137	PWN	0.55	$(0.23 \pm 0.02) \times (0.26 \pm 0.02)$	10.6	7.3±1.4	9

[a] The differential flux normalization in units of $10^{-14}\text{TeV}^{-1}\text{cm}^{-2}\text{s}^{-1}$, assumes a spectral index of 2.3. Only statistical uncertainties are quoted here.

[b] Massive Star Cluster

8.1.1 Uniform Surface Brightness

After the identification of 11 seed sources, a model containing these seed sources and a uniform surface brightness for the entire region was fitted simultaneously to the data. While for the sources a spectral index of 2.3 is assumed, the spectral index assumption for the uniform surface brightness is 2.5. The ΔTS of adding the uniform surface brightness as another free parameter to the source model is 33, i.e. $5.7\,\sigma$ that the uniform surface brightness component is preferred. The fitted surface brightness at 5 TeV is $(1.6 \pm 0.4) \times 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$, which is compatible with the average diffuse flux of $(1.0 \pm 0.2) \times 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$ reported by H.E.S.S. extrapolated to 5 TeV within the same region [58]. However, the uniform surface brightness measured in this dataset is not simply diffuse emission but also a combination of unidentified sources, source extensions, and photon contaminations from sources due to uncertainties in the PSF. As evident in the residual map of Fig 8.2, there are several $3\,\sigma$ regions around known TeV sources that are not detected in this dataset and are contributing to this uniform surface brightness fit. The uniform surface brightness fit is not included in Table 8.1. The contribution to the source differential flux normalization is $< 30\%$ of the smallest reported flux.

8.1.2 Source Detection: 1HWC J1857+023

The source 1HWC J1857+023 is detected at $6.2\,\sigma$ post trials and is $\sim 0.4^{\circ}$ away from both HESS J1857+026 and HESS J1858+020. These two TeV sources were discovered by the H.E.S.S. collaboration during their Galactic plane survey and are $\sim 0.7^{\circ}$ apart. The flux of HESS J1857+026 is approximately an order of magnitude higher than HESS J1858+020 [5], and the differential flux normalization from 1HWC J1857+023 is compatible with the combined flux of HESS J1857+026 and HESS J1858+020. Both of the HESS sources were detected as extended, with HESS J1857+026 as the larger of the two. HESS J1857+026 has energy dependent morphology [13] with two distinct components, MAGIC J1857.2+0263 and MAGIC J1857.6+0297. These sources cannot be resolved with these data from the partial HAWC array. The spectrum reported by MAGIC is for the entire region and is compatible with the differential flux normalization derived from this dataset.

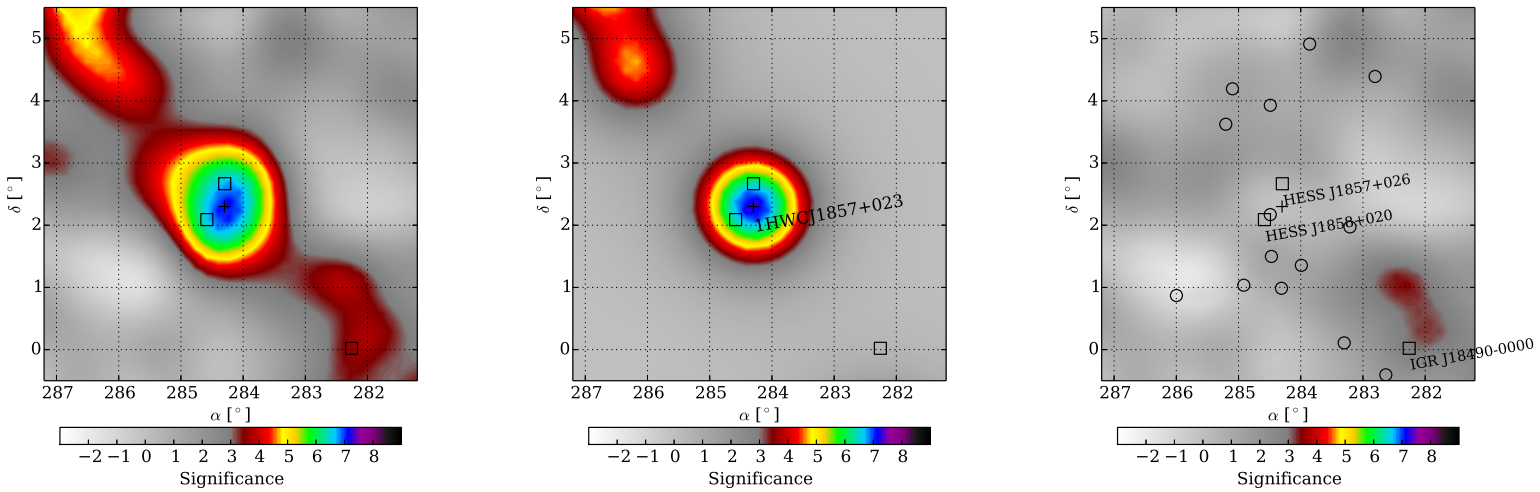

Figure 8.3 Data (left), model (middle), and residual (right) map for Region 2. 1HWC source/candidates are marked with crosses. TeVCat [32] sources and *Fermi* 3FGL [14] sources are marked with squares and circles, respectively.

8.1.3 Source Detection: 1HWC J1838-060

1HWC J1838-060 is detected at $6.1\,\sigma$ post trials and is located in the middle of the known TeV sources HESS J1837-069 and HESS J1841-055. This detection overlaps with the extension of HESS J1841-055, and the differential flux normalization is compatible with that reported by H.E.S.S.[5]. ARGO reported a $5.3\,\sigma$ detection [59] near the position of HESS J1841-055, with a 0.4° source extent and is closest to this detection. The ARGO source extends towards HESS J1837-069 and includes several *Fermi*-LAT sources. The flux reported by ARGO, when converted to differential flux at 7 TeV for comparison, is $\sim 4\times$ the differential flux normalization derived from this dataset.

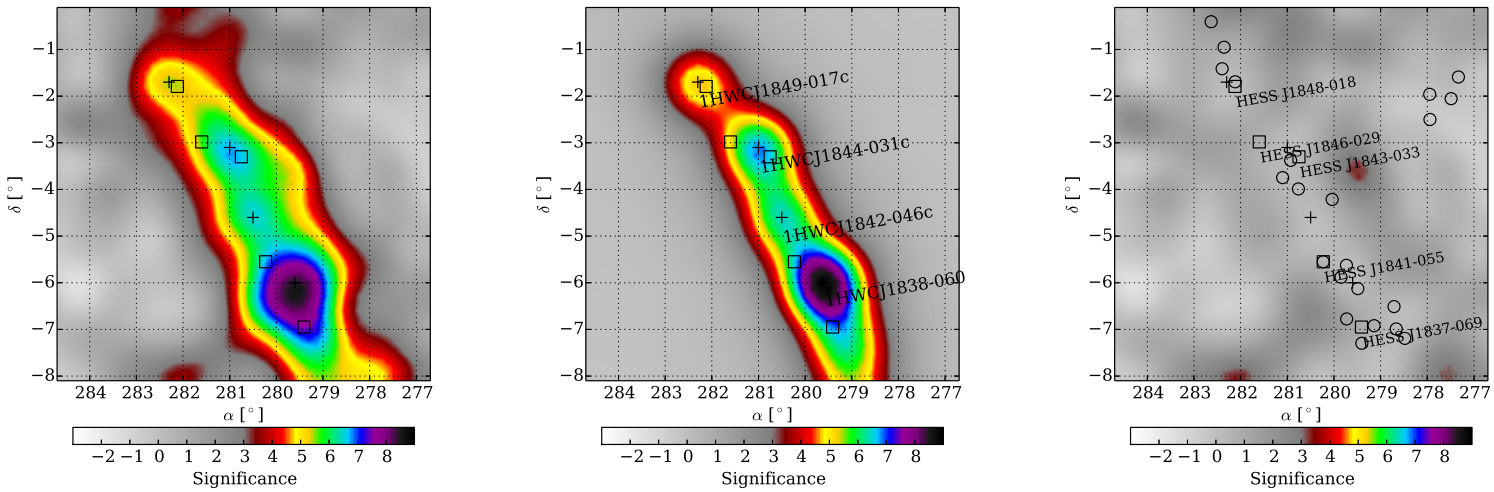

Figure 8.4 Data (left), model (middle), and residual (right) map for Region 3. 1HWC source/candidates are marked with crosses. TeVCat [32] sources and *Fermi* 3FGL [14] sources are marked with squares and circles, respectively.

8.1.4 Source Detection: 1HWC J1825-133

1HWC J1825-133 has a post-trials significance of $5.4\,\sigma$. It is $\sim 0.5^\circ$ to the south of the HESS J1825-137 centroid position, which is an extended PWN with spectral softening as a function of distance from the pulsar towards a southeast direction [60]. The simple power-law flux derived from this dataset is lower than the flux extrapolated from the simple power-law assumption measured by H.E.S.S. However, it is reported in [60] the spectrum is unlikely to be a simple power law and presented several alternative fits. The derived flux normalization from this dataset is most compatible with the energy dependent photon index power law fit by H.E.S.S. There is also a nearby PSR $\sim 0.4^\circ$ away, PSR J1826-1256, seen by *Fermi*-LAT [14] and associated with the Eel nebula [61].

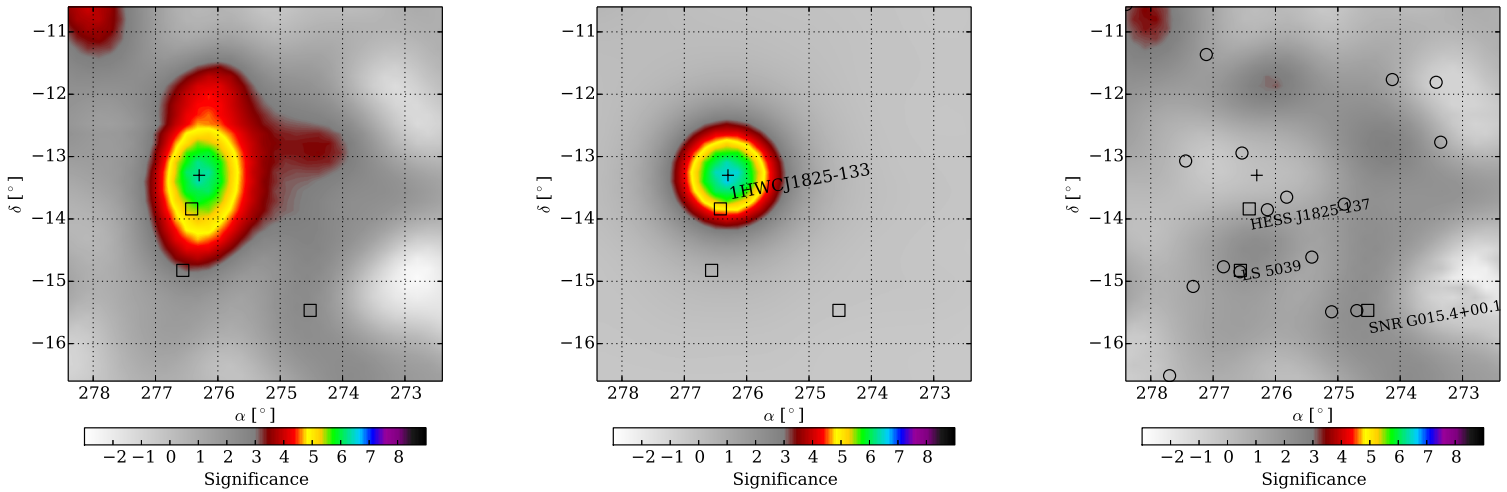

Figure 8.5 Data (left), model (middle), and residual (right) map for Region 5. 1HWC source/candidates are marked with crosses. TeVCat [32] sources and *Fermi* 3FGL [14] sources are marked with squares and circles, respectively.

8.1.5 Source Candidate: 1HWC J1907+062c

1HWC J1907+062c is 4.6 σ post trials with a best-fit position that is compatible with previously reported positions of MGRO J1908+06 (see [9, 12, 62, 63] for example). The differential flux normalization is consistent with the flux measured by H.E.S.S. and VERITAS and in agreement with Milagro given the statistical uncertainties of both instruments. Strong excess near the pulsar PSR J1907+0602 has been reported [12] but also extends toward SNR G 40.5-0.5. The Pass 1 dataset is not able to resolve the spatial morphology of this source.

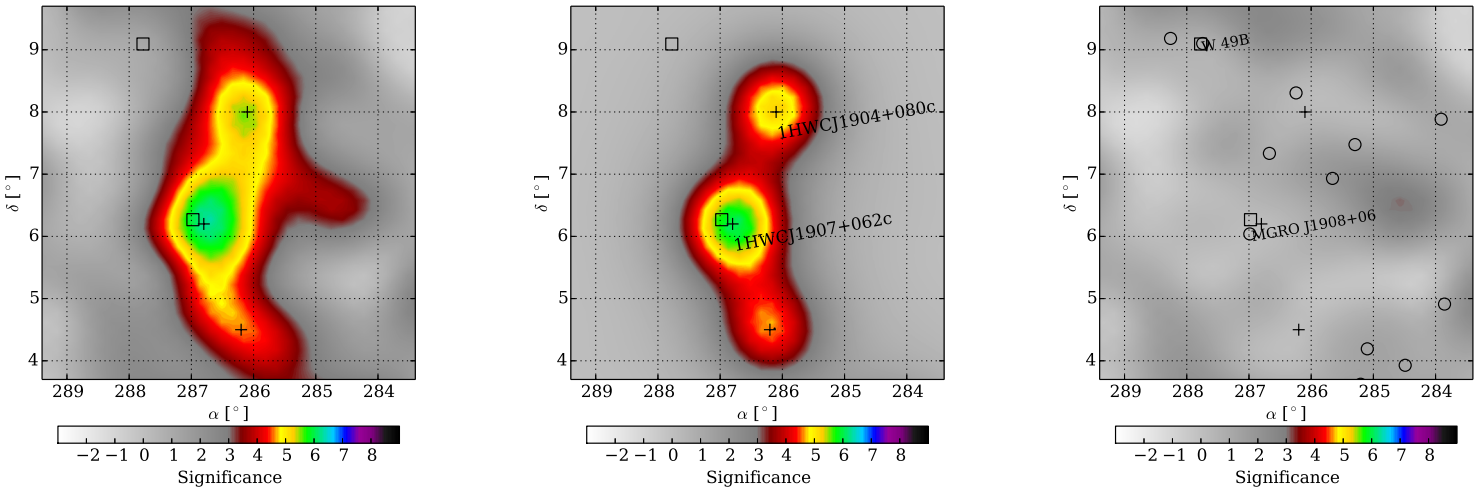

Figure 8.6 Data (left), model (middle), and residual (right) map for Region 1. 1HWC source/candidates are marked with crosses. One cross without source name represents the seed source $< 3\,\sigma$. TeVCat [32] sources and *Fermi* 3FGL [14] sources are marked with squares and circles, respectively.

8.1.6 Source Candidate: 1HWC J1904+080c

1HWC J1904+080c has a post-trials significance of $3.9\,\sigma$. There is currently no previously reported TeV detection near this location. The nearest gamma-ray source is 3FGL J1904.9+0818 at 0.3° away [14]. However, this is a weak detection from the *Fermi*-LAT 3FGL catalog, at $< 5\,\sigma$, with no known association.

8.1.7 Source Candidate: 1HWC J1844-031c

1HWC J1844-031c has a post-trials significance of $4.7\,\sigma$ and is spatially coincident with HESS J1843-033, which is classified as an unidentified source [6]. However, the morphology of this detection appears to extend towards HESS J1846-029, a pulsar wind nebula [7]. A $4.2\,\sigma$ excess has been reported [10], ARGO J1841-0332, associated with HESS J1843-033 despite being 0.7° away, due to the large systematic pointing error at high zenith angle.

8.1.8 Source Candidate: 1HWC J1849-017c

1HWC J1849-017c is detected at $3.7\,\sigma$ post trials and is positionally coincident with the extended source HESS J1848-018, which is possibly associated with the star forming region W43 [8]. The differential flux normalization at 6 TeV from this dataset is $\sim 3.5\times$ the flux reported by H.E.S.S. A index of 2.8 for this source is reported [8] and the spectral index assumption of 2.3 in this analysis would result in a different flux normalization by 20%. More importantly, diffuse emission from this star forming region that contains a molecular cloud could contribute more to the differential flux normalization derived from the Pass 1 dataset than that measured by IACTs, which have a smaller angular integration region.

8.1.9 Source Candidate: 1HWC J1842-046c

1HWC J1842-046c has a post-trials significance of $3.4\,\sigma$ and has no clear gamma-ray association. There is a nearby X-ray SNR G27.4+0.0 (Kes 73) at 0.4deg away and a PWN candidate G27.8+0.6 at 0.6deg away seen by XMM and Fermi-LAT (3FGL J1840.1-0412).

8.1.10 Source Candidate: 1HWC J1836-090c

1HWC J1836-090c is detected at $3.9\,\sigma$ post trials. It is spatially coincident with HESS J1834-087 and the SNR W41 [4, 11, 64]. The differential flux normalization from the Pass 1 dataset at 8 TeV is $\sim 6\times$ higher than the reported flux by H.E.S.S [64]. The source is reported by H.E.S.S. as having a central point-like component and an extended component. A similarly extended component is also seen by *Fermi*-LAT. The region contains a candidate pulsar at the center of the SNR W41, and two scenarios are supported [64]: PWN or SNR interaction with a nearby molecular cloud. The cloud density traced by ^{13}CO appears wider than the H.E.S.S. detection and may contribute to the increased flux seen in the Pass 1 data with HAWC.

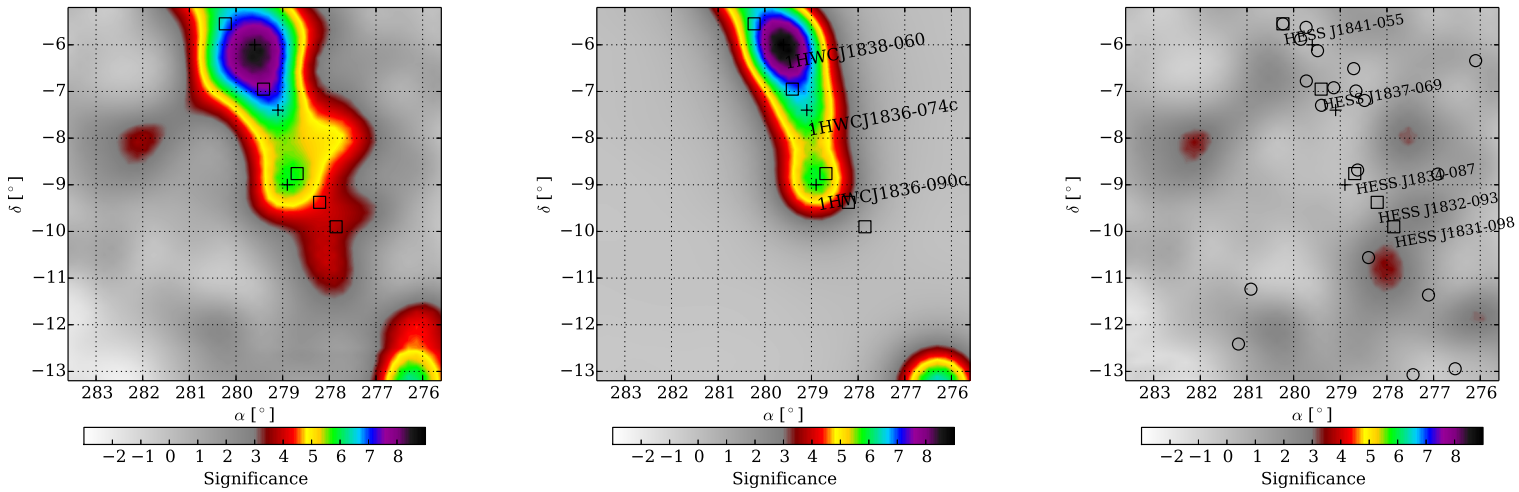

Figure 8.7 Data (left), model (middle), and residual (right) map for Region 4. 1HWC source/candidates are marked with crosses. TeVCat [32] sources and *Fermi* 3FGL [14] sources are marked with squares and circles, respectively.

8.1.11 Source Candidate: 1HWC J1836-074c

1HWC J1836-074c has a post-trials significance of $3.2\,\sigma$, with the nearest TeV PWN, HESS J1837-069 [4], $\sim 0.5^\circ$ away with a compatible differential flux. There is also a GeV source 3FGL J1837.6-0717 [14] that is $\sim 0.3^\circ$ away with no association.

8.2 Systematic Check with HAWC-250 Data

HAWC-250 data, as an independent data set, are used to provide a systematic check for the HAWC-111 results. This is a preliminary data set, on which the systematic uncertainty has not been studied well enough at the time this thesis is written. A similar likelihood analysis method is applied to this data set in the same region. Instead of selecting ROIs based on the significance map, in this data set, ROIs are 3° circles centered at Galactic latitude of 0° and every 5° of Galactic longitude from 15° to 50°. Fig. 8.8 shows the significance map obtained with HAWC-250 data. Eight ROIs are marked with black circles.

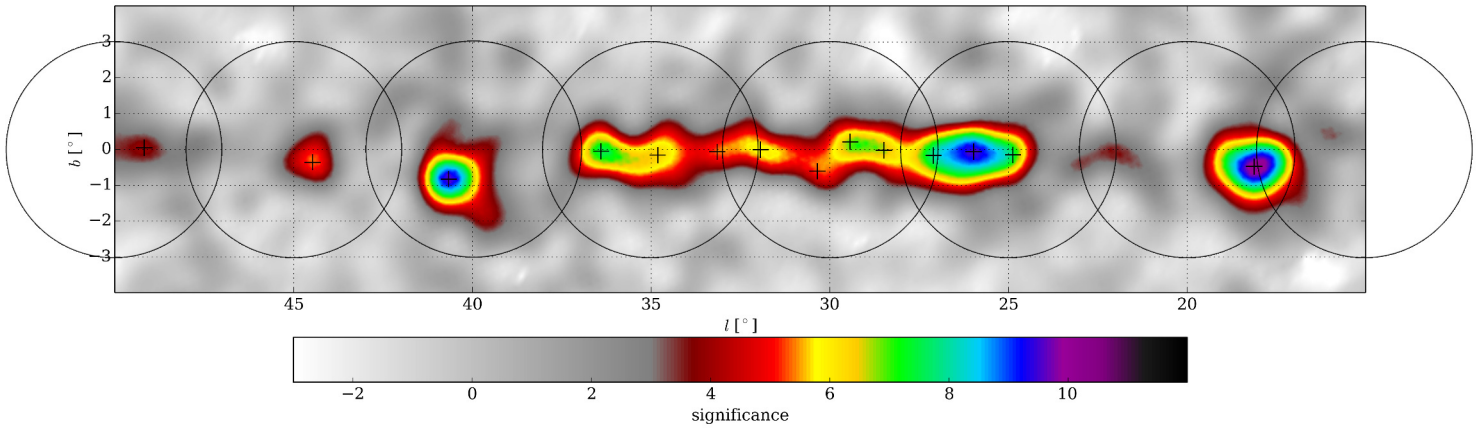

Figure 8.8 HAWC-250 significance map with eight ROIs (large circles). The crosses mark 14 identified sources.

Fourteen seed sources (the positions are shown with the crosses in Fig. 8.8) are identified in eight ROIs. Table 8.3 summarizes the positions and fluxes. Since the study on the pivot energy is yet to be done for HAWC-250 data set, all the differential fluxes reported in Table 8.3 are at a pivot energy of 8 TeV (spectral indices are fixed at 2.3). It is worth noting that the source in Region 8 (near 1HWC J1825-133) appears to be extent in HAWC-250 data. The likelihood fit on the source extent is performed and the result is $0.56 \pm 0.11°$. Eight of ten 1HWC sources/candidates have nearby HAWC-250 seed sources. However, a small offset on source position may lead to large change on the flux in a source-busy region. It is not straightforward to compare the fluxes between two two data sets. To first order, the sum of fluxes in a region, e.g. in each of the five ROIs that is used in HAWC-111 analysis, is similar between these two data sets.

Table 8.3. Galactic Source Detections and Candidates with HAWC-250 Data

Region	Source	ΔTS[a]	RA (°)	Dec (°)	Differential Flux at 8 TeV (10^{-14}TeV^{-1}cm^{-2}s^{-1})	Note
1	a	16.3	290.4±0.2	14.4±0.2	1.3±0.7	
2	a	30.6	288.5±0.1	10.1±0.1	1.9±0.7	
3	a	96.3	287.1±0.1	6.5±0.2	3.8±0.7	
4[b]	a	60.5	284.5±1.0	3.1±1.0	3.1±0.3	
	b	35.5	283.9±1.0	1.6±1.0	2.4±0.2	
	c	26.4	283.0±1.0	0.2±1.0	1.9±0.2	
	d	16.5	282.0±1.0	-1.0±1.0	8.4±0.2	near the edge, same as 5-b
5[b]	a	52.8	281.1±1.0	-3.0±1.0	2.4±0.3	
	b	37.4	282.4±1.0	-0.9±1.0	2.6±0.3	
	c	31.0	280.3±1.0	-5.3±1.0	3.6±0.3	near the edge, same as 6-b
	d	20.1	282.2±1.0	-2.6±1.0	1.9±0.2	
	e	15.5	280.9±1.0	-4.0±1.0	2.4±0.1	
6	a	93.4	279.7±0.4	-6.2±0.6	3.6±1.2	
	b	44.3	280.3±0.4	-5.3±0.3	3.3±1.2	
	c	29.8	279.3±0.4	-7.2±0.3	3.0±1.2	
7	a	132.4	276.4±0.2	-13.4±0.2	11.0±2.4	0.56° extent

Table 8.3 (cont'd)

Region	Source	ΔTS[a]	RA (°)	Dec (°)	Differential Flux at 8 TeV (10^{-14}TeV^{-1}cm^{-2}s^{-1})	Note
8	a	81.9	276.0±0.2	-13.8±0.3	4.9±1.3	near the edge, same as 7-a

[a] ΔTS of a model with one more source (ΔDoF=3) over the previous model.

[b] MINUIT fails to estimate the uncertainty on positions in this region possibly due to too may free parameters in the fit.

8.3 Fitting the Crab Spectrum with HAWC-250 Data

Recently, various parts of reconstruction algorithms have been revisited and improved using Crab as a benchmark source, including core reconstruction, curvature correction, and gamma/hadron separation. In the preliminary results, using 211 days of HAWC-250 data, the PSF measured on the Crab nebula is confirmed to be consistent with the predicted value from the detector simulations. The events with their cores on the array have better PSF than the ones with cores off the array. In this study, only events, coming from the direction of Crab, with their cores on the array are used to perform a likelihood fit on flux normalization and spectral index. The results, $I_0 = 3.34 \pm 0.22 \times 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ at 1 TeV and $\Gamma = 2.60 \pm 0.04$, are in good agreement with the spectrum H.E.S.S. reported [24], $I_0 = 3.45 \pm 0.05 \times 10^{-11}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ at 1 TeV and $\Gamma = 2.63 \pm 0.01$. Fig. 8.9 shows that the excess from Crab, using 70% containment radius, in each f bin compared with the predicted excess using the detector response with the Crab spectrum measured by H.E.S.S. and with the best fit spectrum with HAWC-250 data. The gray band represents 40% of systematic uncertainties, adopted from HAWC-111 analysis, on the predicted excess with the Crab spectrum measured by H.E.S.S..

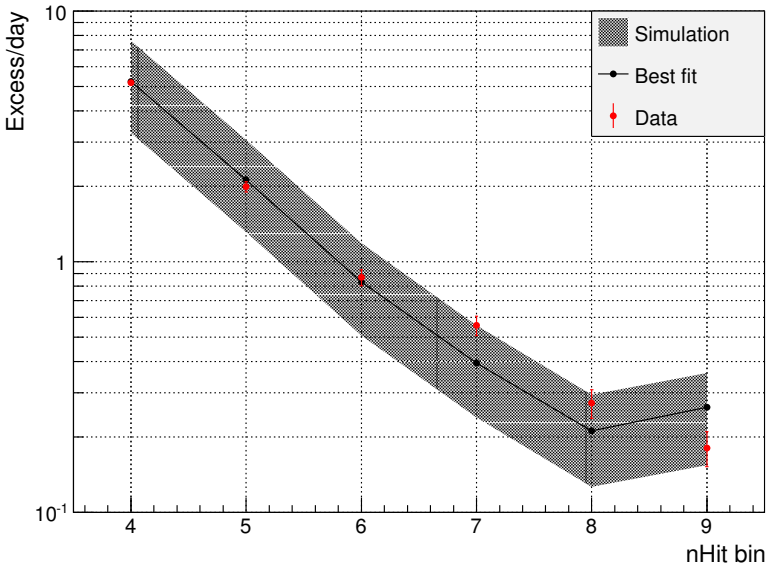

Figure 8.9 Red: excess from Crab, using 70% containment radius, in each *f*. Black: predicted excess using the detector response with the best fit spectrum with HAWC-250 data. Gray band: predicted excess with the Crab spectrum measured by H.E.S.S. with 40% of systematic uncertainties, adopted from HAWC-111 analysis,.

Chapter 9

Conclusion

9.1 Summary

A survey of the inner Galaxy has been presented in the region of Galactic longitude $l \in [15^\circ, 50^\circ]$ and latitude $b \in [-4^\circ, +4^\circ]$ using 283 days of data with the partial HAWC Gamma-Ray Observatory from August 2013 to July 2014. Three sources have been detected at $> 5\,\sigma$ with an additional seven candidate sources detected at $> 3\,\sigma$ after accounting for trials. While associations with previously published IACT detections are not always within the experimental uncertainties, about half of them have differential flux normalizations that are compatible with the previous detections.

A likelihood method similar to the *Fermi*-LAT source-finding algorithm has been

applied to data from an extended air shower array for the first time to properly address challenges arising from source identification. The point sources presented here have differential fluxes $> 20\%$ of the Crab Nebula flux at several TeV.

9.2 Future

HAWC was inaugurated in March, 2015 and the full array is now continuously operating. With more data taken with the full array and recently improved reconstruction algorithms, the significance of the Crab nebula above the background increases from $3.1\,\sigma/\sqrt{\text{day}}$ to $5.5\,\sigma/\sqrt{\text{day}}$. These more sensitive data sets will be used to confirm the new sources/candidates that are discovered with HAWC-111 data. This also will lead to a more sensitive survey of TeV gamma-ray sky and more new source discoveries. For example, in the source-finding algorithm that is presented in this thesis with HAWC-111 data, sources are considered as point sources and all spectral indices are fixed at 2.3. With more sensitive data, it will be possible to consider these properties as free parameters as well.

HAWC data with better angular resolution and higher sensitivity allow more precise spectral and morphological studies on individual sources. The maximum likelihood method that is presented in this thesis is an important tool not only to fit multiple sources simultaneously but also to fit the spectrum and morphology of sources.

HAWC data, with their highest energy reach, provide crucial information on the emission processes at the sources, e.g. if they are of the leptonic or hadronic type, and ultimately on the nature of the sources and on how they accelerate particles. High energy observations are also directly connected to the highest energies which a source is able to accelerate particles to, e.g. Pevatrons. Combining data taken with different instruments is very important, which provides a whole picture of a source spectrum. A Multi-Mission Maximum Likelihood (3ML) [65] framework is currently under development to preform joint fits on spectra and morphologies using data taken with different instruments.

Higher energy gamma-ray air showers, due to their larger footprints, have more chance to have their cores outside of the array. It is very difficult to reconstruct the core positions accurately for these events thus the angular resolution is limited. A sparse outrigger array of HAWC is designed to address this issue. The outrigger array covers a large area ($\sim 4\times$ of the current HAWC array size) with a relatively low fraction of ground covered with detectors. This makes it a great complement of the HAWC observatory to help reconstruct the shower cores that are outside the array. The outrigger array will greatly increase the sensitivity of HAWC for gamma-ray energies $> 10\,\mathrm{TeV}$.

References

[1] Abeysekara, A. U.; others. *ArXiv Astrophysics e-prints: 1509.05401* **2015**.

[2] Aharonian, F.; others. *Reports on Progress in Physics* **2008**, *71*, 56.

[3] Lieberman, M. A. and Lichtenberg, A. J.. *Physical Review A* **1972**, *5*, 1852.

[4] Carrigan, S. and others. In *Proceedings of 33th ICRC*, 2013.

[5] Aharonian, F.; others. *Astronomy and Astrophysics* **2008**, *477*, 353.

[6] Hoppe, S. In *Proceedings of 30th ICRC*, , 2008.

[7] Djannati-Ataĭ, A. In *Proceedings of 30th ICRC*, , 2008.

[8] Chaves, R. C. G.; Renaud, M.; Lemoine-Goumard, M. *American Institute of Physics Conference Series* **2008**, *1085*, 372.

[9] Abdo, A. A.; others. *Astrophysical Journal Letters* **2007**, *664*, L91.

[10] Bartoli, B. and others. *ArXiv e-prints:1311.3376* **2013**.

[11] Albert, J. and others. *Astrophysical Journal Letters* **2006**, *643*, L53.

[12] Aliu, E. and others. *Astrophysical Journal* **2014**, *787*, 166.

[13] Aharonian, F.; others. *Astronomy and Astrophysics* **2008**, *477*, 353.

[14] The Fermi-LAT Collaboration. *ArXiv e-prints: 1501.02003* **2015**.

[15] Ackermann, M. and others. *Astrophysical Journal Supplement Series* **2013**, *209*, 34.

[16] Kraushaar.; Clark. *Physics Review Letters* **1962**, *8*, 106.

[17] Fichtel.; others. Catalog of sas-2 gamma-ray observations Technical report, NASA, **1990**.

[18] Swanenburg, B. N.; others. *Astrophysical Journal* **1981**, *243*, 69.

[19] Hartman, R. C.; others. *The Astrophysical Journal Supplement Series* **1999**, *123*, 79.

[20] Baillon, P.; others. *Astroparticle Physics* **1993**, *1*, 341.

[21] Goret, P.; others. *Astronomy & Astrophysics* **1993**, *270*, 401.

[22] Hillas, A. M.; others. *Astrophysical Journal* **1998**, *503*, 744.

[23] Aliu, E.; others. *Astrophysical Journal Letters* **2014**, *781*, 11.

[24] Aharonian, F.; others. *Astronomy & Astrophysics* **2006**, *457*, 899.

[25] Albert, J.; others. *Astrophysical Journal* **2008**, *674*, 1037.

[26] Amenomori, M.; others. *Physics Review Letters* **1992**, *69*, 2468.

[27] Atkins, R.; others. *Physical Review Letters* **2005**, *95*, 251103.

[28] Abeysekara, A. U.; others. *Astroparticle Physics* **2013**, *50*, 26.

[29] U.S. Standard Atmosphere 1976. COESA. **1976**.

[30] Large Photocathode Area Photomultiplier Tubes. Hamamatsu.

[31] Abeysekara, A. U.; others. *Astrophysical Journal* **2015**, *800*, 78.

[32] Online catalog for TeV Astronomy. Wakely, S. P.; Horan, D. **2008**.

[33] Abdo, A. A.; others. *Astrophysical Journal* **2012**, *753*, 159.

[34] Abdo, A. A.; others. *Astrophysical Journal Letters* **2009**, *700*, 127.

[35] Su, M.; Slatyer, T. R.; Finkbeiner, D. P. *Astrophysical Journal* **2010**, *724*, 1044.

[36] Abeysekara, A. U.; others. *Astrophysical Journal* **2014**, *796*, 108.

[37] Abeysekara, A. U.; others. *Astroparticle Physics* **2015**, *62*, 125.

[38] Heck, D.; others. CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers Technical report, **1998**.

[39] Agostinelli, S.; others. *Nuclear Instruments and Methods in Physics Research* **2003**, *A506*, 250.

[40] Greisen, K. *Progress in Cosmic Ray Physics* **1956**, *3*, 1.

[41] Kamata, K.; Nishimura, J. *Progress in Theoretical Physics Supplement* **1958**, *6*, 93.

[42] W-M, Y. *Journal of Physics G: Nuclear and Particle Physics.* **2006**, *33*, 1–1232.

[43] Atkins, R.; others. *Nuclear Instruments and Methods in Physics Research A* **2000**, *449*, 478.

[44] Salesa, F. In *Proceedings of 34th ICRC*, The Hague, The Netherlands, 2015.

[45] Salesa, F.; others. In *Proceedings of 34th ICRC*, The Hague, The Netherlands, 2015.

[46] H$^{'}$untemeyer, P.; Matthews, J.; Dingus, B. In *Proceedings of 31th ICRC*, łódź, Poland, 2009.

[47] Duffett-Smith, P. *Practical Astronomy with Your Calculator;* Cambridge University Press, 3 ed., 1988.

[48] Górski, K. M.; others. *Astrophysical Journal* **2005**, *622*, 759.

[49] Abdo, A. A.; others. *Astrophysical Journal* **2012**, *750*, 63.

[50] Abdo, A. A.; others. *Physical Review Letters* **2008**, *101*, 221101.

[51] Seon, K.-I. *ArXiv Astrophysics e-prints* **2007**.

[52] Grenander, U. *Probability and Statistics: The Harald Cramr Volume;* 1959.

[53] Li, T. P.; Ma, Y. Q. *Astrophysical Journal* **1983**, *272*, 317.

[54] James, F.; Roos, M. *Computer Physics Communications* **1975**, *10*, 343.

[55] Biller, S. D. *Astroparticle Journal* **1996**, *4*, 285.

[56] Aharonian, F.; others. *Astrophysical Journal* **2004**, *614*, 897.

[57] Albert, J.; others. *Astrophysical Journal* **2007**, *663*, 125.

[58] Abramowski, A.; others. *Physics Review D* **2014**, *90*, 122007.

[59] Aharonian, F.; others. *Astronomy and Astrophysics* **2008**, *477*, 353.

[60] Aharonian, F.; others. *Astronomy and Astrophysics* **2008**, *477*, 353.

[61] Aharonian, F.; others. *Astronomy and Astrophysics* **2008**, *477*, 353.

[62] Aharonian, F.; others. *Astronomy and Astrophysics* **2009**, *499*, 723.

[63] Bartoli, B. and others. *Astrophysical Journal* **2012**, *760*, 110.

[64] Aharonian, F.; others. *Astronomy and Astrophysics* **2015**, *574*, A27.

[65] Vianello, G. and others. In *Proceedings of 34th ICRC*, The Hague, The Netherlands, 2015.

[66] Drury, L. O.; Ellison, D. E.; Aharonian, F. A.; others. *Space Science Reviews* **2001**, *99*, 329.

[67] Aharonian, F. A.; others. *Reports on Progress in Physics* **2008**, *71*.

[68] Atkins, R.; others. *Astrophysical Journal* **2003**, *803*, 811.

[69] V1190A/N & VX1190A/N User Manual. CAEN, **2012**.

[70] Lamb, R. C.; others. In *Proceedings of the International Conference on High Energy Gamma-ray Astronomy*, 1991.

Appendix A

Figure Copyright Permissions

A.1 Fig. 1.1

Image credit: S. Lafebre. Retrieved from Wikipedia at http://en.wikipedia.org/wiki/Cosmic_ray. Used under Creative Commons Attribution-ShareAlike 3.0 (CC BY-SA 3.0) license, https://creativecommons.org/licenses/by-sa/3.0/us/.

A.2 Fig. 1.2

Image credit: NASA/CXC/SAO. Public domain created by federal employees. Retrieved from http://www.nasa.gov/chandra/multimedia/

chandra-15th-anniversary-crab-nebula.html.

A.3 Fig. 1.3

Image credit: NASA/CXC/SAO. Public domain created by federal employees. Retrieved from http://www.nasa.gov/chandra/multimedia/chandra-15th-anniversary-tycho.html.

A.4 Fig. 1.4

Image credit: The H.E.S.S. collaboration. Retrieved from [4]. Permission granted by email from D. Berge.

A.5 Fig. 1.5

Image credit: NASA. Public domain created by federal employees. Retrieved from http://fermi.gsfc.nasa.gov/ssc/observations/types/allsky.

A.6 Fig. 2.2 and 2.3

Image credit: F. Schmidt and J. Knapp. Retrieved from https://www.ikp.kit.edu/corsika/. Permission granted by email from F. Schmidt.

A.7 Fig. 2.5

Image credit: S. Klepser. Retrieved from Wikipedia at https://en.wikipedia.org/wiki/High_Energy_Stereoscopic_System. Used under Creative Commons Attribution-ShareAlike 3.0 (CC BY-SA 3.0) license, https://creativecommons.org/licenses/by-sa/3.0/us/.

A.8 Fig. 2.6

Courtesy of Institute for Cosmic Ray Research, the University of Tokyo, Japan. Retrieved from http://www.icrr.u-tokyo.ac.jp/em/. Permission granted by email from M. Takita.

A.9 Fig. 2.7

Image credit: The Milagro collaboration. Permission granted by email from J. Goodman.

A.10 Fig. 3.1

Image credit: A. Carramiñana. Permission granted by email.

A.11 Fig. 3.2

Image credit: S. BenZvi. Permission granted by email.

A.12 Fig. 3.10

Map data: Google, DigitalGlobe. According to Google: "You may use our maps in internal reports, presentations, proposals, and other related professional documents. We request you still retain attribution to both Google and our data providers."

A.13 Fig. 5.4, 5.5, and 5.6

Image credit: F. Salesa. Permission granted by email.