Advancements in Very-High-Energy Gamma-Ray Astronomy with Applications to the Study of Cosmic Rays

Andrii Petrashyk

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Advancements in Very-High-Energy Gamma-Ray Astronomy with Applications to the Study of Cosmic Rays

Andrii Petrashyk

This work aims to contribute to the study of the origins of cosmic rays, and broadly, to the advancement of both data analysis methods and instrumentation for very-high-energy γ-ray astronomy. First, reviewing the state of γ-ray astronomy, we show how gains in sensitivity can be achieved through sophisticated data analyses and improved instrumental designs. We then develop such an improved analysis method for the Very Energetic Radiation Imaging Telescope Array System (VERITAS) by combining Image Template Method (ITM) with Boosted Decision Trees (BDT), and study its performance, attaining a 30-50% improvement in integral sensitivity over the instrument's standard analysis. Systematic issues in spectral reconstruction that the analysis displays are resolved satisfactorily by imposing a more stringent condition on the selection of its energy threshold. We employ the newly developed analysis to measure the γ-ray energy spectrum of the starburst galaxy M82, and combining our result with a measurement from the Fermi Large Area Telescope (*Fermi*-LAT), we find that a single power law fits the spectrum well between 100 GeV and 10 TeV, with no evidence for a spectral break or a cutoff. We conclude that this is in line with the current understanding that M82 is not a good proton calorimeter. Finally, we detail the design, implementation, and performance of the optical alignment system of the prototype

Schwarzschild-Couder Telescope (pSCT) for the Cherenkov Telescope Array (CTA), a novel two-mirror design that addresses many shortcomings of current instruments.

Contents

List of Figures

List of Tables

Acknowledgements

While this dissertation bears my sole authorship, a great many people, in fact, contributed to the work in it – some directly and intentionally, through their material help and advice (and lending figures), and others more subtly, by sometimes catching me at the ground state and pushing me back up. Here, I would like to acknowledge their role in my finishing this thesis.

First, I would like to thank my academic advisor, Thomas Brian Humensky, for his immense patience, openness, and enthusiasm; but also, for encouraging me to do so many things, and trusting and supporting me through multiple failures; and of course, for fixing an embarrassingly large number of typos in my drafts.

I would also like to thank my senior colleagues at Columbia – Reshmi Mukherjee, Marcos Santander, Daniel Nieto, Qi Feng – for bearing with my questions, sharing their knowledge, and helping me find my place in academia, science, and the world at large.

I would like to express my gratitude to the communities of the Very Energetic Radiation Imaging Telescope Array System (VERITAS) and the Cherenkov Telescope Array (CTA)

(and I'm OK with double-counting the individuals who belong to both): Wystan Benbow, Ralph Bird, Jodi Christiansen, Michael Daniel, Lucy Fortson, Gareth Hughes, Gernot Maier, Cameron Rulten, David Williams, Ben Zitzer and many, many more. Here, I would like to single out Vladimir Vassiliev, the Principal Investigator of the prototype Schwarzschild-Couder Telescope (pSCT), whose scientific rigor and daily vigor, I hope, rubbed off on me; and whose occasional urgent weekend calls made me feel very important.

I am supremely grateful to all my friends at Columbia University and elsewhere (I hope you all know who you are). Yes, you were right, I did finish it; also, text me.

I thank my family for getting me here – to the start, to the finish, and all the way between the two. I hope you now have a better idea of what it was I spent so many years on; but that's irrelevant, I thank you for the unconditional love.

Chapter 1

Introduction

Cosmic rays, first discovered in 1912 by Victor Hess (Hess and Anderson 2013), are the most energetic particles in the universe. They originate in violent astrophysical explosions and get accelerated by shocks and jets to energies beyond 10^{20} eV. They propagate throughout the Galaxy along magnetic field lines, producing highly energetic photons as they interact with the ambient gas and light. The study of cosmic rays through the observation of this γ radiation is the subject of this work.

Cosmic rays are deflected and scrambled by magnetic fields and do not point back to their sources, forming instead a nearly isotropic flux at the Earth. However, the electromagnetic radiation that is created during their interactions with the surrounding environment propagates in straight lines from its place of origin. From the spectrum of this radiation, which typically extends over many wavelengths from centimeter-wave radio to TeV gamma rays, one can reconstruct the flux and spectrum of the progenitor cosmic rays. As different

processes are responsible for electromagentic radiation at different energies, the observed spectra are not sufficient to fully reconstruct and characterize the cosmic ray flux, unless other independent measurements of the interaction environment are made. Nevertheless, accurately quantifying the electromagnetic radiation generated by cosmic rays at all energies is paramount to the study of the processes that produce cosmic rays. This work deals with the measurement of the spectra of γ rays, the highest energy range of such radiation.

Starburst galaxies are a class of galaxies with very high star-formation rates. A galaxy of this type typically contains a core where high gas densities lead to the formation of many massive bright stars that, after quickly burning through their nuclear fuel, collapse and then explode, powering highly energetic supernovae. Shocks of the many expanding supernovae accelerate charged particles to very high energies, resulting in high fluxes of cosmic rays. At the same time, high gas and light densities of the surrounding enviornment provide plenty of target material for cosmic rays to interact with, generating high fluxes of γ radiation. Cores of starburst galaxies, then, are an excellent cosmic laboratory for the study of cosmic rays.

This chapter gives a high-level overview of cosmic rays and a brief survey of the main processes responsible for their acceleration and governing their interactions. For a thorough review see, for instance, Longair 2011. It then discusses a way to study cosmic rays by observing electromagnetic radiation from starburst galaxies, and introduces M82, a starburst galaxy studied in this work. A fuller review of starburst regions and the production of highly energetic particles in them can be found in, for example, Bykov 2014. We conclude the chapter with a concise summary of this work.

1.1 Cosmic Rays

As observed at Earth, the energy spectrum of cosmic rays spans more than twelve orders of magnitude in a near power-law spectrum (Figure 1.1). The two main deviations from a single power law occur at the *knee*, where the spectral index steepens slightly from 2.7 to 3.1, near $\sim 10^{15}$ eV, and at the *ankle*, where the spectrum flattens again to an index of 2.7, at $\sim 10^{18.5}$ eV (Longair 2011). These features may correspond to a transition between different populations of cosmic rays as it is thought that cosmic rays below the knee are Galactic in origin and cosmic rays above the ankle are extragalactic in origin (Hörandel 2004; Aharonian 2004; Beatty and Westerhoff 2009). Indeed, an intuitive way to see this is by looking at the the gyroradius of a proton in the Galactic magnetic field of $\sim 10^{-9}$ T: at $\sim 10^{18.5}$ eV, it is $r_g \approx 3.3 \cdot \frac{(\gamma mc^2/\mathrm{GeV})(v_\perp/c)}{(|q|/e)(B/\mathrm{T})}$ m $\approx 10^{19}$ m, which is on the order of the size of our Galactic arm, so particles with lower energies are confined to the Galaxy, while particles with greater energies escape it.

Cosmic rays, a general term for all high-energy particles with intrinsic mass,[1] can be classified as *primary*, such as those accelerated directly by astrophysical sources, and *secondary*, which are produced in interactions of primary or secondary cosmic rays with interstellar matter, photons, or magnetic fields. With electromagnetic fields being the engine of particle acceleration, all cosmic rays are charged particles. Indeed, at GeV energies, 86% of all

[1]While some use the term "cosmic rays" to also refer to high-energy photons, it seems more common to call particles cosmic rays, and to call photons either X-rays or γ rays. High-energy astrophysical neutrinos also seem to be excluded from the "cosmic rays" umbrella.

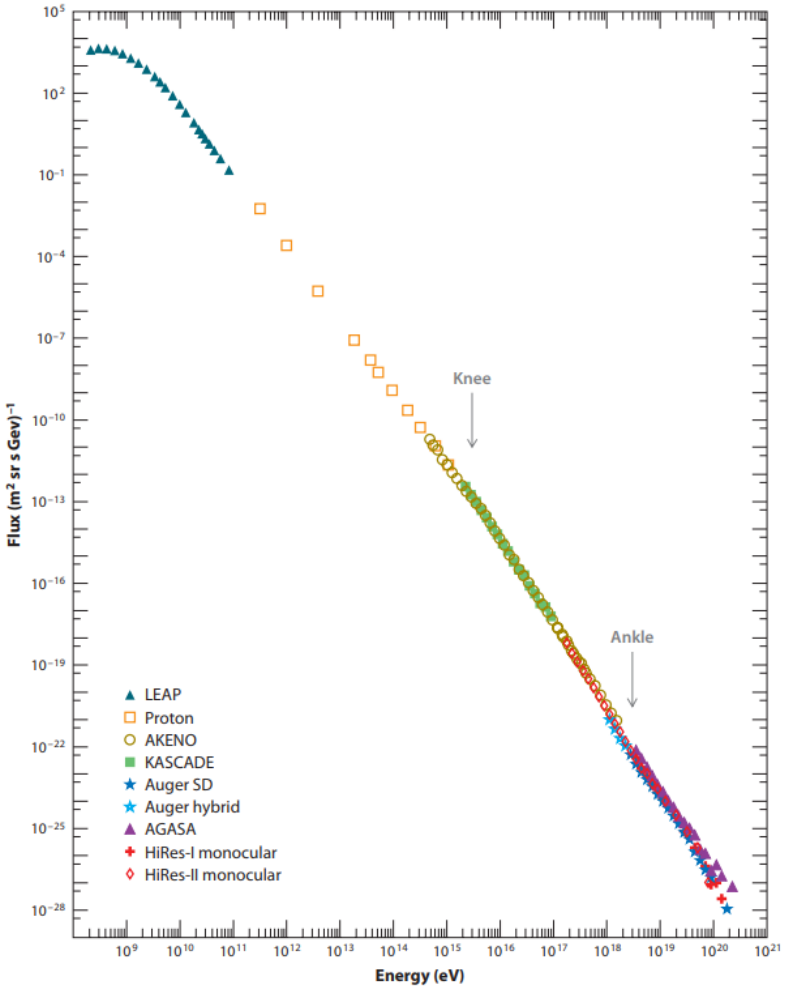

Figure 1.1: Overview of the observed cosmic ray spectrum. The spectrum is a near power-law with an index of $\Gamma = 2.7$. Approximate energies of the breaks in this spectrum, commonly referred to as the knee where the spectrum steepens, and the ankle where it flattens again, are indicated by arrows. Figure from Beatty and Westerhoff 2009.

cosmic rays are protons, 11% helium nuclei, 1% heavier nuclei, and 2% electrons (Longair 2011). There are several mechanisms that could explain the acceleration of charged particles in a magnetized plasma, the main ones being the second- and first-order Fermi acceleration (Longair 2011).

In second-order Fermi acceleration (Fermi 1949), charged particles get reflected by irregularities in the Galactic magnetic field that act as "magnetic mirrors," and are assumed to move randomly with a typical velocity u. It can then be shown that particles reflected from these mirrors gain energy stochastically, with an average gain per reflection being $\langle \delta E/E \rangle = 8/3\,(u/c)^2$.

First-order Fermi acceleration, also referred to as diffusive shock acceleration (DSA) (Bell 1978), is the acceleration of charged particles in strong shock waves. The main idea is the same as in second-order Fermi acceleration, only now the mirrors do not move randomly: instead, there is a shock front of velocity u, from which particles reflect in preferentially head-on collisions. Here, the average energy gain per one round trip across the shock front is $\langle \delta E/E \rangle = 4/3 \cdot u/c$, proportional to the first order in the shock speed.

Both of these processes result in a constant *fractional* increase in particle energy per interaction with a "mirror," and with particles having a fixed probability of remaining in the accelerating region, a simple power-law energy spectrum of a population of particles is established. Indeed, if the fractional energy gain per interaction is β, a particle of initial energy E_0 will reach $E = E_0\beta^n$ after n interactions; if the probability of remaining in the accelerating region is p, then after the same n interactions, only $N = N_0 p^n$ of the original N_0 particles will remain; from here, $\ln(N(E)/N_0) = \ln(E/E_0)\,(\ln p/\ln \beta)$ and $\frac{dN(E)}{dE} \propto \frac{N(E)}{E} \propto E^{-1+\ln p/\ln \beta}$. However, the acceleration process is slow and requires long lifetimes and large sizes of the accelerating regions. As such, non-linear effects that account for the interaction between the shock, the accelerated particles, and the environment, need to be included.

However, we can make some simple estimates pertaining to the sources of acceleration of cosmic rays from an argument made by Hillas (Hillas 1984). With the typical magnetic field in the accelerating region being B, and the typical shock velocity u, the typical accelerating electric fields are then $\sim Bu$; with L as the typical size of the accelerating region, the maximum energy to which such fields can accelerate particles of charge q is then $E_{\max} \sim qBuL$. In order to accelerate protons to energies of $\sim 10^{20}$ eV, BL needs to be around $3 \cdot 10^{11}$ m $\cdot$ T. This relation is shown with a straight line in figure 1.2: if a potential source is to the left of the line, protons there cannot be accelerated to $\sim 10^{20}$ eV; if it is to the right, protons there could be accelerated to energies beyond the ankle.

As cosmic rays propagate through the interstellar medium (ISM), they are deflected and scrambled by magnetic fields, which results in our inability to trace them back to their sources. However, their interactions with ambient magnetic fields and the ISM also lead to the release of electromagnetic radiation that can be used to study the accelerators of cosmic rays. Due to the partial or complete confinement of cosmic rays to their host galaxies, such electromagnetic radiation is a particularly good messenger for the study of acceleration mechanisms in other galaxies.

Energy losses of cosmic rays that produce electromagnetic radiation occur through synchrotron emission, bremsstrahlung, inverse Compton scattering, and pion production. Synchrotron radiation is emitted due to the acceleration of a charged particle moving at an angle to a magnetic field. The rate at which energy is radiated, as well as the characteristic frequency of the radiation, are dependent on the energy density of the magnetic field and the

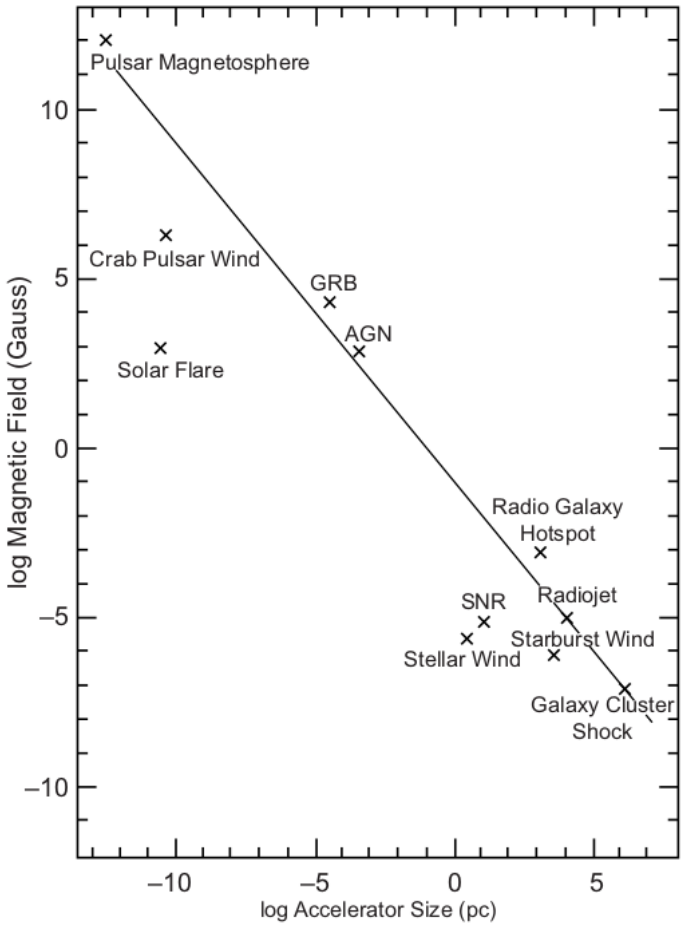

Figure 1.2: The combinations of length-scale and magnetic flux density necessary to accelerate particles to extremely high energies, at the order of 10^{20} eV. Crosses indicate the typical combinations of magnetic flux density and size found in different classes of astronomical objects. Figure from Longair 2011.

mass of the emitting particle: $dE/dt \propto -\gamma^2 U_B/m^2,\ \nu \propto \gamma\sqrt{U_B}/m$, where γ is the Lorentz factor of the charged particle. Therefore, the observed synchrotron emission from galaxies is produced by cosmic-ray electrons, and for electrons with energies of 100 GeV in typical galactic magnetic fields of 10^{-9} T, is emitted in the radio band. In extreme environments, such as forward shocks of supernova remnants, electrons accelerated to energies of $\sim 10^{13}$ eV

in magnetic fields of $\sim 10^{-8}$ T produce synchrotron emission in the X-ray band, as happens in, for instance, Cassiopeia A (Allen et al. 1997; Helder and Vink 2008). Due to their large masses, protons produce measurable synchrotron radiation only in extreme magnetic fields.

Similarly to their interactions with ambient magnetic fields, charged particles also emit energy by upscattering photons of ambient radiation fields. The energy loss rate of this inverse Compton emission looks very similar to the synchrotron radiation above, being proportional to the energy density of the ambient photon field: $dE/dt \propto -\gamma^2 U_{\rm rad}/m^2$. The typical energies of upscattered photons are now much higher, however: $\nu' \propto \gamma^2 \nu_0$. Once again, this process does not affect protons, while inverse Compton scattering of 100 GeV electrons on near-infrared photons results in the production of ~ 50 GeV γ-rays.

Cosmic-ray electrons also produce γ-rays through inelastic interactions with the electric fields of protons and other nuclei in the ISM, a process known as *bremsstrahlung*. The energy loss rate due to bremsstrahlung is proprotional to the density of the surrounding medium: $dE/dt \propto -\gamma n_{\rm ISM}$. This process is not effective for cosmic-ray protons because of their large masses. Therefore, in an environment such as a starburst region, with high densities of the ISM and strong magnetic and radiation fields, very-high-energy electrons quickly cool down by losing their energy through bremsstrahlung, synchrotron and inverse Compton radation.

Cosmic-ray protons are affected by their interactions with other nuclei in the ISM. Above 1.22 GeV, these interactions lead to the production of secondary pions, which are short-lived and quickly decay to γ-rays (neutral pions) or electrons and neutrinos (charged pions) (see Section 2.1.2 for a few more details). Since protons make up the largest portion of the

total cosmic-ray flux, neutral pion decay is a major mechanism of production of γ-rays above several GeV. γ-ray astronomy is therefore a powerful tool for the study of cosmic-ray protons as well as electrons.

At extremely high energies, hadronic cosmic rays can also produce pions in interactions with the ambient light, through a process like $p + \gamma \rightarrow n + \pi^+$. A thorough calculation of the interaction cross-section for such a process shows that ~ 100 EeV protons lose over 20% of their energy per interaction to photo-pion production on photons of the cosmic microwave background (Greisen 1966; Zatsepin and Kuz'min 1966). This effect may limit the highest attainable energy of cosmic rays, known as the GZK cutoff. A suppression of cosmic ray flux above $\sim 10^{19.5}$ eV, consistent with the GZK cutoff, has been measured (Abraham et al. 2010; Sokolsky 2011).

Since the majority of the pions produced in hadronic interactions are charged, their decay leads to the production of neutrinos. The recent detection of a high-energy neutrino associated with the blazar TXS 0506+056 (IceCube Collaboration et al. 2018; Ansoldi et al. 2018) ushered in a new era of multimessenger astronomy, opening up new channels for the study of cosmic rays.

1.2 Starburst Galaxies and M82

Supernova remnants, while seemingly not quite able to accelerate particles past the knee energies (Figure 1.2), are nevertheless major sources of cosmic rays as charged particles

are accelerated to very high energies by the shocks between the expanding shells and the surrounding environment. Starburst galaxies are characterized as being in a state with a significantly higher star-formation rate (SFR) than the long-term average. This elevated rate of star formation, which lasts only a small fraction of the total lifetime of the galaxy, is often triggered by an interaction with another galaxy, such as a close passage or a merger. It leads to the formation of bright, massive O- and B-type stars, which burn through their nuclear fuel within a few million years and explode in core-collapse supernovae. Consequently, the frequent supernovae result in large densities of cosmic rays. In addition, regions of high star formation often contain large amounts of dense molecular gas, while the large numbers of massive stars are accompanied by strong magnetic and radiation fields. As discussed above, this combination of high densities of cosmic rays and extreme environments results in high fluxes of γ-rays, making starburst galaxies excellent astrophysical sources for the study of cosmic ray acceleration, propagation and interaction.

The thermal spectrum of massive OB stars peaks in the ultraviolet (UV) range. This UV radiation is absorbed by the surrounding dust and molecular gas, and re-radiated in the far infrared (FIR) band. Thus, if the density of the surrounding dust and gas is high enough to absorb all of the UV radiation, the luminosity in the FIR range is proportional to the number of OB stars in the galaxy, and therefore, to the star-formation rate: $FIR \propto SFR$ (Kennicutt 1998). On the other hand, synchrotron emission of cosmic-ray electrons in ambient magnetic fields results in non-thermal radio radiation. The supernova rate is correlated with the SFR, which implies that the density of cosmic-ray electrons – and the luminosity of non-thermal

radio radiation – is also correlated with the SFR, and therefore, with the FIR luminosity. This well-known FIR-radio correlation has been extensively measured for many galaxies with wide ranges of SFRs (Lacki, Thompson, and Quataert 2010).

Starburst galaxies make the FIR-radio correlation all the more interesting, since a similar argument can be applied to the observed γ-ray flux from them. Indeed, the density of hadronic cosmic rays is correlated with the supernova rate, and consequently, the SFR. If energy losses of these hadronic cosmic rays dominate over advective or diffusive losses, the cosmic ray energy is efficiently converted into γ radiation, whose luminosity then scales with the SFR and the FIR luminosity. Recent studies have indeed found a positive correlation between the SFR of galaxies and their γ-ray luminosity (Rojas-Bravo and Araya 2016). This is demonstrated in Figure 1.3, where the γ-ray luminosities of a sample of star-forming galaxies observed with the *Fermi*-LAT follow a power-law correlation with their SFR.

The considerations of a relationship between the γ-ray luminosity and the SFR have led to the concept of calorimetry (Pohl 1994): a galaxy is a proton calorimeter if protons above the pion production threshold convert all of their energy into pions and, subsequently, γ-rays. For galaxies below this calorimetric limit, the γ-ray luminosity scales with the calorimetric fraction $f_{\rm cal}$, the fraction of the energy of cosmic rays above 1.22 GeV that gets converted to pions. A system that is not fully-calorimetric indicates the presence of a strong energy-independent loss mechanism, such as a wind, meaning that high-energy cosmic rays are advected away before losing a substantial amount of their energy.

To date, only two starburst galaxies without dominant active galactic nuclei (AGNs)

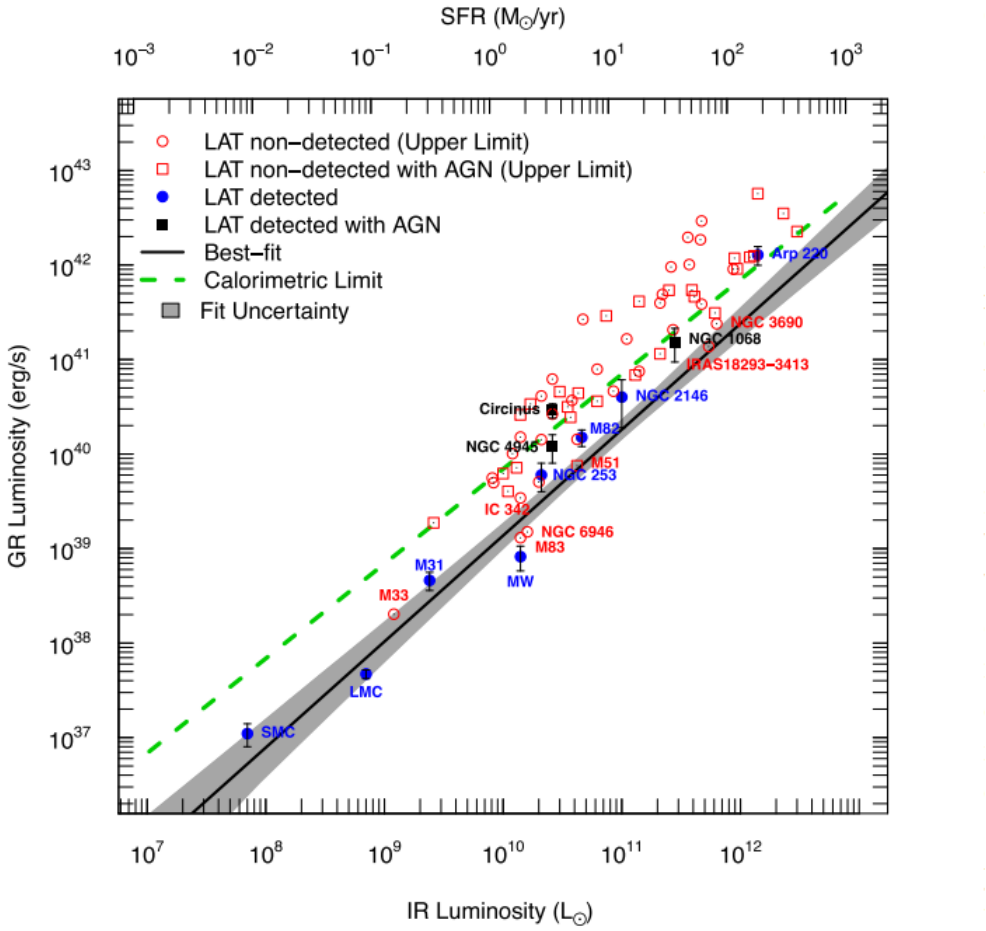

Figure 1.3: γ-ray luminosity (0.1-100 GeV) versus total infrared luminosity (8-1000 μm) for star-forming galaxies. The best-fit power-law correlation calculated for these data is shown with its 95% confidence level interval. Figure from Rojas-Bravo and Araya 2016.

have been detected in the γ-ray band at both GeV and TeV energies – NGC 253 and M82 (Abdo et al. 2010; VERITAS Collaboration et al. 2009; Acero et al. 2009; Ackermann et al. 2012; H. E. S. S. Collaboration et al. 2018). At GeV energies, both of them were detected by the *Fermi*-LAT, and NGC 253 was detected by H.E.S.S., while M82 by VERITAS, at TeV energies. The Southern location of NGC 253 makes it unobservable from the site of

VERITAS, while the Northern location of M82 makes it inaccessible to view by H.E.S.S.

M82 (Figure 1.4) is the prototypical starburst galaxy, whose increased SFR was triggered by a close encounter with the more massive neighboring galaxy M81, during which tidal forces funneled matter into the core of M82. The galaxy, located at a distance of ~ 3.9 Mpc (Sakai and Madore 1999), is estimated to have an SFR of $\sim 10\,M_\odot\,\mathrm{yr}^{-1}$ in its starburst core (Gao and Solomon 2004), greater than that of the entire Milky Way, $\sim 1\,M_\odot\,\mathrm{yr}^{-1}$ (Robitaille and Whitney 2010), fed by the high density of molecular Hydrogen of $1 - 6 \cdot 10^4\,\mathrm{cm}^{-3}$ (Naylor et al. 2010). The supernova rate in the core of M82 has been estimated to be $\sim 0.1\,\mathrm{yr}^{-1}$ (Fenech et al. 2008), about 5 times that of the Milky Way, $\sim 0.02\,\mathrm{yr}^{-1}$ (Tammann, Loeffler, and Schroeder 1994), resulting in a high density of cosmic rays in its core, with current estimates of the total cosmic-ray energy density of $\sim 250\,\mathrm{eV\ cm}^{-3}$ (Persic and Rephaeli 2014). The high density of hot ionized gas in the core, $\sim 180\,\mathrm{cm}^{-3}$ (de Cea del Pozo, Torres, and Rodriguez Marrero 2009), and the high energy density of thermal radiation, $\sim 1000\,\mathrm{eV\ cm}^{-3}$ (Yoast-Hull et al. 2013), provide plenty of target material for cosmic rays and result in a high γ-ray luminosity.

M82 is well-known for its galactic wind, seen in Figure 1.4 extending off the galactic plane. Estimates on the velocity of this outflow range from $\sim 600\,\mathrm{km\ s}^{-1}$ (Strickland, Ponman, and Stevens 1997; Shopbell and Bland-Hawthorn 1998) to $\sim 2200\,\mathrm{km\ s}^{-1}$ (Strickland and Heckman 2009). This variation in speeds indicates that the wind has a complex structure and may be driven by a variety of mechanisms. While the traditional intrepretation of the M82 wind has been that of a hot outflow powered by shock heating from supernovae (Chevalier

Figure 1.4: A multiwavelength image of the starburst galaxy M82 by the Hubble Space Telescope. The photo is 10 arcmin across. The red in the image represents hydrogen and infrared light, indicating starburst activity. The blue and greenish-yellow color represent visible wavelengths of light.

and Clegg 1985), cosmic rays could also be a driving factor (Breitschwerdt, McKenzie, and Voelk 1993). The wind is a major source of losses of high-energy cosmic rays, controlling the energy-independent advective timescale. How calorimetric the galaxy is, and its γ-ray luminosity, depend, therefore, on the speed of the wind. Of course, the wind is not the only parameter, the other important factors being the acceleration efficiency of the injected cosmic rays, and the average ISM density, which determines the energy-dependent energy loss timescale.

Detailed simulations of cosmic-ray production and acceleration, as well as γ-ray emission in M82, that are constrainted by the observed radio spectrum have been performed in, for

instance, de Cea del Pozo, Torres, and Rodriguez Marrero 2009, Paglione and Abrahams 2012 and Yoast-Hull et al. 2013. Existing models view the starburst core as an injector of cosmic rays and deal with its averaged properties, modeling the advective, diffusive, and radiative losses of cosmic rays, and their interaction with a uniform environment. The radio observations are usually taken to constrain the magnetic field strength and the energy density of cosmic-ray electrons, although a degeneracy exists between the magnetic field strength and the speed of the galactic wind: the same effect can be achieved either by decreasing the magnetic field strength or by increasing the wind speed. Indeed, an increased wind speed means a shorter advective timescale, and a decreased number of cosmic-ray electrons in the starburst region, leading to a decrease in the amount of synchrotron radiation (Yoast-Hull et al. 2013). This degeneracy highlights the importance of γ-ray observations: while the magnetic field has no direct effect on the γ-ray emission, increased advection also results in higher losses of high-energy protons, affecting the γ-ray flux.

The current understanding is that both cosmic-ray electrons and protons are important contributors to the γ-ray flux of M82 at GeV energies, with pion decay being the dominant γ-ray emission mechanism at TeV energies (**Ohm2016RF250Mfoo**; Yoast-Hull et al. 2013). While electron energy losses have a shorter timescale than the advective losses, resulting in the electrons quickly cooling down and radiating away all of their energy, for protons, energy-independent advective losses are comparable to radiative losses, especially at the highest energies. Thus, while M82 is an electron calorimeter, it is not a complete proton calorimeter (de Cea del Pozo, Torres, and Rodriguez Marrero 2009; Yoast-Hull et al. 2013;

Rojas-Bravo and Araya 2016). Lacki et al. 2011 review previously published models and compare them to the measurement in Abdo et al. 2010, concluding that the models are largely consistent, and arguing that the γ-ray emission is predominantly hadronic in origin, with an estimated calorimetric fraction of 0.3 to 0.4, the uncertainty driven by the uncertainties on the supernova rate, supernova acceleration efficiency, and the distance to the galaxy. In this work, we update the VERITAS TeV spectrum of M82 and briefly discuss its possible implications in Chapter 5.

The upcoming Cherenkov Telescope Array[2] (CTA) (CTA Consortium 2017) is scheduled to collect 100 hours of observations of M82 in the first 3 years of its operations. With CTA, it may be possible to spatially resolve the starburst core at TeV energies, as well as to distinguish the diffuse starburst emission from the γ-ray flux produced by individual sources (CTA Consortium 2017).

1.3 About This Work

This work aims to contribute to the study of the origins of cosmic rays, and broadly, to the advancement of both data analysis methods and instrumentation for very-high-energy γ-ray astronomy. First, for the near term, we improved the sensitivity of VERITAS, an existing γ-ray observatory, through the development and implementation of advanced data analysis methods. Second, for the long term, we developed the mirror alignment control software

[2] `https://www.cta-observatory.org/`

as part of the construction of a prototype of a novel telescope for the Cherenkov Telescope Array, a next-generation instrument that is expected to come online in the next several years. Our sincere hope is that these contributions leave a lasting impact in the scientific community by enabling future research.

In the following, Chapter 2 reviews the Imaging Atmospheric Cherenkov Technqiue that is widely utilized to study γ-rays at the highest energies. It also gives an overview of VERITAS, the instrument employed in this work, and introduces the Cherenkov Telescope Array. Chapters 3 and 4 present our improvements to VERITAS data analaysis, while Chapter 5 uses the improved analysis to study M82. Finally, Chapter 6 introduces a prototype medium-sized Schwarschild-Couder Telescope for the Cherenkov Telescope Array and details the design and implementation of its optical alignment system.

Chapter 2

Imaging Atmospheric Cherenkov Technique, VERITAS and CTA

The study of cosmic rays presented in this work relies on the detection of gamma rays, very-high-energy (VHE; E $>$ 100 GeV) photons that are produced through interactions of cosmic rays with the surrounding medium and ambient photons and, undisturbed, reach the Earth's atmosphere. Apart from being first-hand witnesses to particle acceleration and propagation processes in the most extreme environments, gamma rays are a unique probe of intergalactic magnetic and extragalactic background photon fields. Moreover, limits of the Standard Model can be challenged by testing violations of Lorentz invariance as gamma rays propagate across the visible Universe. Finally, VHE signatures of dark matter annihilation are searched for through γ-ray observations of the densest clumps of unaccounted-for gravity.

As we will discuss shortly, the atmosphere of the Earth is opaque to VHE radiation.

Thus, seemingly the easiest way to study gamma rays is to put a satellite-based detector in an orbit around the Earth. However, as discussed in Section 1.1, astrophysical processes that produce highly energetic cosmic rays and the accompanying gamma rays result in steeply falling power-law spectra, leading to very low photon fluxes and making the study of gamma rays above ∼100 GeV impossible by satellite-born instruments. For instance, the brightest known γ-ray source, the Crab Nebula, gives off only about 6 photons per m^2 per year at the Earth above 1 TeV. Luckily, there is a clever trick – leveraging the atmosphere as an intrinsic part of the detector and making use of the way gamma rays interact with it, it is possible to build ground-based instruments with effective collecting areas exceeding 10^5 m^2.

The first studies of this method, dubbed the Imaging Atmospheric Cherenkov Technique, were performed by Jelley and Galbraith in the 1950s (Galbraith and Jelley 1953). The detection of the Crab Nebula in 1989 by the Whipple 10m Telescope (Weekes et al. 1989) was the first major proof of the technique, facilitated by the development of methods to record *images* of the Cherenkov emission from atmospheric particle cascades initiated by VHE cosmic and gamma rays. Imaging Atmospheric Cherenkov Telescopes (IACTs) that rely on this technique provide the most sensitive view of the VHE γ-ray sky through a wide field of view and good angular and spectral resolutions. Since the Crab Nebula discovery, the number of known VHE sources has rapidly increased, approaching 200 different objects classified into 7 different classes, with many yet unidentified (Figure 2.1).

This chapter summarizes the technique used by IACTs and gives a short overview of the

[1] `http://tevcat2.uchicago.edu/`

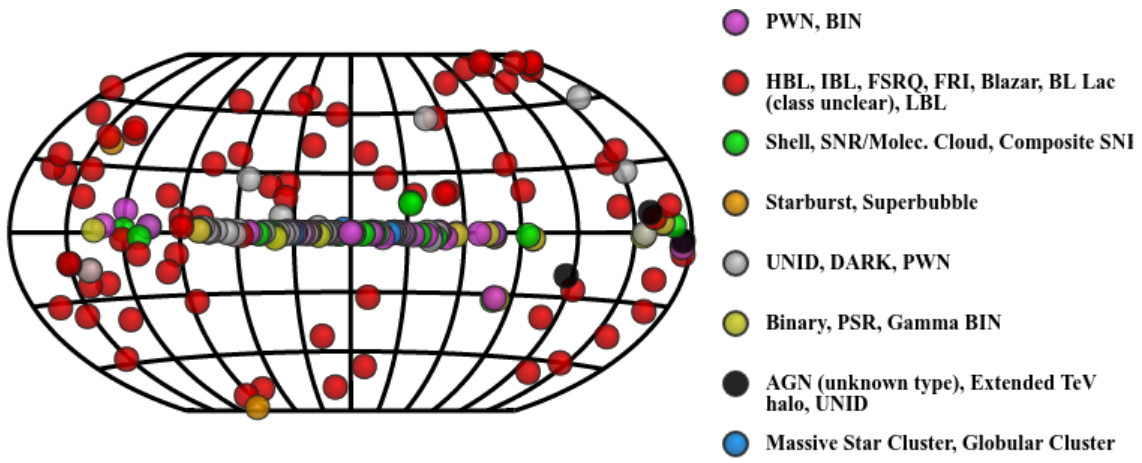

Figure 2.1: The current view of the TeV sky, catalogued at TeVCat[1].

history of VHE γ-ray astronomy. It then reviews the hardware of VERITAS, the telescope array utilized in this work, as well as IACT data-analysis methods, with an emphasis on VERITAS. Finally, we talk about the limitations of VERITAS and motivate the need for advanced analysis methods, as well as current plans for the Cherenkov Telescope Array (CTA), the next-generation IACT instrument.

2.1 The Imaging Atmospheric Cherenkov Technique

A VHE particle entering the atmosphere (from now on, a *primary*), will interact with air molecules and produce an *extensive air shower*, a cascade of relativistic charged particles, which, traveling faster than the speed of light in air, will emit Cherenkov radiation. The collective Cherenkov emission from many particles in the shower will produce a pool of Cherenkov light on the ground, whose intensity and morphology depend on the energy and nature of the primary. Placing a telescope within this pool of light and pointing it in the

direction of the shower, it is possible to record an image of the projection of the Cherenkov emission on the ground. Being able to accurately image Cherenkov light of an extensive air shower from multiple points of view with an array of telescopes is at the core of the stereoscopic Imaging Atmospheric Cherenkov Technique. Understanding the technique is cruical to this work, and here we dedicate a significant amount of time to its explanation.

2.1.1 Electromagnetic showers

A gamma ray in the field of a nucleus pair produces if it has enough energy, $E_0 > 2 \times 0.511$ MeV:

$$\gamma \rightarrow e^+ + e^- \tag{2.1}$$

The presence of a nucleus is necessary for the conservation of both momentum and energy during pair production. The electron and positron produced in this process will share the energy of the primary photon, and for a primary with energy of several GeV, will be highly relativistic, and will further radiate VHE photons through Bremsstrahlung, which will further pair produce, and so on (Figure 2.2).

Bremsstrahlung radiation length χ_0, the distance over which an electron radiates away $1/e$ of its initial energy, is 36.7 g cm^{-2} in the atmosphere, and it is 7/9 of the mean free path of a gamma ray for pair production. The overall thickness of the atmosphere to sea level is 1030 g cm^{-2}, so this *pair production* $\leftrightarrow$ *Bremsstrahlung radiation* loop will happen many times, resulting in the generation of a shower of highly energetic electrons and positrons.

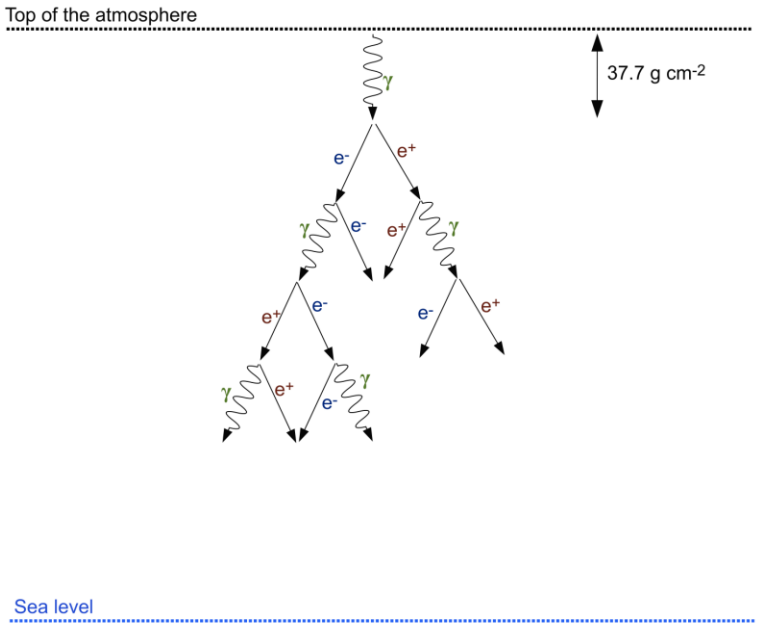

Figure 2.2: Gamma-initiated electromagnetic cascades in the atmosphere. Figure from Gammell 2004.

The similarity in radiation lengths between the two processes allows for a simple analytic model of the shower development (Heitler 1954), in which the total number of particles in the shower – electrons, positrons and photons – doubles every $\chi_0 \ln 2$: the probability of pair production, or conversely, gamma emission through Bremsstrahlung, is equal to 1/2 at traveled path χ given by

$$\frac{1}{2} = \exp\left(-\frac{\chi}{\chi_0}\right) \;\Rightarrow\; \chi = \chi_0 \ln 2 \equiv X_0$$

Therefore, for a cascade initiated by a primary gamma of energy E_0, an electron-positron pair is produced, on average, every X_0 traveled. Assuming, for simplicity, that the pair evenly

splits the primary's energy, the cascade will have at this point two particles of energy $E_0/2$. After a distance of sX_0, the shower will contain 2^s particles of energy $E_0/2^s$ each. Shower development continues until electrons reach the critical energy below which ionization losses dominate, $E_c = 84$ MeV – this is the point of maximum shower development, $X_{\max}$, with the total number of particles in the shower being

$$N_{\max} = \frac{E_0}{E_c},$$

and the number of radiation lengths reached when this happens is

$$s_{\max} = \frac{\ln\left(E_0/E_c\right)}{\ln 2}.$$

After this, the number of particles in the shower rapidly decreases because of ionization losses, which dominate once electrons become non-relativistic. Particles in an electromagnetic cascade also undergo multiple Coulomb scatterings, which distribute their velocities over a small angular range and generates a relatively narrow lateral extent of the shower. Several features of γ-ray-initiated showers are worth noting:

- The shower consists only of electrons, positrons and photons – there are no pions, muons or other particles.
- Particles in the shower have small lateral momenta resulting in a narrow lateral spread of the shower.

- The total number of particles in the shower is directly proportional to E_0, the energy of the primary photon.
- The depth of the shower maximum, $X_{\max}$, is proportional to the logarithm of E_0, the energy of the primary photon.

As we shall describe below, the first two properties help distinguish γ-ray primaries from hadronic cosmic rays. And it should be apparent that the last two features help reconstruct the initial energy E_0 of the primary. A few more details of the development of electromagnetic cascades in the atmosphere are given in Section 3.2.

2.1.2 Hadronic showers

The vast majority of VHE particles striking the atmosphere are hadronic cosmic rays (CRs) – charged, highly relativistic protons and nuclei. These particles also produce extensive air showers in the atmosphere, but the cascade development in this case is more complex: hadronic interactions fragment nuclei in the air and lead to the production of secondary nucleons, as well as charged and neutral pions. The pions quickly decay to secondary particles – neutral pions decay into gamma rays, while charged pions produce muons and neutrinos:

$$\text{CR + atmospheric nuclei} = \begin{cases} \pi^0 \rightarrow \gamma + \gamma \\ \pi^+ \rightarrow \mu^+ + \nu_\mu \\ \pi^- \rightarrow \mu^- + \bar{\nu}_\mu \end{cases} \tag{2.2}$$

Neutral pion-produced gamma secondaries can then trigger secondary electromagnetic sub-showers, while the long-lived high-energy muons will often reach the ground losing energy only through Cherenkov radiation (Figure 2.3). The pions produced in CR-initiated showers will generally have relatively large transverse momenta, resulting in a larger lateral spread of the cascade compared to a purely electromagnetic one. This difference in the longitudinal and lateral development of electromagnetic and hadronic showers helps determine the nature of the primary.

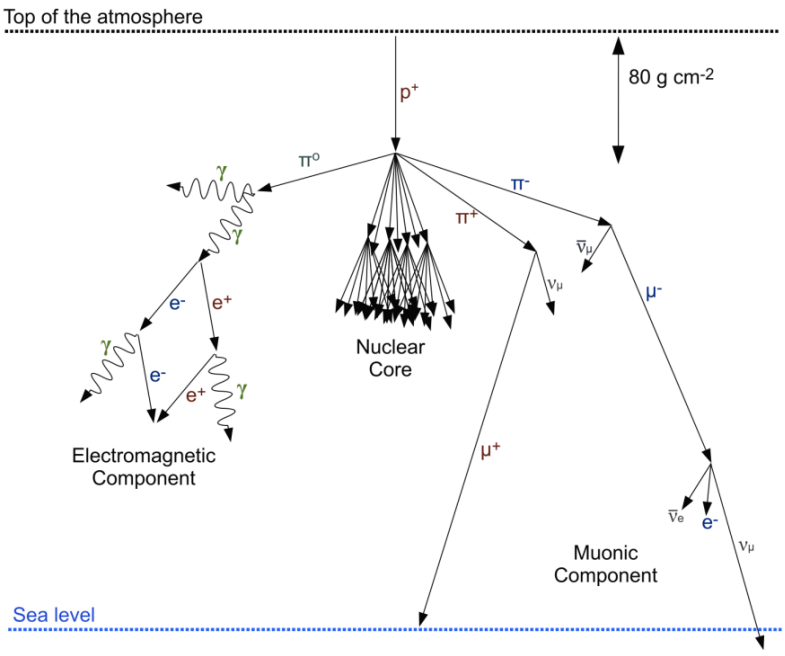

Figure 2.3: Hadron-initiated particle cascades in the atmosphere. Figure from Gammell 2004.

Of course, not all showers can be distinguished – some hadronic showers will look very similar to electromagnetic ones, and vice versa. In particular, hadronic showers that produce high-energy neutral pions early in their development dump most of their energy into elec-

tromagnetic sub-showers, and are virtually indistinguishable from gamma rays (Maier and Knapp 2007). Moreover, electron- and positron-induced showers are purely electromagnetic and, apart from a small difference in the depth of the first interaction in the atmosphere, are indistinguishable from gamma ray-induced showers. Such showers form the irreducible background for IACTs that needs to be estimated and accounted for.

2.1.3 Cherenkov radiation

The relativistic charged particles produced in atmospheric cascades will emit observable electromagnetic radiation. This effect, first extensively studied by Cherenkov, Frank and Tamm, is due to the effect a quickly-moving charged particle has on a dielectric medium: the molecules in the dielectric medium get sequentially polarized and relaxed, briefly radiating as this happens (Jelley 1963) (Figure 2.4). Since the polarization is spherically symmetric about the propagating particle, the radiation emitted by the medium is incoherent and has no net observable effect. This symmetry is broken when the particle travels faster than the phase velocity of light in the medium. When this happens, regions of coherent polarization form, their radiative pulses interfere constructively and result in net radiation being emitted in the direction of travel of the particle. Such emission is called Cherenkov radiation and can be observed as a narrow cone of light.

This intuitive explanation is made more concrete by looking at the Liénard-Wiechert

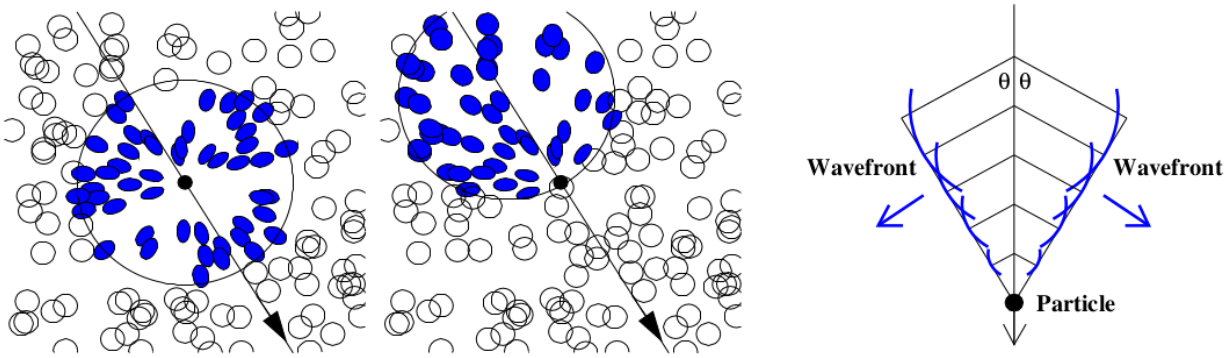

Figure 2.4: An intuitive explanation of Cherenkov emission. *Left:* Cherenkov emission as a result of coherent polarization in a medium. *Right:* The Huygens construction for the Cherenkov emission angle. Figure from Cogan 2006.

potentials $\vec{A}(\vec{r},t)$ and $\phi(\vec{r},t)$ of a moving charge q:

$$\phi(\vec{r},t) = \frac{1}{4\pi\epsilon_0 r}\left(\frac{q}{1-\hat{r}\cdot\vec{v}/c}\right)_{t_{\rm ret}}, \quad \vec{A}(\vec{r},t) = \vec{\beta}(t_{\rm ret})\frac{\phi(\vec{r},t)}{c}. \tag{2.3}$$

Here, $\vec{r}$ is the vector from the particle traveling at velocity $\vec{\beta}$ relative to the observer, and $t_{\rm ret}$ is the retarded time. While there is no radiation from a charge moving uniformly in vacuum, in a dielectric medium with refractive index $n > 1$ there exists a thin cone along which expressions (2.3) become singular. The opening angle θ_c of this cone is defined by the condition

$$1 - \hat{r}\cdot n\vec{\beta} = 0, \quad n\beta\cos\theta_c = 1, \quad \cos\theta_c = \frac{1}{\beta n}, \tag{2.4}$$

where this time, the speed of light in a medium is nc. This calculation can be intuitively seen from the Huygens construction, as shown in (Figure 2.4). It implies that Cherenkov

light is emitted whenever

$$\beta\, n(\lambda) > 1. \tag{2.5}$$

For clarity, we write the refractive index here as a function of wavelength. At X-ray frequencies, which are greater than all resonant frequencies of atoms in the air, the refractive index of light becomes slightly less than 1, and no Cherenkov radiation is emitted there. Condition (2.5) also defines the minimum energy of Cherenkov light-emitting particles, which for electrons in air is ∼21 MeV (Jelley 1963). A more accurate calculation of the Cherenkov angle accounts for recoil and has been shown (Tamm 1939) to be

$$\cos\theta_c = \frac{1}{\beta\, n(\lambda)} + \frac{E_\gamma}{2\beta\, E}\left(1 - \frac{1}{n(\lambda)^2}\right) \approx \frac{1}{\beta\, n(\lambda)}. \tag{2.6}$$

Here, E_γ is the energy of the emitted photon and E the energy of the charged particle. In air, $n \approx 1$, $E_\gamma <$ a few keV (no X-ray emission) and $E \gtrsim 21$ MeV, hence the last approximate equality, reproducing (2.4). In the atmosphere, with $n = 1.00028$ at sea level, Cherenkov light from highly relativistic particles with $\beta \approx 1$ is emitted in a thin cone of a maximum opening half-angle of ∼1.35°.

The number of Cherenkov photons emitted in the wavelength interval $d\lambda$ by a particle of charge Z traveling with velocity β over the path dx is given by Frank and Tamm (Frank and Tamm 1937):

$$\frac{d^2N}{dx\,d\lambda} = \frac{2\pi\alpha Z^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right), \tag{2.7}$$

where α is the fine structure constant and $n(\lambda)$ is the refractive index of light.

Integrating equation (2.7) over all wavelengths for which condition (2.5) holds, and over the total path of travel of a particle, gives the total number of photons emitted by a single particle within its Cherenkov cone.

2.1.4 Cherenkov light pool

As an extended shower initiated by a VHE gamma or cosmic ray develops in the atmosphere, every charged particle in it emits Cherenkov light in the direction of its travel. But showers have some lateral extent and momenta of the particles have non-zero components perpendicular to the shower axis, so the collective Cherenkov radiation from a cascade forms a pool of light on the ground. Thus, integrating expression (2.7) over all particles in the cascade and accounting for atmospheric absorption, we can compute the total number of photons produced in an electromagnetic cascade that are observed at a given point on the ground. The full utility of this exercise will be discussed in Section 3.2, but it is important to know some general features of the Chernkov light pool in order to understand the operation of Imaging Atmospheric Cherenkov Telescopes.

The photon density of the Cherenkov light pool on the ground varies with the energy of the primary from one to hundreds of photons per m^2. The refractive index of the atmosphere decreases with altitude, causing a focusing effect as the Cherenkov angle changes. This results in the pool having a disk-like structure with a radius of $\sim$130 m (Figure 2.5).

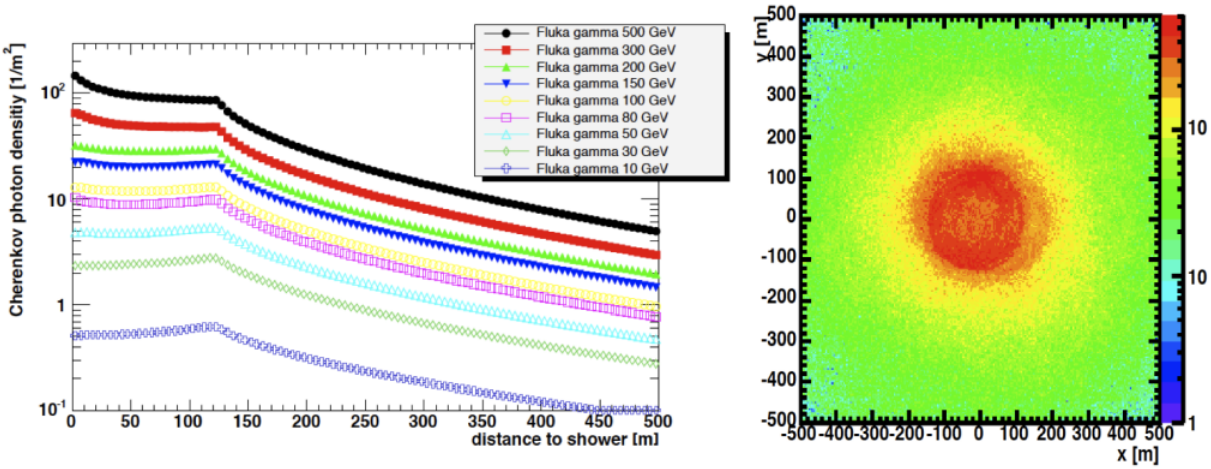

Figure 2.5: A simulated Cherenkov light pool. As the refractive index of the atmosphere increases with depth, the Cherenkov angle widens, causing "rings" of light emitted at different depths to overlap, a focusing effect that produces a light pool of radius $\sim$ 130 m. Figure from Gernot Maier.

Per the Frank-Tamm formula (2.7), the photon yield in the pool is inversely proportional to λ^2. The spectrum is therefore dominated by blue to UV emission, peaking around 350 nm (Figure 2.6) – shorter wavelengths get absorbed in the atmosphere, particularly the ozone layer in the lower parts of the stratosphere (97-99% of medium-frequency UV gets absorbed there), and do not reach the ground unless they are generated close to it by penetrating muons. The photons in the pool arrive within a few nanoseconds, forming a short pulse of light. While the averaged photon yield from all air showers constitutes only about 0.01% of the total night sky background (NSB) light, a single shower can outshine the brightest objects in the brief duration of its Cherenkov flash.

The size of the Cherenkov light pool and its duration are utilized by IACTs to image the particle showers that generate them (Figure 2.7). Placing a telescope with a camera

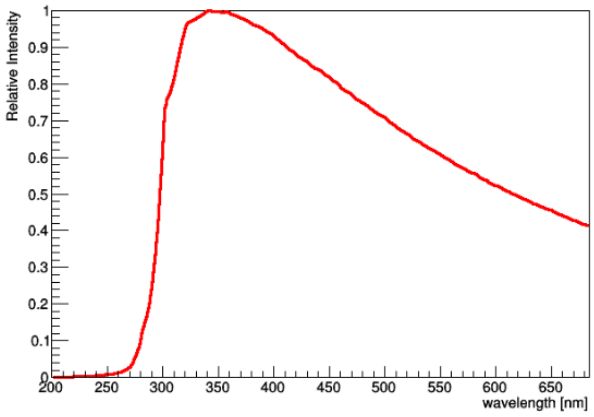

Figure 2.6: A simulated spectrum of Cherenkov light from showers initiated by gamma rays, as seen by Cherenkov telescopes on the ground. Photomultiplier tubes in TeV instruments target the prominent peak around $\sim$ 350 nm. Figure from Bird 2015.

capable of resolving nanosecond pulses within the light pool, one can effectively image the projection of the Cherenkov emission from extended air showers on the ground as an ellipse in the camera (Section 2.4) (Figure 2.8). The faintness of a single Cherenkov flash places a lower limit on the energy of primary particles whose showers can be detected; at the same time, determining the nature of the primary of an extensive air shower relies on resolving the morphology of its Cherenkov pool, hence the increasing duration, size and uniformity of the light pool at high primary energies places an upper limit. We shall discuss the details of IACT technology and its limitations in the rest of this chapter.

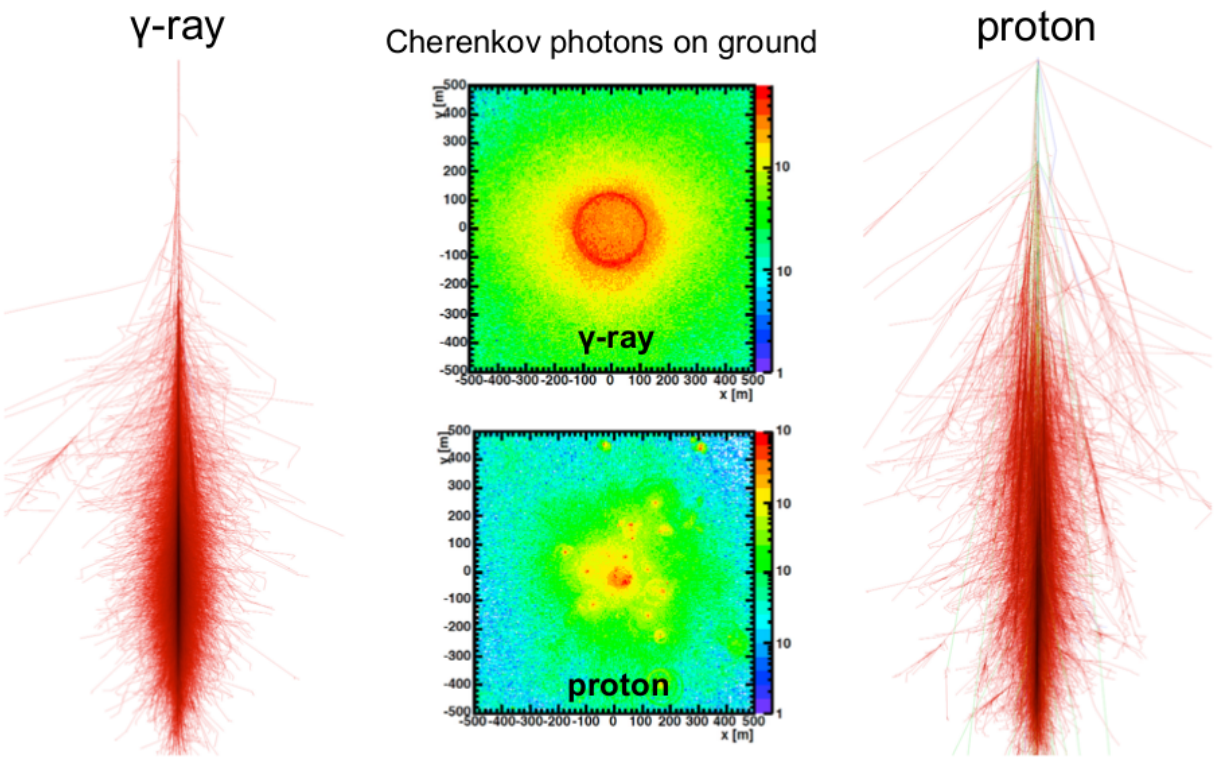

Figure 2.7: Simulated particle showers initiated by γ-ray and proton primaries of the same energy. The hadronic shower has a much larger lateral spread. The difference in the morphology of the resulting Cherenkov light pools is utilized by IACTs to distinguish between the primaries. Figure from Gernot Maier.

2.2 Quick History

The history of γ-ray astronomy is a surprisingly thrilling read (see, for instance, Weekes 2005; Hillas 2013; Mirzoyan 2013 and references therein), and here we present only a very brief overview of the development of the field.

The first mention of Cherenkov light emitted by cosmic rays is likely due to P.M.S. Blackett (Cronin 2011; Galbraith and Jelley 1953), who pointed out that perhaps 0.01% of the night-sky light is attributable to it. This rather benign remark was noticed by Bill

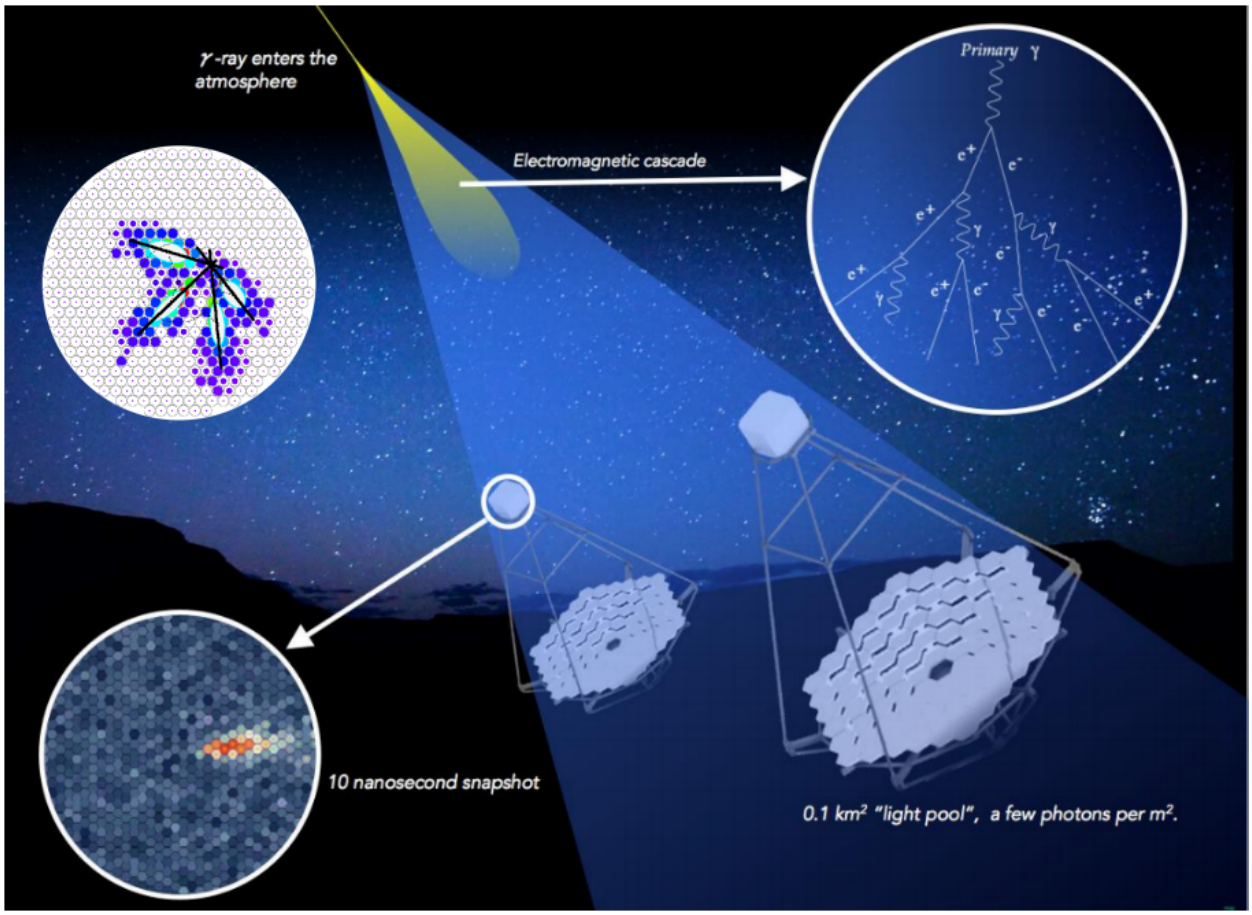

Figure 2.8: An overview of the Imagining Atmospheric Cherenkov Technique, with a shower as seen by the four VERITAS telescopes in the top left. Original figure from `https://cta-observatory.org`.

Galbraith and John Jelley, who realized that while the net Cherenkov light flux is exceedingly low, it might be possible to detect flashes from individual air showers. They tested out this idea within a week using a 5 cm diameter photomultiplier tube (PMT) in the focus of a 25 cm parabolic mirror, all mounted in a garbage can and coupled to an amplifier (Galbraith and Jelley 1953). Every two minutes, they observed oscilloscope triggers exceeding the night-sky background noise, confirming their theory. Soon after, a prediction of a strong

flux of TeV gamma rays from the Crab Nebula (Cocconi 1960) increased the interest in such experiments and motivated an instrument of a similar design in the Crimea (Hillas 2013; Mirzoyan 2013). These early instruments used World War II-era searchlights as mirrors and single PMTs as detectors (Figure 2.9) and were not very sensitive. A rapid growth in the number of candidates for TeV gamma ray emission precipitated the development of the first purpose-built gamma ray telescope in 1968, the Whipple Observatory 10 m Telescope on Mount Hopkins in Arizona (Figure 2.9) (Weekes 2005).

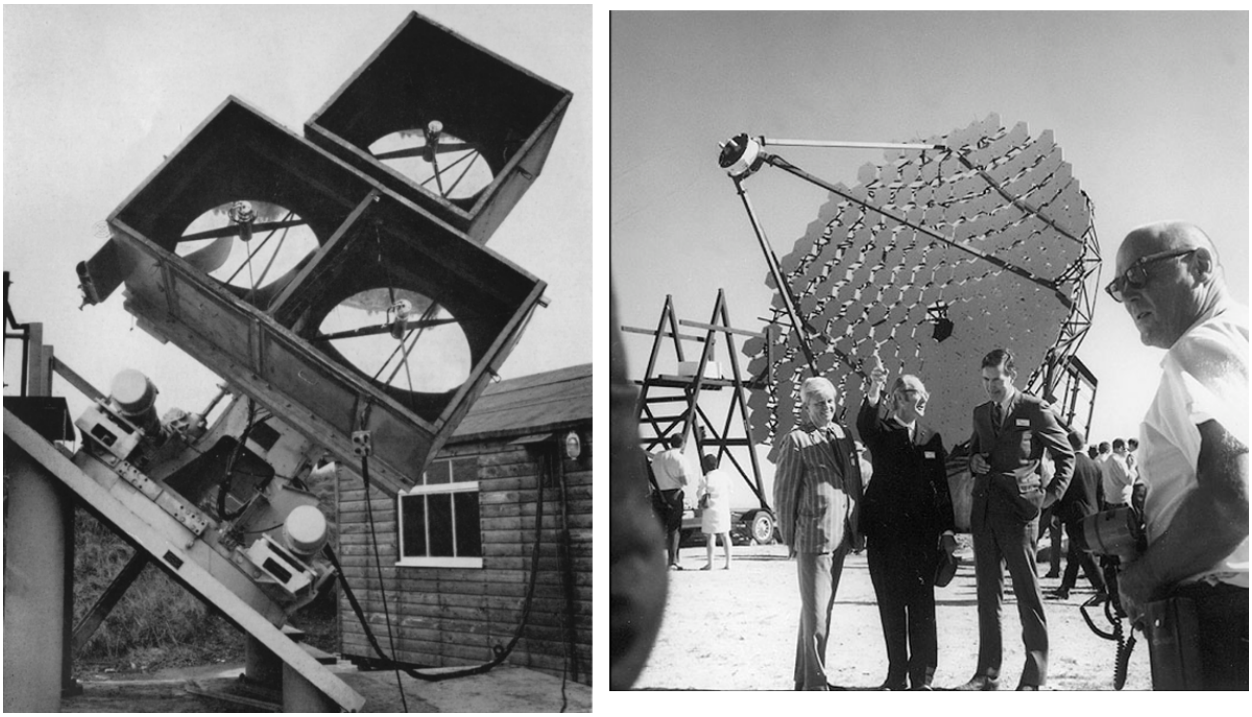

Figure 2.9: The first generation of TeV instruments. *Left:* Jelley's "light receiver" c. 1961; figure from Hillas 2013. *Right:* The unveiling of the 10-m Gamma Ray Reflector at the opening of the Mount Hopkins Observatory in 1968, with SAO director Fred Lawrence Whipple in the middle. The telescope and the observatory were renamed in his honor in 1981. Photo from `https://www.mmto.org`.

The first generation of instruments was quite simple, and did not provide the means of identifying gamma rays among the overwhelming background of cosmic rays. A detection of a gamma ray flux required a way to discard the cosmic ray background. Such a method came with the idea of using matrices of PMTs as cameras to image the Cherenkov light from air showers (Weekes and Turver 1977) and the upgrade of the Whipple 10 m to include a 37-pixel camera. A series of careful studies with Monte Carlo simulations led A. M. Hillas to the identification of image parameters that could be used for signal background discrimination (Hillas 1985). These *Hillas parameters* were used by the Whipple collaboration to credibly detect the TeV emission from the Crab Nebula for the fist time in 1989 (Weekes et al. 1989).

By the late 1990s, the limitations of a single instrument, such as the Whipple Telescope, and the benefits of an array of multiple detectors, had become apparent. The efforts for the third generation instruments, therefore, focused on the construction of arrays of Imaging Cherenkov Telescopes with more elaborate pixelated cameras. These efforts led to the construction of CANGAROO-III in Australia (Kubo et al. 2004), H.E.S.S. in Namibia (Hinton and the HESS Collaboration 2004), MAGIC on the Canary Islands (Cortina et al. 2009) and VERITAS in Southern Arizona (Holder et al. 2008). Apart from CANGAROO-III, which ceased its operations in 2011, these third-generation instruments are currently the most sensitive gamma ray experiments and are the major driving force behind the recent rapid increase in the number of known γ-ray sources and the understanding of VHE astrophysical processes.

Despite, or rather because of, the recent explosion of discoveries, γ-ray astronomy still

holds immense potential for future advancements of science. With the deveopment and construction of the Cherenkov Telescope Array (CTA) currently underway (Section 2.5; Chapter 6), the field is seeing another generational shift. And with the ability to combine γ-ray observations with data from gravitational wave and neutrino observatories, we are entering a whole new era of multimessenger astronomy.

2.3 The VERITAS Telescope Array

The scientific study undertaken in this work utilized data from VERITAS [2], an array of four 12 meter Imaging Atmospheric Cherenkov Telescopes located at the Fred Lawrence Whipple Observatory in Southern Arizona (Figure 2.10) (Holder et al. 2008).

Figure 2.10: The VERITAS Telescope Array.

The relatively wide field of view of VERITAS of $\sim 3.5°$, and the need for a large aperture,

[2] `https://veritas.sao.arizona.edu/`

result in a relatively small f/D ratio. To handle off-axis optical aberrations in such a system, the Davies-Cotton (DC) design (Davies and Cotton 1957) with 350 hexagonal tessellated reflectors is used for the mirrors. A downside of this design is the lack of isochronicity, inducing a spread in photon time arrivals. While parabolic mirror figures fix this problem, they are much more complicated and expensive to produce, since mirror facets of different shapes and sizes are needed. On the other hand, aplanatic two-mirror designs solve the issue of aberrations, while also being isochronous and reducing the plate scale, making it possible to use advanced finely pixelated camera designs. Such approach is used in the prototype Schwarzschild-Couder telescope, detailed in Chapter 6.

In order to adequately resolve Cherenkov shower images and cover a wide field of view, VERITAS employs a large camera of 499 photomultiplier tubes (PMTs) with a photon detection efficiency of 20-30%. The PMTs have nanosecond-scale response times, large detection areas and very clean signal amplification, and are capable of detecting single photons. The dead space between them is recovered with Winston cones (Figure 2.11). The relatively large size of the PMTs places a rather relaxed constraint on the quality of the alignment of the mirror figure – an optical PSF of $\sim 0.1^\circ$ is sufficient, allowing the telescopes to rely on the rigidity of their structures, with mirror alignment performed only about once per year. It should be noted that the PSF for gamma ray imaging is better than the optical PSF because the arrival directions of primary gammas are reconstructed from several stereoscopic images.

Cherenkov flashes arrive randomly and last only a few nanoseconds, but continuously taking snapshots of the sky at GHz rates in the hopes of recording them is completely

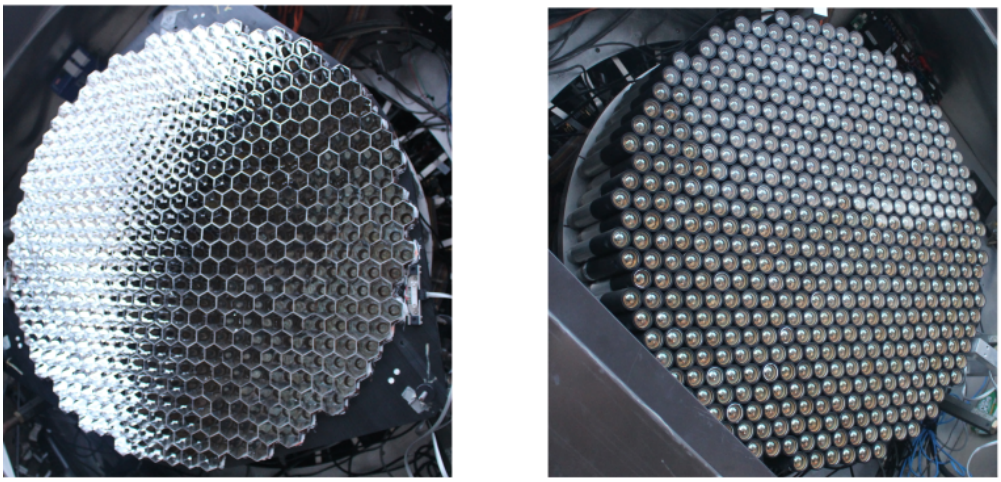

Figure 2.11: The VERITAS camera. *Left:* The Winston cones minimize the dead space between the PMTs. *Right:* The camera without the Winston cones.

impractical. Therefore, a way to *trigger* the camera in order to record PMT outputs only in a short time window around a candidate Cherenkov event is needed. In the case of VERITAS, the trigger system consists of three levels of hardware triggers. This multi-level design is driven by the need to trigger on the faintest flashes while keeping sane rates of data acquisition by excluding false event candidates due to the NSB (Cogan 2006; Weinstein 2007; Zitzer 2013). The data acquisition system employs 500 megasample-per-second flash analog-to-digital converters (FADCs) (Hays 2007), which continuously digitize the analog PMT signal every 2 ns with an 8-bit (256 digital counts per channel) dynamic range. The last 32 μs of the FADC readout are stored in a memory buffer, awaiting an array-level event trigger signal. When such a trigger occurs, the FADC digitization is paused and the appropriate block of memory is read out from the buffer. The 3-level trigger system of VERITAS proceeds as follows:

- The first stage, the L1 trigger, is a *pixel-level trigger* and occurs when the signal from an individual PMT exceeds a pre-determined threshold. More specifically, each PMT is connected to a constant fraction discriminator (CFD), which finds the maximum of the PMT pulse. This stage is triggered whenever the running maximum exceeds an experimentally set threshold value. The possibility of NSB fluctuations triggering this first stage places a lower limit on CFD thresholds; an upper limit is determined by the desire to include the faintest possible Cherenkov events.
- The L2 trigger is a *telescope-level trigger* and occurs when multiple neighboring pixels trigger simultaneously. The idea here exploits the fact that unlike NSB noise, which is uncorrelated between pixels, Cherenkov showers are compact in their angular and temporal extent. Level 2 triggers when 3 neighboring PMTs are triggered within 5 ns of each other (Zitzer 2013).
- The third and final stage, the L3 trigger, extends the idea of the L2 trigger to the whole array – this is the *array-level* trigger and it occurs when at least 2 telescopes trigger within a short time frame. The specific time interval is programmable and is usually set to 50 ns. When this stage is triggered, an *event* occurs, and a signal is sent to the FADCs to read out and record 16 ns (8 samples) of PMT output from the relevant range in the memory buffer. This trigger vetos single-telescope events (in particular, muons) and assures that the recorded events have a high reconstruction quality (Weinstein 2007).

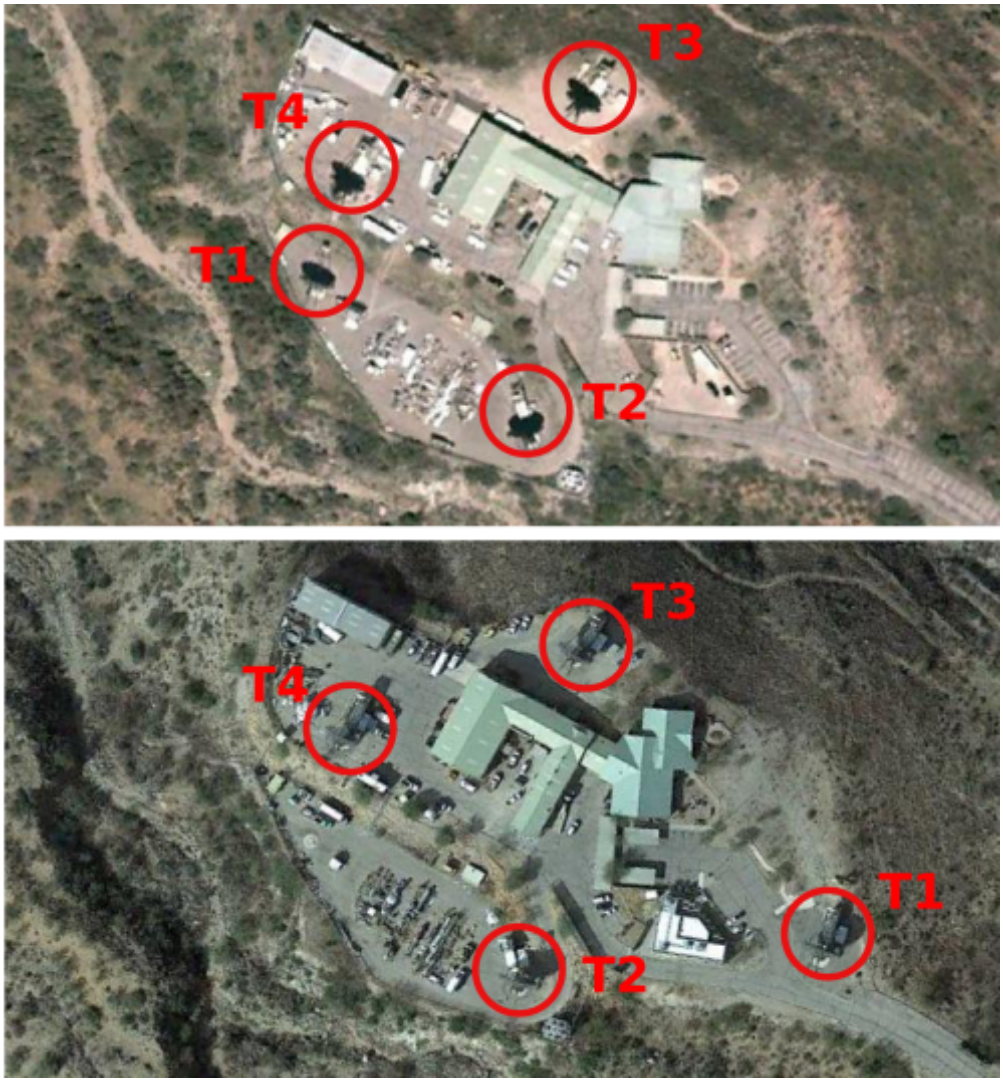

Figure 2.12: The evolution of the layout of the VERITAS array. *Top:* The original array, before the relocation of T1 in 2009. *Bottom:* The current array configuration. Figure from Bird 2015.

Since first coming online in 2007, VERITAS has undergone two major upgrades. In the first upgrade in the summer of 2009, T1 was relocated from its initial position near T4 (Figure 2.12). The initial proximity of T1 and T4 resulted in many showers looking very similar in the two telescopes, and the relocation improved the angular reconstruction and the overall

sensitivity of the array by increasing the average baseline. In fact, after the relocation, T1 ended up being the furthest from its neighbors, and therefore the most important in event reconstruction. The second upgrade in the summer of 2012 replaced the camera PMTs with higher quantum efficiency ones, lowering the energy threshold and further increasing the overall sensitivity. The three epochs of the instrument are referred to as: Old Array (OA or V4) – the period before the T1 move; New Array (NA or V5) – the period after the T1 relocation but before the camera upgrade; Upgraded Array (UA or V6) – after the camera upgrade and the current configuration. When performing data analysis, care must be taken to properly account for differences in the instrument response and systematics between the three epochs.

Stereoscopic imaging of Cherenkov showers, facilitated by a sophisticated trigger system, dramatically improves angular resolution of gamma primaries and increases sensitivity compared to single images. While VERITAS uses only four telescopes, studies have shown (Bernlöhr et al. 2013) that the sensitivity of IACT arrays increases roughly as the square root of the number of telescopes thanks to the increased image multiplicity of each event – this is the motivation behind the Cherenkov Telescope Array, discussed in Section 2.5.

2.4 Data Analysis for VERITAS

In order to extract meaningful physical information, the snapshots of Cherenkov light recorded by instruments such as VERITAS need to be processed and analyzed. The analysis

chain from PMT currents to a physical spectrum is long and involved, and its efficiency plays a major role in the sensitivity of the whole array. Due to its complexity, the analysis is usually broken into several stages, with each subsequent stage moving further away from raw hardware-specific data and closer to physically relevant products. To ensure consistency and accuracy, VERITAS has a policy of double-checking all of its results with two independently developed analysis packages. In this work, we use, and contribute to the development of, one of the packages, VEGAS (Cogan 2007). In their analyses, VERITAS packages rely on the ROOT data analysis framework[3] (Brun and Rademakers 1997).

In order to observe each target at its maximum possible elevation, VERITAS observations are conducted in relatively short windows of 20-30 minutes, known as *runs*. Each data run needs to be downloaded and analyzed separately, with physical analysis products combined at the very last stage. In VEGAS, data analysis is broken into 5 stages:

- *Calibration calculation.* The main purpose of this stage is to identify bad pixels, determine gain and timing parameters for each PMT and calculate the noise levels of the fluctuating NSB. To be able to do this, a supplementary *flasher* run, which uniformly illuminates the camera at a fixed rate, is taken each night for calibration purposes (Hanna et al. 2010).
- *Calibration application.* As the name suggestes, this stage applies the calculated calibration to data runs, performs timing correction (Holder 2005) and image cleaning

[3] `https://root.cern.ch/`

(Reynolds et al. 1993). Moreover, it performs the calculation of Hillas parameters for each telescope image.

- *Shower Reconstruction.* This stage of the analysis uses the cleaned up and parameterized images of individual telescopes to construct a stereoscopic view of each shower. With this, it then reconstructs the physical parameters of the shower primary, its direction, energy and core location – the location of the hypothetical impact of the primary on the ground. While the majority of the recorded events are initiated by cosmic rays rather than γ-ray primaries, these reconstruction techniques are applied to all events. Some showers aren't seen by all telescopes, while others are too faint or too far from the array. Such events fail *reconstruction quality criteria* set at this stage and are eliminated from the analysis.
- *Event Selection.* This is where the overwhelming background of cosmic rays is reduced and the signal-to-noise ratio is increased. At this stage of the analysis, the reconstructed shower parameters are used in conjunction with parameters that characterize the morphology of the Cherenkov pool to eliminate as many hadron-like events as possible. While it's not possible to eliminate all background events (Section 2.1), this stage pares them down significantly, with the background rejection efficiency of $\sim$ 90%. The aggressivness of event selection at this stage is controlled by user-defined *event selection cuts*.
- *Results extraction.* At this final stage, events that passed the previous selection cuts are used to measure the quantities of physical relevance, such as fluxes or their upper limits,

and spectra. Here, skymaps of excess events are plotted, significances of observations calculated, lightcurves constructed.

The first two stages of the analysis are referred to as the *lower* stages, due to their specificity, closeness to the hardware and a relatively low flexibility from the perspective of the end user. We won't be discussing them further – a full description of these stages of the analysis can be found in, for instance, Cogan 2006. The last three stages are known as the *upper* stages, and a high degree of familiarity and user control are required here. Chapter 3 discusses the improvements to shower reconstruction and event selection contributed by this work. Here, we give a concise review the upper stages of VERITAS analysis.

2.4.1 Shower Reconstruction

Analysis of VERITAS data relies on a large number of Monte-Carlo simulations. These simulations follow γ-ray primaries and the showers they initiate in the atmosphere, compute the Cherenkov emission from the showers, and produce a simulated waveform in each pixel, accounting for the mirror shape and reflectivity, shadowing by the telescope structure, response of light cones and PMTs, and electronic effects, such as stretching and attenuation of the pulse shape in the cables from the camera to FADCs, and FADC response. Every data image is parameterized, and its parameters, and the telescope pointing conditions, are then fit to a precomputed lookup table (more on this, as well as on a different reconstruction method, in Chapter 3) in order to reconstruct its energy and determine how likely it is to

be a signal event.

After image cleaning, the remaining image of a shower is parameterized according to the procedure due to M. Hillas (Hillas 1985). Each shower is seen in a telescope camera as an ellipse, whose width corresponds to the lateral spread of the shower, and whose length measures the longitudinal development of the shower. The computed Hillas parameters (Figure 2.13) are essentially the zeroth, first and second moments of the image used to characterize this ellipse:

- *nTubes* is the total number of pixels composing the image.
- *Size* is the total integrated charge in all of the pixels, or equivalently, the total intensity of the image.
- *Width* is the minor axis of the shower ellipse, a measure of the lateral development of the shower.
- *Length* is the major axis of the shower ellipse, a measure of the longitudinal development of the shower.
- *Distance* is simply the distance from the center of the ellipse to the center of the camera, and is a measure of the shower impact distance.

The Hillas parameters are used to reconstruct the direction, the core location and the energy of the primary, as well as to calculate mean shower parameters, a process known as *stereo reconstruction*. First, the direction is reconstructed in a straightforward geometric way by placing the individual telescope images into a single camera plane and finding the

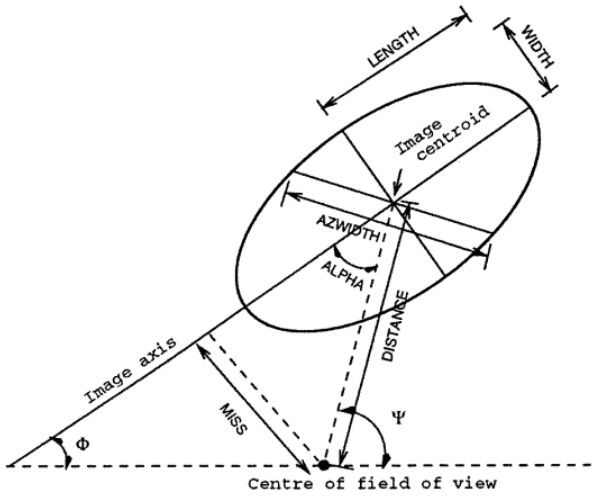

Figure 2.13: Hillas parameterization of shower images. This paremeterization is computed for every telescope, and the results are averaged into array parameters. Figure from Fegan 1997.

intersection of their major axes (Figure 2.14, left). Similarly, the core location, the point at which the primary would have hit the ground, is reconstructed by transforming the images into the ground plane and finding the intersection of their major axes in these new coordinates (Figure 2.14, right). This method is exceedingly simple, but suffers from a loss of precision if not all telescopes detected an event; if images are clipped, which happens when the shower core is far from the array; when the determination of the axes of the image is ambiguous, which happens both for very small (low energy) and very large (high energy) images; or if several of the images have almost parallel major axes, which happens at elevations lower than 55°. More sophisticated analyses or bigger arrays with a higher image multiplicity are required to increase the accuracy of angular reconstruction.

To reconstruct the energy of the primary, we exploit the fact that for a given impact

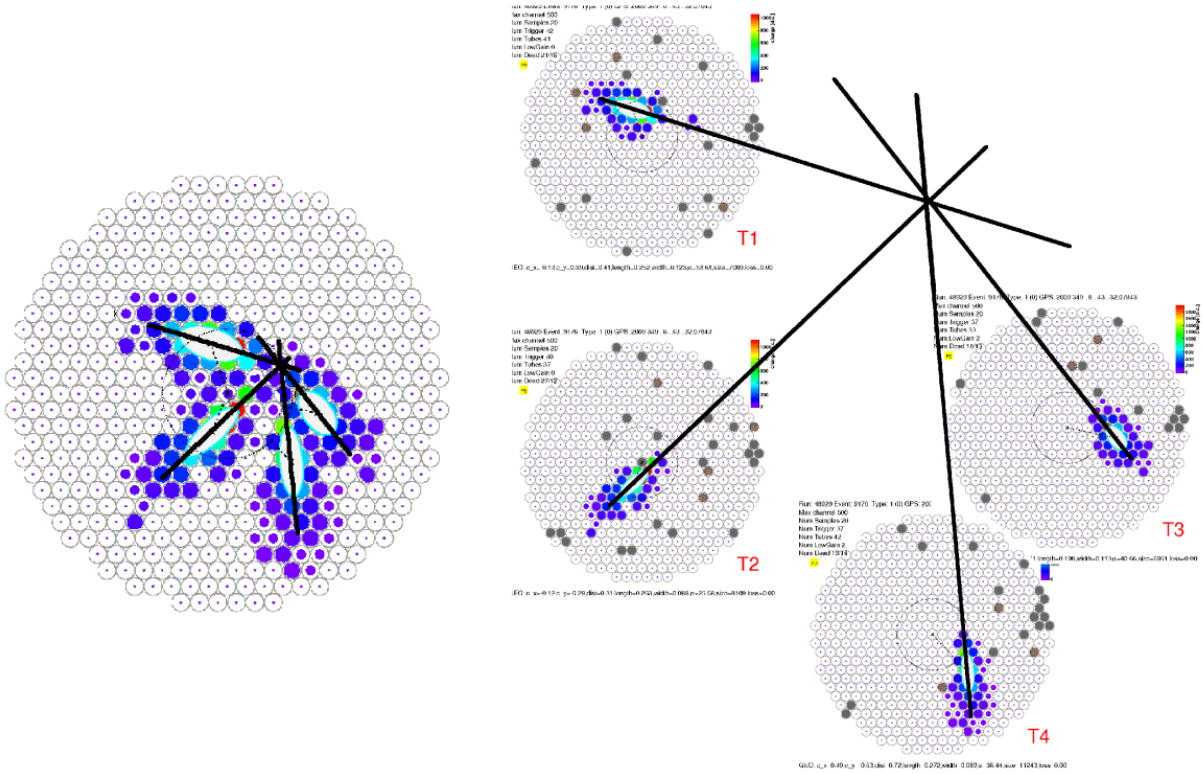

Figure 2.14: A simple geometric approach to reconstructing the direction and core location of the shower primary. *Left:* direction reconstruction. *Right:* core location reconstruction. Figure from Andrew McCann.

distance from a telescope, the image size is approximately linear in the primary energy (Section 2.1). To account for the dependence of this linear slope on the impact distance, lookup tables are used. These lookup tables are simply histograms filled with impact position and image size for each telescope, and the corresponding energy of the simulated primary. The accuracy of such reconstruction depends on the linear relationship between image size and impact distance, and the accuracy with which impact distance is calcualted, and thus suffers from the same problems as the angular reconstruction mentioned above. Thus, at low ($\lesssim$ 100 GeV) and high ($\gtrsim$ 30 TeV) primary energies, the energy reconstruction suffers from a high

bias. The estimated energy of the primary is a weighted average of energies reconstructed by each telescope. The variance of this estimate determines the energy resolution. It should be clear that once again, more sophisticated analysis techniques that don't rely on approximate linearity or larger arrays are needed to decrease the energy bias and resolution.

Finally, the so-called stereo parameters are calcuated for each event. These are:

- The mean scaled width (MSW) and mean scaled length (MSL) – weighted averages of the width and length Hillas parameters:

$$\mathrm{MSW} = \frac{1}{N_{\mathrm{tel}}} \sum_{k=1}^{N_{\mathrm{tel}}} \frac{w_k}{\langle w_{\mathrm{sim}}(El_k, I_k, s_k)\rangle}, \quad \mathrm{MSL} = \frac{1}{N_{\mathrm{tel}}} \sum_{k=1}^{N_{\mathrm{tel}}} \frac{l_k}{\langle l_{\mathrm{sim}}(El_k, I_k, s_k)\rangle}$$

  Here, the individual telescope parameters are weighted by their expected simulated values for a given elevation El_k, impact distance I and image size s.
- The RMS values of the above – how much per-telescope scaled width and length differ from the array means.
- The height of the shower maximum, derived from the image widths, impact distances, and the primary energy (see, for example, Patterson and Hillas 1983).

In order to mitigate the issues leading to the loss of accuracy described above, the decision to include the image from a given telescope into the shower stereo reconstruction is based on the quality of that image. To avoid images that are too small, a requirement on the minimum number of pixels in the image is imposed, $nTubes \geq 5$; to exclude clipped images,

a requirement on the distance of the image centroid to the camera is set as $Dist \leq 1.43$. The lower limit on the image multiplicity is usually set to 2 or 3. Finally, three different levels of cuts are imposed on the *size* parameter. These are *soft, medium,* and *hard* and their choice depends on the needs of the specific analysis. The full impact and meaning of the *size* cut will be explored in Chapter 3.

2.4.2 Event Selection

Having reconstructed the shower direction, core and energy, VEGAS performs event selection by placing simple box cuts on the mean scaled parameters. The idea here is that gamma rays result in showers with a different distribution of these parameters than hadrons (Section 2.1), see Figure 2.15. Specifically, hadron showers have a much broader distribution of these parameters, and placing a restrictive cut on the maximum allowed value of the mean scaled parameters will eliminate the majority of background while keeping most of the signal.

This approach to gamma-hadron separation has worked remarkably well for IACTs. However, due to its simplicity, it is not very efficient and is one of the main determinants of instrumental sensitvity. In particular, this procedure is not very flexible – the selection of an optimal cut has to be done through trial and error on a test source. It is also not very tolerant of nonlinear relationships and correlations between the parameters or their dependence on other shower characteristics, such as energy or direction. Finally, there are other useful parameters overlooked by the simple box cuts. We improve upon the VERITAS

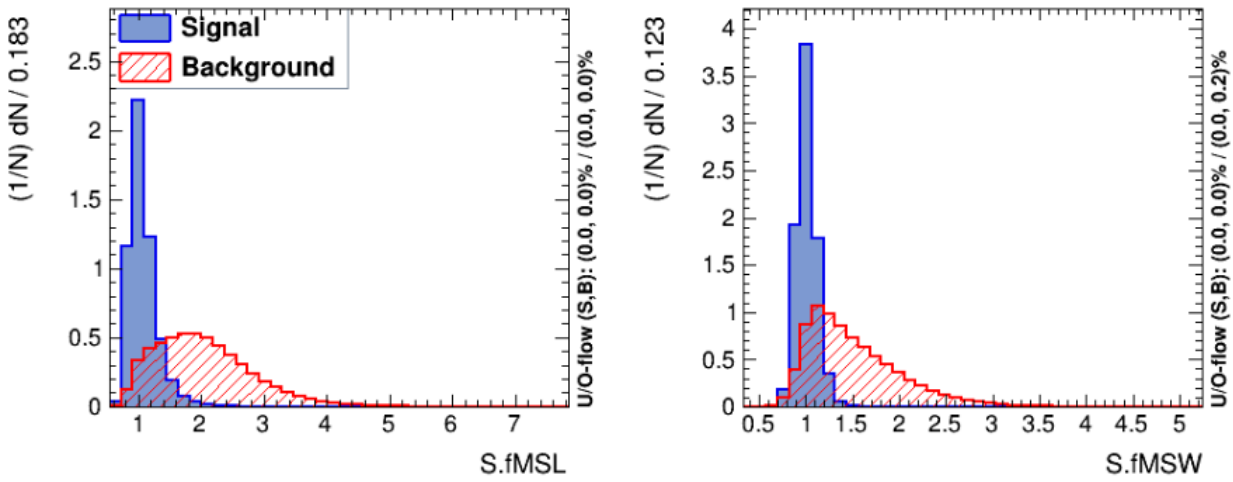

Figure 2.15: Gamma-hadron separation is based on the difference in distributions of shower parameters. *Left:* the distributions of mean scaled length are compared between gamma-initiated showers in blue and hadron-initiated showers in red. *Right:* the same distributions for mean scaled width.

gamma-hadron separation in Chapter 3.

2.4.3 Putting it all together

Now that much of the background has been eliminated, it is time to calculate physically relevant quantities, the last stage of VEGAS analysis. For VERITAS, the quantities of interest are the distribution of gamma rays in the sky, the total gamma ray flux, and its energy spectrum. Event selection does not eliminate all of the background, so any measured gamma ray flux is detected as excess above the expected background. Therefore, correctly estimating this background and the statistical signficance of the gamma flux is of high importance.

On the face of it, it should be straightforward to estimate the background flux – exclude

events in the direction of the source of interest and count up the remainder. However, due to a variety of factors, the camera acceptance varies with distance from its center – this means that regions of the camera further from the center have a lower effective exposure. If we denote the ratio to the exposure of the control region to that of the siginal region as α, we have

$$N_{\mathrm{Excess}} = N_{\mathrm{On}} - \alpha N_{\mathrm{Off}}, \tag{2.8}$$

where N_{On} and N_{Off} are the event counts from the signal (On) and control (background or Off) regions, respectively, and N_{Excess} is the number of excess gamma events above the expected background. To test the statistical significance of such an excess, we can employ a likelihood ratio test with the null hypothesis $\{H_0 : E[N_{\mathrm{On}}] = \alpha E[N_{\mathrm{Off}}]\}$, modeling the event counts as coming from a Poisson process. The widely used result, known as the Li Ma equation 17 (Li and Ma 1983), computes the significance of an excess in standard deviations of a normal distribution as

$$S(N_{\mathrm{On}}; N_{\mathrm{Off}}; \alpha) = \sqrt{2}\left\{ N_{\mathrm{On}} \ln\left[\frac{1+\alpha}{\alpha}\frac{N_{\mathrm{On}}}{N_{\mathrm{On}}+N_{\mathrm{Off}}}\right] + N_{\mathrm{Off}} \ln\left[(1+\alpha)\frac{N_{\mathrm{Off}}}{N_{\mathrm{On}}+N_{\mathrm{Off}}}\right]\right\}^{1/2} \tag{2.9}$$

N_{On} and N_{Off} are both the results of a direct measurement of the number of events; however, due to the varying camera acceptance, the exposure ratio α is not an obvious parameter and its accurate determination is important. There are two approaches to measuring α (Figure 2.16). In both, the On region is defined in the sky coordinates with a directional cut, denoted

as θ^2, with θ measuring the angular distance of an event from the center of the field of view. For point sources, θ^2 is typically $0.01\,\mathrm{deg}^2$, reflecting the PSF of the instrument.

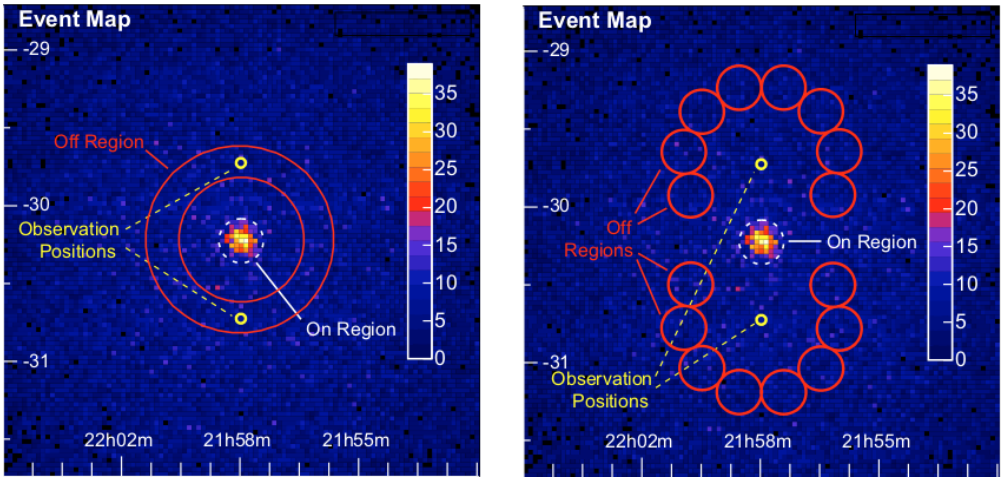

Figure 2.16: The choice of Off regions and determination of the α parameter. *Left:* the Ring Background Method – a ring of constant thickness is selected around the source position. This ensures a large number of events for background estimation but requires the estimation of a varying camera acceptance. *Right:* the Reflected Regions method – the Off regions are obtained by rotating the On region about the center of the field of view to ensure constant acceptance for both On and Off regions. In this case, α is the reciprocal of the number of Off regions. Figure from Berge, Funk, and Hinton 2007.

The Ring Background Method selects a ring of constant thickness around the source position. The acceptance varies within this ring, and needs to be corrected for. This is done with the assumption of azimuthal independence, by segmenting the field of view into narrow rings and setting the acceptance for a given radial distance θ as the fraction of all events that fall within that ring. Clearly, a region around the tentative source needs to be excluded from this calculation, and in order to be able to estimate the acceptance at the source location, the source needs to be offset from the center of the field of view – in

other words, the telescope needs to point with a slight offset from the source, known as a *wobble*. VERITAS telescopes usually wobble 0.5° around the source, alternating between North, South, East and West. The ring background method works well when there are many events that allow for an accurate determination of the acceptance. If there are not enough events to accurately calculate the acceptance, a library of acceptance curves from past observations can be used. With the camera acceptance being energy-dependent, this method introduces too much uncertainty in the measurement of spectra and is usually used only for the determination of the overall significance of the source, as well as to plot maps of excess events and their significance.

The Reflected Regions Method also requires a wobble in telescope pointing. Here, the source region is rotated about the camera center to create a number of Off regions that are located at the same radial distance from the center of the field of view as the On region. This ensures that the camera acceptance is the same for all regions and results in a very straightforward estimation of $\alpha = 1/R_{\mathrm{Off}}$, with R_{Off} being the number of the Off regions. This value of α is independent of energy and can therefore be used to accurately measure the gamma ray flux and spectrum. Its use, however, is limited to relatively compact sources, when at least several reflected regions can be constructed.

In order to reconstruct the spectrum the excess events have to be binned by energy. The gamma ray flux in each energy bin is then the number of excess events per unit time and per unit detection area. The total live time is the total exposure adjusted for the dead time of the data acquisition system, and minus any time cut out due to, say, bad weather. The

total area of the detector is the total effective area (EA) at the top of the atmosphere around the array such that gamma rays falling within it get detected. This area is a function of observing conditions, the primary energy and the analysis cuts, and is once again computed with simulations: a number N_{thrown} of showers are thrown over a region A_0 around the array large enough that any additional showers outside of it won't trigger the instrument. These showers are then analyzed through the full analysis chain, with the correct analysis cuts, resulting in N_{passed} events. The effective area is then

$$A_{\text{eff}}(E) = A_0 \frac{N_{\text{passed}}(E)}{N_{\text{thrown}}(E)},$$

where we make explicit the dependence on energy. Because the effective area, as well as the reconstructed excess events, are binned, all with finite energy resolution, the conversion between excess events and the gamma flux is not a simple scaling, as energy-dependent migration between bins needs to be accounted for. For the full details of this conversion, see, for instance, Mohanty et al. 1998.

This concludes the overview of VERITAS and its analysis. Several directions for improvement of the Imaging Atmospheric Cherenkov Technique should have become apparent by now, and this work explores both the instrumental and the analysis-based paths.

2.5 Beyond VERITAS: The Cherenkov Telescope Array

The Cherenkov Telescope Array[4] can be thought of as the next natural evolutionary step in the deveopment of γ-ray astronomy. The limited field of view of the current instruments makes it difficult to conduct surveys or study extended sources; the small sizes of the arrays and the modest mirror areas restrict them to a relatively narrow range of energies; both the small sizes of the arrays and the relatively coarse cameras limit the angular resolution and background rejection capabilities; the use of PMTs confines observations to clear moonless nights; finally, the reliance on proprietary data formats and in-house analysis packages substantially hampers broader collaboration. The Cherenkov Telescope Array (Figure 2.17) aims to address all of these limitations (CTA Consortium 2017) by building an open, proposal-driven observatory with large arrays of IACTs on two sites, the Northern site in La Palma in the Canary Islands in Spain and the Southern site in the Atacama Desert in Chile. The operations of the CTA are expected to begin in 2022, with the construction finishing some time in 2025.

The CTA plans to achieve its performance goals by combining several major design changes. First, it will employ much larger arrays than those currently in operation. In order to optimize cost allocation, several sizes of telescopes will be used. For the core energy range of 100 GeV to 100 TeV, 40 Medium-Sized Telescopes (MSTs) of 10-12 m in diamater will

[4] `https://www.cta-observatory.org/`

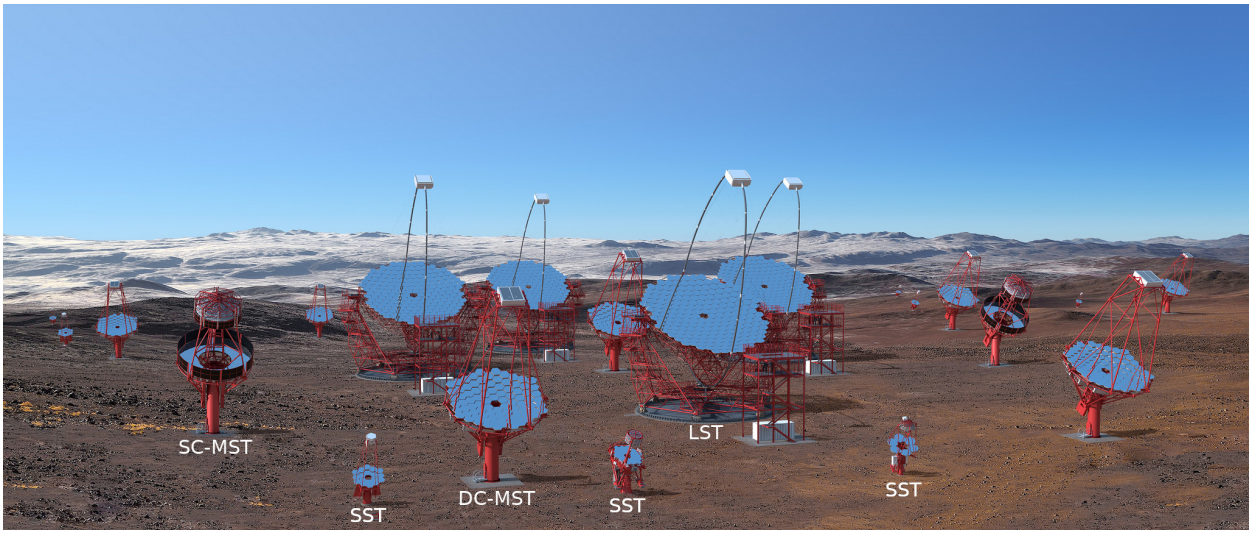

Figure 2.17: A rendering of the Cherenkov Telescope Array showcasing different types of telescopes in the array. The Schwarzschild-Couder Medium-sized Telescope is on the left. Three different proposals for Small-Sized Telescopes are not labelled individually. Original figure from `https://www.cta-observatory.org`.

be distributed between both sites; an additional 8 23-meter Large-Sized Telescopes (LSTs), 4 at each site, will extend the core energy range down to about 20 GeV; 70 Small-Sized Telescopes (SSTs) at about 4 meters in diameter will further extend the energy range of the Southern array to 300 GeV.

Second, the CTA telescopes will use much more finely pixelated cameras with wide fields of view – $> 4.5°$ for the LSTs, $> 7°$ for the MSTs, and $> 8°$ for the SSTs. This, combined with the large number of telescopes, will make it possible to conduct surveys much more efficiently than current-generation instruments. In particular, different sub-arrays will be able to have different pointing directions, observing multiple sources and wide patches of the sky simultaneously. The optical support structures and pointing systems of the telescopes

within the array have been optimized for rapid slewing, which, together with the wide field of view will enable the CTA to detect transient events, such as gamma ray bursts.

Finally, massive ongoing simulation and data analysis efforts will increase the analysis efficiency of the CTA compared with current instruments, as well as standardize analysis tools and make them open to the public. Specifically, techniques deemed advanced within current experiments, such as the use of image templates and machine learning approaches (Chapter 3), will be standard within the CTA. The improved analyses will make full use of the finely-pixelated cameras and the large image multiplicity.

The large size of the CTA collaboration and the need to optimize cost allocation has led to several proposals regarding the design and configuration of the medium- and small-sized telescopes. Specifically, two possible designs for the medium-sized telescope are currently evaluated – the 12 m Davies-Cotton design (DC-MST), which is analogous to currently operating VERITAS and H.E.S.S. telescopes, and the 10 m Schwarzschild-Couder design (SC-MST), a novel two-mirror design that constitutes a radical departure from current practices. The SC-MST has numerous advantages over the traditional DC-MST, but incurs higher per-telescope costs and requires a complicated optical alignment system.

Much of this work focused on the design, development and implementation of the optical alignment system of the prototype SC-MST, currently starting its very first operations at the VERITAS site. We will return to the discussion of this instrument in Chapter 6.

Chapter 3

Improving VERITAS Analysis

As Chapter 2 pointed out, traditional IACT data analysis methods, while very robust and easily interpretable, are relatively simple and inefficient. Several approaches to improving their efficiency can be taken, and the goal of this chapter is to do just that.

In this work, we focus on two major areas of improvement of IACT analysis – event reconstruction and event selection. Event selection in IACT data (Section 2.4) relies on the fact that appropriately parameterized, gamma-initiated showers look different, on average, from hadron-intiated showers. The typical approach is to find two or three parameters that characterize showers, find their distributions for the two populations of showers, and find optimal ranges of these parameters, such that, discarding all events outside of these ranges gets rid of a large number of background events while retaining the majority of the signal. This method, called *box cuts*, does not take into account any interactions between the parameters of choice or the multidimensional nature of the full parameter space, however

small. An obvious improvement would be to instead find an optimal region in the full *multidimensional space* of shower parameters. This is precisely the idea behind the machine learning method of *boosted decision trees (BDTs)*, which we develop and implement here.

BDTs have been used for data analysis by the VHE astroparticle physics community for almost a decade now, with the first studies of BDTs for gamma-hadron separation in IACT data by the H.E.S.S. collaboration in 2009 (Ohm, Eldik, and Egberts 2009). EventDisplay, one of the analysis packages employed by the VERITAS collaboration, has been using BDTs since 2015 (Krause, Pueschel, and Maier 2017), achieving better background rejection and higher sensitivity than simple box cuts. While VEGAS has relied on simple box cuts up until now, this work brings it to parity with, and in fact, surpasses, EventDisplay, offering a complete and robust implementation of BDT analysis for gamma-hadron separation.

Event reconstruction, described in Section 2.4, is largely geometric, relying on first and second moments of cleaned shower images. This, by construction, discards a lot of available camera data, reducing all the pixel information to only a handful of averaged parameters. Further, energy reconstruction relies on the assumption that for a given impact distance, the primary energy is only a function of the image size, a relationship encoded in lookup tables derived from simulations. Making better use of camera pixel data and constructing a more realistic model of the energy dependence on shower parameters is the goal of the *image template method (ITM).*

In ITM (Le Bohec et al. 1998; de Naurois and Rolland 2009; Parsons and Hinton 2014), a model of atmospheric shower development, as well as a detector model, are used to calculate

the distribution of Cherenkov light in the camera for a gamma primary with a given direction, core location, energy and, importantly, the depth of the first interaction. These computed distributions are used to construct a collection of *templates*, which are simply the expected, or predicted, images for a given primary. In the reconstruction step, the template that best agrees with the real image is found, resulting in a simultaneous fit of all 6 parameters of the primary. While this work relies on an existing implemention of ITM (Christiansen and VERITAS Collaboration 2017), it contributes to its development and validation within VEGAS.

We also take the natural step of combining the two improvements into a new analysis, *ITM+BDT*, using ITM-derived parameters in the construction of the BDT model. In the following, we start with overviews of boosted decision trees and the image template method. After these initial reviews, we go into details of the implementation of the combined ITM+BDT method, postponing the discussion of its performance and limitations to the next chapter.

3.1 Boosted Decision Trees

We start with a description of boosted decision trees. Tree-based methods for regression and classification involve partitioning the parameter space of predictor variables into small simple regions. Such partitioning process, called *training*, also decides the predicted value of the target variable as the mean or the mode of the training observations in the region

into which it falls. While the theory behind using decision trees for regression is very similar to that of classification (and in fact, it is often easier to understand regression with BDTs), here we focus on classification only, as our problem is the assignment of events to one of two classes, gammas or hadrons.

3.1.1 Simple Trees

Decision trees classify instances by sorting them down the tree structure from the *root*, the top node, to some *leaf*, a terminal node. In a simple decision tree (Figure 3.1), each node specifies a test of some *attribute* of the instance under classification, and the branches descending from the node specify all possible values of the attribute. In practice, with attributes taking on continuous values, a *splitting criterion* is chosen, creating a binary choice for the possible values of the attribute, less or greater than the splitting criterion.

As already mentioned, we can intrepret this decision tree structure as a segmentation, or partitioning, of the space of attribute values into non-overlapping *rectangular* regions, with a classiffication label applied to each region as a whole. To better illustrate this crucial point, Figure 3.2 displays a simple decision tree over two attributes X_1 and X_2. At the root node, an attribute is chosen to perform a test on. For this attribute, a splitting criterion is then chosen, creating two branches corresponding to the two possible choices of the value of the attribute. The process is then repeated for each of the branches, until terminal leaves are reached, where the exact classification label is assigned to the tested instance. It is

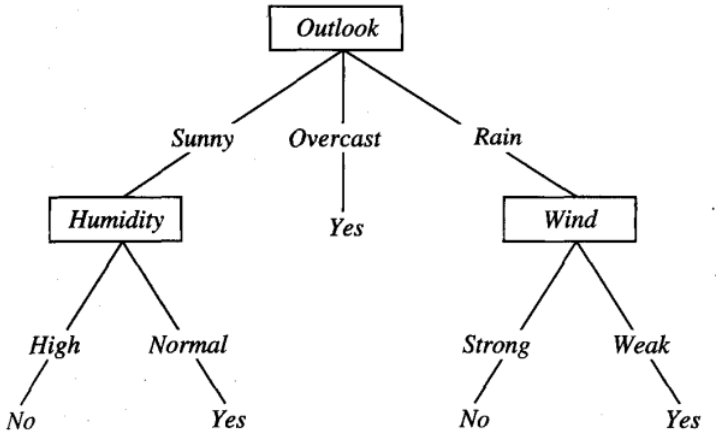

Figure 3.1: A simple decision tree deciding whether one should play tennis based on the current weather. An example is classifed by traversing its attributes through the tree structure until a leaf node is reached. The result of the classification is the value in the final leaf node. Figure from Mitchell 1997.

important to note that the same attribute can be tested multiple times. The partitioning of the parameter space can be very fine, and thus follows non-linear boundaries between the event classes; the interleaving of attributes at different nodes also models the interactions between them.

The process of constructing a decision tree, called *training*, then follows a very simple top-level algorithm. Given a *training* data set of labeled observations with labels S and B and known values of m attributes $A_1, A_2, \ldots, A_m$:

- divide the parameter space of attributes into N distinct non-overlapping hyperrectangular regions $R_1, R_2, \ldots, R_N$;
- for every region R_i, label it with the majority label of the training observations that

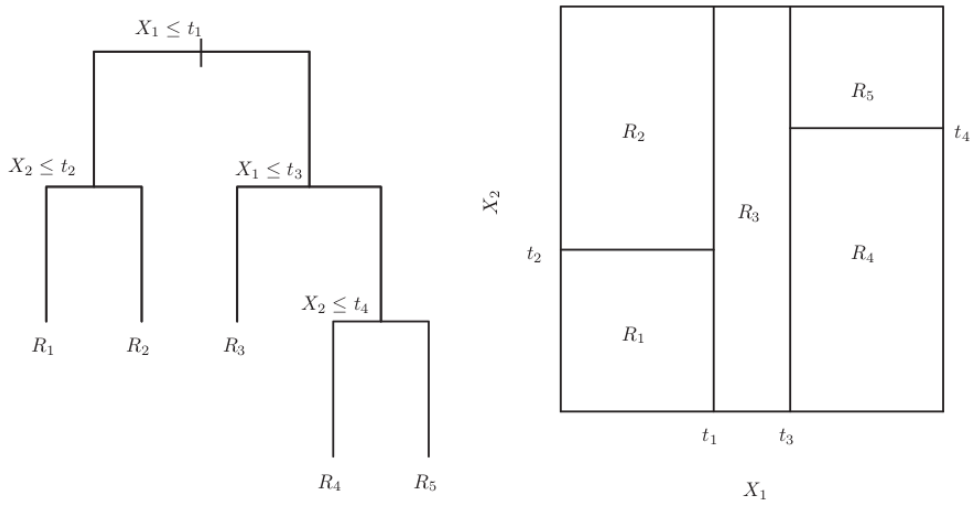

Figure 3.2: *Left:* a simple decision tree with two attributes X_1 and X_2 and 5 labels. Each node applies a splitting criterion to one of the attributes. *Right:* the partitioning of the parameter space of attributes corresponding to the tree on the left. Each rectangular region has a distinct label, the regions are non-overlapping, and the regions cover the whole parameter space. Figure from James et al. 2014.

reside within it.

Having done this, to classify a new observation, we simply need to find the region R_j into which it falls, and assign to it the label of that region. Three questions in this training process remain unanswered: how to choose the attribute to test at each node, how to choose the splitting criterion for each attribute, and when to stop adding further nodes.

Suppose that in the process of training we have reached a node d with k training observations, and these observations have *weights* assigned to them, $w_1, w_2, \ldots, w_k$. The full utility of the weights will become apparent in a moment, but for now, we can think of the weights as reflecting our confidence that the labels assigned to the observations are correct – when we start out building the tree, we have no reason to think that the assigned labels

could be wrong, and so all the weights are equal to 1. We further denote the fraction of the observations that have the label S as p (and the fraction with the label B is then $1-p$). We then define the *Gini index* of this node as

$$\mathrm{Gini}_d = p(1-p)\sum_{i=0}^{k} w_i. \tag{3.1}$$

The quantity $p(1-p)$ by itself is a measure of *purity* of the node – it is equal to 0 when all observations in the node have the same label, and reaches its maximum value of 0.25 when there are equal numbers of events with each label. Intuitively then, the Gini index is the amount of purity of the whole tree contributed by this node. We choose the attribute A_0 to test at this node, and its splitting criterion S_0, in such a way that the resulting partition maximizes the purity gain:

$$\mathrm{Gain}(d; A, S) = \mathrm{Gini}_d - \mathrm{Gini}_{d\to A\leq S} - \mathrm{Gini}_{d\to A>S}, A_0, S_0 \quad = \arg\max_{A,S} \left\{\mathrm{Gain}(d; A, S)\right\}. \tag{3.2}$$

(3.3)

Here, the first equation defines the purity gain as the difference between the Gini index of the parent node and the sum of the gini indices of its daughters, with the subscript $d \to A \leq S$ denoting the daugher for which the value of attribute A is less than or equal to S, and $d \to A > S$ denoting the one whose attribute A is greater than S. The second equation

defines the optimal splitting to be such an attribute A_0 and its value S_0 that maximize the purity gain functional. If the node is completely pure, $\text{Gini}_d = 0$, no such maximization is possible, and we terminate the tree here, making this node a leaf node. In the training process then, every possible splitting criterion of every attribute is checked at each node, making it possible for multiple nodes to test the same attributes, and even with the same splitting criteria! In practice, for attributes with continuous values, the optimal values of splitting criteria are searched only on discrete grids.

It should be clear that this process will result in a tree that is completely pure and will correctly classify all the training observations. This is not a good thing in practice, as irreducible error in training data will force a decision tree to *overtrain* and have a large model variance – small changes in the training data will result in large changes of the resulting partitioning. We can assess the classification performance of this tree on a *test* data set, but it should be clear that this fully grown tree will result in high misclassification error rates on test data, since the created partitions will be highly specific to the training sample. One possible solution to this problem is *pruning* – discarding complete tree branches, or equivalently, merging adjacent partitions. Taking this process to its absurd conclusion, we can discard all tree branches, keeping only the root node. This is the opposite of a fully grown tree, a *stump*, and while it will have a very low variance, it will exhibit a high *bias*, simply classifying all future data with the majority label of the training sample.

Optimizing this bias-variance tradeoff between stumps and fully grown trees is an important topic in machine learning, and one popular approach to it is cross-validation. In k-fold

cross-validation, the training data set is split into k buckets, and the training is performed k times, each time choosing a different bucket as the test subsample and training on the remaining $k-1$ buckets. The misclassification error is then estimated as the average of the k test errors, and the optimal pruning depth (or some other parameter controlling the bias-variance tradeoff) is found by minimizing this quantity. A detailed discussion of cross-validation can be found in, for instance, Hastie, Tibshirani, and Friedman 2009, Chapter 7.

3.1.2 Boosting and Bagging

Simple decision trees have a low bias and a high variance, and increasing their accuracy entails minimizing their variance by pruning fully grown trees. But we could also do it the other way round – starting from stumps with a high bias and a low variance, we could try to increase the overall performance by decreasing bias. Very simple classifiers like stumps or shallow trees are considered *weak learners*, in that their error rates are only slightly lower than those of random guessing. The idea of boosting came from a question by Kearns and Valiant (Kearns and Valiant 1989), who pondered whether an ensemble of weak learners could form a single strong learner. The affirmative answer to this question (Schapire 1990) laid the groundwork for the development of highly performant boosting algorithms, in particular *AdaBoost* (Freund and Schapire 1997), utilized in this work, which earned its inventors the 2003 Gödel Prize.

In boosting, an ensemble E of shallow trees is grown sequentially, with each new tree

trained on the *residuals* of the model formed by summing the existing trees. That is, each consecutive tree focuses on partitioning the region of the parameter space where the model formed by combining the preceding trees had a large classification error. This way, the overall bias of the combined model is iteratively reduced, and the process terminates when the increase in variance becomes comparable to the reduction in bias. However, when classifying a new observation x with this model, the output is not a disrete label, but a *score*, giving the "average label" of the ensemble of trees:

$$\widetilde{C}(x) = \sum_{t \in E} C_t(x), \tag{3.4}$$

where $C_t(x)$ is the label assigned to the observation by tree t. In our case, this BDT score is a measure of how signal-like an event is and it can be used to define a simple cut that maximizes the separation between the two event classes. We describe this cut optimization procedure in Section 3.3.1.

The notion of residuals isn't obvious in the context of classification. This is where the idea of event weights comes into play – each training observation x_i is assigned a weight w_i that reflects how often the event was *misclassified* by the trees in the current ensemble. The larger the weight, the more difficulty the current model has correctly classifying this event. These events then contribute more to node purity and its gain in equations 3.1 and 3.3, making sure that the test attributes and splitting criteria of the subsequent nodes – and trees – are chosen to maximize the separation between these misclassified observations.

Let us make this idea concrete. For an observation x_i, we will denote the *ground truth* $T(x_i) = 1$ when x_i belongs to class S and $T(x_i) = -1$ when $x_i \in B$. We will similarly denote the classification that the tree t assigns to this event as $C_t(x_i) = 1$ if the tree assigns x_i to class S and $C_t(x_i) = -1$ if the tree assigns it to B. For a perfect tree, $C_t(x_i) = T(x_i), \forall i \in 1 \dots N$, and an observation is misclassified when $C_t(x_i) \neq T(x_i)$. At every step of the tree-building process, we endow each of the N observations with a weight, $w_i^{(t)}$, and we set all the weights equal to $1/N$ before constructing the first tree: $w_i^{(0)} = 1/N, \forall i \in 1 \dots N$. After the tree t is constructed, the weights of all events misclassified by *that tree* are multiplied by the boost factor $\exp \alpha_t$, and all weights are re-normalized such that their sum is equal to 1:

$$w_i^{(t+1)} = \begin{cases} \frac{w_i^{(t)} \exp \alpha_t}{\sum_i^N w_i^{(t+1)}}, & C_t(x_i) \neq T(x_i) \\ \frac{w_i^{(t)}}{\sum_i^N w_i^{(t+1)}}, & C_t(x_i) = T(x_i) \end{cases} \tag{3.5}$$

In *AdaBoost* (Freund and Schapire 1997), the boost factor α_t is the logodds of the error rate of the tree t:

$$\alpha_t = \beta \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right), \tag{3.6}$$

where β is the user-defined *learning rate*, and ϵ_t is the weighted fraction of the misclassified events:

$$\epsilon_t = \frac{\sum_{C_t(x_i) \neq T(x_i)} w_i^{(t)}}{\sum_i^N w_i^{(t)}}. \tag{3.7}$$

Boosting is a remarkably powerful tool and BDT classifiers have a low bias while remaining robust against overtraining. However, while simple trees are easily interpretable, boosting obscures the interpretability of attribute importance and makes model interpretation difficult. The measure of attribute importance used in this work is the average purity gain increase due to splits at this attribute. While this measure is *accurate* (the importance of all features sums up to 1), it is *inconsistent* – the importance of an attribute *may decrease* if a model is changed in a way that relies *more* on this attribute. Therefore, model interpretation has to be done with caution, or consistent measures of importance, such as tree SHAP (Lundberg, Erion, and Lee 2018) need to be used in future studies.

In practice, to further avoid overtraining with large ensembles of trees, we can combine this method with another powerful tool from machine learning called *bootstrap aggregating*, or *bagging*. In bagging, each tree is trained only on a subset of all events, and this subset is sampled from the original training dataset *with replacement*. Sampling with replacement is crucial, as it guarantees that subsamples are representative of the whole population. Bagging not only reduces overtraining, but improves the overall stability of the model and decreases its bias. For details on bagging and why it works, refer to the seminal paper by Leo Breiman (Breiman 1996).

Overall, the BDT algorithm introduces interactions between predictors and partitions the predictor parameter space into many discrete regions, and should therefore perform much better for event classification purposes than simple linear box cuts. The result of classification with BDTs is a score ranging from -1 to 1, which facilitates an efficient and

robust cut optimization procedure (see Section 3.3.1).

3.2 Image Template Method

While BDTs should certainly improve event selection, a major roadblock to improving the performance of VERITAS is the earlier event reconstruction stage. As already mentioned at the beginning of this chapter, event reconstruction can be improved by making better use of camera pixel data and performing a simultaneous fit of all shower parameters by finding the image in a library of simulated templates that best agrees with the current observation. This image template method (ITM) was first introduced for the CAT telescope (Le Bohec et al. 1998), and later developed for the EventDisplay package within VERITAS by Stephane Vincent (Vincent 2016), whose realization of the method was then reimplemented in VEGAS by Jodi Christiansen (Christiansen and VERITAS Collaboration 2017; Christiansen 2019). In this work, we contribute to the implementation and validation of the method in VEGAS, but because the bulk of the development was done in the mentioned references, we give only a rather brief overview of its principles.

The image of a shower in an IACT camera plane, or in other words, the distribution of camera pixel intensities, is a function of the shower direction (azimuth and elevation), impact distance (and more rigorously, its core location on the ground, X and Y), energy of the primary, and the height, or depth, of the first interaction of the primary in the atmosphere. If we could simulate showers with all possible sets of these parameters, we could

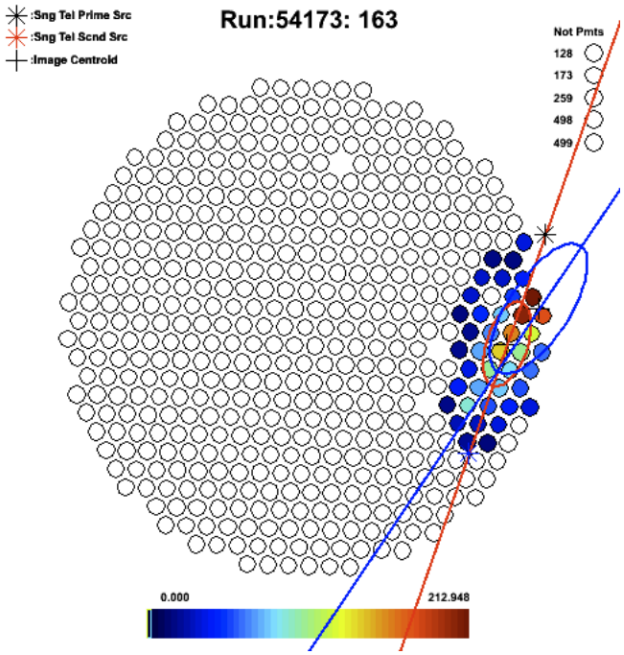

Figure 3.3: An example of a clipped shower image in the VERITAS camera. The standard geometric reconstruction is shown in red, and the 1σ contour of a more complicated 2D Gaussian fit is shown in blue. The Gaussian fit gives a better estimate of the image centroid, its major axis, and its total size than the geometric fit. A likelihood fit of this image to a template library does not need to make any distributional assumptions, while not taking any longer computing time. Figure from Jon Dumm.

then propagate the showers through a telescope simulation and obtain a set of all possible images of gamma-intiated showers, from now on referred to as templates. A likelihood fit of an observation to a library of these templates can then reconstruct the most likely shower parameters, and has several obvious advantages over geometric reconstruction methods. Most importantly, this approach to parameter estimation is *robust* and *resistant* to missing information, a situation that happens when a shower image is clipped from being close to the camera edge (Figure 3.3) or when the field of view contains a bright star. Such a fit,

quantified in a χ^2-like goodness-of-fit value, also provides another measure of how γ-like the primary is, as hadron-initiated showers are expected to result in poor fits to simulated images. This goodness-of-fit value can be used in training a BDT model for event selection.

The development of the distributions of charged particles in atmospheric showers is well-studied (Hillas 1982; de Naurois and Rolland 2009), with known semi-analytical models for the angular, lateral and energy-dependent longitudinal distributions of charged particles in atmospheric showers derived from Monte-Carlo simulations. For completeness, we note the following interesting features of these distributions:

- The depth of the first interaction is the dominant source of shower-to-shower fluctuations. This is the reason for its inclusion as a parameter in template generation and the event reconstruction fit.
- The angular distribution around the the mean angle $\langle w \rangle$ is virtually independent of charged particle energy and shower age.
- The Heitler model for the number of particles in the shower (Section 2.1) is applicable only at very small shower depths. The slightly more complicated Greisen formula (Wilson and Greisen 1956; Schiel and Ralston 2007) that takes into account particle extinction remains a very good approximation.

With known distributions of charged particles, the intensity of Cherenkov light from a shower in the telescope camera can be calculated by an eight-fold integral over the following quantities (Le Bohec et al. 1998; de Naurois and Rolland 2009):

- Cherenkov photon wavelength λ;
- Cherenkov photon azimuthal angle $\phi_{\rm ph}$ around the direction of the electron in the shower that emits it;
- electron position w.r.t. its direction (x and y, the lateral distribution);
- direction of the Cherenkov-emitting electron w.r.t. the telescope (w and ϕ, the angular distribution of charged particles in the shower);
- electron energy E;
- shower depth ξ (the longitudinal distribution).

Instrumental effects, such as the light-collection efficiency, the quantum efficiency of the PMTs and the electronic response of the camera, as well as atmospheric effects, have to be taken into account in this integral and are derived from detector simulations. Calculating this integral is computationally cheap and results in well-behaved image templates that, stemming from a smooth model, circumvent the shower-to-shower fluctuations present in full simulations. Such a semi-analytical approach to template generation was undertaken in the first implementation of ITM (Le Bohec et al. 1998), as well as in its later refinement by the H.E.S.S. collaboration (de Naurois and Rolland 2009).

The semi-analytical model was used in these early implementations because templates derived from full simulations need to be constructed by averaging a large number of simulated images, in order to smooth over shower-to-shower fluctuations. This, combined with the large size and dimensionality of the parameter space, made the *brute force* approach prohibitive

computationally. However, the increase in computing power over the past decade relaxed this restriction and allowed to forgo the semi-analytic model, resulting in more accurate templates for primaries with energies above 10 TeV and showers with impact distances greater than 300 m (Parsons and Hinton 2014). VERITAS uses this more computationally expensive but also more accurate procedure (Vincent 2016).

VERITAS uses the program CORSIKA (Heck et al. 1998) to simulate showers, which outputs the light distribution on the ground and the arrival direction of each photon. For each shower, the photons falling onto the telescope mirror are followed individually to the camera, including the atmospheric density profile and optical absorption, and some instrumental effects, such as shadowing and the quantum efficiency of the PMTs. The images are generated on-axis, that is, for a source at the center of the field of view, with the longitudinal development of the shower directed along the x-axis of the camera frame. For a perfect telescope, a change in the offset and direction result in a rotation and a translation of the image in the camera frame, transformations that are applied to the *real image* in the fit procedure. For a more realistic telescope, the broadening of the optical point spread function away from the camera center needs to be taken into account, but this was not done for the current templates. Throughout this work, the telescope and detector model as implemented in GrISUDet[1] was used.

Templates were generated for 9 first interaction depths, from $0\chi_0$ to $5\chi_0$ ($\chi_0 = 36.7$ g cm^{-2} being the radiation length in air), 90 energies, from 30 GeV to 30 TeV, and

[1] `http://www.physics.utah.edu/gammaray/GrISU/`

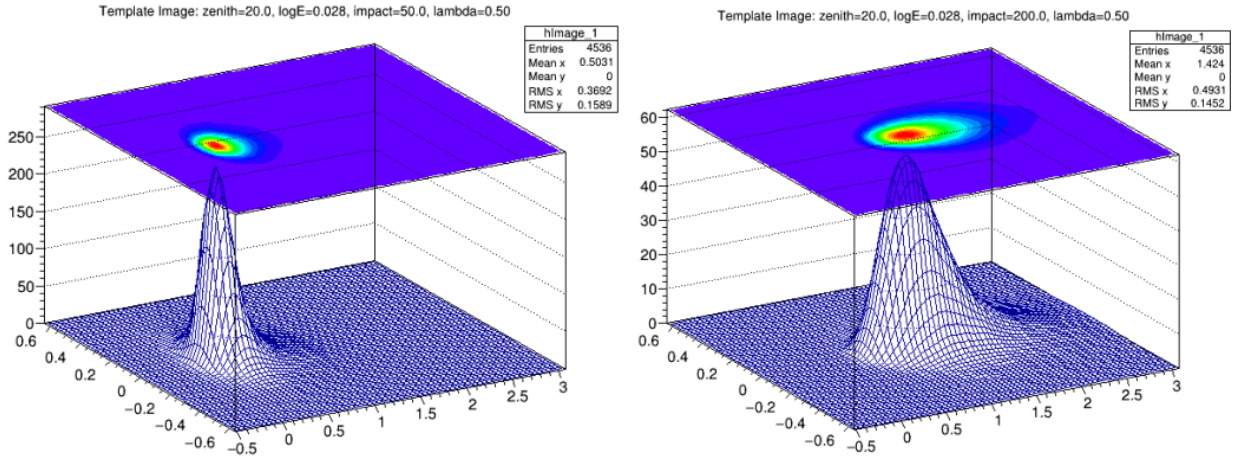

Figure 3.4: Two generated image templates for a primary with energy of ~ 1 TeV, zenith angle of 20° and first interaction depth of $\chi_0/2$. *Left:* the simulated shower has an impact distance of 50 m. *Right:* the simulated shower has an impact distance of 200 m. The larger impact distance results in a broader distribution of lower intensity. Figure from Jodi Christiansen (Christiansen 2019).

50 impact distances, from 0 to 500 m. Interpolation between these templates allows production of an image template for any shower parameters within the simulated ranges. The Earth's magnetic field at the location of VERITAS, which affects the lateral spreads of showers, was not included in the simulations. Figure 3.4 shows two image templates of primaries with energy of about 1 TeV, 20° zenith angle and first interaction depth of $\chi_0/2$, but with different impact distances of 50 m and 200 m respectively. In the case of VEGAS, only NA ATM21 templates were produced, and the other epoch-atmosphere combinations derived from them through nonlinear calibration factors.

In order to find the template image that best fits the current camera image, we define and minimize the following χ^2-like quantity that depends on shower parameters

$\Theta = (\theta_{\rm zen}, \phi_{\rm az}, X_{\rm core}, Y_{\rm core}, E, \xi)$:

$$\tilde{\chi}^2(\Theta) = \frac{1}{n_{\rm pix}} \sum_{i=0}^{n_{\rm pix}} \frac{\left(N_i^{\rm (cur)} - N_i^{\rm (tmplt)}(\Theta)\right)^2}{\sigma_{{\rm ped},i}^2 + (1+\sigma_\gamma^2) N_i^{\rm (cur)}}. \tag{3.8}$$

Here, the sum runs over all the $n_{\rm pix}$ pixels in the camera, and its numerator is the squared difference between the actual, $N_i^{\rm (cur)}$, and the simulated , $N_i^{\rm (tmplt)}(\Theta)$, photon counts in pixel i. The denominator is the expected variance of this difference under the null hypothesis of the actual camera image being formed by a gamma-initiated shower. There, $\sigma_{{\rm ped},i}^2$ is the variance of the *pedestal* of pixel i, its charge distribution in the abscence of a Cherenkov signal. This value, which accounts for the night-sky background fluctuations, is recorded by VERITAS during data runs with dedicated 1 Hz artificial triggers and is calculated for each pixel in the calibration stages of the analysis. The second term in the denominator is just the variance of the recorded Poisson counts in the pixel, with the σ_γ term modeling the width of the single photoelectron peak (the PMT resolution), estimated as 0.4 for the VERITAS PMTs. VEGAS uses MINUIT[2] (Hatlo et al. 2005) to find the shower parameters Θ that minimize the expression (3.8). Parameters obtained from standard geometric reconstruction (Section 2.4.1) are used as the initial guess in the fitting procedure.

While standard geometric shower reconstruction requires the image to be cleaned of noise, leaving only the Cherenkov pool ellipse, ITM fitting does not need any such cleaning. However, in practice, uncleaned images result in much larger file sizes and significantly slower

[2] `http://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/`

fits. For practical purposes, therefore, VEGAS uses cleaned images with *one extra ring* – standard cleaning is applied to extract the shower image, and then one ring of pixels around the ellipse is restored.

The minimized $\tilde{\chi}^2$, termed the ITM goodness-of-fit, is expected to follow a narrow distribution with values close to 1 for gamma-initiated showers. The hadronic showers, on the other hand, are expected to have a much wider spread with a larger mean and a far broader tail on the right side. Moreover, the values of shower parameters obtained from the $\tilde{\chi}^2$ fit should be in a good agreement with their geometrically reconstructed counterparts for gamma-initiated showers, while hadronic showers should exhibit large differences between the two reconstruction techniques. This predicted behavior can be exploited by using $\tilde{\chi}^2$ and the differences between ITM and geometric parameter estimates as training attributes in a BDT model, described next.

3.3 The Improved VERITAS Analysis

Now that we are familiar with the theory behind boosted decision trees and the image template method, we move on to the implemention and performance of a combined ITM+BDT analysis method that leverages both of these approaches.

3.3.1 Implementation

We use the ITM analysis described above for angular and energy reconstruction, and boosted decision trees are then used in the event selection stage. We relied on the Toolkit for Multivariate Analysis[3] (TMVA, Hoecker et al. 2007) and its implementation of BDTs in the training process. The following 12 attributes were used for BDT training:

- *MSL* and *MSW*, the mean squared length and width. These are the usual Hillas parameters that have proven to have a high separating power in standard box cuts analysis, see Section 2.4.
- MSL_{RMS} and MSW_{RMS}, the between-telescopes root-mean-squared values of the above. These incorporate array-level stereo information and are expected to be distributed narrowly for γ-like events, with wider distributions for hadronic showers.
- H_{max}, the height of the shower maximum. For a fixed primary energy E, H_{max} is once again expected to have a narrow distribution for signal events, with a wider distribution for the background.
- E_{RMS}/E, the ratio of the between-telescopes RMS of the *geometrically* reconstructed energy to the reconstructed energy. This ratio once again encodes how well a given event was reconstructed, and is therefore expected to have a wider distribution for background showers.
- θ_{RMS}, the between-telescopes variance of the *geometrically* reconstructed direction.

[3] `https://root.cern.ch/tmva`

Just as the above, this is expected to have a wider distribution for background showers.

- D_{core}, the distance of the shower core to the center of the array, and $\langle S_{\text{max2}} \rangle$, the average charge in the second-brightest pixel. These attributes do not have any discriminating power on their own, and the idea behind them is to track shower energy, which to first order depends on impact distance and image size. Tracking some measure of shower energy in this way without reliance on the reconstruction procedure enhances the discriminating power of the other attributes. For example, tree splits on H_{max} that happen below splits on these two attributes will be equivalent to tree splits on H_{max} for showers of similar energy. At first glance, $\langle S_{\text{max2}} \rangle$ is a rather awkward parameter – indeed, it would appear the total image size $\langle S \rangle$ is both a more intuitive and a more representative measure of shower energy. The reasoning behind using $\langle S_{\text{max2}} \rangle$ is that it's an *order statistic*, and is therefore more robust and resistant to shower fluctuations downward than the mean. Using the second-brightest pixel avoids possible issues with saturation of the brightest pixel, while being less computationally expensive than tracking the median.
- $\tilde{\chi}^2$, the ITM goodness-of-fit. As explained at the end of section 3.2, this is expected to have a narrow distribution with mean close to 1 for γ-like events, and a broader distribution with a larger mean for hadronic showers.
- ΔD_{core}, the difference in the reconstructed core location between the geometric and ITM reconstruction methods. We expect reconstruction results to agree well between the two methods for gamma events, resulting in narrow distributions close to zero.

Hadronic events are expected to have worse disagreement between the two techniques.

- $\Delta\theta$, the difference in the reconstructed primary direction between the geometric and ITM reconstruction methods. The behavior and effect are similar to the above.

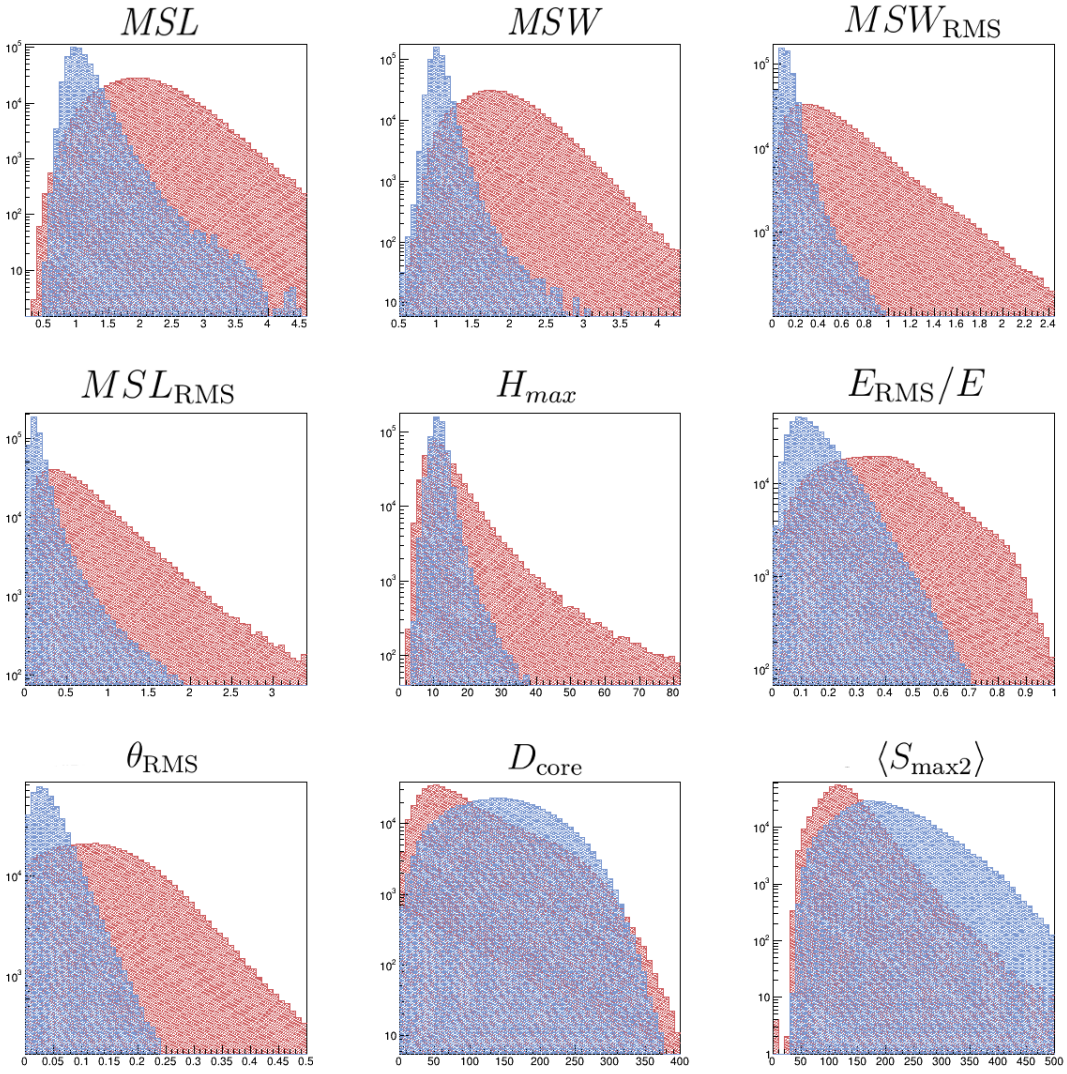

Figure 3.5: Distributions of attributes used in BDT training, for a test dataset with primaries of energies 282 GeV $\leq E \leq$ 501 GeV and elevations in the range $60° \leq El \leq 70°$. The training attributes are as described in the main text – these are the first 9 attributes that do not rely on ITM. Background events are in red, signal events in blue.

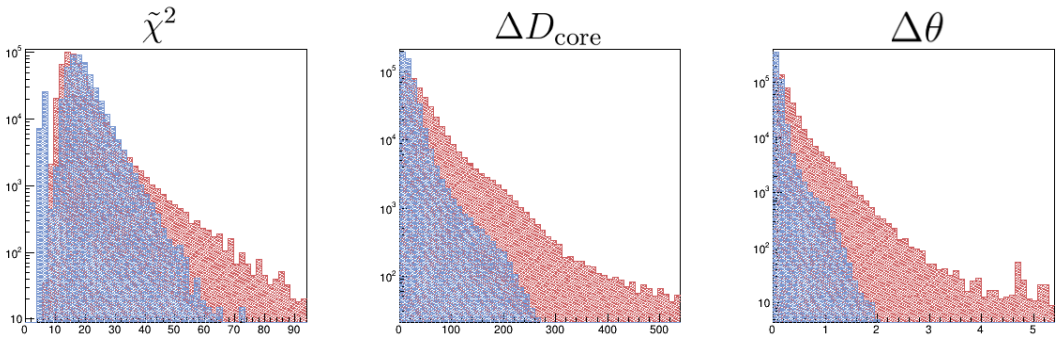

Figure 3.6: Distributions of attributes used in BDT training, for a test dataset with primaries of energies 282 GeV $\leq E \leq$ 501 GeV and elevations in the range $60° \leq El \leq 70°$. The training attributes are as described in the main text – these are the 3 attributes derived from ITM. Background events are in red, signal events in blue.

Figures 3.5 and 3.6 show the distributions of the training attributes in the *test* dataset. The distributions exhibit the expected behavior described above.

The BDT model was trained with the following parameters:

- $N_{\text{trees}} = 400$ – the number of trees grown.
- $\beta = 0.5$ – the learning rate in Equation 3.6.
- BaggedFraction = 0.6 – the BDT model incorporated bagging, with each tree being trained on BaggedFraction of total events, sampled from all events with replacement.
- $Depth = 50$, $PruneStrength = -1$ – each tree was fully grown and pruned. While general practice suggests that growing shallow trees works best with BDTs, we found no performance difference between shallow trees and fully grown trees with aggressive pruning. This makes sense intuitively – aggressive pruning discards most nodes, leaving behind a shallow tree. Our reason for choosing fully grown trees with pruning is given

below.

VERITAS data includes 3 array epochs (Section 2.3), as well as two atmospheric profiles, winter and summer atmospheres, denoted ATM21 and ATM22 respectively. A separate model needed to be trained for each epoch-atmosphere configuration, resulting in 6 individual models.

Summer atmosphere spans the months of May-October, and is interrupted by a 3-month-long period of downtime due to the summer monsoon in Southern Arizona. Because of a short period of engineering runs immediately following the summer shutdown, VERITAS normally collects less than 3 months of ATM22 data each year. Because the "Old Array" and "New Array" epochs lasted only 2 and 3 years respectively, very little OA ATM22 and NA ATM22 data are available.

For the signal training dataset, we employed the same CORSIKA shower simulations as the ones used for template generation. For the background training dataset, real VERITAS data were used – a mix of dark matter targets and point sources, with known sources of γ-ray emission masked out.

In order to ensure we were training on the relevant data, all training data were chosen to pass *preselection* cuts, a set of quality and sanity requirements. We require that each training event is reconstructed by at least 3 telescopes; that signal events are within $\theta = 0.223^\circ$ of the simulated source; and that background events are outside of $\theta = 0.25^\circ$ of tentative point sources. No minimum size preselection criterion was imposed as the requirement of a

minimum of 3 reconstructed images was stringent enough on its own. No extended sources were included in the background training dataset. Moreover, only known and published point sources were included in the training dataset.

Our knowledge of shower development suggests that the training attributes vary the most with the zenith angle and energy of the primary. In order to improve training performance, the training dataset was binned by energy and zenith, with training performed separately in each zenith-energy bin. The binning was chosen for each epoch, with the goal of being as fine as possible while maintaining 500,000 training events in each bin. This last condition could not be fulfilled for OA and NA ATM22 training data. Furthermore, data in each bin was split into training and test subsets, with the test subset used to test the performance. When available, the number of test events was set to 500,000; when there were not enough training events, the full dataset was split in a 3:1 ratio, with 3 parts used for training and 1 part used for testing.

The zenith-energy binning used in training is shown in Tables 3.1-3.3. The increasing energy binwidth with rising zenith reflects the incerasing thickness of the atmosphere and the difficulty in reconstructing high-zenith events, effects that increase the energy threshold and decrease the overall event rates. The increasing sensitivity of VERITAS with each consecutive upgrade is also reflected in finer energy binning.

The output of the trained model is a set of *weights* files that can be used to map the attributes of a test event to a BDT score, with a score of $+1$ signifying signal and a score of -1 – background. The training output also includes *the signal and background efficiencies* as

Zenith, deg	[0, 20]	[20, 30]	[30, 40]	[40, 55]
Energy, GeV	[90, 178] [178, 316] [316, 1000] [1000, 80000]	[90, 224] [224, 398] [398, 1000] [1000, 80000]	[90, 316] [316, 1000] [1000, 8000]	[90, 501] [501, 1000] [1000, 80000]

Table 3.1: Zenith-energy binning of training data for the OA epoch.

Zenith, deg	[0, 20]	[20, 30]	[30, 40]	[40, 55]
Energy, GeV	[80, 158] [158, 282] [282, 501] [501, 1000] [1000, 80000]	[80, 158] 158, 282] [282, 501] [501, 1000] [1000, 80000]	[80, 316] [316, 501] [510, 1000] [1000, 80000]	[80, 501] [501, 1000] [1000, 80000]

Table 3.2: Zenith-energy binning of training data for the NA epoch.

Zenith, deg	[0, 20]	[20, 30]	[30, 40]	[40, 55]
Energy, GeV	[63, 141] [141, 282] [282, 501] [501, 1000] [1000, 80000]	[63, 141] [141, 282] [282, 501] [501, 1000] [1000, 80000]	[63, 224] [224, 398] [398, 1000] [1000, 80000]	[63, 398] [398, 1000] [1000, 80000]

Table 3.3: Zenith-energy binning of training data for the UA epoch.

fuctions of the BDT score for both the training and test subsets. These efficiencies are simply the complements of the cumulative distribution functions (CDF) of the BDT score. The Kolmogorov-Smirnov test was used to compare the equality of these distributions between the test and training subsets, with a lack of equality indicating overtraining. There was no evidence of overtraining in the final models.

The set of training attributes was not arrived at immediately. The model selection process involved testing a large number of models with attributes chosen in a forward selection procedure – a minimal set of attributes was expanded by adding one or two extra features;

the performance of the two models was compared and the additional attributes kept if the performance increase was deemed substantial. The figure of merit used was the receiver operating characteristic (ROC) curve, displaying the background rejection as a function of signal efficiency. An example of such a comparison is shown in figure 3.7. There, the red curve corresponds to a model with the first 9 training attributes, excluding the ITM-derived ones (see, for instance, figure 3.6), while the black curve includes all 12 training attributes and clearly results in a higher separation efficiency.

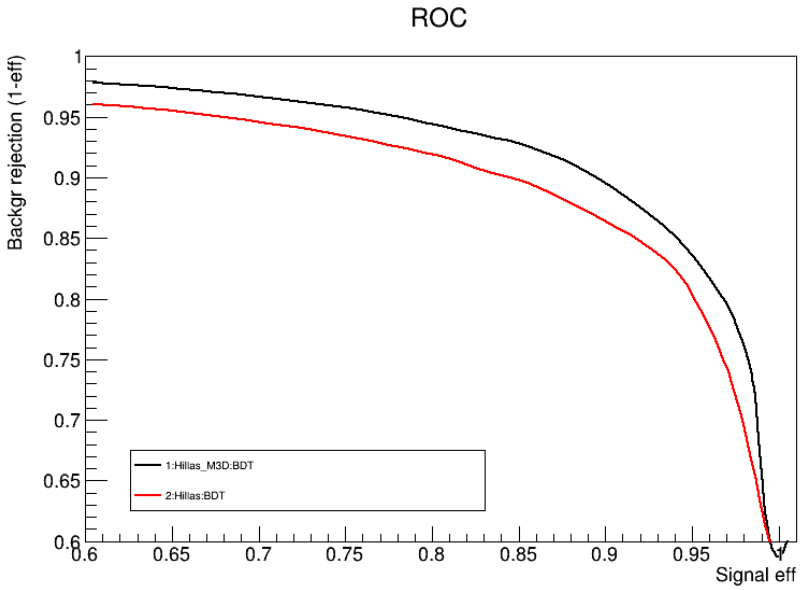

Figure 3.7: An example of ROC curve comparison used in BDT model evaluation and selection. Here, the black curve is the final model; the red curve is the same model excluding the last three ITM-derived attributes shown in figure 3.6. The curves shown are for the test dataset of OA ATM21, $20° \leq$ Zenith $\leq 30°$, 90GeV $\leq E \leq$ 224GeV.

The model parameters were chosen by testing the performance of a set of different parameters on an initial minimal model, and kept frozen for the rest of the model-building

process. To choose the number of trees, the cross-validation misclassification rate of the model was monitored, and the tree-building process was terminated shortly after the error rate stopped decreasing. The learning rate and bagging fraction did not make a noticeable difference in the tested ranges, from 0.2 to 0.8 for the learning rate and from 0.4 to 0.8 for the bagging fraction; therefore, the midpoints of these intervals were chosen for these parameters.

Finally, fully grown trees with aggressive pruning performed on par with shallow trees. However, the resulting distributions of BDT scores were markedly different between the two approaches. Signal and background efficiency curves for both cases are shown in Figure 3.8. Shallow trees resulted in narrower distributions of BDT scores, with no events reaching the endpoints of the $[-1, 1]$ interval. This is because no individual event was correctly classified by *all* shallow trees. Such narrow distributions of scores resulted in steep efficiency curves. As a consequence, the cut optimization procedure, described next, was much less stable for shallow trees. For this reason, deep trees with pruning were chosen in our model.

Cut optimization

In order to perform event classification, a threshold value of the BDT score needs to be chosen, such that all events with a score less than the threshold are deemed background, while all events with a score equal to or greater than the threshold are considered signal. Choosing the correct value of this cut on the BDT score is a crucial part of the model-building process.

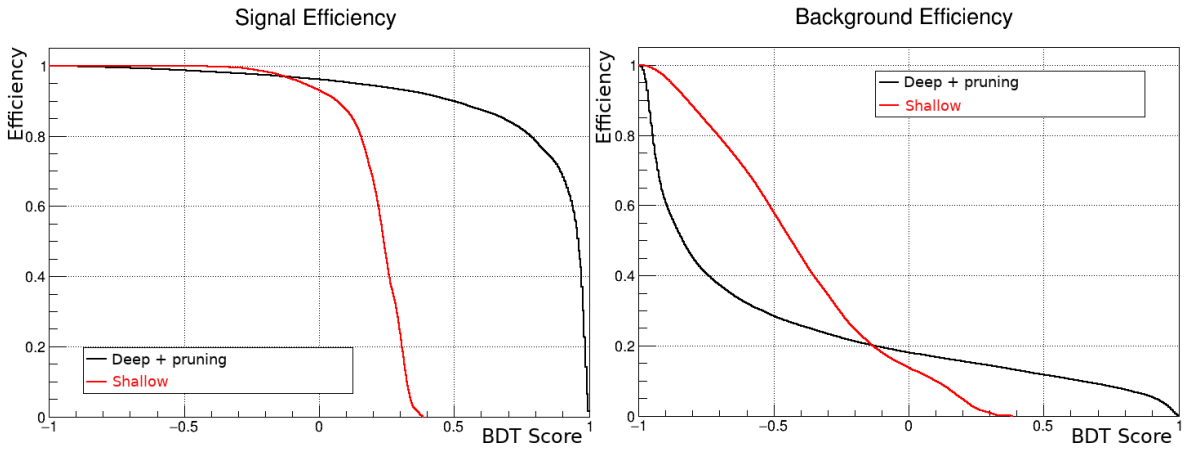

Figure 3.8: Signal and background efficiency curves for deep and shallow trees. Deep trees with pruning are shown in black, shallow trees in red. Shallow trees result in much narrower distributions of BDT scores, with much steeper efficiency curves. While the resulting ROCs are comparable, the more gradual efficiencies of deep trees lend themselves to a more stable cut optimization procedure. The curves shown are for the test dataset of UA ATM21, $20° \leq \text{Zenith} \leq 30°$, $63\text{GeV} \leq E \leq 141\text{GeV}$.

In order to understand the cut optimization procedure, let's take a step back and revisit what is meant by *sensitivity* of our experiment. Given n measured events from an "On" region, and b events from an "Off" region, we try to answer, with a confidence level $1 - \alpha$, whether n is greater than b. Specifically, we calculate the probability of obtaining a number of events equal to or greater than n by pure chance, given an expected background rate b. If this probability is less than the chosen value α, we claim that we have detected a signal at a confindence level $1 - \alpha$ (or a significance α); if, on the other hand, we find that this probability is less than α, we say that the null hypothesis H_0 that n arose from the background expectation b by pure chance cannot be discarded at a significance α.

For events that come from a Poissonian process, as is our case, in order to be able to

discard the null hypothesis H_0, n has to be equal to or greater than some critical value $x_c(b, \alpha)$, defined through:

$$P\left(t \geq x_c(b,\alpha) \mid b\right) \equiv \sum_{t=x_c}^{\infty} P(t \mid b) \leq \alpha. \tag{3.9}$$

Here, $P(t \mid b)$ is the probability of obtaining t events from a Poisson distribution with rate b, and the inequality follows the definition of a significant detection given above.

Similarly, given a signal of strength s, we consider the probability $F_\beta \equiv 1-\beta$ of detecting this signal, which is the probability of rejecting the null hypothesis at a confidence level $1-\alpha$. This probability is the *power* of our hypothesis test. Simply speaking, given a background rate b, in order to detect a signal at a confidence level $1-\alpha$, we need to measure $n(b+s) \geq x_c(b,\alpha)$ with probability at least F_β. Concretely we detect a signal s with power F_β when

$$P\left(n \geq x_c(b,\alpha) \mid b+s\right) \geq F_\beta = 1-\beta. \tag{3.10}$$

The *sensitivity* is then the smallest signal $s = s_{\text{th}}$ such that the equation 3.10 holds. While n, the number of detected events, is featured explicitly there, the sensitivity s_{th} does not depend on it, it's purely a function of the background rate b, the confidence level $1-\alpha$ and the desired power F_β. Figure 3.9 makes this more intuitive (Lundberg et al. 2010): given a background rate b, we find the critical value x_c for which the null hypothesis is rejected at

a confidence level $1-\alpha$; the sensitivity is then the signal $s_{\rm th}$ on top of the background that gives $n \geq x_c$ counts with probability F_β.

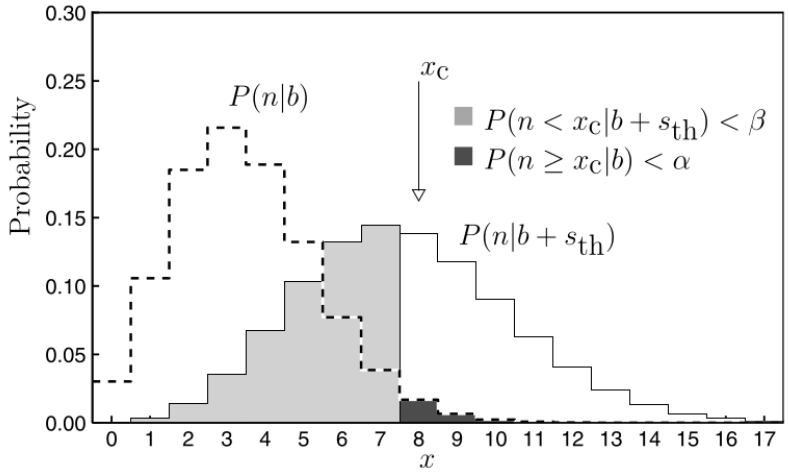

Figure 3.9: An intuitive illustration of the definition of sensitvity. Given a background rate b and a predefined power F_β, the sensitivity $s_{\rm th}$ is the signal that is discovered with probability F_β at a confidence level $1-\alpha$. Figure from Lundberg et al. 2010.

The cut optimization procedure, then, aims at finding the threshold value of BDT score that minimizes the sensitivity $s_{\rm th}$: for a given background rate b, find the minimum source strength such that after imposing BDT cuts, a detection at a confidence level $1-\alpha$ happens with probability F_β:

$$\widetilde{C}_{\rm opt} = \arg\min_{\widetilde{C}} \left\{ s_{\rm th}(\widetilde{C}) : x_c(b \cdot \epsilon_{\rm Bkg}(\widetilde{C}) + s_{\rm th}(\widetilde{C}) \cdot \epsilon_{\rm Sig}(\widetilde{C}), \beta) \geq x_c(b \cdot \epsilon_{\rm Bkg}(\widetilde{C}), \alpha) \right\}. \quad (3.11)$$

Here, $\epsilon_{\rm Bkg}(\widetilde{C})$ and $\epsilon_{\rm Sig}(\widetilde{C})$ are the background and signal efficiencies as functions of the BDT score $\widetilde{C}$, obtained during model training. This last expression in braces for the con-

dition of having a detection is a straightforward interpretation of equation 3.10 and figure 3.9. In the cut optimization procedure, the range $\widetilde{C} \in [-1, 1]$ is scanned, and the sensitivity $s_{\text{th}}(\widetilde{C})$ is found for this range; the optimal BDT score is chosen as the value that minimizes $s_{\text{th}}(\widetilde{C})$. The critical values x_c are computed with the *TRolke2* class (Lundberg et al. 2010) implemented in ROOT.[4]

It should be noted that while the sensitivity $s_{\text{th}}(\widetilde{C} \,|\, b, \alpha, \beta)$ depends on the background rate, as well as the chosen confidence level and power, the optimal cut is only a function of the signal and background efficiency curves in the absence of systematic uncertainties. We do not include any systematic uncertainties in the cut optimization procedure in this work. For this reason, the optimal cuts can be found at computationally convenient values of $b = 100, \alpha = 0.05$ and $\beta = 0.5$.

In practice, the optimal cuts defined in equation 3.11, dubbed *detection cuts*, do not have the best performance for known strong sources. This is because *minimizing the visibility threshold* is achieved through background suppression, which, in the case of strong sources, discards more signal events than is needed to *maximize the detection significance*. For this reason, a separate optimization procedure was used for known sources. In this procedure, the optimal cut is found by maximizing the LiMa signficance in equation 2.9:

$$\widetilde{C}_{\text{opt}} = \arg\max_{\widetilde{C}} \left\{ S\left(A(n-b) \cdot \epsilon_{\text{Sig}}(\widetilde{C}) + b \cdot \epsilon_{\text{Bkg}}(\widetilde{C});\, b\tau \cdot \epsilon_{\text{Bkg}}(\widetilde{C});\, 1/\tau \right) \right\}. \tag{3.12}$$

[4] `https://root.cern.ch/doc/master/classTRolke.html`

Here, to avoid confusion with the confidence level α, we replaced the ratio of the On region exposure to the Off region exposure with its reciprocal, $\tau = \epsilon_{\rm Off}/\epsilon_{\rm On}$ (see section 2.4.3); $b\tau$ is then the number of measured Off events; n is the number of measured On events. Thus, $b\tau \cdot \epsilon_{\rm Bkg}(\widetilde{C})$ is the post-cuts background rate; while $(n-b) \cdot \epsilon_{\rm Sig}(\widetilde{C})$ are the post-cuts excess events. The value and location of the maximum of expression 3.12 depend on all of the parameters, n, b and τ; to control the dependence on the number of On events, we introduce an extra normalization parameter A.

This *source cuts* optimization procedure relies on having the knowledge of On and Off event rates and the exposure ratio τ. The Crab nebula, a VHE standard candle, was used to estimate event rates – for this, a large number of Crab runs were processed without any event selection cuts; On and Off event rates were computed and scaled with the *observation time for optimization*, $T_{\rm obs} = 10$ hours; the normalization parameter A was then used as a measure of the test source strength, in Crab units (CU). The exposure ratio τ was found to have only a mild influence on the final optimized cuts, and was chosen as a reasonable typical value for analyses for sources at intermediate zenith angles, $\tau = 5$. The source cuts were optimized for two values of $A = 0.2$ and $A = 0.05$, as these are typical strengths of sources VERITAS considers *strong* and *weak*, respectively.

It should be noted that the optimization procedure defined by equation 3.11 is equivalent to that in equation 3.12, under a very special arrangement: for $F_\beta = 0.5$, fixing $S = S_\alpha$, the significance corresponding to the confidence level $1-\alpha$, we continuously increase the normalization A from zero, until the maximum of 3.12 reaches S_α. The value of A is then the

sensitivity, in Crab units, as defined by 3.10; the found value of the optimal cut is equivalent to the one found through 3.11.

The optimization procedure can be extended further to find the optimal directional cut θ^2. For this, the above procedure is extended by measuring the On and Off rates as functions of θ^2, with the optimization repeated over a sufficiently large range of the paramter, modifying τ accordingly. The optimal θ^2 found this way will correspond to the angular extent of a point source at a given zenith and energy. However, because the angular resolution is energy-dependent, and including this dependence in spectral reconstruction is far from trivial, a compromise was made when selecting the directional cut – the optimal θ^2 cut is found at energies $\sim 500\ GeV \leq E \leq 1\ TeV$, where the majority of well-reconstructed events seen by VERITAS reside.

	Hard	Medium	Soft
OA	1000	400	200
NA	1000	400	200
UA	1200	700	400

Table 3.4: Size cuts as defined in VEGAS. The values are in FADC counts.

One final note needs to be made about VERITAS event quality selection. As mentioned in section 2.4.1, VERITAS uses three levels of quality cuts on the size parameter, hard, medium and soft size cuts. Table 3.4 shows the values of these cuts for different instrumental epochs, as defined in VEGAS. The size cuts are imposed on each image before shower reconstruction – if the image does not pass the size cut, it is not included in the reconstruction stage. Size cuts aid reconstruction by discarding "small" events, which have either a lower energy or a

larger impact distance. For a given impact distance, lower energy events are discarded, with stricter size cuts effectively raising the energy threshold, and because the background rates fall steeper with energy than spectra of typical astrophysical sources, imposing medium or hard size cuts aids with background rejection for weaker sources with harder spectra.

We did not include any size cuts in the model training, as the benefits of having more training events were found to outweigh difficulties in reconstructing smaller events – in fact, imposing medium or hard size cuts on training data led to overtraining due to a lack of a sufficient number of training events. However, size cuts needed to be included in the optimization procedure, as they define the typical event rates.

Additional testing of optimized cuts for the hard size cuts showed no improvement compared to cuts optimized for medium size cuts. This is likely for two reasons: first, while estimates of event rates were obtained on data that passed size cuts, efficiency curves were computed during model training on data with no size cut imposed. Thus, the efficiency curves were not representative of the data used for optimization. Soft and medium size cuts affect cut optimization only at the lowest energies, while hard cuts encroach on higher energy bins. Second, BDT analysis with medium cuts resulted in such a substantial sensitivity gain, while maintaning a lower energy threshold, that the use of hard cuts for background rejection, at the expense of raising the energy threshold, was rendered ineffectual. The performance of the ITM+BDT analysis is discussed in the next chapter.

The BDT cuts were optimized in each zenith-energy bin for each of the three array epochs and two atmospheric profiles, with separate cut optimizations performed for medium and

soft size cuts. Detection cuts and two sets of source cuts, with normalizations at 0.2 CU and 0.05 CU, were computed for each epoch-atmosphere-size configuration. The full collection of 36 sets of optimized cuts is given in Appendix.

Chapter 4

Performance of the Improved Analysis

Combining the image template method for event reconstruction with boosted decision trees for event selection has proven to be an effective way to increase the sensitivity of VERITAS. While each of the techniques gives a moderate improvement of sensitivity on its own, the compound effect of ITM+BDT is significantly more substantial, resulting in a $\sim 30\%$ increase of the overall detection significance for analyses with medium size cuts, and a $\sim 50\%$ increase for analyses with soft size cuts.

Spectral reconstruction with ITM+BDT, however, reveals some problems. Specifically, analyses with soft size cuts tend to reconstruct harder spectra than expected; the hardening is caused by an underestimation of the flux at the lowest energies, typically in the first one or two energy bins. This underestimation is due to a poorly behaved energy reconstruction bias as well as an overestimation of the BDT cut efficiency during the construction of effective areas. The issue can be mitigated or avoided altogether by increasing the energy threshold

of the analysis, although an optimal strategy for such threshold selection needs to be studied in more detail. Until such fuller studies are performed, ITM+BDT with soft size cuts should only be used for source discovery analyses. Spectral reconstruction with medium size cuts, on the other hand, behaves as expected, agreeing with standard analysis, and is ready to be used in scientific studies, as done for the starburst galaxy M82 in Chapter 5.

This chapter details the performance of the ITM+BDT analysis method. First, we look at the sensitivity gain in comparison to the standard analysis; we also check how much of this gain comes from the use of ITM for event reconstruction alone, with standard event selection, and BDTs for event selection alone, with standard event reconstruction. We then turn to spectral reconstruction, investigating its problems and limitations, and making sure ITM+BDTs can be used for the analysis of M82 data. We conclude the chapter with suggestions on future studies and improvements of the method.

4.1 Detection Sensitivity

We begin the discussion of perfomance by looking at detection sensitivity of a number of *validation* sources. While the ROC curves and efficiencies obtained during training can inform us about the relative performance of different models, they cannot be used to judge real-world performance of an analysis. Therefore, we test the analysis by running it on a number of pre-selected validation sources. Care was taken not to include any data used for BDT training into these validation analyses.

4.1.1 Angular Resolution

Using ITM for event reconstruction results in a substantial improvement of angular reconstruction, and correspondingly, tighter θ^2 cuts for point sources. This alone should result in a decreased number of background events and larger values of the ratio of exposures τ, increasing overall significance. Figure 4.1 compares the 68% containment radius of a simulated point source between geometric and ITM event reconstructions. The results are shown as a function of energy for several zenith bins, and are averaged over ATM21 simulations of all 3 telescope epochs. ITM reconstruction clearly results in a superior angular resolution.

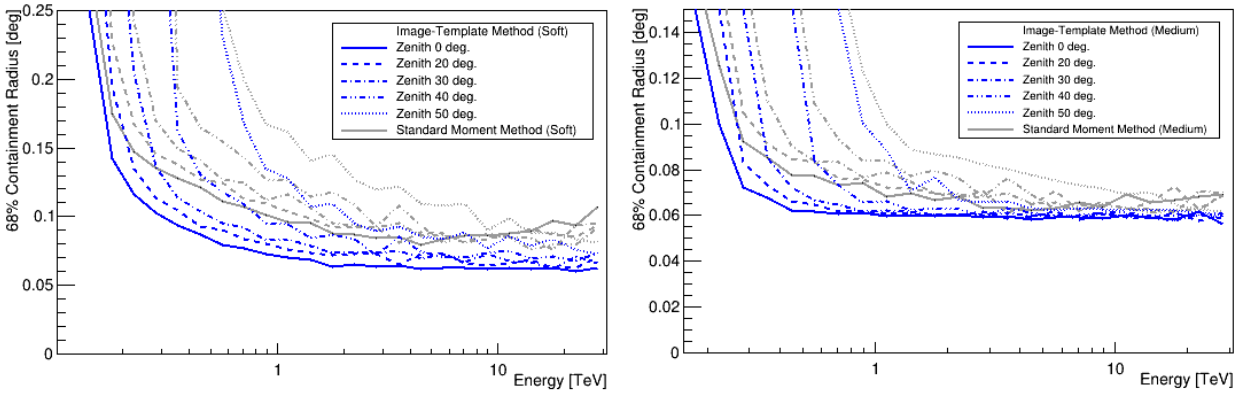

Figure 4.1: Improvements to angular resolution stemming from the use of ITM event reconstruction. The 68% containment radius of a simulated point source is shown as a function of energy for several zenith bins. ITM reconstruction is in blue, standard geometric reconstruction in grey. *Left:* soft cuts. *Right:* medium cuts. Note the change in the vertical scale between the two plots. Figure from Christiansen 2019.

Angular resolution degrades at lower energies, near the energy threshold of the array trigger, as the size of the shower – and the number of bright signal pixels – decreases and the relative prominence of background fluctuations increases. The energy threshold increases

with larger zenith angles due to the increased thickness of the atmosphere, and angular resolution degrades at higher zeniths due to the lower effective telescope multiplicity. As expected, this decrease in telescope multiplicity has less of an effect on ITM reconstruction than on the geometric one.

We conclude from the plots above that at energies far from the trigger threshold, θ of 0.1° for soft cuts and 0.07° for medium cuts covers a point source for zenith angles up to 45°, although some relaxation of the cut may be useful between 35° and 45°. Angular resolution degrades rapidly at zenith angles above 45°, and this work does not consider those cases.

We independently confirm these values of the optimal θ^2 cuts by performing the cuts optimization procedure outlined in Section 3.3.1 on a number of Crab runs. An example of the resulting θ^2 plots is shown in Figure 4.2. Here, the distribution of γ-like events as functions of θ^2 are shown for the On (blue) and Off (red) regions. The θ^2 at which the two distributions meet is then taken as the optimal value of the directional cut.[1] The ITM reconstruction procedure, shown on the left, clearly results in a narrower peak in the On region than the geometric reconstruction on the right. For the zenith range shown, the θ^2 cut optimization was done for events in the energy range of 400 GeV to 1 TeV. The procedure was repeated for Old and New Array data, as well as for medium cuts, although the lack of statistics at the earlier array epochs made the optimization procedure somewhat unstable. Therefore, the results obtained for Upgraded Array were used for the earlier epochs as well.

[1]The "meeting point" of the two distributions was determined by finding the maximum of the LiMa significance (equation 2.9) as a function of θ^2, accounting for the change in α as θ^2 increased.

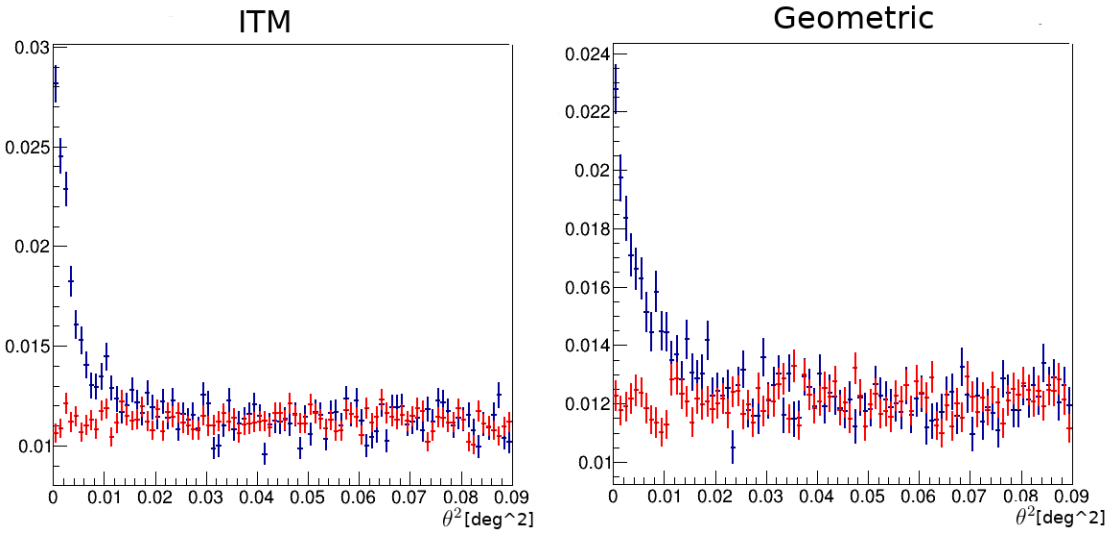

Figure 4.2: Angular distributions of On and Off events for ITM and geometric event reconstruction. Distribution of γ-like events in the On region is in blue, while the Off region events are in red. The θ^2 at which the On and Off distributions meet is taken as the optimal value of the directional cut. The plots are shown for ATM21 UA Crab data, analyzed with soft cuts, in the range of observational zeniths between 30° and 40°, for events of energies in the range $400GeV \leq E \leq 1TeV$. *Left:* ITM reconstruction. *Right:* geometric reconstruction.

This data-driven optimization confirmed the earlier findings: for ITM event reconstruction, the optimal directional cut should be taken at $\theta^2 = 0.01\,\mathrm{deg}^2$ for soft cuts; for medium cuts, $\theta^2 = 0.005\,\mathrm{deg}^2$ for data below 35° in zenith and $\theta^2 = 0.006\,\mathrm{deg}^2$ for data above 35° in zenith. While this last distinction for medium cuts below and above 35° in zenith may have been due to insufficient data and statistical fluctuations, the need for a less restrictive cut above 35° is also visible in Figure 4.1, and is expected from theoretical considerations. As before, no extrapolations are made for data taken at large zenith angles above 45°. The θ^2 cuts optimized for the ITM analysis are much tighter than the standard analysis cuts used

by VERITAS, $\theta^2 = 0.03\,\mathrm{deg}^2$ for soft and $\theta^2 = 0.01\,\mathrm{deg}^2$ for medium cuts.

4.1.2 Significance

We compare the performance of the ITM+BDT analysis with the standard analysis by analyzing a number of sources. The Crab Nebula, a strong stable source, is used as the primary reference. We bin the Crab data by the zenith angle of observations. We further select a number of strong and weak sources for each array epoch. Tables 4.1 and 4.2 show the improvement in detection significance for the zenith-binned Crab data. There, 10° zenith denotes data below 15° in zenith; 20° corresponds to data in the range $15° - 25°$; 30° includes data from 25° to 40°; and $> 40°$ are data above 40° in zenith, typically up to 55°.

Tables 4.3 and 4.4 show the sensitivity improvements for a number of other sources, both weak and strong. The blazars Mrk 421 and Mrk 501 are strong sources, analyzed with cuts optimized for 0.2 CU gamma flux, while the rest are weak, analyzed with the cuts optimized for 0.05 CU.

It is immediately clear that the ITM+BDT analysis offers a substantial sensitivity improvement at both soft and medium cuts. The improvement is greater for soft cuts, reflecting the presence of a large number of smaller showers that are more difficult to reconstruct and classify with simpler methods. Breaking the analysis down into separate BDT and ITM components also provides insight into where the performance improvements come from. It is clear that both BDT and ITM analyses alone perform better than the standard analysis,

	OA ATM21								***OA ATM22***							
	BOX		BDT		ITM		ITM+BDT		BOX		BDT		ITM		ITM+BDT	
10° Zen	**32.5**		**36.9**		**39.3**		**45**		**25.0**		**27.6**		**26.8**		**31.2**	
	787	256	823	178	692	101	559	18	457	138	426	99	346	57	296	12
20° Zen	**26.0**		**29.7**		**29.6**		**34.8**		**30.7**		**33.3**		**34.8**		**38**	
	469	138	482	87	375	50	339	11	612	147	581	87	498	55	398	10
30° Zen	**28.3**		**30**		**33.7**		**36.5**		**8.3**		**9.4**		**11.7**		**10.5**	
	568	171	502	93	494	68	382	14	66	30	56	13	63	9	29	0
>40° Zen	**20.2**		**20.1**		**23.5**		**26.5**		**21.0**		**18.6**		**17.6**		**16.2**	
	307	97	257	59	250	35	210	8	319	98	252	78	169	37	140	30
	NA ATM21								***NA ATM22***							
	BOX		BDT		ITM		ITM+BDT		BOX		BDT		ITM		ITM+BDT	
10° Zen	**88.7**		**98.9**		**101.5**		**114.2**		**43.0**		**48.0**		**49**		**54.7**	
	5390	1513	5961	1328	4224	501	3733	140	1160	270	1264	221	930	93	828	25
20° Zen	**61.1**		**67.8**		**68.9**		**77.8**		**34.7**		**39**		**41.7**		**47.4**	
	2501	702	2745	618	4250	475	1711	64	845	241	889	174	718	85	645	23
30° Zen	**28.9**		**32.8**		**33.7**		**38.1**		**21.5**		**23.3**		**25.8**		**27.7**	
	579	186	598	127	466	63	386	14	296	72	292	49	266	29	216	7
>40° Zen	**25.4**		**29.8**		**31.25**		**34.3**		**4.4**		**3.76**		**7.25**		**6.3**	
	482	158	496	90	421	56	348	14	16	3	10	2	17	0.9	12	0.4
	UA ATM21								***UA ATM22***							
	BOX		BDT		ITM		ITM+BDT		BOX		BDT		ITM		ITM+BDT	
10° Zen	**92.4**		**103**		**108.2**		**121.7**		**55.9**		**61.1**		**67.6**		**76**	
	4281	767	4870	716	3654	256	4184	214	1716	293	1953	298	1535	102	1760	83
20° Zen	**85.0**		**93.4**		**96.9**		**108.1**		**51.6**		**56.2**		**60.9**		**66.9**	
	3155	456	3475	405	2664	154	2968	118	1421	248	1609	253	1255	86	1384	69
30° Zen	**48**		**50.9**		**55.3**		**59.6**		**31.7**		**33.4**		**38.3**		**41.6**	
	1145	185	1222	175	952	60	983	41	562	97	566	77	498	33	521	22
>40° Zen	**32.7**		**35.1**		**39.7**		**42.3**		**25.8**		**28.7**		**29.9**		**31**	
	584	94	622	82	513	30	523	20	417	94	424	62	327	62	309	17

Table 4.1: Comparison of analysis methods on zenith-binned Crab data with *soft* cuts. See main text for the zenith ranges. *BOX* is the standard analysis; *BDT* uses geometric event reconstruction and BDTs for event selection; *ITM* uses ITM for event reconstruction and box cuts for event selection; *ITM+BDT* combines both ITM for reconstruction and BDT for event selection, with the BDT model including extra ITM training attributes. Each two-row cell includes the overall significance, in bold in the top row, as well as the number of Excess and Off events at the bottom; Off events are scaled by alpha, giving the number of background events at the point source. The BDT analyses were done with cuts optimized for strong sources.

	OA ATM21								***OA ATM22***							
	BOX		BDT		ITM		ITM+BDT		BOX		BDT		ITM		ITM+BDT	
10° Zen	**33.6**		**35.7**		**40.1**		**42.8**		**23.9**		**25.5**		**26**		**28.8**	
	364	22	407	23	400	13	421	9	190	11	214	12	177	7	203	6
20° Zen	**26.4**		**28.4**		**30.1**		**32**		**28.1**		**29.5**		**33.4**		**35.5**	
	214	11	244	11	216	6	234	5	253	13	272	13	274	7	294	6
30° Zen	**28.9**		**28.9**		**32.7**		**32.8**		**9.0**		**9.5**		**10.3**		**8.6**	
	267	15	263	14	268	8	250	5	31	3	33	2	30	1	17.6	0.4
>40° Zen	**20.4**		**20.2**		**22.9**		**23.1**		**19.5**		**19**		**15.6**		**16.6**	
	143	9	132	7	135	5	127	3	123	7	103	4	72	4	67	2
	NA ATM21								***NA ATM22***							
	BOX		BDT		ITM		ITM+BDT		BOX		BDT		ITM		ITM+BDT	
10° Zen	**94.7**		**99.8**		**100.4**		**107.1**		**43.6**		**47.5**		**49.2**		**53.2**	
	2743	134	3222	193	2431	67	2722	69	565	24	697	34	571	13	627	11
20° Zen	**62.9**		**66.8**		**68**		**72.7**		**34.9**		**38.5**		**39.95**		**44.4**	
	1217	64	1433	88	1118	31	1237	30	402	25	483	28	412	14	471	11
30° Zen	**31.7**		**33.2**		**33.7**		**35**		**21.9**		**23.1**		**25.1**		**25.4**	
	298	17	326	18	266	8	271	7	146	7	166	8	155	4	153.5	3.5
>40° Zen	**26.5**		**28**		**30.75**		**31.2**		**5.0**		**4.5**		**5.64**		**4.7**	
	232	14	249	14	236	7	228	5	8	0.5	7	0.6	8	0.2	6	0.2
	UA ATM21								***UA ATM22***							
	BOX		BDT		ITM		ITM+BDT		BOX		BDT		ITM		ITM+BDT	
10° Zen	**87.9**		**96.1**		**98.7**		**108.5**		**55.5**		**60.7**		**61.9**		**67.9**	
	2008	76	2445	100	2005	38	2473	51	860	28	1065	40	853	14	1062	21
20° Zen	**79.7**		**84.4**		**89.4**		**97**		**50.8**		**53.6**		**56.2**		**60.1**	
	1509	47	1704	54	1532	24	1836	31	722	24	846	35	705	13	834	18
30° Zen	**43.2**		**45.9**		**48.8**		**51.6**		**30.3**		**32.1**		**35**		**37.1**	
	503	19	573	22	505	9	565	11	271	10	305	12	275	5	310.7	5
>40° Zen	**30.9**		**32.6**		**34**		**37.2**		**22.2**		**22.74**		**27**		**27.6**	
	265	8	300	10	270	6	309	5	167	10	164	7	171	4	174	3

Table 4.2: Comparison of analysis methods on zenith-binned Crab data with *medium* cuts. See main text for the zenith ranges. *BOX* is the standard analysis; *BDT* uses geometric event reconstruction and BDTs for event selection; *ITM* uses ITM for event reconstruction and box cuts for event selection; *ITM+BDT* combines both ITM for reconstruction and BDT for event selection, with the BDT model including extra ITM training attributes. Each two-row cell includes the overall significance, in bold in the top row, as well as the number of Excess and Off events at the bottom; Off events are scaled by alpha, giving the number of background events at the point source. The BDT analyses were done with cuts optimized for strong sources.

	OA							
	BOX		BDT		ITM		ITM+BDT	
1ES 1218 +304	**28**		**30**		**30.7**		**38.8**	
	2494	5885	2226	3927	1782	2488	1114	422
Mrk 501	**23.3**		**27.7**		**26.3**		**35.5**	
	1044	1381	1093	979	785	568	644	126
	NA							
	BOX		BDT		ITM		ITM+BDT	
1ES 1218 +304	**37.5**		**43.2**		**45.9**		**58.5**	
	4903	12693	4975	9415	3596	4370	2427	839
Mrk 501	**84.9**		**96.4**		**104.6**		**124.5**	
	6588	3109	7145	2528	5538	1062	5015	303
	UA							
	BOX		BDT		ITM		ITM+BDT	
1ES 1218 +304	**15.2**		**18.2**		**18.5**		**23.2**	
	797	2043	874	1724	590	727	615	436
PKS 1424 +240	**14.4**		**16**		**18.3**		**19.5**	
	1708	11617	1796	10235	1285	4082	1040	2377
OJ 287	**2.48**		**4.57**		**5.43**		**6.1**	
	120	1943	211	1771	153	679	135	405
Mrk 421	**305**		**314**		**333**		**353**	
	33065	1636	33999	1454	28177	565	29852	392

Table 4.3: Comparison of analysis methods on several known sources with *soft* cuts. *BOX* is the standard analysis; *BDT* uses geometric event reconstruction and BDTs for event selection; *ITM* uses ITM for event reconstruction and box cuts for event selection; *ITM+BDT* combines both ITM for reconstruction and BDT for event selection, with the BDT model including extra ITM training attributes. Each two-row cell includes the overall significance, in bold in the top row, as well as the number of Excess and Off events at the bottom; Off events are scaled by alpha, giving the number of background events at the point source. The BDT analyses of the Markarians were done with cuts optimized for strong sources; the rest were analyzed with cuts optimized for weak sources.

but achieve their respective higher sensitivities in different ways: ITM decreases the number of background events through better angular reconstruction; BDT, depending on the choice

	OA							
	BOX		BDT		ITM		ITM+BDT	
M87	**7.6**		**9.8**		**11.1**		**12.7**	
	243	852	306	795	278	500	262	314
Cas A	**10.4**		**11.6**		**11.9**		**13.1**	
	220	361	213	255	187	186	162	103
Mrk 501	**28.4**		**29.2**		**29.7**		**32.7**	
	478	125	507	133	399	68	422	54
	NA							
	BOX		BDT		ITM		ITM+BDT	
IC 443	**6.5**		**7.1**		**6.6**		**8.2**	
	120	286	136	299	83	131	88	88
Mrk 501	**97**		**103.6**		**108.6**		**118.3**	
	3378	282	4062	394	3253	151	3689	145
	UA							
	BOX		BDT		ITM		ITM+BDT	
Cas A	**15**		**15.5**		**18.7**		**19.3**	
	300	289	300	268	293	158	269	116
Mrk 421	**250**		**267**		**279**		**299**	
	14702	158	16751	192	15107	83	17561	109

Table 4.4: Comparison of analysis methods on several known sources with *medium* cuts. *BOX* is the standard analysis; *BDT* uses geometric event reconstruction and BDTs for event selection; *ITM* uses ITM for event reconstruction and box cuts for event selection; *ITM+BDT* combines both ITM for reconstruction and BDT for event selection, with the BDT model including extra ITM training attributes. Each two-row cell includes the overall significance, in bold in the top row, as well as the number of Excess and Off events at the bottom; Off events are scaled by alpha, giving the number of background events at the point source. The BDT analyses of the Markarians were done with cuts optimized for strong sources; the rest were analyzed with cuts optimized for weak sources.

of the optimized cuts, either increases the number of excess events, while slightly lowering the background rates (when dealing with strong sources), or suppresses the background (when dealing with weak sources). Combining the two approaches, ITM+BDT typically manages

to achieve a dramatic suppression of the background, while also finding more excess events in the case of strong sources.

The performance gain decreases slightly with the increasing zenith angle of observations, likely the result of the increasing energy threshold and the corresponding decrease in the number of lower-energy showers, where the significance gain is the most prominent.

The OA ATM22 BDT models at zenith angles above 30° seem to perform worse than the standard analysis. As mentioned previously, this is a special case as there is a very small amount of summer data available for the OA epoch, especially at larger zeniths. The lack of training data likely resulted in overtraining of the BDT model that went unnoticed as the training and test samples had a high degree of temporal correlation. This poor performance, however, is a non-issue as there is simply no OA ATM22 data at these zenith angles, other than the very limited sample used for training.

While these tables show the performance gain for known sources, we would also like to see if the visibility threshold itself is lowered. That is, in a given observation time, what is the difference between the minimum source strengths detectable with standard analysis and ITM+BDT? Figures 4.3 and 4.4 show just that – Figure 4.3 shows the *differential* sensitivity at a given energy, while Figure 4.4 shows the *integral* sensitivity above a given energy. In the figures, the *least detectable source* after 50 hours of observations is shown as a function of energy for both analyses. Just as with known sources, the sensitivity gain for soft cuts is greater than for medium cuts. The sensitivity depends solely on the background rejection efficiency, and we see that in relative terms, the background rejection efficiency improves

towards higher energies, an effect we observed during BDT training as well. The ostensibly poor performance of medium cuts Old Array model at energies above 2 TeV is actually due to very low statistics in the measurement procedure and is unlikely to be representative of actual performance. *Overall, the ITM+BDT analysis sees a 40-60% gain in integral sensitivity for soft cuts and a 15-50% improvement for medium cuts.* Such a substantial upgrade is very much welcome when hunting for transient sources, especially in the age of multimessenger astronomy.

To generate Figure 4.3, a number of Crab runs were processed with the standard analysis and with the ITM+BDT analysis using the *detection cuts*. For each analysis, the On and Off event rates were then measured and the number of excess as well as background events resulting from a 50-hour observation computed. Using the scaled background events, equation 3.10 was then used to determine the visibility threshold at the 5σ confidence level. Finally, the visibility threshold was divided by the number of excess events to directly obtain the sensitivity in Crab units, circumventing any confounding factors that might result from computing effective areas. In order to obtain the energy-dependent curves, the measurement was performed in 5 energy bins, with quadratic spline interpolation between them. Figure 4.4 was generated in the same way, with the only difference being that the numbers of On and Off events were computed above a given energy; here, 9 energy bins were used.

Table 4.3 demonstrates the benefit of the ITM+BDT analysis in the case of a blazar flare. The *OJ 287* row near the bottom corresponds to an early stage of the February 2017 flare of the blazar – while the standard analysis did not see anything, ITM+BDT conclusively

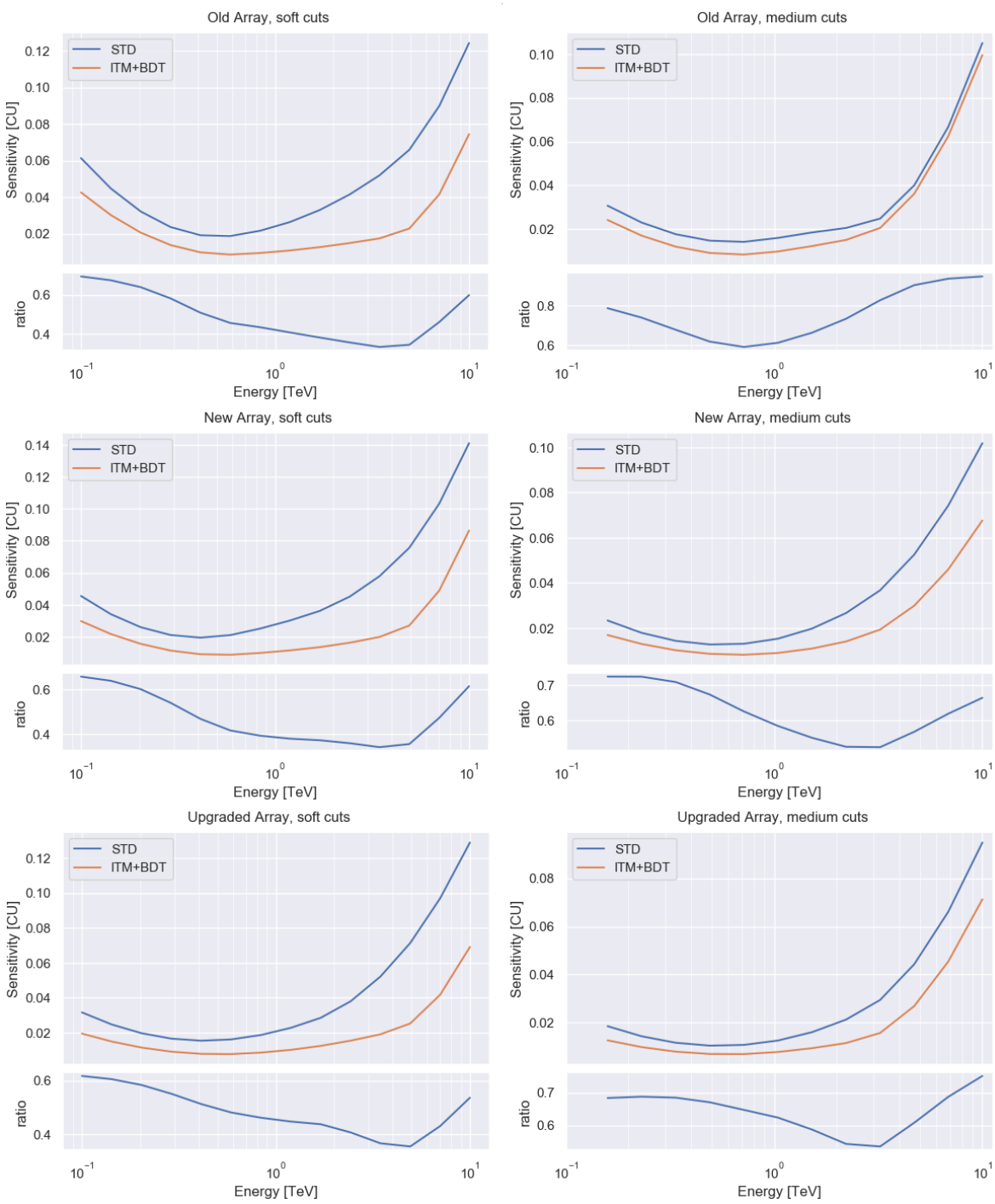

Figure 4.3: Comparison of differential sensitivities of 50h observations between standard analysis and ITM+BDT. Soft cuts are on the left, medium cuts on the right. Array epochs go in chronological order from top to bottom. The sensitivities, shown in the top subplot of each panel, are in Crab units at each energy. The bottom subplot of each panel shows the ratio of the sensitivities. See main text for full details.

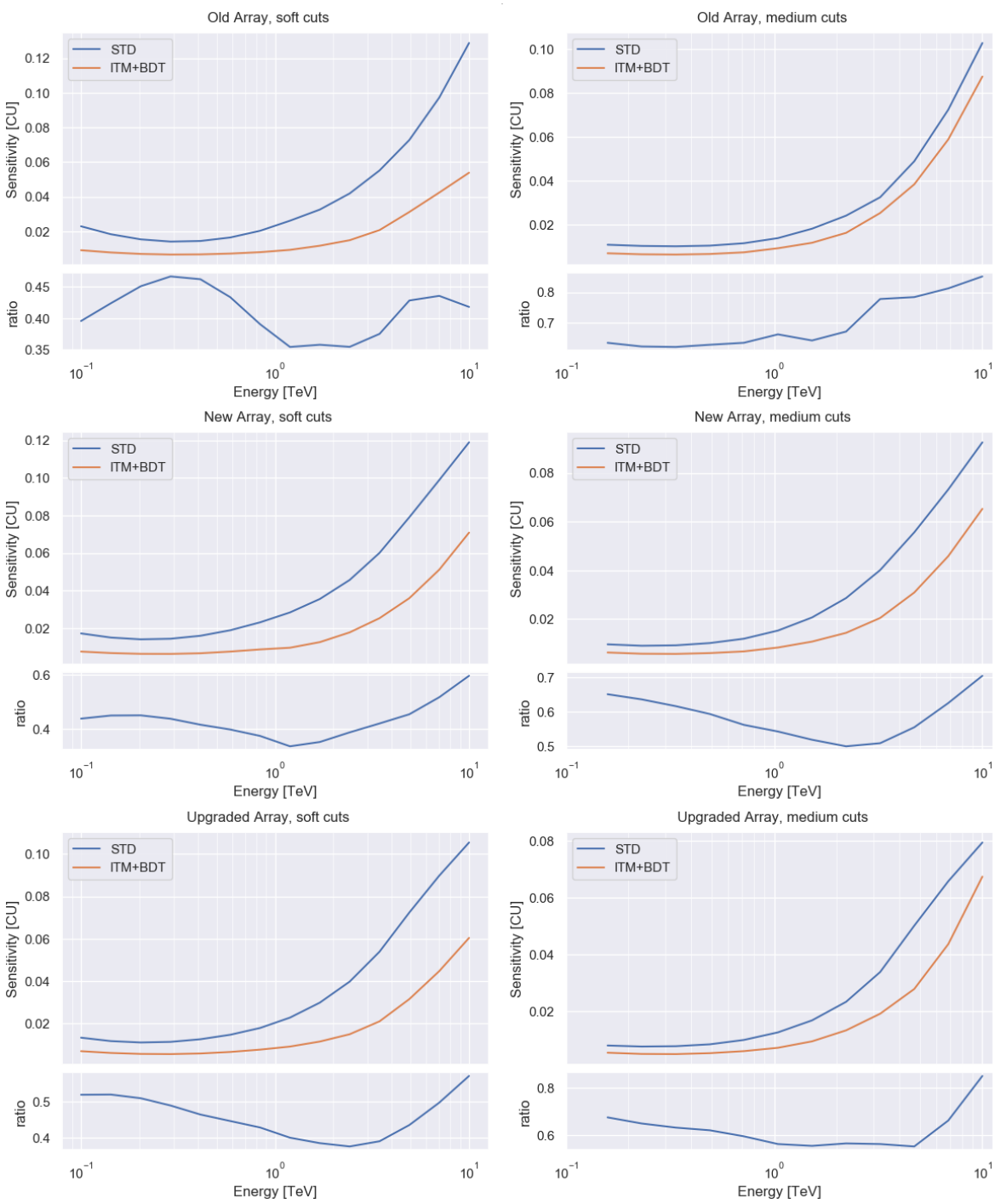

Figure 4.4: Comparison of integral sensitivities of 50h observations between standard analysis and ITM+BDT. Soft cuts are on the left, medium cuts on the right. Array epochs go in chronological order from top to bottom. The sensitivities, shown in the top subplot of each panel, are in Crab units at each energy. The bottom subplot of each panel shows the ratio of the sensitivities. See main text for full details.

detected a source. In reality, BDT-based analysis with EventDisplay was similar to the second column shown in the table, suggestive of a flare. This prompted further observations of the blazar, ultimately resulting in a conclusive detection with both analysis packages. A transient event of a shorter timescale would have been only visible with ITM+BDT.

Finally, in order to make sure no strange artifacts appear in the field of view, sky maps generated by the ITM+BDT analysis were examined. Figure 4.5 shows Crab skymaps for each epoch, for data analyzed with medium cuts. The panels on the right are significance distributions with a region of 0.3° in radius around the source excluded – these are close to normal, as they should be. Note that the significance maps are correlated, hence the relatively broad source region.

4.2 Spectral Reconstruction

We now turn to spectral reconstruction. In order to measure the energy spectrum, energy-binned event rates need to be converted to energy flux. This conversion is done by means of effective areas (EAs) described in Section 2.4.3. Generating EAs is a computationally expensive process, as a large set of shower simulations for all possible zeniths, azimuths and levels of the night sky background (NSB) noise need to be processed in the same manner as the data for which a spectrum is desired. In this case, EAs were generated for soft and medium cuts for both atmospheres and all array epochs for the standard, ITM (ITM used for event reconstruction, with box cuts for event selection), and ITM+BDT strong and weak

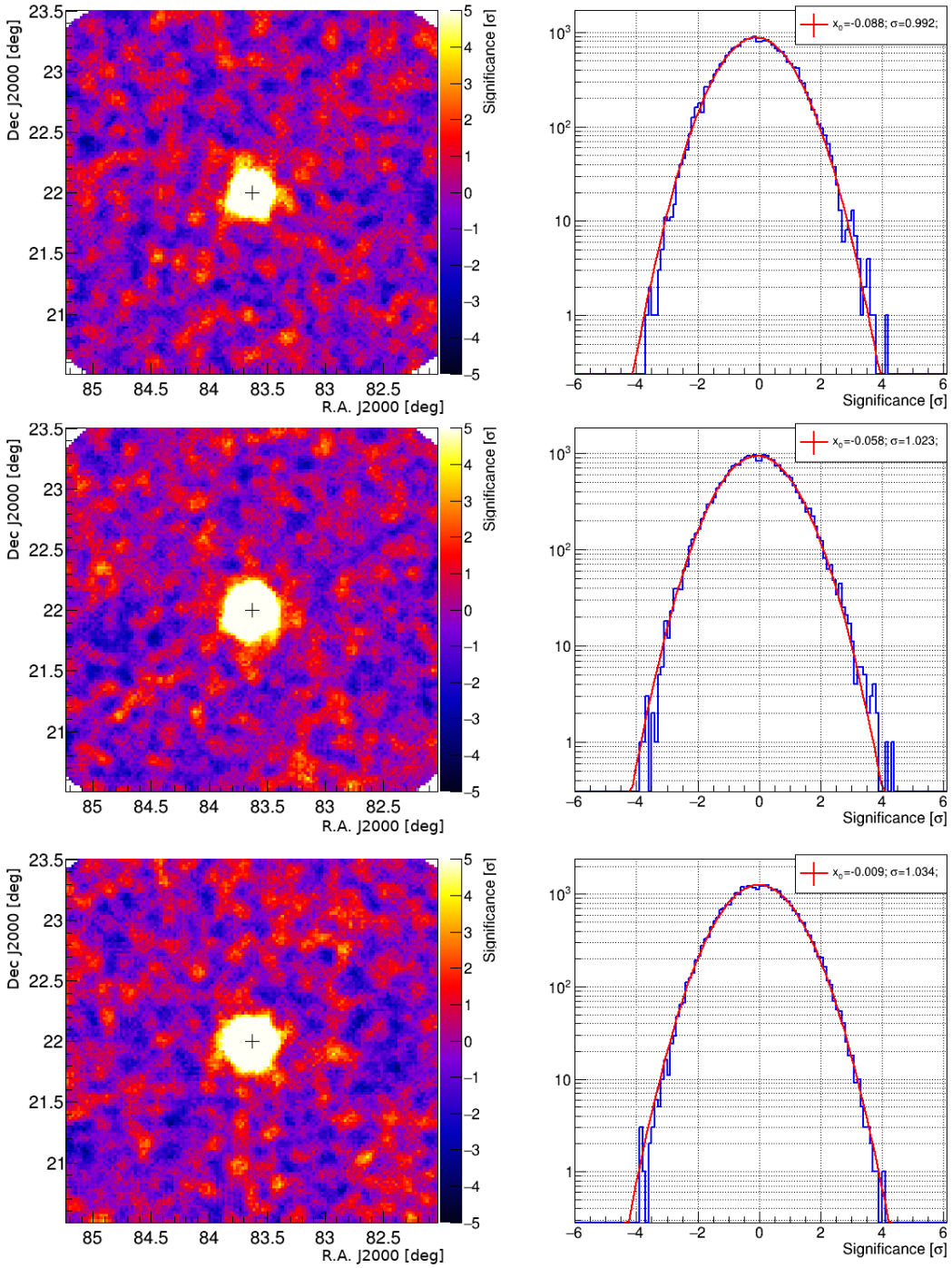

Figure 4.5: Crab skymaps with medium cuts ITM+BDT analysis. The array epochs go in chronological order from top to bottom. *Left:* significance maps. *Right:* significance distributions with the source regions excluded. These are expected to be normal.

source cuts analyses. For the same validation sources as in the previous section, spectra were reconstructed with 5 bins per energy decade. Only sources that were expected to show simple power-law spectra were used, and all results with 3 or fewer significant points were ignored.

Immediately, a discrepancy between the standard analysis and ITM/ITM+BDT analyses is apparent. Figure 4.6 shows Crab spectra for all 4 analyses, for both soft and medium cuts, with Crab data for all zenith ranges pooled together. The standard analysis is in black, ITM+BDT strong, ITM+BDT weak and ITM are in red, blue and green, respectively.

The lowest energy bins of ITM and ITM+BDT analyses show significantly lower fluxes than their standard analysis counterparts, resulting in systematically harder spectra. The effect is especially pronounced for soft cuts ITM+BDT analyses for the old and new array epochs. While soft cuts ITM analysis here behaves better than the ITM+BDT analysis for old and new array epochs, its discrepancy in flux at low energy is just as pronounced for upgraded array data, and for all medium cuts spectra.

There are two possible reasons for such a discrepancy in the measured flux. First, a large bias in reconstruced energy near the energy threshold could push some lower-energy events out of their true energy bins, either completely discarding them by making them appear subthreshold (when the bias is negative), or effectively distributing their flux into higher energy bins (when the bias is positive). Second, EAs could be overestimated because of subtle differences between real gamma events and simulated ones. An effective area is the fraction of detected showers out of all thrown showers, and is therefore a measure of *efficiency*

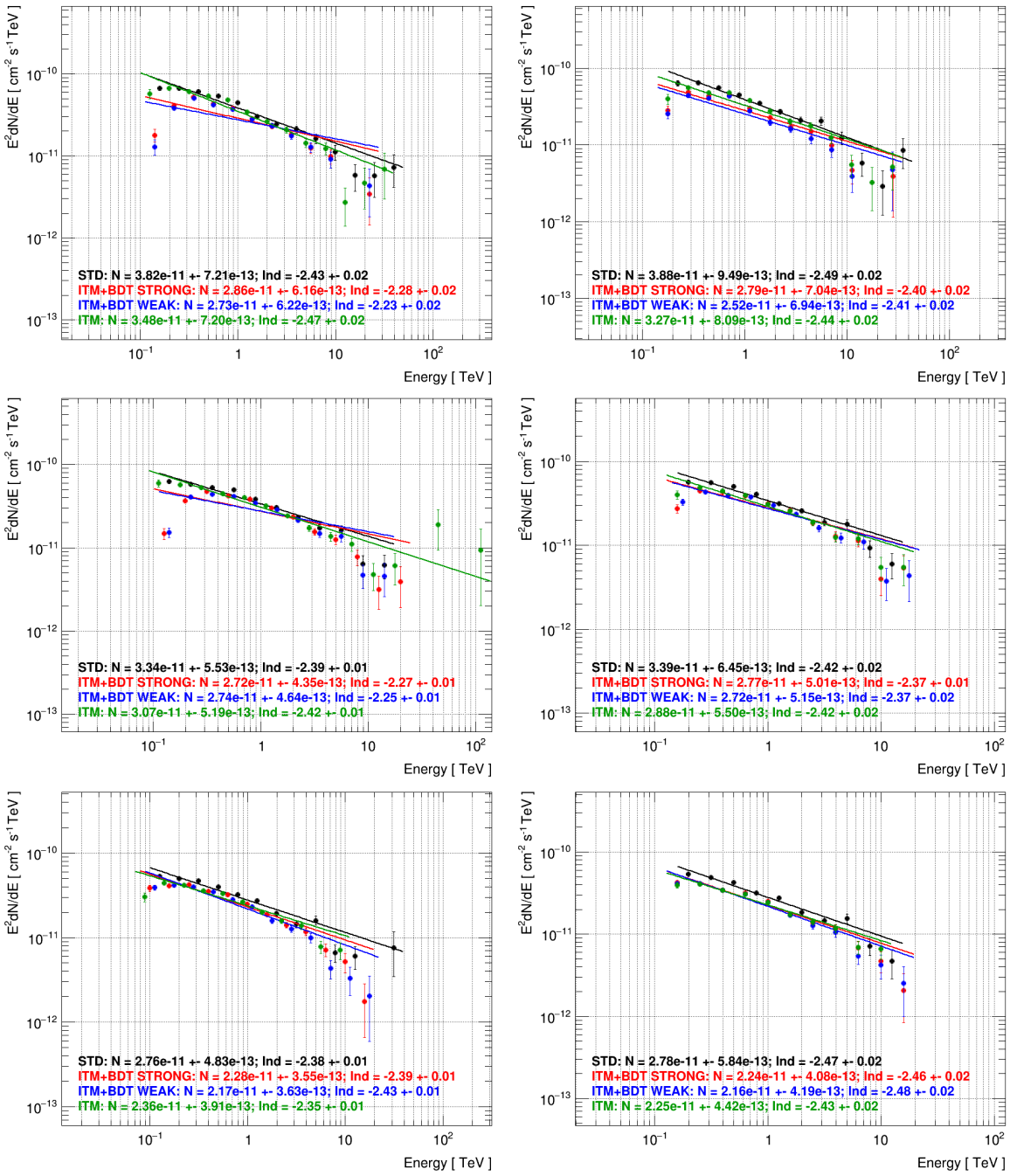

Figure 4.6: Reconstruction of energy spectra of the Crab Nebula. ATM21 data are shown, with array epochs in chronological order from top to bottom. Standard analysis is in black, while the red, blue and green points correspond to the ITM+BDT *strong*, ITM+BDT *weak*, and ITM analyses. Soft cuts analysis is on the left, medium cuts on the right.

of the instrument and the analysis chain. If, for whatever reason, the true efficiency for real gamma events is lower than the computed one for simulations, the EA is overestimated, and the flux measured on real data, therefore, underestimated. These two effects could work together, of course – they could either compound or cancel each other out, depending on how the true efficiency differs from the simulated one. To hunt down the low-energy issue of ITM+BDT spectral reconstruction, we investigate both of these effects.

The effect of energy bias on spectral reconstruction is well-known for VERITAS. In order to combat this, VERITAS analyses choose a safe energy range, ignoring all events that fall below the lower-energy threshold or above the upper one. While initially VEGAS chose its safe energy range to be the region in which the energy reconstruction bias is lower than a pre-determined threshold of 10%, studies showed that the effect of large uncertainties in EAs near that threshold had a large negative impact on the measured spectra. Therefore, currently, the safe energy range of the standard analysis is chosen to be the range in which the relative uncertainty of the EA is $\leq 20\%$.

We revisit the energy reconstruction bias by first attempting to reconstruct spectra of simulations. Figure 4.7 shows the reconstructed simulated spectra for all 4 analyses, for both soft and medium cuts, for ATM21 simulations thrown at 30° zenith. The standard analysis is in black, ITM+BDT strong, ITM+BDT weak and ITM are in red, blue and green, respectively. The simulated spectrum has an index of -2, therefore the energy spectra should be horizontal lines. This is mostly true, apart from the lowest and highest energies. The low-energy behavior is particularly interesting, with a prominent break. For this particular

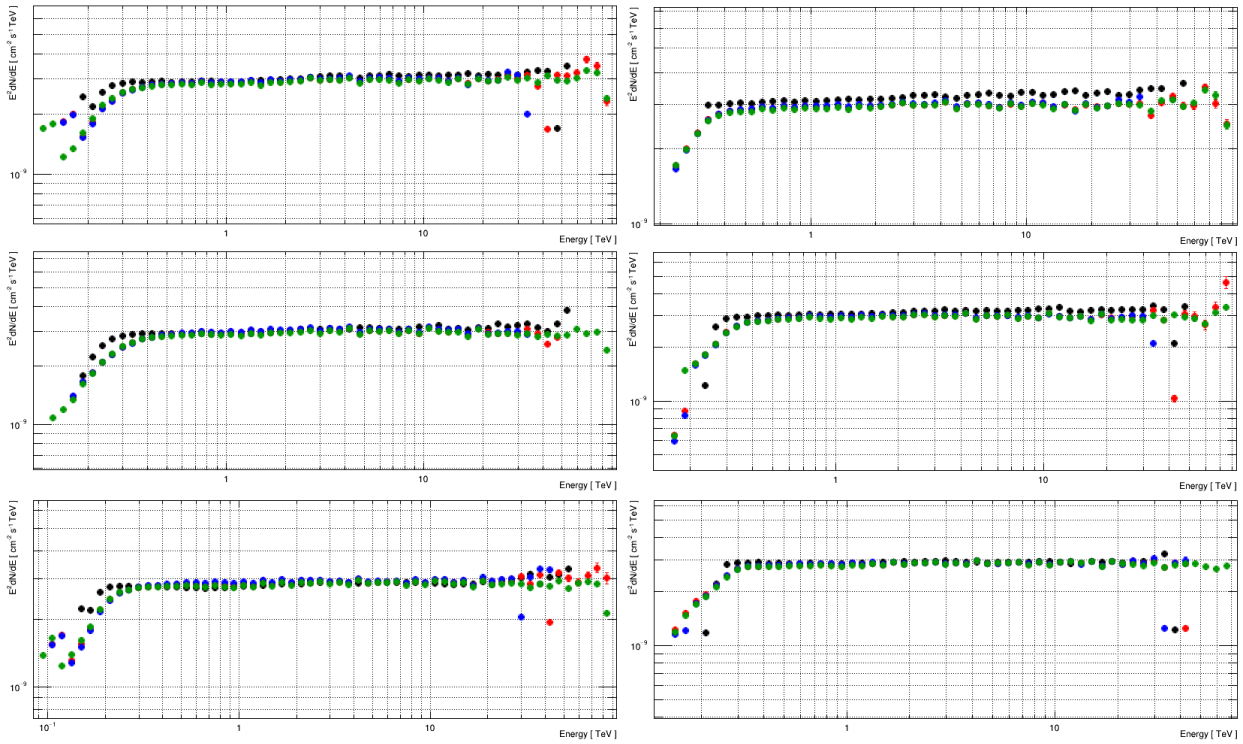

Figure 4.7: Reconstruction of energy spectra of simulations. ATM21, zenith 30° showers were simulated with a spectral index of -2.0, hence the reconstructed spectra should be horizontal lines. Standard analysis is in black, while the overlapping red, blue and green points correspond to the ITM+BDT *strong*, ITM+BDT *weak*, and ITM analyses. Soft cuts analysis is on the left, medium cuts on the right. Array epochs are in chronological order from top to bottom. See main text for a discussion.

case, the standard analysis lower-safe-energy threshold is set at around 250 GeV for the soft analysis (to be exact, the thresholds are 282, 251 and 224 GeV for OA, NA and UA epochs, respectively) and around 320 GeV for the medium analysis (to be exact, the thresholds are 355, 316 and 316 for OA, NA and UA epochs, respectively), right above the break in each case. The onset of the spectral break in the ITM and ITM+BDT analyses, however, clearly occurs at lower energ, while the safe energy threshold, as determined from EA uncertainty, is usually slightly *lower* than the standard analysis (for soft cuts, it's set at 246, 276 and 196

GeV for OA, NA and UA, respectively; for medium cuts, it's at 310, 310 and 246 GeV for OA, NA and UA, respectively). This is troubling, as for ITM-based analayses, the reconstruced simulated fluxes near these thresholds are already substantially lower than they should be, 10-20% lower, in fact. The same behavior is observed at other zeniths.

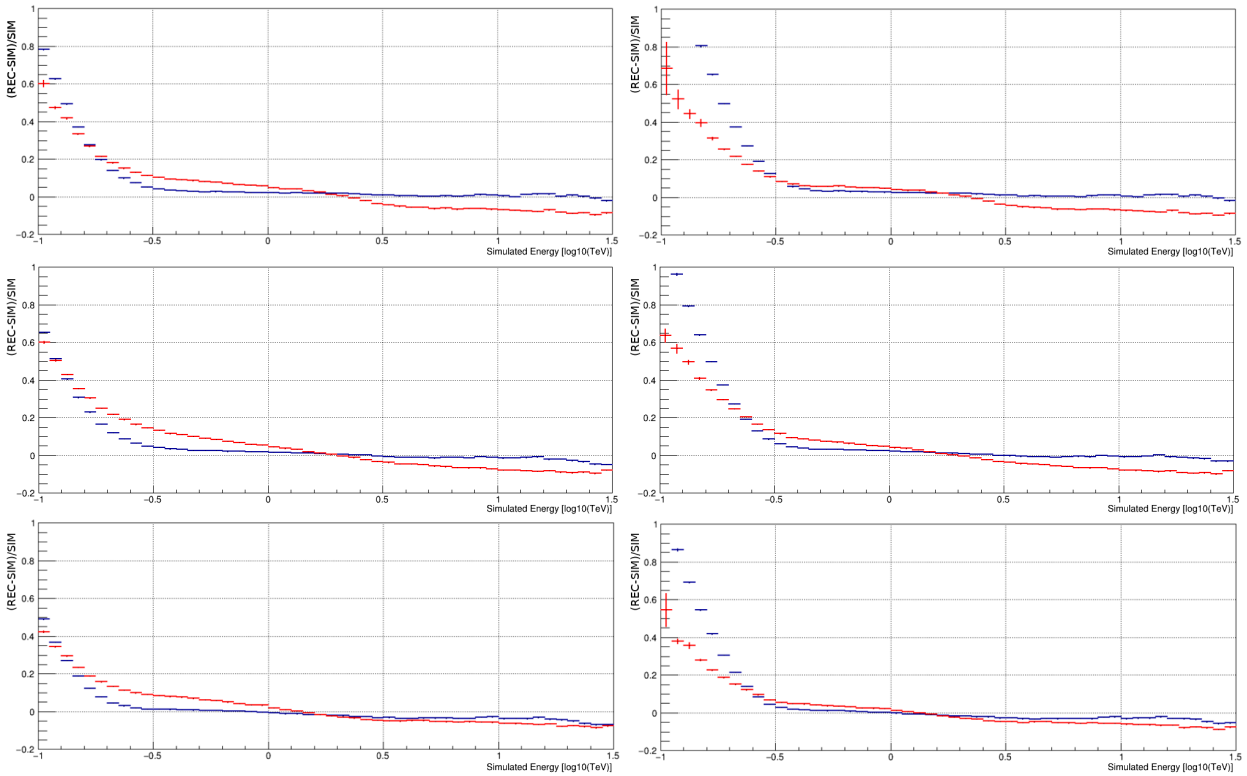

Figure 4.8: Comparison of the energy reconstruction bias between standard analysis and ITM. The horizontal axis is the simulated energy, while the vertical axis is the difference between the reconstructed and the true energies, as a fraction of the true energy. ATM21, zenith 30° simulated showers are shown, with array epochs in chronological order from top to bottom. Standard energy reconstruction is in blue, while the red points correspond to the ITM analysis. Soft cuts analysis is on the left, medium cuts on the right. See main text for a discussion.

We further look at the reconstructed energy bias directly. Figure 4.8 shows the energy bias for standard (in blue) and ITM (in red) event reconstructions for the same simulations

as shown in Figure 4.7. The bias, plotted on the vertical axis, is the ratio $(E_{\mathrm{rec}} - E_{\mathrm{sim}})/E_{\mathrm{sim}}$, where E_{sim} is the true simulated energy, and E_{rec} is the reconstructed one. As expected from the use of lookup tables, the standard energy bias is very close to zero throughout, only behaving poorly at the lowest energies, where shower-to-shower fluctuations, interacting with the trigger and analysis threshold, result in only atypically bright showers being detected by the array. The ITM energy reconstruction, however, looks a little troubling throughout the energy range, especially for soft cuts, with a slight slope in the bias. This slope is the result of convolving the true energy spectrum with a detector response of finite resolution (Christiansen 2019), and would have the effect of slightly tilting the whole spectrum towards softer indices. It is, however, rather benign, as it confines the bias to a narrow range of around 10% and can be corrected for by the VEGAS spectral reconstruction machinery by means of migration matrices. On the other hand, the relatively early onset of the more prominent bend upwards at lower energies near the trigger threshold, again very visible at soft cuts, is the cause of the early spectral break seen in the simulations. Again, the same behavior is observed at other zeniths.

It is clear then that setting the safe energy range away from this nonlinear bend in the energy reconstruction bias is crucial to the correctness of spectral reconstruction. But what role, if any, does the EA efficiency over- or underestimation play in spectral reconstruction? While it is far from clear how to approach the efficiency of multidimensional box cuts, BDT-based event selection provides us with a very useful proxy, the BDT score and its cumulative distribution function (CDF).

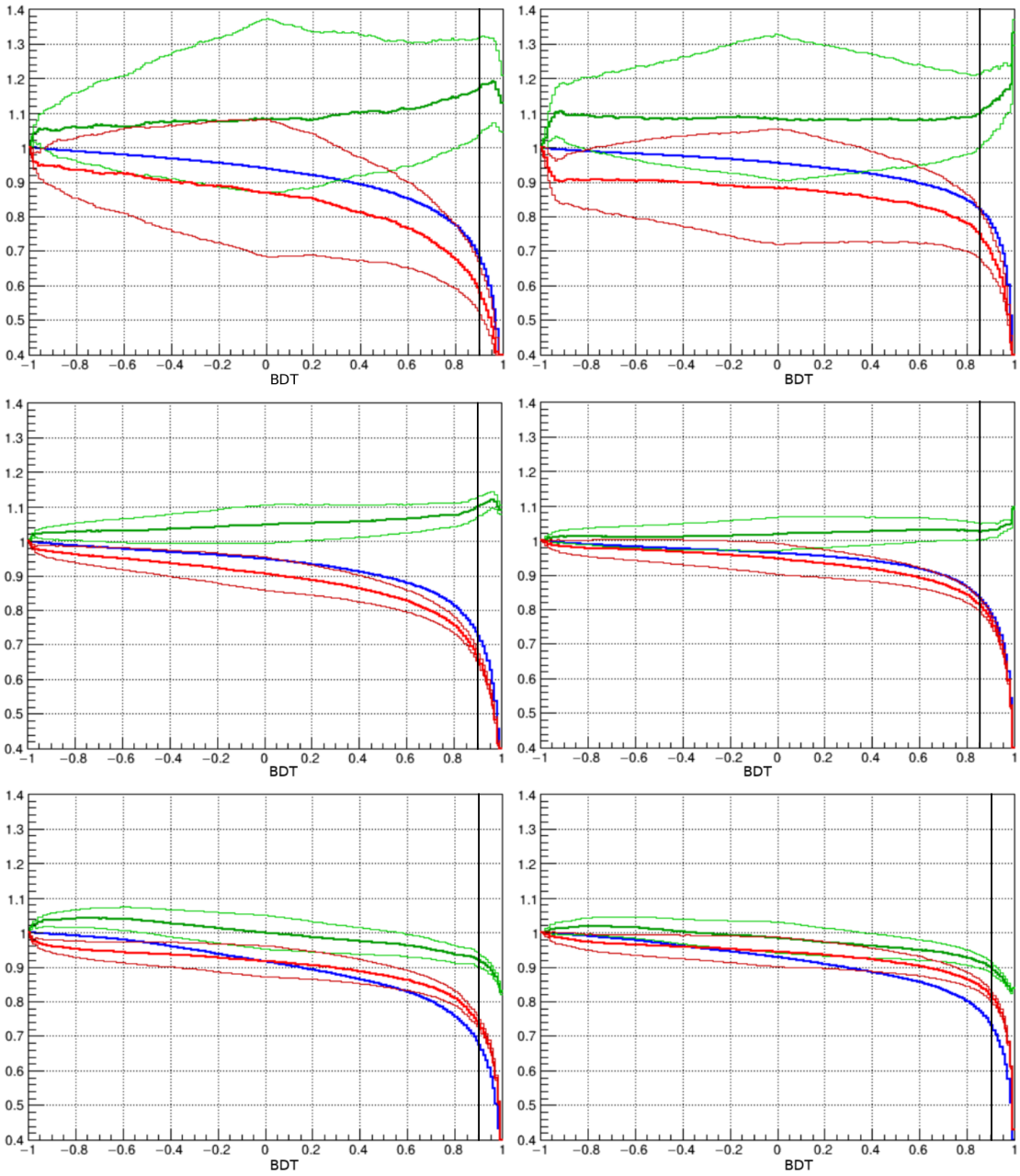

Figure 4.9: Comparison of BDT score CDFs for γ-ray events between simulations and data below 500 GeV. The CDFs of simulations are in blue, while those of data in red. The ratio of simulation to data is in green. Typical BDT cut values are shown with thicker black lines. The thin lines show naive errors that track the statistical uncertainties of the underlying distributions. Soft cuts are on the left, medium cuts on the right, telescope epochs in chronological order from top to bottom. See main text for a full explanation and discussion.

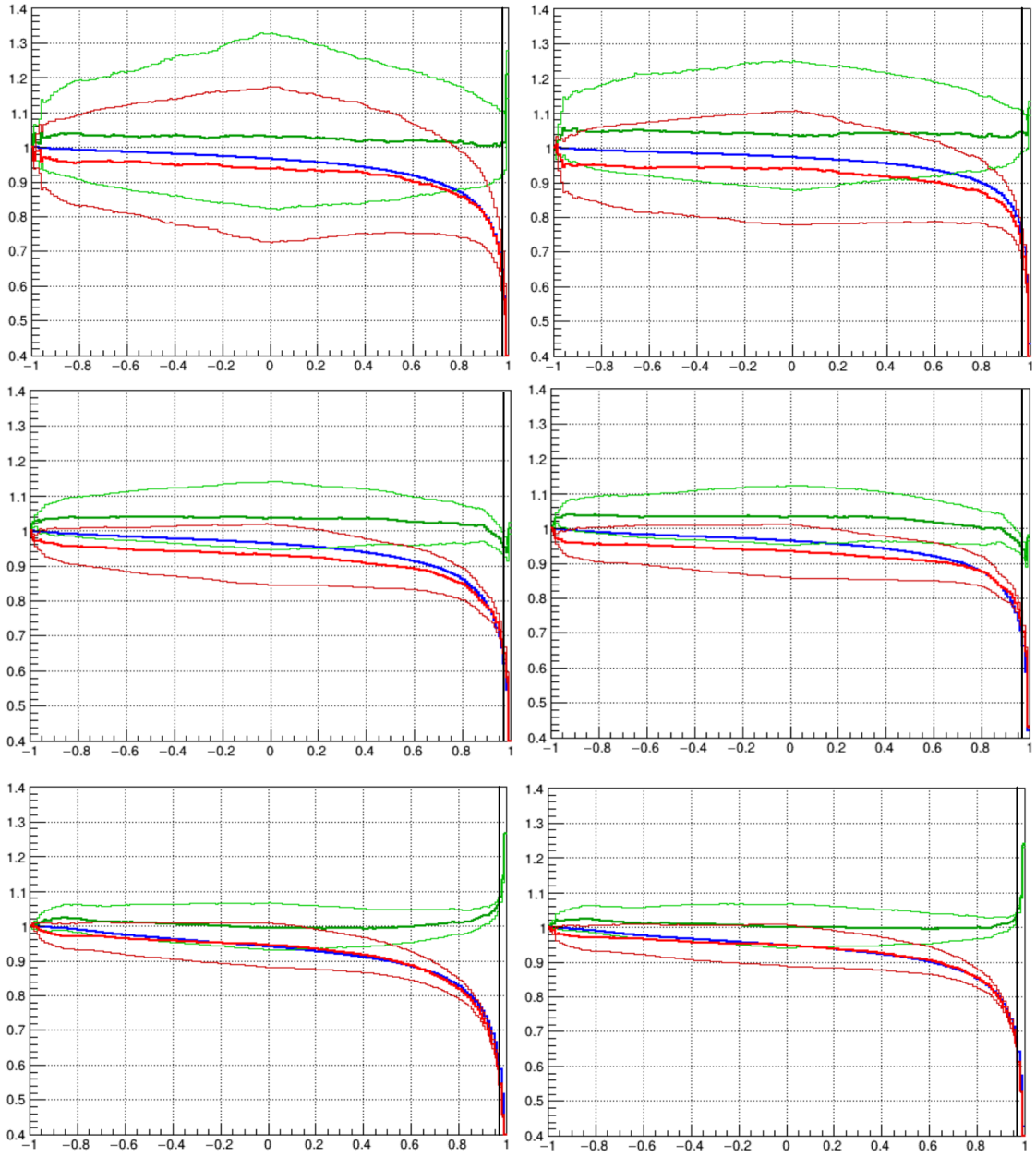

Figure 4.10: Comparison of BDT score CDFs for γ-ray events between simulations and data between 500 GeV and 1 TeV. The CDFs of simulations are in blue, while those of data in red. The ratio of simulation to data is in green. Typical BDT cut values are shown with thicker black lines. The thin lines show naive errors that track the statistical uncertainties of the underlying distributions. Soft cuts are on the left, medium cuts on the right, telescope epochs in chronological order from top to bottom. See main text for a full explanation and discussion.

Figures 4.9 and 4.10 compare the CDFs of BDT score between simulations and data. Here, the value of CDF at BDT score s is defined as the integral from s to 1 of the underlying probability distribution of BDT scores. Figure 4.9 shows this comparison for events below 500 GeV, while 4.10 shows it for events between 500 GeV and 1 TeV. Here, the blue curve is the CDF of BDT score for simulations, while the red curve is for data events. The data distributions were taken from collections of Crab runs – the Off region distribution was subtracted from the On region distribution, accounting for the exposure ratio α; this difference gives the distribution of BDT score of the γ-ray excess. The green curve is the ratio of the simulated CDF to the data-derived one. The thin lines represent naive statistical errors that, while not being a rigorous measure of error, provide insight into the statistics of the procedure: defining the "edges" of a histogram bin as being one RMSE away from the bin mean, the lower-bound CDF was computed by running the CDF integral along the bottom edge of the underlying distribution until the mid-point, and along the top edge after the midpoint; the upper-bound was obtained similarly, only first running the integral along the top edges and then the bottom ones.

It is apparent from these figures that there is a clear discrepancy between the simulated and true BDT CDFs at lower energies. This translates into a difference in cut efficiency between simulations and data: the relevant BDT cut values are at around 0.9 for the data in Figure 4.9 and around 0.95 for Figure 4.10, meaning that, for instance, for OA soft cuts analysis, signal extraction efficiency below 500 GeV (Figure 4.9, upper-left panel) is almost 20% higher for simulations than for data, and this, in turn, means that for this energy range, the

EAs are overestimated by almost 20%. This discrepancy is much less prominent for medium cuts, being at around 10%, and almost entirely vanishes above 500 GeV. Interestingly, the efficiency, and the EA, for UA data below 500 GeV is in fact *underestimated*. Of course, this does not tell us anything about the behavior of the standard analysis efficiency and how the two compare, but this difference in CDFs stems from differences between simulated and real events that are the result of simulation inaccuracies. Similar differences in CDFs exist between the underlying training shower parameters, such as the mean-scaled width or length, or the ITM goodness-of-fit.

With this, we have an explanation for the behavior of ITM and ITM+BDT spectra at low energies, and some idea of how to fix it. The underestimation of flux is the result of both high positive energy bias and overestimated signal efficiency. Both effects are more prominent for soft cuts, resulting in particularly poor performance for soft analyses. For UA data, the BDT cut efficiency is in fact underestimated, working against the energy bias and making the discrepancy there less obvious. However, this only masks the underlying issues, as there is no real reason why the two effects should cancel, or come close to canceling, each other.

Our proposed solution is to incorporate a criterion on the energy reconstruction bias into the safe energy determination procedure, that is, we define the safe energy range to be the range where both the energy reconstruction bias and the EA uncertainty are below predetermined thresholds. We stick with the original VEGAS threshold of 20% for EA uncertainty, and after a few empirical tests of reconstructed spectra, settle on a 15% threshold

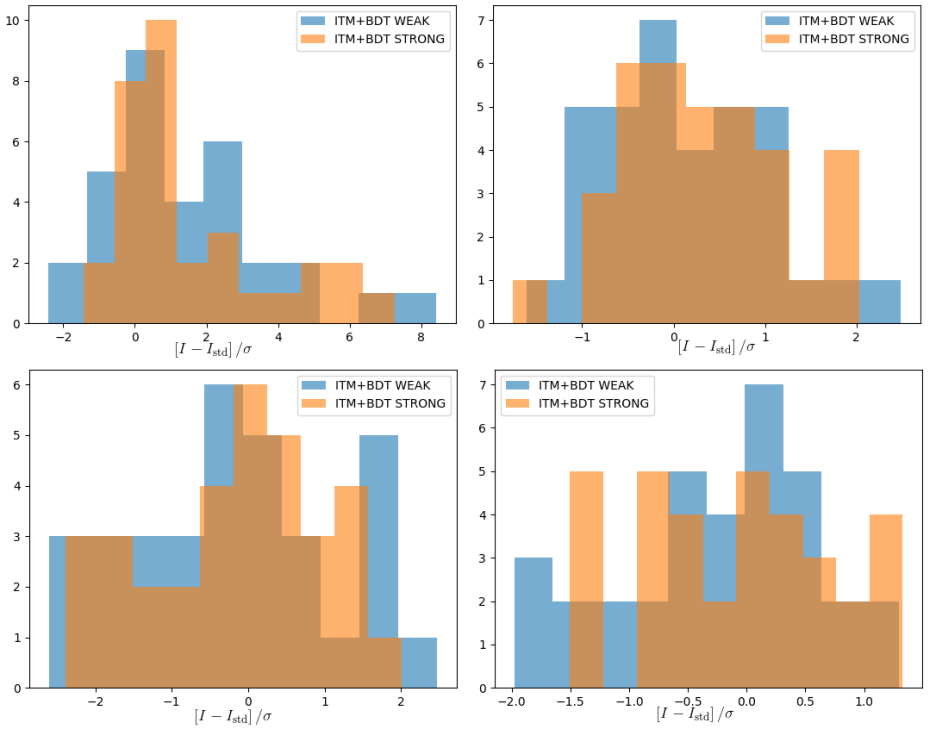

Figure 4.11: Reconstructed spectral indices for the validation sources. In orange (blue), the distribution of differences in indices between ITM+BDT strong (weak) and standard analyses, divided by the pooled standard error. Soft cuts are on the left, medium cuts on the right. Results before and after the energy threshold correction are at the top and bottom, respectively. Note the different x-axis ranges between the panels. See main text for a discussion.

for energy reconstruction bias. This value, however, needs to be investigated more rigorously. This new raised safe energy threshold discards both the energy bias ankle and lower-energy events where the disagreement between data and simulations is most prominent.

Figure 4.11 demonstrates the effect of applying this correction to the spectra of the validation sources. Here, we show the differences in the reconstructed spectral indices between ITM+BDT and the standard analysis, divided by the pooled standard errors: the histograms

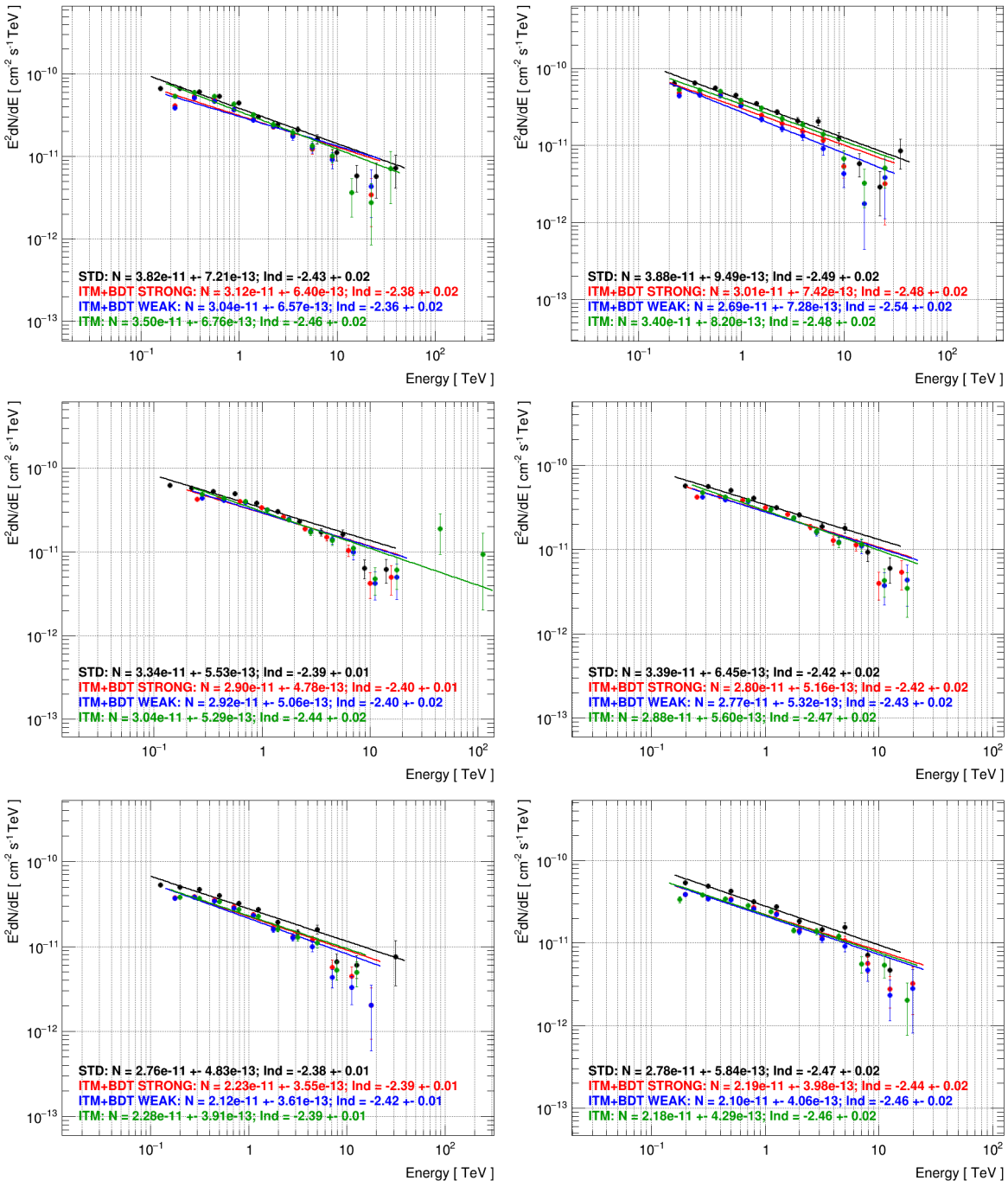

Figure 4.12: Reconstruction of energy spectra of the Crab Nebula, including a criterion on reconstructed energy bias in determining the energy threshold. ATM21 data are shown, with array epochs in chronological order from top to bottom. Standard analysis is in black, while the red, blue and green points correspond to the ITM+BDT *strong*, ITM+BDT *weak*, and ITM analyses. Soft cuts analysis is on the left, medium cuts on the right.

show the distributions of $(I_{\mathrm{ITMBDT}} - I_{\mathrm{std}})/\sigma$, where σ is the pooled statistical error from the estimation of the two indices. This measure, while not being exactly the same, is similar to a t-statistic and should be symmetric, fitting roughly into the interval $[-2, 2]$.[2] The top two histograms are the distributions before applying the threshold correction, while the bottom two are after the correction. The soft analysis before the correction displays a long right tail, meaning that many of the spectra are measured as harder than they should be, as is expected from underestimating high-leverage low-energy points. The threshold correction mostly fixes this, but the resulting distribution is still wider than we'd expect. The medium cuts spectra look decent even before the correction, and the correction further trims the right tail of the distribution, just as is desired. Figure 4.12 shows the same Crab spectra as in Figure 4.6, only after the energy threshold correction. The agreement in indices for medium cuts is very good, while soft cuts still display troubling low-energy behavior.

The difference in flux normalizations is another issue that we've avoided talking about until now. Figure 4.13 shows the differences in the reconstructed spectral normalizations of the validation sources between ITM+BDT and the standard analysis, after the threshold correction for ITM+BDT. Here, we show the difference in the flux normalizations as a fraction of the standard analysis normalization: $(N_{\mathrm{ITMBDT}} - N_{\mathrm{std}})/N_{\mathrm{std}}$. In the absence of

[2] An intuitive justification of this may be given as follows. With ~ 40 estimates in the histogram, we expect ~ 2 of them outside of the 95% central confidence interval of the distribution from which the slope estimates are sampled. With the linear regression estimates performed on 5-6 points, this sampling distribution follows Student's t-distribution with 3-4 degrees of freedom, the t-multiplier of which is around 3. Clustering and serial correlation effects will make the true distribution somewhat narrower, and while estimating the true value of this correction factor is intractable within the scope of this work, taking the lower bound of 2 for the t-factor, which corresponds to a normal distribution, is reasonable.

systematic differences, this should be a narrow symmetric distribution. However, both soft and medium cuts distributions are shifted to the left, with longer than expected left tails. The shifts to the left by about 0.2 imply that ITM+BDT gives fluxes of around 20% lower than the standard analysis. This value is within the systematic uncertainty of VERITAS flux normalization, but it is unclear whether the effect is the result of EA overestimation or energy underestimation. That is, is the effect the result of ITM+BDT spectra being shifted to the left rather than downwards?

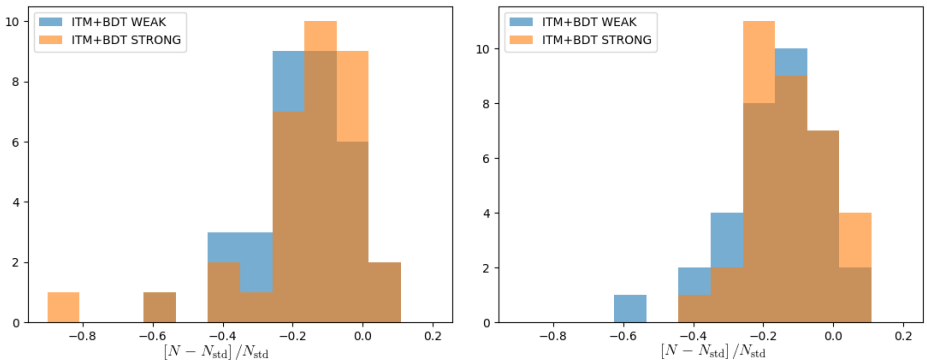

Figure 4.13: Reconstructed spectral normalizations for the validation sources, after the energy threshold correction. In orange (blue), the distribution of differences in normalizations between ITM+BDT strong (weak) and standard analyses, divided by normalization of the standard analysis. Soft cuts are on the left, medium cuts on the right. See main text for a discussion.

To check this possibility, we look at the curved spectra of Mrk 421 and fit them to a power law with an exponential cutoff (Figure 4.14) – disagreement in cutoff energies for the same index and normalization then would be a clear indicator of problems with energy reconstruction, or, as it were, a horizontal shift between the spectra. However, we can draw no such conclusion from the case of Mrk 421: while the agreement between the UA spectra

is excellent, NA and OA spectra disagree on both cutoffs and indices, meaning that no simple shift in either direction is responsible for the observable differences. Extra care must therefore be taken when studying curved spectra with ITM+BDT.

4.3 Summary and Outlook

The ITM+BDT analysis provides a remarkable improvement in sensitivity, both for soft and medium size cuts analyses. The large gain at soft cuts, coupled with a lowered energy threshold, is particularly useful in the context of multimessenger astronomy and hunting for transient sources.

However, the method is not without issues. Its out-of-the-box spectral reconstruction showed significant problems that surfaced a number of systematic issues. Particularly, the energy reconstruction of ITM has a high bias at low energies and displays a sloping trend across the whole energy range; while the differences between data and simulated showers plague BDT event selection efficiency. The two effects together dramatically affect the low-energy spectral reconstruction.

In this work, both issues were corrected for by a means of a correction introduced to the selection of the safe energy range of spectral reconstruction. For medium cuts, the spectral indices of the reconstructed spectra with medium cuts agree within the statistical errors, while the generally lower flux normalization of ITM+BDT is within VERITAS systematics. Results for upgraded-array data generally agree very well between the different analyses,

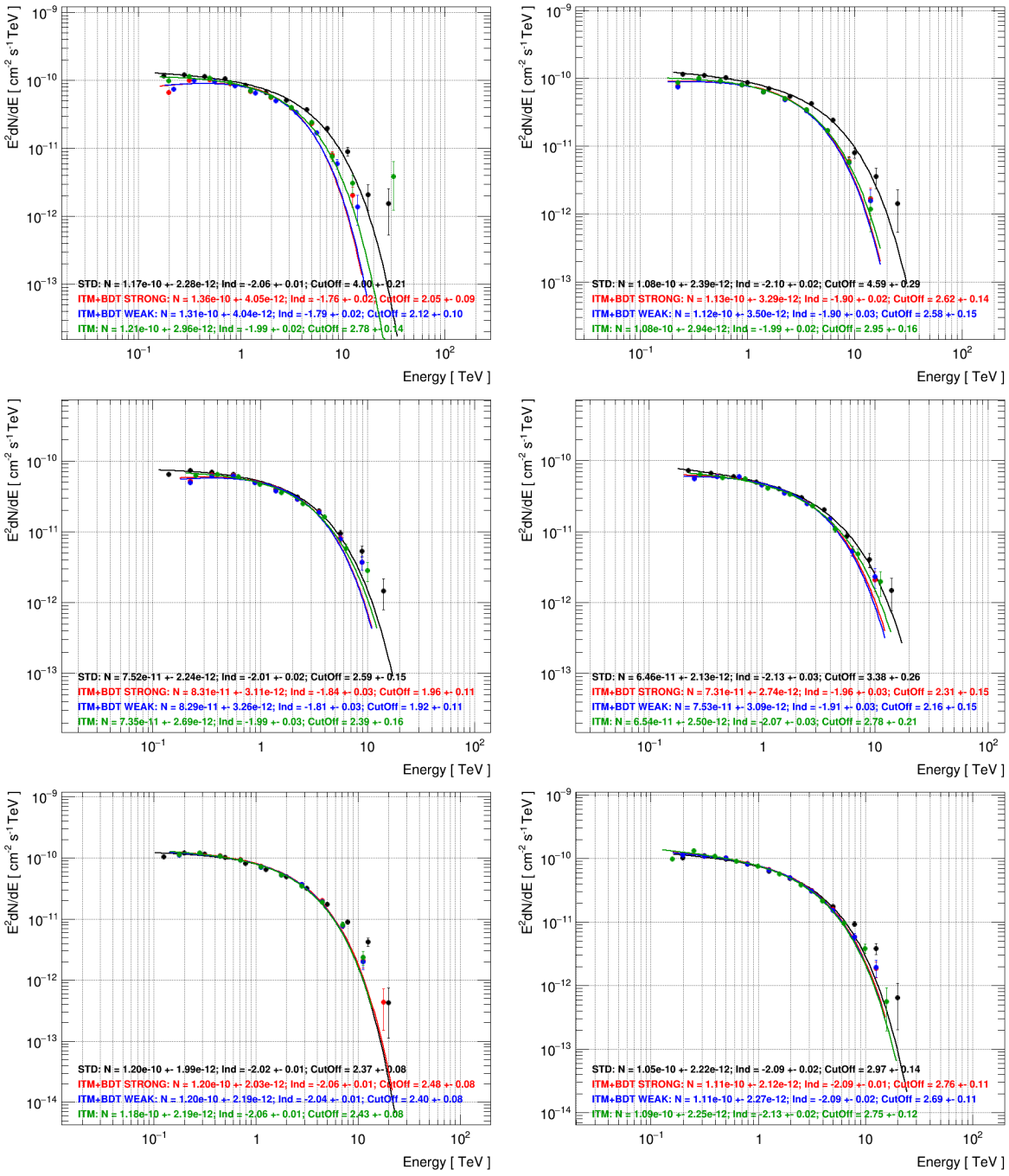

Figure 4.14: Reconstruction of energy spectra of Mrk 421, including a criterion on reconstructed energy bias in determining the energy threshold. ATM21 data are shown, with array epochs in chronological order from top to bottom. Standard analysis is in black, while the red, blue and green points correspond to the ITM+BDT *strong*, ITM+BDT *weak*, and ITM analyses. Soft cuts analysis is on the left, medium cuts on the right.

possibly the result of a better detector simulation for that epoch. For soft cuts, flux underestimation at lower energies persists even with the raised energy threshold, the result of data-simulation discrepancies that lead to the overestimation of signal efficiency. ITM+BDT soft cuts analyses therefore should not be used for spectral reconstruction.

Despite, or rather becasue of, the spectral reconstruction shortfalls of soft cuts analysis, ITM+BDT should be studied further. In particular, its energy reconstruction bias should be understood and corrected for, possibly through the use of intermediary lookup tables. The need for the production of separate sets of ITM templates for each epoch-atmosphere combination should be investigated. A better understanding of the BDT model is also desired, particularly the noticeable low-energy difference in BDT score distributions between data and simulations. In this work, 3-telescope events with no image size cut were used for BDT training, and models that include 2-telescope events while imposing an image size cut should be investigated in the future. The agreement between data and simulations should be taken into account during the selection of training parameters, possibly excluding attributes for which true and simulated distributions disagree significantly. Finally, a rigorous study of the energy threshold and spectral reconstruction near it needs to be performed.

Chapter 5

M82 as seen with VERITAS

We are now in a position to apply the ITM+BDT method of Chapters 3 and 4 to the analysis of TeV γ-ray emission from an individual source. The high sensitivity of ITM+BDT allows us to study TeV emissions from sources substantially weaker than seen comfortably with the standard analysis. The starburst galaxy M82, introduced in Section 1.2, is just such a source. While VERITAS detected it a decade ago (VERITAS Collaboration et al. 2009), the measured spectrum was poorly constrained. Moreover, the detection at the 5σ level was right at the visibility threshold of VERITAS, gaining little benefit from subsequent observations.

In this chapter, we use the ITM+BDT analysis method to study M82 with VERITAS. We first give an overview of the data used in the study and specifics of the analysis. We then present the measured M82 VHE γ-ray spectrum between 250 GeV and 8 TeV, compare it to other measurements, and discuss its implications. A theoretical modeling treatment of the measurement is beyond the scope of this work.

5.1 Data Selection and Analysis

The original TeV detection of M82 was based on 137 hours of observations with the OA configuration of the VERITAS array. Subsequently, another ∼50 hours of NA and ∼30 hours of UA observations were added to this dataset between 2009 and 2014. The higher sensitivity of the NA and UA configurations, together with the lowered energy threshold of the UA epoch, were expected to help obtain a more accurate spectral measurement, extending it to both higher and lower energies. Incorporations of these additional data, however, were difficult to leverage due to the weakness of the source and the relatively high zenith angle of observations of 40°. Subsequent analyses required to be optimized specifically for the source, with the use of very stringent image size selection cuts discarding the available lower-energy data (Ratliff 2015).

When the ITM+BDT analysis started showing promise, we decided to revisit M82, analyzing the data with medium size cuts and the ITM+BDT cuts optimized for weak sources. The preliminary analysis resulted in a 6σ detection, with a spectral measurement extending from ∼400 GeV to ∼8 TeV. This extension of the spectrum renewed our interest in the source, and prompted an observational campaign in the 2018-2019 season, adding another ∼30 hours of UA data to the already massive dataset.

Here, we analyze all of the currently available VERITAS M82 data, totalling 251 hours: the original OA dataset, the subsequent NA and UA observations between 2009 and 2014, and the latest UA observations from the 2018-2019 season. Table 5.1 summarizes the dataset

	Exposure, min	N_{Excess}	N_{On}	N_{Off}	α	γ rate, min^{-1}	BG rate	σ
OA	8847	76	434	13061	0.02738	0.0085	0.0404	3.9
NA	3056	55	223	6011	0.02788	0.0180	0.0548	4.0
UA	3145	46	178	4752	0.02786	0.0146	0.0421	3.7
Total	15048	177	835	23824	0.02756	0.0118	0.0437	6.5

Table 5.1: Summary of the M82 data and analysis results, broken down by array epoch. The exposure is the observation live time corrected for the array deadtime. The background estimation follows the Ring Background Method (see Section 2.4.3). The last column is the significance computed from Equation 2.9. See main text for more details.

and the analysis results. The observations were performed with M82 between 45° and 53° in elevation, resulting in an average zenith angle of $\sim 40°$. Only data runs with all 4 telescopes present in the array were analyzed: while the original analysis of VERITAS Collaboration et al. 2009 included observations with only 3 telescopes present in the array, the systematic discrepancies between the standard and ITM+BDT analyses (Section 4.2) deterred us from using these data until studies of ITM+BDT efficiency on 3-telescope data are conducted.

We use the medium size quality selection cuts (Table 3.4) with ITM event reconstruction and BDT event selection. Because of the high zenith angle of the data, we require that each event be reconstructed with at least 3 telescopes. We apply the $\theta^2 = 0.006$ deg^2 directional cut discusesd in Section 4.1.1 and the BDT cuts optimized for weak sources (see Appendix).

5.2 Results

Table 5.1 summarizes the measured excess of TeV γ-ray events from M82 and its significance. The presented analysis sees a total of 177 excess events, with an overall detection significance

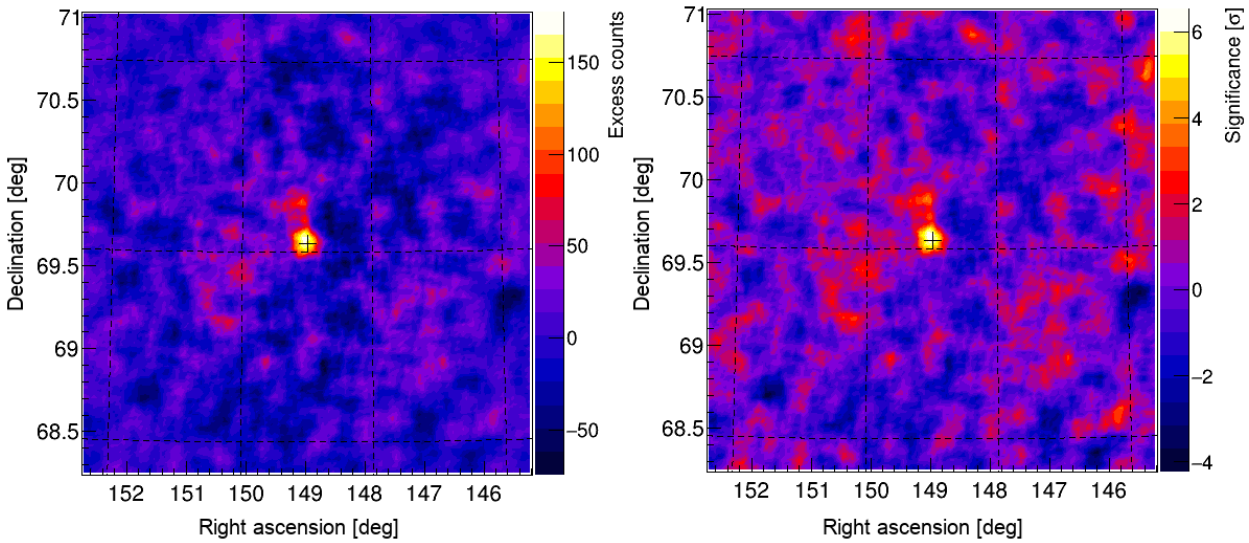

Figure 5.1: VHE skymaps of the M82 region. The source location is marked with a cross. The Ring Background Method (see Section 2.4.3) is used for background estimation. *Left:* The measured excess (color scale) of γ-like events above the estimated background for a region centered on M82. *Right:* The significance of the measured excess (color scale) as computed from Equation 2.9.

of 6.5σ, when estimating the background with the Ring Background Method (Section 2.4.3). When the background estimate is performed with the Reflected Regions Method, we obtain a slightly lower 160 excess events and a significance of 6σ.

The differences seen in raw γ and background event rates between array epochs are expected. Raw γ and background rates increase when moving to the NA epoch due to its larger overall effective area. The subsequent slight decrease of both rates when moving to UA is the effect of tighter event selection cuts at that epoch that are needed to suppress the higher raw background rates associated with its more sensitive PMTs. Figure 5.1 shows the TeV skymaps of the M82 region. Both excess counts and their significance are shown.

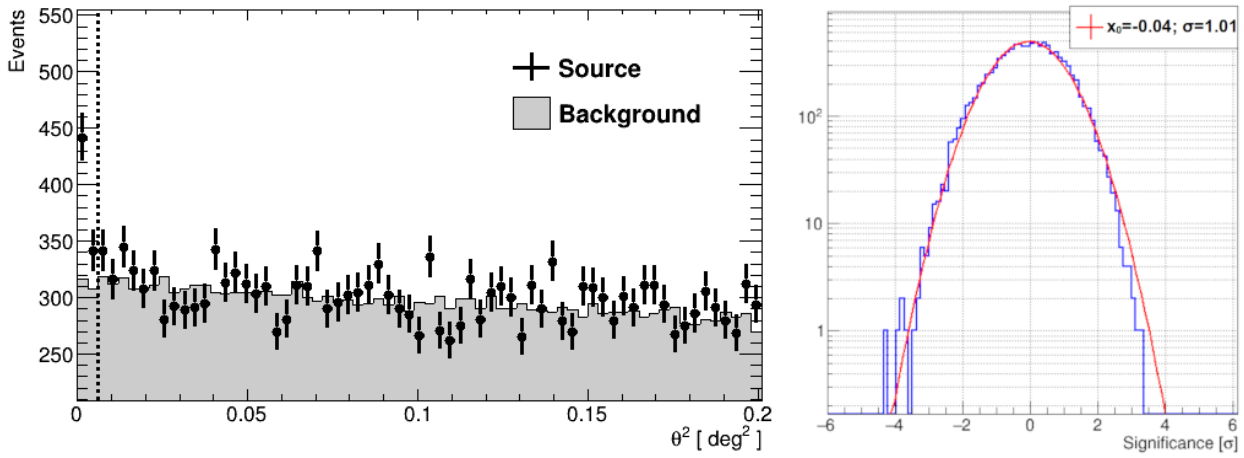

Figure 5.2: *Left:* the distribution of γ-like events around the source and background regions as a function of θ^2. The background regions are defined as in the Reflected Regions Method. The dashed vertical line marks the directional cut that defines the extent of the source region; we expect γ-like events around the source to the right of this line to follow the same distribution as the events around the background regions. *Right:* the significance distribution of the measured excess, with the source region excluded. We expect the excess outside of the source region to be due to statistical noise, and therefore this distribution is expected to be normal.

The slight trail of excess events above the starburst is not statistically significant, and the emission is consistent with coming from a point source. We further check these points by looking at the θ^2 distribution of γ-like events and the significance distribution of the excess. In the θ^2 distribution (Figure 5.2, *left*), background regions are defined as in the Reflected Regions Method (Section 2.4.3), and the dashed vertical line marks the directional cut that defines the extent of the source region. As we expect, γ-like events around the source to the right of this line follow the same distribution as the events around the background regions, with no indication of source events spilling outside of the θ^2 cut, and no differences in the source and background acceptances. With the source region excluded, the measured excess

is expected to come purely from statistical noise, and the distribution of its significance is therefore expected to be normal. This is exactly what we see (Figure 5.2, *right*), with no spurious features in the background excess.

E, GeV	$E_{\rm Lo}$	$E_{\rm Hi}$	Φ, cm^{-2}GeV^{-1}s^{-1}	$\Delta\Phi$	$N_{\rm Excess}$	$N_{\rm On}$	$N_{\rm Off}$	α	σ
334	251	447	2.03e-15	2.15e-15	11.69	124	1760	0.06381	1.05
594	447	794	3.68e-16	2.06e-16	37.53	282	3831	0.06381	2.27
1057	794	1413	1.35e-16	3.56e-17	51.31	158	1672	0.06381	4.47
1879	1413	2512	2.39e-17	8.86e-18	24.48	74	776	0.06381	3.13
3341	2512	4467	1.08e-17	3.13e-18	23.92	46	346	0.06381	4.26
5941	4467	7943	2.69e-18	1.14e-18	10.81	20	144	0.06381	2.96
10570	7943	14125	$<$ 7.74e-19	-1	-1.06	2	48	0.06381	-0.63

Table 5.2: The VHE γ-ray spectrum of M82 as measured with VERITAS. The energy is in GeV and the flux and its error are in events per cm^2 per GeV per second. The last point is an upper limit at the 99% confidence level. The background estimation follows the Reflected Regions Method (see Section 2.4.3). The last column is the significance computed from Equation 2.9.

Table 5.2 and Figure 5.3 summarize our measurement of the VHE spectrum of M82. Only events that are above the safe energy threshold are included into spectral reconstruction, and the background estimates come from the Reflected Regions Method. The spectrum extends from 250 GeV to almost 8 TeV. The lowest-energy spectral point is significant only at the 1σ level, but we nevertheless include it in the spectral fit. The significance of spectral points is the significance of the γ excess in the corresponding energy bin, while the error on the flux incorporates the individual effective area of each event and its uncertainty. This is the reason for the large uncertainty on the flux of the lowest-energy spectral point. We also check that the measured fluxes and spectral indices are consistent between the three array

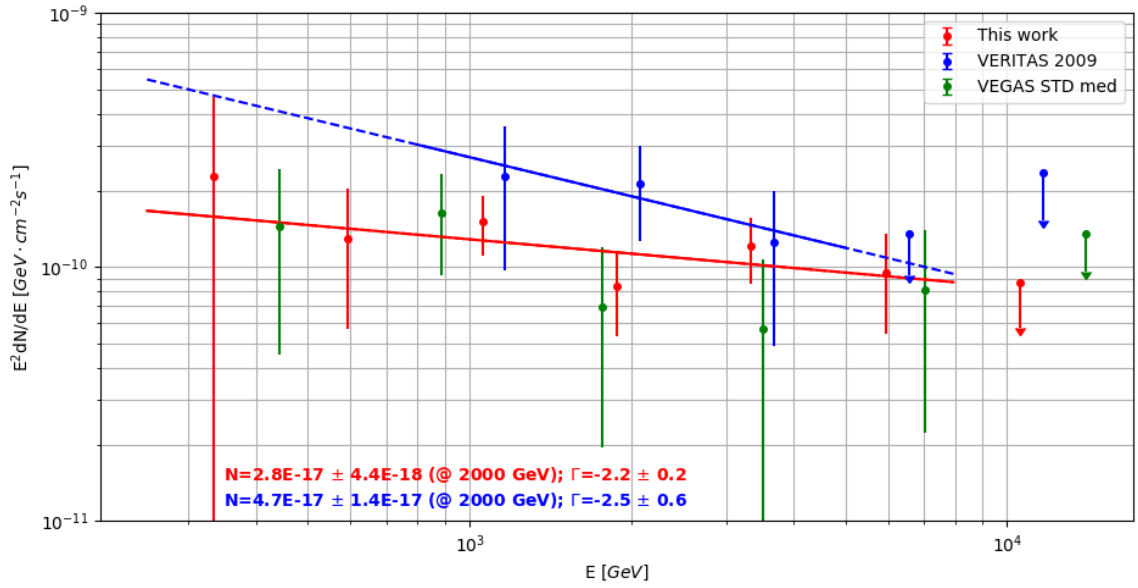

Figure 5.3: The VHE γ-ray spectrum of M82 as measured with VERITAS. The measurement performed in this work is shown with red points, the solid red line is a least-squares power-law fit to it. The VERITAS M82 spectrum and the fit to it from VERITAS Collaboration et al. 2009 are shown with blue points and the solid blue line, while the dashed blue line extends the fit to the energy range covered by the spectrum in red. The green points correspond to the spectrum as measured with the standard VEGAS analysis with medium size cuts; notice the wider energy binning. The flux normalizations are given in $\mathrm{cm}^{-2}\mathrm{GeV}^{-1}\mathrm{s}^{-1}$. See main text for details.

epochs, with no significant instrumental effects. Figure 5.4 shows the integral fluxes above 560 GeV, as well as the spectral indices, for data separated by array epochs, and for the entire dataset. Due to the different energy thresholds of different epochs, the energy of 560 GeV for integral flux comparison was chosen as the energy of the first common significant spectral bin. All three epochs agree with each other and the combined measurement, albeit, due to the low signficances, the errors on the measurements of individual epochs are quite large.

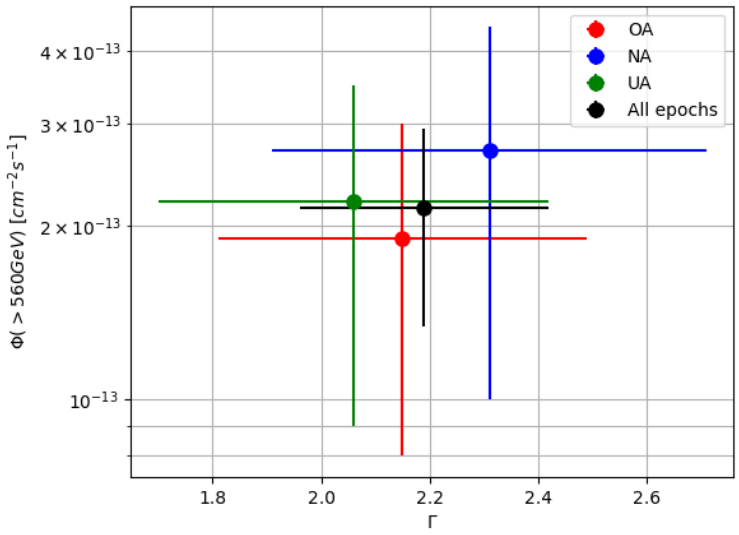

Figure 5.4: The measured M82 fluxes and spectral indices for different array epochs. The integral flux above 560 GeV is on the vertical axis, and the spectral index is on the horizontal axis. The red, green, and blue points correspond to the OA, NA, and UA epochs. The black point is the measurement from the entire dataset.

A simple least-squares fit of the spectrum to a power law results in a spectral index $\Gamma = 2.19 \pm 0.23$ and a normalization factor at 2000 GeV of $N_0 = (2.82 \pm 0.44) \cdot 10^{-17}\,\mathrm{cm}^{-2}\mathrm{GeV}^{-1}\mathrm{s}^{-1}$. The reduced χ^2 of the fit is equal to $\chi^2_{\mathrm{red}} = 0.41$, showing clear signs of overfitting and a low discriminating power of the data points, as expected from the large uncertainties. The total integral flux above 250 GeV is estimated to be $\Phi(> 250\ \mathrm{GeV}) = (5.6 \pm 2.1) \cdot 10^{-13}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, making M82 the weakest TeV γ-ray source seen by VERITAS. The differential upper limit shown between 8 and 14 TeV is at the 99% confidence level and lies right above the best-fit line to the measured spectrum.

While the spectral index agrees with the measurement in VERITAS Collaboration et al. 2009 (shown in Figure 5.3 with blue points, and its fit to a power law with a blue line),

the normalization and the total flux measured here are lower than previously reported. The integral flux above 700 GeV calculated from our measurement, $\Phi(> 700 \text{ GeV}) = (1.6 \pm 0.6) \cdot 10^{-13}\,\text{cm}^{-2}\text{s}^{-1}$, is less than half of the value reported in VERITAS Collaboration et al. 2009, $(3.7 \pm 1.1) \cdot 10^{-13}\,\text{cm}^{-2}\text{s}^{-1}$. This difference, however, while significant at the 1σ level, is driven by the difference of $(1.9 \pm 1.5) \cdot 10^{-17}\,\text{cm}^{-2}\text{GeV}^{-1}\text{s}^{-1}$ in the differential flux measurements at 2 TeV. The discrepancy is only slightly larger than one standard deviation, and becomes $(1.2 \pm 1.5) \cdot 10^{-17}\,\text{cm}^{-2}\text{GeV}^{-1}\text{s}^{-1}$, within statistical errors, after a correction for a possible systematic difference of 20% in flux normalizations between ITM+BDT and the standard analysis (Section 4.2).

When using the Reflected Regions Method for background estimation, the standard analysis with medium size cuts does not produce a significant excess at the location of M82 ($\sigma = 3.7$), while the Ring Background Method detects the starburst with a significance of 5.4σ. We can nevertheless extract a spectrum from this analysis if we bin the resulting excess in wider energy bins. This spectrum, shown for comparison in Figure 5.3 with green points, agrees well with the ITM+BDT measurement.

5.3 Discussion

The presented measurement is the most precise measurement of the TeV γ-ray flux of M82 to date. And yet, being a featureless power-law with large uncertainties on the parameters, it does not on its own offer much physical insight into the physical processes responsible for

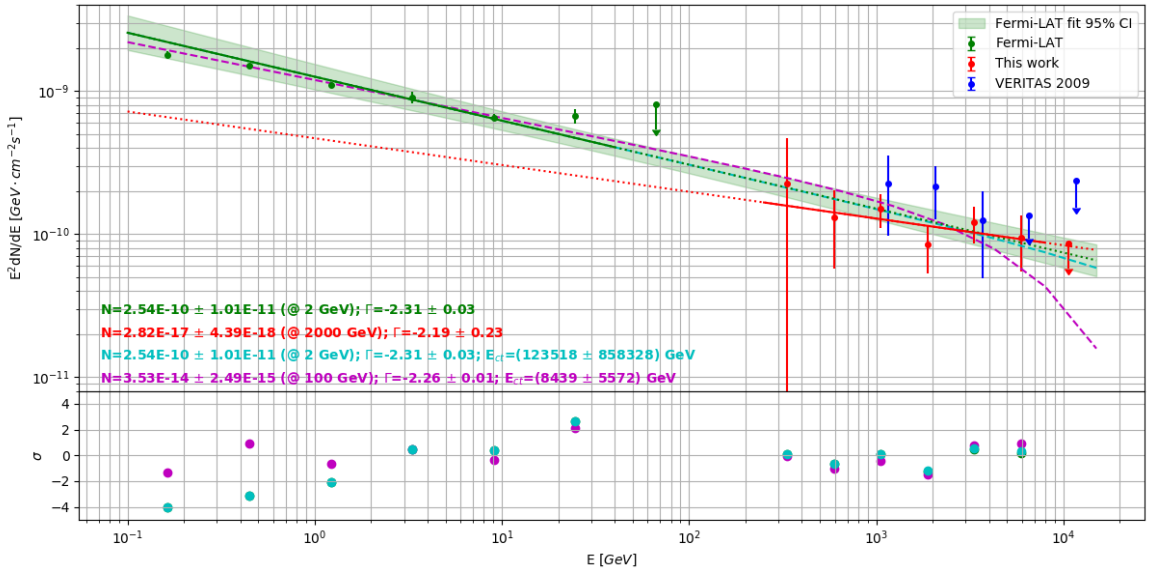

Figure 5.5: Fitting the M82 γ-ray spectrum. The measurement performed in this work is shown with red points, the solid red line is a least-squares power-law fit to it, and the dashed red line extends this fit to *Fermi*-LAT energies. Pass-8 *Fermi*-LAT analysis, from a private correspondence with Marcos Santander, is shown in green, the solid green line is a power-law binned likelihood model fit to the *Fermi*-LAT events, and the dashed green line extends this fit to VERITAS energies. The VERITAS M82 spectrum from VERITAS Collaboration et al. 2009 is shown with blue points. The dashed magenta line is a least-squares fit of a power law with an exponential cutoff to the green and red points. The dashed cyan line is a least-squares fit of a power law with an exponential cutoff to the green and red points where the index and normalization are fixed to the Fermi fit parameters, and only the cutoff energy is varied. The flux normalizations are given in $\mathrm{cm}^{-2}\mathrm{GeV}^{-1}\mathrm{s}^{-1}$. See main text for a discussion.

the detected emission. To gain such insight, the measurement needs to be incorporated into broadband fits, a task beyond the scope of this work. Nevertheless, we try to find some clues about the physics of γ-ray production in M82, by turning to the *Fermi*-LAT.

In the following, we include an analysis of the Pass-8 *Fermi*-LAT data of M82 up to October 2018. The results of this standard analysis, provided by Marcos Santander, agree

with the results published in Ackermann et al. 2012 and give much more accurate estimates of the spectral parameters. Compared with the published Pass-7 results, this measurement includes 7 more years of data, and uses the LAT 4-year Point Source Catalog (3FGL) and an updated Galactic diffuse model, as well as improved instrument model and response functions, resulting in better event reconstruction at both low and high energies. A binned likelihood fit of a power law to the data gives a spectral index $\Gamma = 2.31 \pm 0.03$ and a normalization factor at 1.0152 GeV of $N_0 = (1.21 \pm 0.05) \cdot 10^{-9}\,\mathrm{cm}^{-2}\mathrm{GeV}^{-1}\mathrm{s}^{-1}$. The resulting spectrum is shown in Figure 5.5 with green points, and the spectral fit with a solid green line. The dashed green line extends this fit to VERITAS energies. The same figure shows the VERITAS spectrum measured here with red points, the fit to it with a solid red line and the extension of the fit to *Fermi*-LAT energies with a dashed red line. Finally, the published VERITAS spectrum from VERITAS Collaboration et al. 2009 is shown with blue points.

First, we note that the extension of the *Fermi*-LAT fit to VERITAS energies fits our measurement very well, with the resulting goodness-of-fit $\chi^2_{\mathrm{red}} = 0.50$. Of course, as we noted earlier, this also simply means the VERITAS measurement isn't all that restrictive, but the fit to the published VERITAS data is actually quite poor, despite the spectrum in this case having only 3 points and large uncertainties: there, $\chi^2_{\mathrm{red}} = 2.38$. Furthermore, an extension of the *Fermi*-LAT fit to the VERITAS energy range gives a flux of $(7.4 \pm 0.4) \cdot 10^{-13}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ above 250 GeV and $(1.9 \pm 0.1) \cdot 10^{-13}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ above 700 GeV, agreeing well with our measurement. Finally, the upper limit at 10 TeV is consistent with the extrapolated *Fermi*-LAT spectrum.

In order to investigate a possible steepening of the γ-ray spectrum between the *Fermi*-LAT and VERITAS energies, or a cutoff at TeV energies, we fit two power laws with exponential cutoffs. The first one, shown with a dashed cyan line in Figure 5.5, fixes the spectral index and normalization obtained from the *Fermi*-LAT analysis and only varies the cutoff energy. The second one, shown with a dashed magenta line, fits all three parameters to both the *Fermi*-LAT and VERITAS points; this fit uses a simple least-squares fit to the binned *Fermi*-LAT fluxes and therefore does not utilize the full power of a model likelihood fit to the GeV data. A non-zero cutoff found by either of these procedures would suggest a possible break in the spectrum driven by the VERITAS data. However, neither fit finds any such evidence: the first one does not find a statistically significant cutoff parameter; the second one, while finding a cutoff at 8.4 ± 5.6 TeV, does so with a $\chi^2_{\rm red} = 1.40$, only marginally better than an equivalent simple power law fitted to both the *Fermi*-LAT and VERITAS points, whose $\chi^2_{\rm red} = 1.5$. It is worth mentioning here that the *Fermi*-LAT and VERITAS measurements have very different instrumental uncertainties, with the 20-30% systematic uncertainty of VERITAS dominating the $\sim$ 5% uncertainty of the *Fermi*-LAT. However, with no indication of a cutoff, we do not explore the instrumental differences further.

As already mentioned, the full theoretical treatment of our measurement is beyond the scope of this work. We can, however, try to investigate how much the combined *Fermi*-LAT and VERITAS spectra constrain radiative models, or at least compare them with past model predictions. In Figures 5.6 and 5.7, we attempt to fit naive radiative models to the

observed spectra. For this purpose, we use *Naima*,[1] a Python package for computation of non-thermal radiation from relativistic particle populations. The code is capable of fitting radiative models to X-ray, GeV and TeV spectra through MCMC sampling, and computing the parameters of the underlying particle populations; the environmental parameters of the radiative models, however, need to be fixed. For our exploratory purposes, we chose the following parameters: the distance to M82, $L = 3.9$ Mpc (Sakai and Madore 1999); the density of the ISM used as target material for hadronic interactions and bremsstrahlung, $n_{\mathrm{ISM}} = 180\,\mathrm{cm}^{-3}$ (de Cea del Pozo, Torres, and Rodriguez Marrero 2009); the energy density and temperature of the thermal photon field used as target for inverse Compton interactions, in addition to the CMB, $U_{\mathrm{rad}} = 10^3\,\mathrm{eV\,cm}^{-3}$ and $T = 8000$ K, respectively (Yoast-Hull et al. 2013). For the cosmic ray spectra, we chose power laws with exponential cutoffs at 100 TeV, mirroring the models in de Cea del Pozo, Torres, and Rodriguez Marrero 2009, but leaving the indices and normalizations for both proton and electron populations free and independent. Our fits, thus, place no restrictions on the total electron energy, a vital piece of information that could be gained from radio data.

Figure 5.6, left, shows a simple, purely hadronic model. It is apparent that, while pion decay alone can fit the TeV spectrum reasonably well, it deviates significantly from the GeV flux points. This is expected, and necessitates the inclusion of electrons into the radiative model. The right panel of Figure 5.6 shows the fit of such a model to our observations: here, pion decay, bremsstrahlung and inverse Compton emission all contribute to the γ-ray flux. As

[1] `https://naima.readthedocs.io/en/latest/index.html`

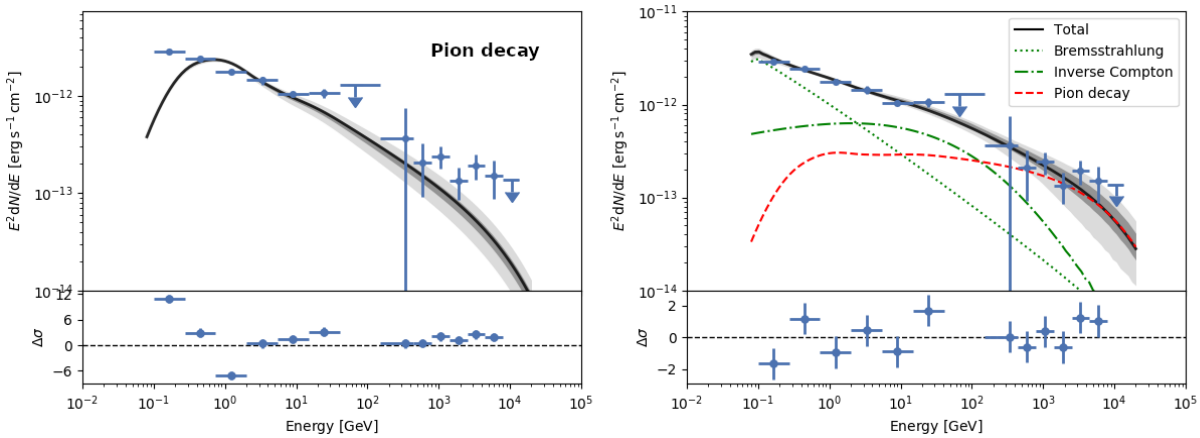

Figure 5.6: Fitting the observed γ-ray spectrum of M82 with naive radiative models. *Left:* only pion decay is included as the source of γ-rays. *Right:* pion decay, bremsstrahlung and inverse Compton emission all contribute to the γ-ray flux. See main text for details and a discussion.

expected from a model with completely free normalizations and indices, with no restrictions on the total electron energy, it fits very well. In fact, this omission of normalization of total electron energy makes it possible to fit the spectrum quite well with a purely leptonic model (Figure 5.7). The fit, however, gives a reasonable estimate of the total energy density of high-energy protons: the fitted proton spectral index is $\Gamma_p = 2.15 \pm 0.08$, and the total energy of protons above 10 MeV is $E_p = (1.7 \pm 1.2) \cdot 10^{65}$ eV, which, assuming a cylindrical starburst region of 200 Mpc in radius and 100 Mpc in height (Yoast-Hull et al. 2013), results in the proton energy density of $(425 \pm 300)\,\mathrm{eV\,cm^{-3}}$, agreeing reasonably well with the estimates of $\sim 250\,\mathrm{eV\,cm^{-3}}$ in Persic and Rephaeli 2014 and VERITAS Collaboration et al. 2009. The large uncertainty on this estimate is due to the lack of restrictions on the total electron energy, which in our fit ended up having a rather unreasonably large density of $\sim 600\,\mathrm{eV\,cm^{-3}}$ above 10 MeV.

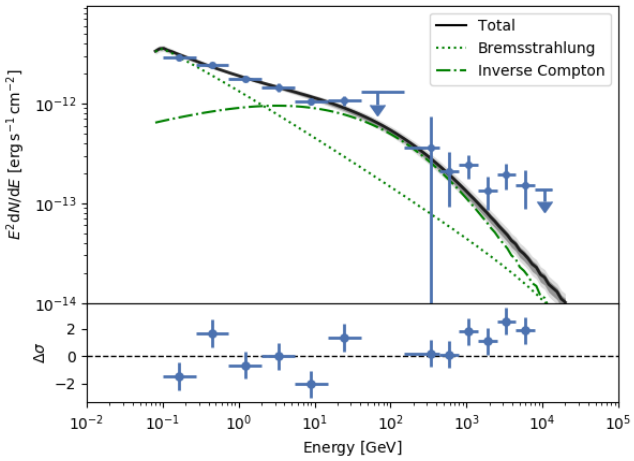

Figure 5.7: Fitting the observed γ-ray spectrum of M82 with a purely leptonic radiative model. Only bremsstrahlung and inverse Compton emission contribute to the γ-ray flux. See main text for details and a discussion.

Finally, Figure 5.8 compares the models of de Cea del Pozo, Torres, and Rodriguez Marrero 2009 (the yellow band, exploring a range of different supernova explosion rates and proton injection indices) and Yoast-Hull et al. 2013 (the dashed magenta line; their best-fit model to the radio emission is shown) with our measurement. The published VERITAS spectrum is also shown. Both models rely on the available radio data. However, Yoast-Hull et al. 2013 do not account for the inverse Compton emission from secondary electrons, and estimate that including this contribution will increase the total γ-ray flux by a factor of ~ 3; unlike the model of de Cea del Pozo, Torres, and Rodriguez Marrero 2009, which assumes a cutoff in the injection spectra at 100 TeV, their model makes no such assumptions, with injection spectra extending to the knee.

Overall, the γ-ray spectrum of M82 presented in this work is fitted well with a single power law between the energies of 100 Gev and 10 TeV. The measurement is in a good

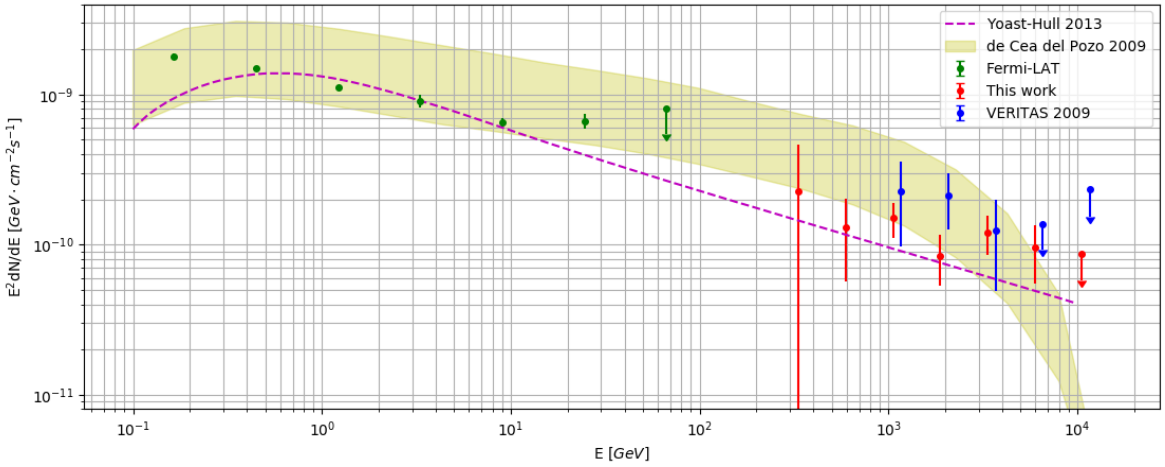

Figure 5.8: Comparison of the masured γ-ray spectrum of M82 with theoretical models. The yellow band is the model from de Cea del Pozo, Torres, and Rodriguez Marrero 2009, reflecting the uncertainties in the supernova rate and the maximum cosmic-ray energy. The dashed magenta curve corresponds to the model in Yoast-Hull et al. 2013 that best fits the observed radio spectrum; inverse Compton emission from secondary electrons is not included in this model. See main text for a discussion.

agreement with past theoretical models that fit the synchrotron radio emission. The GeV emission is dominated by bremsstrahlung, while both pion decay and inverse Compton interactions contribute to the TeV emission. With no apparent spectral break or cutoff at TeV energies, the measurement suggests that energy-independent transport and losses dominate in M82, and although on its own it does not establish whether the TeV emission is dominated by hadronic interactions, it is consistent with the picture that M82 is a good electron calorimeter, but not a good proton calorimeter (Lacki et al. 2011; Yoast-Hull et al. 2013; Rojas-Bravo and Araya 2016). A broadband fit with this new more restrictive measurement needs to be performed in order to better constrain the physical parameters of the starburst region.

Chapter 6

The Prototype Schwarzschild-Couder Medium-Sized Telescope for the CTA

We introduced the Cherenkov Telescope Array in Chapter 2, mentioning that much of this work was dedicated to the construction of the prototype Schwarzschild-Couder medium-sized telescope (pSCT), specifically to the development and implementation of its optical alignment system. The majority of the work consisted of the unglamorous business of writing code, building and installing hardware, and performing various measurements for the purposes of calibration and characterization of the optical system of the telescope. Here, we give an overview of the pSCT design, describe the pSCT optical alignment system and review the implementation of its software. Nieto et al. 2017, Nieto et al. 2015 and Rousselle et al. 2015 include more technical details on the hardware design, as well as the development and construction process, also giving a choronological view of the work. The optics are described

in Vassiliev, Fegan, and Brousseau 2007, while a concise overview of the whole instrument can be found in Vassiliev and pSCT Project 2017.

6.1 Overview

As reviewed in Chapter 2, traditional IACT designs, such as VERITAS, are constructed with single-mirror optical systems, with the segmeted mirror realized in either a parabolic or a Davies-Cotton (DC) (Davies and Cotton 1957) configuration. Due to the relatively narrow field of view (FoV), poor angular resolution and inferior off-axis performance, such designs, however, do not reap the full benefits of large telescope arrays like the CTA (Wood et al. 2016). The optical system of the medium-sized Schwarzschild-Couder IACT (SC-MST) (Vassiliev, Fegan, and Brousseau 2007) was developed by the CTA-US collaboration[1] with the goal of achieving performance limited only by the physics of atmospheric cascades rather than by instrumental deficiencies.

The design provides a wide FoV of 8°, and thanks to image de-magnification by the secondary mirror, facilitates the use of a high-resolution camera with 0.067° pixels. The SC-MST camera has 11,328 pixels, a factor of ~ 6 more than cameras in DC-based medium-sized telescope designs (DC-MST), while being almost one-third the size. Figure 6.1 demonstrates the difference between images captured by the two cameras. Moreover, such a small pixel size is compatible with novel silicon photomultipliers (SiPMs), which, with integrated electron-

[1] `http://cta-psct.physics.ucla.edu/institutions.html`

ics, are both less expensive and more robust than traditional PMTs. The highly pixelated camera and the synchronous optical system that's free of spherical and comatic aberrations allow to take advantage of the higher angular resolution of an array of SC-MSTs, whose footprint would be significantly larger than the area of a single Cherenkov light pool (Wood et al. 2016). While an array of DC-MSTs would also benefit from a larger overall footprint, the improved angular resolution of SC-MSTs, and their better characterization of shower morphology, would aid background rejection through improved event reconsruction and γ-hadron discrimination. We witnessed the positive effect of both of these improvements on the overall array sensitivity with ITM+BDT in Section 4.1, and Wood et al. 2016 demonstrate the resulting superior performance of an array of SC-MSTs compared to its DC-MST counterpart.

Finally, the low-energy performance of the SC-MST, with an effective mirror area half the size of the DC-MST, is aided by a completely different electronics operation regime and trigger system, which decouple trigger pixels and imaging pixels and allow them to be optimized independently. Unlike the PMT-based camera of the DC-MST, the SC-MST operates in a single-photon counting mode, substantially lowering the energy threshold and providing a continuous calibration signal in the data stream. The CAT collaboration was the first to operate an IACT in the Poisson regime and developed new analysis methods that took advantage of it (Barrau et al. 1997). More recently, the FACT collaboration has demonstrated novel analysis methods relying on single-photon extraction with a SiPM-based camera (Müller et al. 2017).

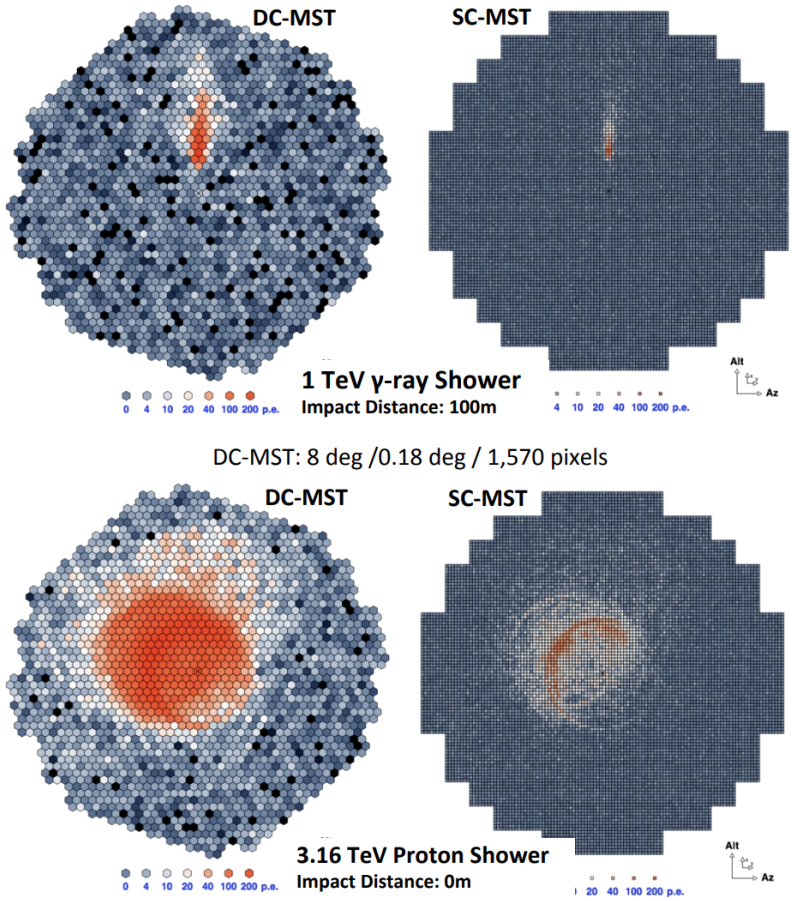

Figure 6.1: Simulated gamma- and proton-initiated showers as seen by the DC-MST (left) and the SC-MST (right). *Top:* a 1 TeV gamma-initiated shower. *Bottom:* a 3 TeV proton shower. Notice the apparent lack of morphological features in the DC-MST image.

The SC-MST is arguably an advance of the IACT technology on par with the introduction of multichannel PMT-based cameras by the Whipple 10m and the arrangement of IACTs into arrays by HEGRA in the late 80s. However, its numerous benefits come at the cost of highly curved aspheric mirrors, whose manufacture and alignment present major technological challenges. In order to investigate the feasibility of constructing and operating

an array of SC-MSTs, a consortium of US institutions was organized in 2012 to construct a prototype Schwarzschild-Couder telescope (pSCT) with a 9.7-m aperture (Figure 6.2).

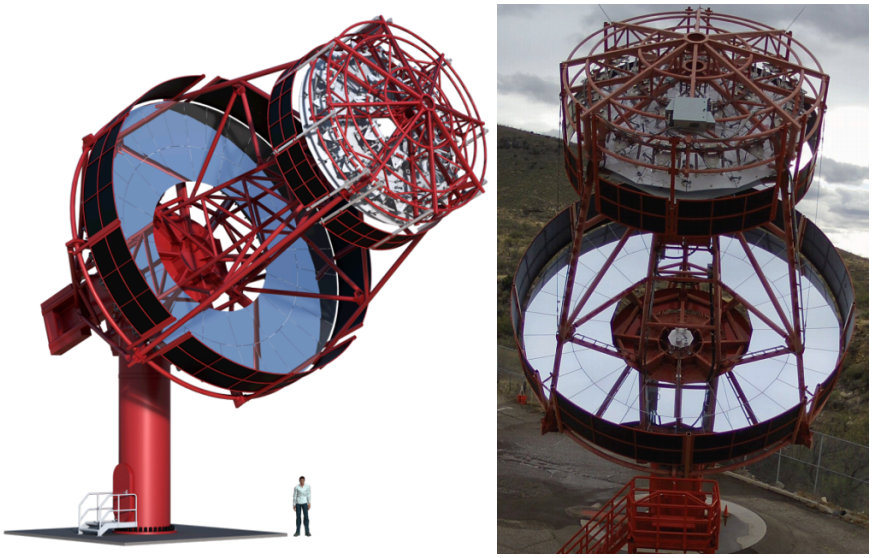

Figure 6.2: The prototype Schwarzschild-Couder medium-sized telescope for the CTA. The baffles around mirrors are necessary for sunlight protection during daytime, and their design was optimized to minimize the amount of stray light reaching the γ-ray camera during normal operations. *Left:* a CAD model of the telescope. *Right:* the pSCT in January 2019, entering the commissioning phase. The primary mirror surface shows panel misalignment near the bottom. The primary optical table is visible in the center of the primary dish. The secondary PEDB is the large white box at the bottom of the secondary.

The construction of the pSCT started in late 2015 at the old site of the VERITAS T1 telescope at the Fred Lawrence Whipple Observatory in Southern Arizona (Figure 2.12). Its optical support structure was built by August 2016 and the mirrors and camera were finally assembled and installed by August 2018. The telescope was inaugurated and entered commissioning in January 2019.

6.2 The Optical Alignment System

To minimize production costs, the mirrors of the pSCT are constructed of many mirror panels (Figure 6.3), with a 14-mm gap between the panels. The 9.7-m diameter primary mirror (M1) is segmented into 48 panels arranged in two rings – an inner ring (P1) of 16 panels and an outer ring (P2) with 32 panels. The 5.4-m diameter secondary mirror (M2) is segmented into 24 mirror panels, also in two rings – an 8-panel inner ring (S1) and a 16-panel outer ring (S2). For reference, the physical dimensions of each panel type are given in Table 6.1. In order to achieve a point spread function (PSF) compatible with the pixel size of the high-resolution camera, the mirror panels need to be aligned with a precision at the sub-mm and sub-mrad levels, both locally (panel-to-panel) and globally (Vassiliev, Fegan, and Brousseau 2007). The exact alignment requirements are listed in Table 6.2, where we note that for a mirror panel of ~ 1 m in size, a rotational misalignment of 0.1 mrad corresponds to a ~ 0.1-mm misalignment at its edge. In addition, for accurate source localization, the tilt of the mirror figures needs to be controlled to within 5 arcsec. While these tolerances are very loose compared to an optical telescope of a similar size, they are far stricter than those of current IACT optical systems and are challenging to achieve for an open-air telescope with a light optical support structure (OSS). These alignment requirements necessitate continuous monitoring of the segmented mirror surfaces and their automated mechanical alignment, functions performed by a specially developed optical alignment system.

The pSCT optical alignment system consists of two major subsystems. The panel-to-

Mirror panel	P1	P2	S1	S2
Number of panels	16	32	8	16
Radius max, m	3.4	4.83	1.60	2.71
Radius min, m	2.19	3.40	0.39	1.60
Diagonal, m	1.61	1.64	1.35	1.38
Panel aera, m^2	1.33	1.16	0.94	0.94

Table 6.1: Definitions of the pSCT mirror panels. P1 (P2) form the inner (outer) ring of the the primary mirror. Similary, S1 (S2) compose the inner (outer) ring of the secondary mirror.

	Primary mirror	Secondary mirror
Global alignment	Value	Value
Translation $\perp$ to optical axis	10 mm	10 mm
Translation $\parallel$ to optical axis	17 mm	5 mm
Tilt	15 mrad	0.15 mrad
Panel-to-panel alignment	Standard deviation	Standard deviation
Translation $\perp$ to optical axis	2.2 mm	1.1 mm
Translation $\parallel$ to optical axis	17mm	4 mm
Rotation around tangent axis	0.1 mrad	0.2 mrad
Rotation around radial axis	0.1 mrad	0.3 mrad
Rotation around normal axis	16.2 mrad	118 mrad

Table 6.2: Precision of alignment of the pSCT optical system required to achieve a PSF compatible with the pixel size of the high-resolution camera. Note that for a mirror panel of $\sim$ 1 m in size, a rotational misalignment of 0.1 mrad corresponds to a $\sim$ 0.1-mm misalignment at its edge.

panel alignment system (P2PAS) measures and corrects for misalignments between neighboring panels, and provides a continuous monitoring of the alignment of the optical surfaces. The global alignment system (GAS) is designed to continuously measure relative positions of the primary and secondary mirrors and the camera focal plane, and to detect large-scale

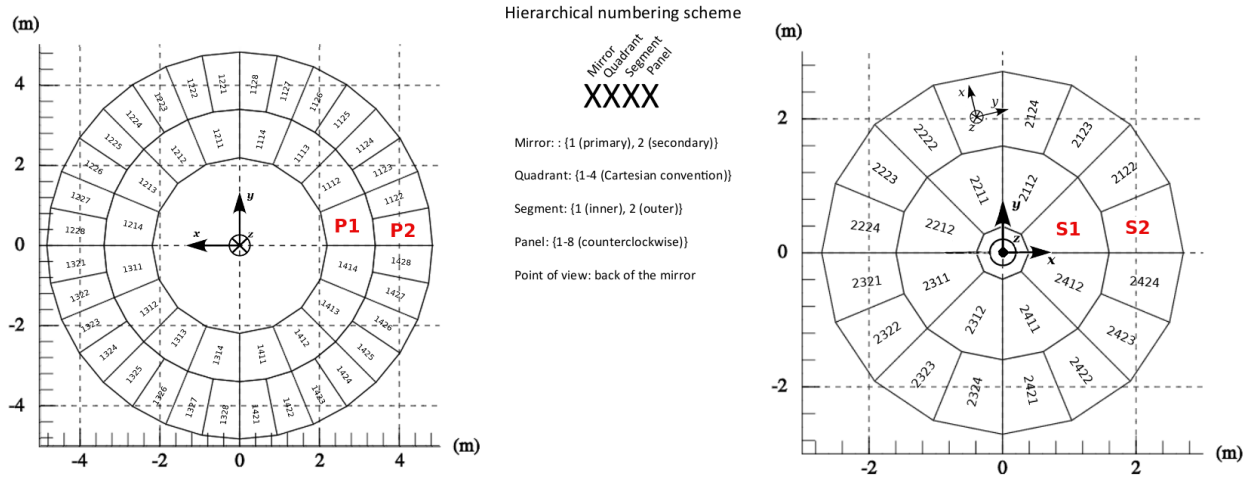

Figure 6.3: The mirror segmentation and panel numbering scheme of the pSCT. The primary is on the left, the secondary on the right. We adopt a hierarchical panel numbering scheme where each panel is denoted by a 4-digit number, as shown in the middle between the two mirrors. The telescope reference frame (TRF) is shown with axes at the centers of mirrors; the panel reference frame of panel 2221 is shown with axes on that panel. The 4 panel types are highlighted in red.

spatial perturbations of the mirror figures. Power and network connectivity are supplied to all components with specially designed and built power and Ethernet distribution boxes (PEDBs) attached to the back of each mirror (Figure 6.2). All hardware components, apart from the mirror panels themselves, were designed and assembled by collaborating institutions. Columbia and the University of California, Los Angeles, split the responsibility of constructing the P2PAS, and installing and aligning the mirrors, while the University of Iowa was responsible for the GAS components.

6.2.1 The Panel-to-Panel Alignment System

The P2PAS is implemented by placing each mirror panel on six linear actuators arranged in a Stewart platform (SP) configuration, and equipping it with a collection of mirror panel edge sensors (MPES). This bundle of a mirror panel, a Stewart platform and edge sensors, together with a mirror panel controller board (MPCB) and a mounting triangle, is denoted as a mirror panel module (MPM) (Figure 6.4). Stewart platform joints and edge sensors were installed on mirror panels with precisely machined steel attachment pads (Figure 6.4) that were glued to the backs of the panels with specially designed jigs. The locations of the attachment pads were computed with a high accuracy, accounting for mirror curvature. The MPCB, a Gumstix[2]-based microcomputer, is used to control the actuators, read out the MPES, and monitor the temperature of each MPM assembly. The aluminum mounting triangle interfaces the mirror panel with the OSS, while also providing space to house the MPCB. Special mounting fixtures were used to connect the MPMs to welded brackets on the OSS, achieving initial positioning accuracy of 10 mm and 1° (Figure 6.4).

The linear actuators of the Stewart platform can be controlled to within 3 μm, making it possible in practice to control panel position and orientation in all 6 degrees of freedom to within 5 μm and 3 arcsec, respectively. As a cost-saving measure, the actuators contain only a single encoder that resets every 200 steps, making it necessary to keep track of the encoder cycle in a database. Moreover, the actuator range of motion is limited to about

[2] `https://www.gumstix.com`

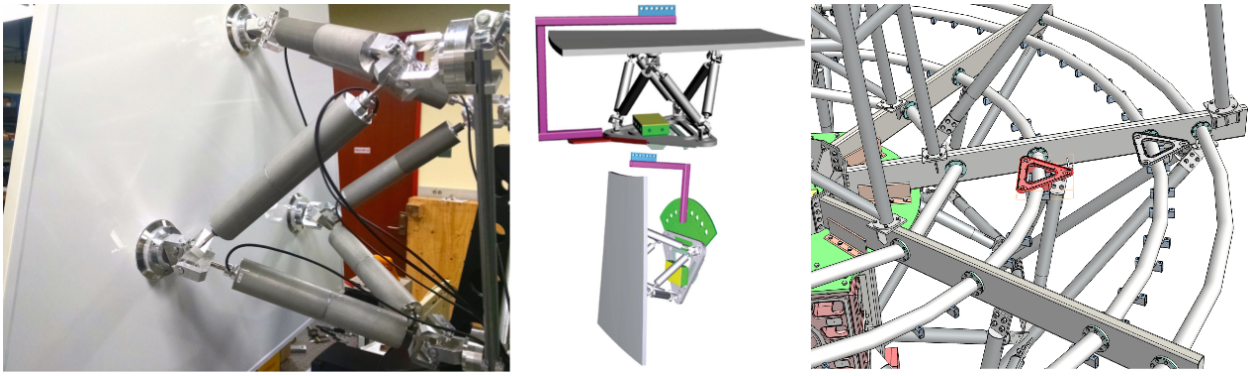

Figure 6.4: The mirror panel module and its interface with the telescope OSS. *Left:* the back of an MPM showing the Stewart platform and the attachement pads glued to the back of the mirror panel. *Middle:* a CAD drawing of the special mounting fixtures used to install MPMs on the OSS. *Right:* a CAD drawing showing the brackets and mounting triangles used to interface MPMs with the OSS.

70 mm. The edge sensors were also designed under budgetary constraints, and consist of a small diode laser and a CCD web camera (Figure 6.5), both off-the-shelf components, housed in a specially designed alumimum enclosure. The two halves of the MPES are placed opposite each other across the edge of two panels; images of the laser spot in the camera are then used to determine the relative displacement of the two panels; in order to protect sensor components and enable daytime operations, the two MPES sides are connected with a flexible plastic tube. Each MPES provides monitoring of two degrees of freedom, and in order to minimize the degeneracy between the degrees of freedom monitored by different sensors, a special placement pattern was chosen for the MPES (Figure 6.6): with all sensors along the same edge being mutually orthogonal, it is possible to measure edge misalignment of under 50 μm. This mutual orthogonality of sensors is also embedded in their design (Figure 6.5) – all sensors are identical, with the laser side, the connecting tube, and the camera

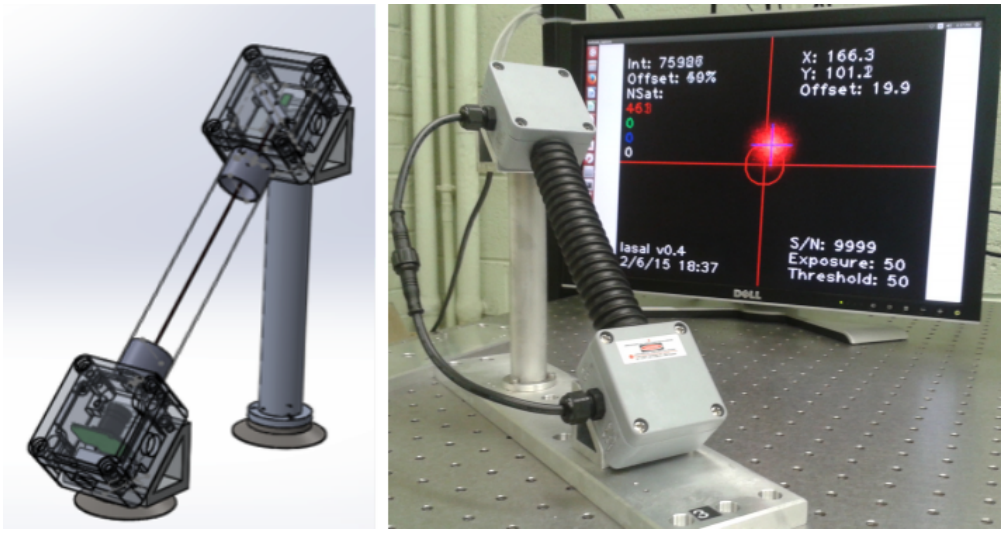

Figure 6.5: The mirror panel edge sensor. *Left:* a CAD drawing of the MPES design. The camera is in the enclosure on the lower left, while the laser is on the top right. Notice the attachment pads similar to those used for actuators *Right:* an assembled sensor being tested.

side forming an equilateral right triangle. Due to geometric constraints, the field of view of MPES cameras is rather limited, resulting in a usable range of around 10 mm – if edge misalignment exceeds this range, the sensor becomes non-operational. The total number of constraints provided by all the MPES of a single mirror is greater than the number of degrees of freedom of panel motion, thus providing some redundancy and making the P2PAS robust under possible sensor failures.

The idea behind the panel alignment procedure is described in more detail in Section 6.3.3. Every edge between every two adjacent panels was aligned in the lab, and the aligned sensor readings, together with the actuator response matrix, which relates the change of sensor readings to actuator displacement near the aligned position, were measured and recorded. While the actuator response matrix can be simulated, the use of the measured matrix was

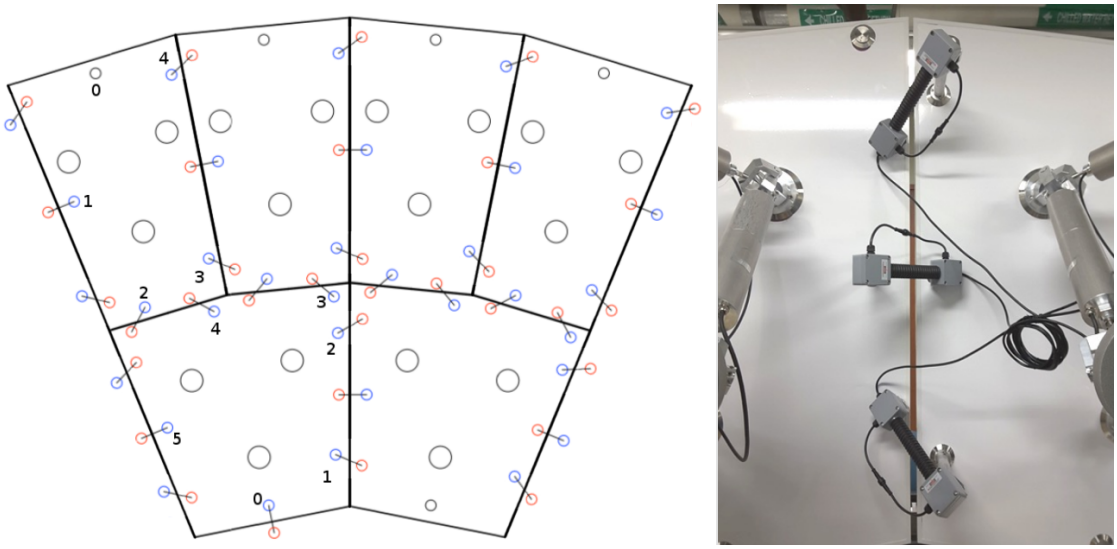

Figure 6.6: The arrangement of MPES. *Left:* a schematic showing the arrangement and numbering of sensors on the back of the primary mirror. Note that the sensors numbered 0, initially planned to interface the OSS, were not installed, but their attachment pads are available for a possible future addition. *Right:* two P1 panels with the MPES across their edge are being calibrated.

preferred due to the practical need to be able to perform edge calibration in the field after replacing failed components, with actuators far from their ideal lab lengths. Having measured MPES readings along an edge, the actuator displacements needed to return the edge to its aligned state can then be found simply by inverting this matrix. In practice, since the matrix was measured near the aligned sensor readings for small actuator displacements of only 1 mm, and because sensor response is non-linear, the alignment process was iterative (Figure 6.7, right), although the convergence of alignment was quite fast, with typically only 2-3 iterations needed.

In order to align and calibrate an edge in the lab, a coordinate measuring machine was

Figure 6.7: MPM calibration. *Left:* a P1-P2 edge is being calibrated. The half-spheres used to measure the exact position of a mirror panel are highlighted with yellow arrows. The spheres are temporarily fixed to the same exact reference points that were used to glue the Stewart platform attachment pads. The coordinate measuring machine and its probe can be seen on the left. Pink plastic bags cover the calibrated MPES for protection during transportation. *Middle:* a calibrated P1-P1 edge. Note the continuity of the aligned surface. *Right:* the convergence of edge misalignment in a laboratory setup. The plot shows readings from a triad of MPES along a P1-P1 edge (as in Figure 6.6, right) during an MPM alignment convergence test. Solid (dotted) lines represent the offset from the target position in the x (y) local coordinate system of MPES. The vertical line marks the change of scale along the y axis. For this test, the initial edge misalignment was 4 mm.

used (Figure 6.7, left). Three plastic half-spheres were attached with a special jig to the mirror surfaces of two MPMs precisely at the locations corresponding to the reference points used to glue the SP attachment pads. The MPMs were then installed on an optical table side-by-side, and the coordinate measuring machine, probing the locations of the half-spheres, was used to find their locations in the lab reference frame. The alignment software (see Sections 6.2.3 and 6.3.1) was then used to move the panels to their ideal lab coordinates, known from a simulation of the mirror surface.

Figure 6.8: The global alignment system of the pSCT. *Left:* a schematic representation of the GAS. *Middle:* the secondary mirror dish with the optical table in the center. The mirror panels are covered with a protective white film that was in place during transportation, installation and alignment of the mirrors. *Right:* a closeup of the secondary optical table. The laser spot used to monitor the optical axis alignment is highlighted with a red arrow right below the window to the PSD. The watertight cylindrical white enclosures house the CCD cameras and the autocollimator.

6.2.2 The Global Alignment System

The GAS (Figure 6.8) consists of two optical tables in the centers of the mirrors (Figures 6.2 and 6.8), whose CCD cameras detect relative motion of the optical components by imaging strings of LEDs on 3 uniformly spaced panels of the opposite mirror and on the focal plane of the γ-ray camera. Each optical table is placed onto a Stewart Platform identical to the ones used for mirror panels. The GAS utilizes monochrome GigE cameras from The Imaging Source;[3] thin short fiber optic cables are fixed to the LEDs in order to achieve dim point-like light sources. Additionally, autocollimators measure mirror tilt, while a position-sensitive device (PSD) on the secondary optical table, in conjunction with a laser on the primary, maintain the optical axis. Finally, a sky camera on the secondary optical table facilitates

[3] `https://www.theimagingsource.com`

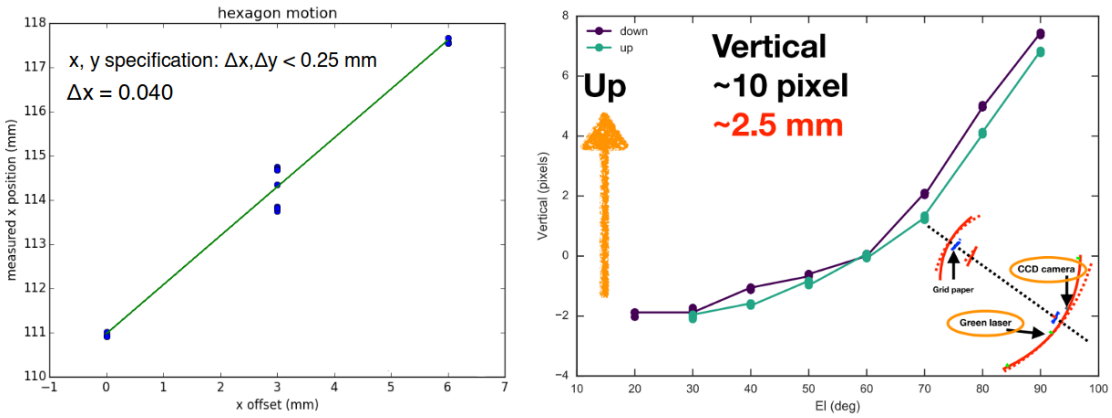

Figure 6.9: First tests of the GAS. *Left:* a GAS CCD is used to reconstruct the motion of a panel along its x axis by imaging a string of LEDs along the edges of the panel. The horizontal axis shows the real position of the panel, while the reconstructed position is along the vertical axis; the panel motion is reconstructed to within 40 μm; some hysteresis of the Stewart platform is apparent from the spread of the points in the middle. Similar tests were performed for panel motion along the other axes. *Right:* The optical axis monitoring capabilities are used to study structural deformations of the OSS under telescope motion. The plot displays the misalignment of the laser used to monitor the optical axis as a function of telescope elevation. The inlay in the bottom right corner explains the experimental setup used to obtain the measurements.

pointing correction by imaging the night sky.

Each component of the GAS was carefully calibrated in the lab, with positions of mirror panel LEDs measured to within a micron with a setup similar to that used to align and calibrate edges. First field tests of the GAS have been performed, giving satisfactory results, and its optical axis monitoring capabilities were also used to perform first studies of mechanical deformations of the OSS under telescope motion (Figure 6.9).

6.2.3 Alignment Control Software

The alignment system software, developed throughout the construction of the telescope, was instrumental to the assembly, calibration, installation and alignment process of the mirrors. The software, written in *C++*, follows a server-client architecture, with each of the 72 MPMs, as well as each of the optical tables, acting as a server, and a central control computer being their client (Figure 6.10). The central computer also acts as an aggregating server of all the underlying data, providing a higher level of abstraction and connecting the alignment system to future array-level controls. The client-server communication was implemented through *OPC UA*,[4] an industry standard communication protocol for industrial automation. The choice of this protocol was made by the wider CTA collaboration with the aim of standardizing and abstracting communication to and control of the many different telescopes and their subsystems in the large array. All hardware intialization data, such as mapping tables, calibration constants, response matrices and so on, as well as logging data, such as actuator motion and sensor readout history, are stored in and accessed through a modular database, enabling easy replacement of hardware components.

The code base comprises three parts: the server code that runs on each of the MPMs and optical tables; the client code that runs on the central computer; and the common code that's shared between the client and the server. Both the server and the client follow the same architectural design. The top-level *OPC UA* abstraction layer serves all underlying data and

[4] `https://opcfoundation.org/about/opc-technologies/opc-ua/`

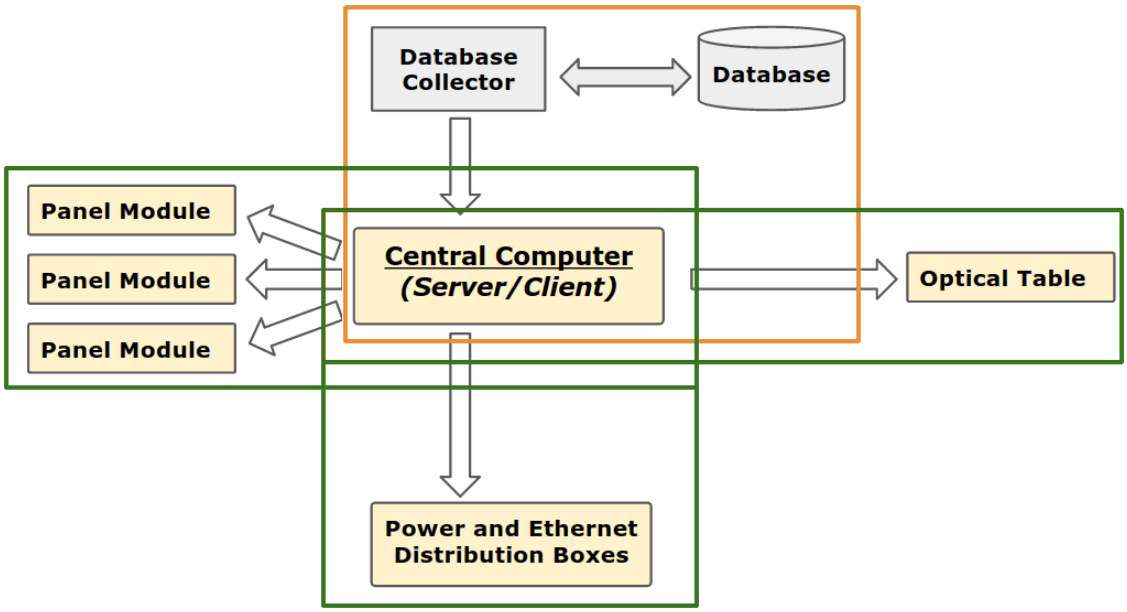

Figure 6.10: A schematic of the alignment software structure. The central computer in the middle acts as a client of the alignment system, and as a server to telescope- and array-level controls. The alignment system components are framed in green. Database communication, framed in orange, behaves as a plug-in component.

methods to an external client: in the case of MPM servers, the external client is the central computer; in the case of the central computer, it is the array-level client or a user client, such as a GUI. This layer accesses the underlying data and methods through a middleman, a communication interface. The communication interface talks to the hardware by abstracting and standardizing data access and manipulation between the many components. In the case of MPM servers, the access to hardware is direct, while in the case of the client, it is a series of requests to the relevant servers. Finally, the bottom layer is hardware-specific code, such as direct readout of GigE or web cameras, or serial port access for PSD readout and actuator control.

The software implements full control of the alignment system. It allows low-level access to individual hardware components, such as individual sensors, actuators, and CCD cameras, as well as more abstract access to edges, panels, mirror sectors or whole mirrors, and even the telescope positioner. The current alignment use case is limited to explicit requests to move individual panels, or align individual edges or sectors of a mirror, with no continuous automatic alignment performed yet. Moreover, while the GAS is fully accessible and controllable through the software, it has not yet been incorporated into the alignment procedures (Section 6.3.3). Thanks to the use of *OPC UA*, a generic GUI client is used to talk to the aggregating server-client of the central alignment computer, while a full-featured GUI specific to the pSCT is being developed separately. The code base has been handed over to the CTA-US collaboration for further development and future maintenance.

6.2.4 Mirror Installation and Initial Alignment

Finally, a few words recounting the process of mirror installation and intial alignment. Due to lab storage constraints, the MPMs were assembled, calibrated and installed on the OSS in quadrants, 12 panels at a time. The panels were installed with a forklift with the fixtures shown in Figure 6.4. Due to the inaccuracies of the bracket welding, this resulted in only a rough placement of panels, with edge sensors outside of their usable range. Moreover, in order to have enough safe working space to install the last several panels, all panels had to be individually spread out. Thus, after the installation, the mirror surface was in a very

misaligned state, and the procedure described in Section 6.3.3 could not be performed.

A major part of the installation process, then, entailed initial alignment of the mirror figures, bringing all sensors into their working range. This was done sequentially, one edge at a time – a starting panel on the inner ring was fixed, its neighbor moved by eye to where the sensors were in working range, the sensors then connected, and the edge aligned by means of the response matrix. Small angular tilts of the chosen starting panel quickly added up to large translations several neighbors down the mirror ring, requiring actuator motions that were outside of the working range. Thus, the process had to be repeated several times, correcting for this initial misplacement.

Small angular misalignments, as well as gravity- and temperature-induced OSS deformations, resulted in large mirror misalignments upon reaching the last edge that closes the ring. That is, with all edges aligned sequentially starting from some initial panel, the last edge had a large enough misalignment where it was far outside the working sensor range, and in the case of the primary mirror, so misaligned that the last panel simply had no space to fit between its neighbors. To overcome this, a systematic misalignment vector was introduced into the alignment procedure: each edge of the same type was assumed to have the same systematic error in the measurement of its aligned position. Such systematic errors were expected to originate from a variety of unresolved uncertainties in the laboratory setup used for edge calibration; inaccuracies in the jig used to place the plastic half-spheres on the mirror surface, uncertainties in the placement and calibration of the coordinate measuring machine, and the gravitational sag of the optical table were all thought to have contributed.

Because the calibration setup was reproduced exactly for every edge, these uncertainties were expected to be identical for every edge of the same type. Measuring the misalignment along the last edge with a caliper, and propagating it backwards to the starting panel, this systematic vector was iteratively estimated and corrected for, ultimately bringing all sensors in their working range and then applying the procedure of Section 6.3.3 to achieve a more accurate alignment. In the process, the tilts and translations of the mirror figure as a whole had to be corrected for to get all actuators back into their working range.

After this, the outer rings were aligned to the inner ones through the P1-P2 and S1-S2 sensors, first moving panels by eye to get to the sensor working range, and then aligning 3-panel sectors as described in Section 6.3.3. Ultimately, sub-mm alignment accuracy was achieved on both mirrors through this very tedious process, and further corrections will be performed once the GAS is incorporated into the alignment procedure.

The process also tested the robustness of the hardware and the foresight of the original design of the alignment system. Several edge sensors and actuators broke down due to a combination of assembly errors and environmental exposure to the harsh sun and powerful rains of Southern Arizona. The most common assembly errors were loose cable glands and seals, improper electrical isolation, and incorrect or insufficient use of epoxy to secure edge sensor components. Actuator replacement was straightforward, while sensor replacement relied on the redundancies of the alignment system inherent in the design. The failed components were successfully replaced and recalibrated in the field.

6.3 Theoretical Considerations

We conclude with a few details of the pSCT optical alignment system from an analytical perspective. The components of the system are positioned relative to each other in the telescope reference frame (TRF) with its origin in the center of the primary mirror. The z axis is defined by the optical axis and points towards the center of the secondary. When the z axis points to the geographical North, the x axis points towards the geographical West and the y axis is directed to zenith (Figure 6.3). All coordinates are measured in millimeters and angles in radians.

Further, for each mirror panel, we define its panel reference frame (PRF) as follows. The barycenter of the radius vectors of the three attachment pads of the SP is given the coordinates $(0, 0, 600)$. The $x-y$ plane is parallel to the plane of the three pads in the initial panel position, when all 6 actuators have the length of 427.919 mm, with the z axis in the direction of reflected light and the x axis along the axis of symmetry of the mirror panel (Figure 6.3). The peculiar non-zero coordinates assigned to the barycenter of the pads are chosen so that the origin of the PRF lies roughly at the center of the mounting triangle. Therefore, the whole frame of reference is fixed to the mounting triangle and unaffected by actuator motions.

6.3.1 Stewart Platform Simulation

An important piece of code in the alignment system is the simulation of the Stewart platform on which each mirror panel sits. This computation is used to find the coordinates of each panel in the PRF and TRF, and move panels individually, and was crucial to the calibration, installation, and initial alignment of mirror panels.

We first describe the *inverse* kinematic transformation – given the location and orientation of the mirror panel, this computes the lengths of the actuators. This is easy: for each actuator, its length is simply the distance between the attachment pad on the mounting triangle and that on the mirror panel. Neglecting straightforward corrections due to the thicknesses of the joint, the panel and the pad itself, this can be written simply as

$$l_i = \left\| \vec{P}_i(\vec{t}, \vec{\xi}) - \vec{B}_i \right\|, \tag{6.1}$$

where $\vec{B}_i$ are the coordinates of pad i on the mounting triangle and $\vec{P}_i(\vec{t}, \vec{\xi})$ are the coordinates of pad i on the panel as defined by the panel location $\vec{t}$ and orientation $\vec{\xi}$, both given in the PRF. Both $\vec{B}_i$ and $\vec{P}_i$ are defined as points on a circle of radius $R_b = 320$ mm, $2\pi/3$ radians apart. $\vec{B}_i$ reside on the mounting triangle and have fixed coordinates:[5]

[5]It is apparent that B_i are equal for $i = 5, 0$, $i = 1, 2$ and $i = 3, 4$. This just means that these pairs of actuators are attached to the same pads on the base triangle side; similarly, actuators $(0, 1)$, $(2, 3)$ and $(4, 5)$ are attached to the same pads on the mirror panel side. In practice, this lets us have only 3 attachment pads to accommodate all 6 actuators. We keep the general notation here because no such restrictions on the attachement configuration are needed, or indeed, assumed, in our simulation.

$$\begin{aligned}\vec{B}_i &= \left(R_b \cos\frac{\pi i}{3}, R_b \sin\frac{\pi i}{3}, 0\right)^T \quad i = 0, 2, 4; \\ \vec{B}_i &= \left(R_b \cos\frac{\pi(i+1)}{3}, R_b \sin\frac{\pi(i+1)}{3}, 0\right)^T, \quad i = 1, 3, 5.\end{aligned} \tag{6.2}$$

$\vec{P}_i$ reside on the mirror panel and therefore transform with it. In the PRF, these pad coordinates are given by

$$\vec{P}_i(\vec{t}, \vec{\xi}) = \vec{t} + R_y(\beta)R_z(\gamma)R_x(\alpha)\vec{\tilde{P}}_i. \tag{6.3}$$

Here, we choose the order of rotations as $x \to z \to y$. $\vec{\tilde{P}}_i$ are the coordinates of the panel-side pads with respect to the mirror surface, given by

$$\begin{aligned}\vec{\tilde{P}}_i &= \left(R_b \cos\frac{\pi(i+1)}{3}, R_b \sin\frac{\pi(i+1)}{3}, 0\right)^T \quad i = 0, 2, 4; \\ \vec{\tilde{P}}_i &= \left(R_b \cos\frac{\pi i}{3}, R_b \sin\frac{\pi i}{3}, 0\right)^T, \quad i = 1, 3, 5.\end{aligned} \tag{6.4}$$

The *forward* kinematic transformation, which defines the location and orientation of the panel given the lengths of the actuators, is, of course, the inverse of the above. Essentially, given actuator lengths l_i, we want to find the panel coordinates $\vec{a}_0 = (\vec{t}^T, \vec{\xi}^T)^T = (x_0, y_0, z_0, \alpha_0, \beta_0, \gamma_0)^T$ that minimize

$$\chi^2(\vec{a}) = \sum_{i=0}^{6} f_i(\vec{a}) = \sum_{i=0}^{6} \left[\left(\vec{P}_i(\vec{t}, \vec{\xi}) - \vec{B}_i\right)^2 - l_i^2\right]. \tag{6.5}$$

In practice, since the minimum of this expression is always equal to zero, we can employ the Newton-Raphson iterative method to find it, with the Jacobian explicitly calculated using the transformation 6.3.

6.3.2 Moving the Mirror as a Whole

While moving individual panels is quite useful, it is also necessary to be able to move an arbitrary sector of the mirror as a rigid body. The PRF is defined as the reference frame tied to the mounting triangle of the panel if the mirror structure were perfectly rigid. Thus, the PRF moves with the telescope but is not affected by local perturbations to the mirror figure. We denote the coordinates of an arbitrary point in the PRF of panel n as $\vec{\rho}_n^w$, where n = MQSP is the mirror, quadrant, segment, panel identifier of the panel and w = P1, P2, S1, S2 is the type of the panel (Figure 6.3). We can transform these PRF coordinates to the telescope reference frame (TRF) as follows:

$$\vec{r}_n^w = R_z(\phi_n^w)\left(\vec{T}^w + R_y(\theta^w)\vec{\rho}_n^w\right), \tag{6.6}$$

and the reason for keeping the subscript on the TRF coordinates will become clear momentarily. Here, ϕ_n^w is the angle between the x-axis of the TRF and the projection of the x-axis of the PRF onto the $x-y$ plane of the TRF; and θ^w is the angle between the z axes of the PRF and the TRF.

The mounting pads of a panel can then be located in the TRF as $\vec{r}_{k,n}^w$, where $k = 1, 2, 3$

is the pad number. If the mirror *as a whole* is perturbed, these pad coordinates will be changed by

$$\delta\vec{r}^{\,w}_{k,n} = R(\delta\vec{\xi})\vec{r}^{\,w}_{k,n} + \delta\vec{T}. \tag{6.7}$$

Here, $\delta\vec{T}$ is the translational and $R(\delta\vec{\xi})$ is the rotational perturbation to the mirror figure.

To first order, this change in the TRF is related to the perturbation of the pad coordinates in the PRF by

$$\delta\vec{\rho}^{\,w}_{k,n}(\delta\vec{T}, \delta\vec{\xi}) = R_y(-\theta^w)\left[R_z(-\phi^w_n)\delta\vec{r}^{\,w}_{k,n} - \vec{T}^w\right]. \tag{6.8}$$

The procedure to move the mirror figure as a whole by $(\delta\vec{T}, \delta\vec{\xi})$ is then as follows. For each panel:

- get the PRF pad coordinates from the current actuator lengths using the forward Stewart platform simulation as described in Section 6.3.1;
- move the pad coordinates in the PRF by means of (6.6)-(6.8);
- finally, obtain the target actuator lengths by performing the inverse Stewart platform simulation as described in Section 6.3.1.

The alignment software allows to apply the same procedure to an arbitrary subset of mirror panels, moving them in unison and preserving the intial figure.

We will use $\delta\vec{\tilde{\rho}}^{\,w}_{k,n}(\vec{L}^n)$ to denote the displacement of a pad from its ideal aligned position

as a function of motions of actuators of panel n. When the mirror figure is aligned, this displacement consists of the displacement of the mirror as a whole plus the perturbation of the mounting bracket at the location of panel n:

$$\delta\vec{\tilde{\rho}}^{\,w}_{k,n} - \delta\vec{\rho}^{\,w}_{k,n} = \vec{\Delta}^{\,w}_{k,n}, \tag{6.9}$$

the difference between pad coordinates as computed from actuator lengths and pad coordinates as computed from telescope coordinates, is the contribution to the pad displacement due to the non-ideal placement of the mounting bracket. We can define a global measure of this non-idealicity of bracket placement as

$$\chi^2 = \sum \left(\vec{\Delta}^{\,w}_{k,n}\right)^2 = \sum \left[\delta\vec{\tilde{\rho}}^{\,w}_{k,n}(\vec{L}^n) - \delta\vec{\rho}^{\,w}_{k,n}(\delta\vec{T}, \delta\vec{\xi})\right]^2, \tag{6.10}$$

where the sum runs over all pads of all panels of a single mirror. It is clear that this expression reaches its minimum when the mirror figure is perfectly aligned. Or, conversely, for an aligned mirror, the minimum of equation 6.10 is reached when $(\delta\vec{T}, \delta\vec{\xi})$ correspond to the true mirror coordinates.

Thus, when the mirror is aligned, minimizing (6.10) with respect to $\delta\vec{T}$ and $\delta\vec{\xi}$ will give local perturbations due to non-ideal bracket placement. While the bracket non-idealicities are not taken into account by the software at the moment, this minimization fit is used to measure and correct for the displacement of the mirror as a whole in the TRF.

6.3.3 Panel Alignment

For an edge $E^{j,k}$ defined by two panels j and k, let $\vec{\sigma}^{j,k}$ be the current readings of its edge sensors and $\vec{\tilde{\sigma}}^{j,k}$ be the aligned, or target, sensor readings. Let $\vec{L}^j$ and $\vec{L}^k$ be the actuator lengths of panels j and k, respectively. Finally, let $R^{j,k}$ be the response matrix of the sensors along $E^{j,k}$ to the motion of the actuators of panel k, while $R^{k,j}$ be their response to the motion of panel j. That is, $R^{j,k}_{si} = \delta\sigma^{j,k}_s / \delta L^k_i$. In order to align the edge, we need to change the lengths $\vec{L}^j$ and $\vec{L}^k$ in such a way that

$$\vec{\sigma}^{j,k} + R^{j,k}\delta\vec{L}^k + R^{k,j}\delta\vec{L}^j = \vec{\tilde{\sigma}}^{j,k} \;\leftrightarrow\; R^{j,k}\delta\vec{L}^k + R^{k,j}\delta\vec{L}^j = \vec{\tilde{\sigma}}^{j,k} - \vec{\sigma}^{j,k} = \delta\vec{\sigma}^{j,k}. \tag{6.11}$$

With 6 actuators per panel, this equation has 12 degrees of freedom, and we only have 2 or 3 sensors per edge, giving 4 or 6 constraints. Thus, the system is underdetermined and requires extra constraints to be solved. These extra constraints come in the form of restrictions on actuator motions. In the case of an edge with 3 sensors, we simply keep one of the panels fixed, setting $\delta\vec{L}^j = 0$, and easily solving the above for $\delta\vec{L}^k$.

We can easily generalize this simple setup to a more complicated situation, when the edge has only 2 sensors, such as a P1-P2 or an S1-S2 edge. In that case, there aren't enough constraints even when fixing one of the panels, and such simple alignment is not possible. In that case, the alignment procedure needs to include additional sensor readings, and is only possible for the whole 3-panel sector. Let such a 3-panel sector consist of panels j, k and m.

The alignment problem then becomes

$$R^{m,j,k}\delta\vec{L}^k + R^{k,m,j}\delta\vec{L}^j + R^{j,k,m}\delta\vec{L}^m + R^{k,m,j}\delta\vec{L}^j + R^{m,j,k}\delta\vec{L}^k + R^{j,k,m}\delta\vec{L}^m = \vec{\tilde{\sigma}}^{j,k,m} - \vec{\sigma}^{j,k,m}. \tag{6.12}$$

Here, $\vec{\sigma}^{j,k,m}$ is the vector of all sensor readings between panels j, k and m. The response matrices are now larger and are defined by $R^{m,j,k}_{si} = \delta\sigma^{j,k,m}_s/\delta L^k_i$. A 3-panel sector has 7 sensors, resulting in 14 readings. Fixing one of the panels, say m, sets $\delta\vec{L}^m = 0$, and results in an overdetermined problem with 12 degrees of freedom and 14 constraints:

$$R^{m,j,k}\delta\vec{L}^k + R^{k,m,j}\delta\vec{L}^j + R^{k,m,j}\delta\vec{L}^j + R^{m,j,k}\delta\vec{L}^k = \vec{\tilde{\sigma}}^{j,k,m} - \vec{\sigma}^{j,k,m}. \tag{6.13}$$

This is easily solvable for $\delta\vec{L}^j$ and $\delta\vec{L}^k$.

We can view the above extension as the alignment of a sector of two panels j and k to the sensors the sector shares with the panel m. The internal sensors between j and k act as constraints that maintain the rigidity of the sector. It should be clear from this simple exercise that further generalizing the alignment problem to an arbitrary collection of panels and sensors is quite straightforward. Let Ω be an arbitrary set of panels that form a continuous mirror sector and Υ be an arbitrary set of sensors along the boundary of the sector. Further, let Ψ be the set of sensors internal to the sector. In order to have a sufficient number of constraints, we need to require $|\Upsilon| \geq 3$ as well as $|\Psi| \geq 3(|\Omega| - 1)$. These conditions can be satisfied even when a number of sensors, both internal and external

to the sector, are malfunctioning. We can then write the alignment problem concisely as follows:

$$\sum_{\substack{p\in\Omega\\1\le i\le 6}} R_{s_c,p_i}\delta L_{p_i} = \widetilde{\sigma}_{s_c} - \sigma_{s_c} \quad \forall s \in \Upsilon \cup \Psi,\ c = 1,2. \tag{6.14}$$

Here, p runs over the panels and i runs over the actuators of each panel; s corresponds to a sensor in the set of all external and internal sensors, and c to a single coordinate of the 2D sensor reading. The response matrix R is constructed just like before, $R_{s_c,p_i} = \delta\sigma_{s_c}/\delta L_{p_i}$, and is relatively sparse because only a small number of sensors are affected by an individual actuator. The overdetermined system is solved for δL_{p_i} in a straightforward way. This flexible sector alignment can be performed on an arbitrary number of panels, as long as there is a sufficient number of constraints. That is, it is impossible to "close the loop" with this method – when including all panels of a single mirror, the system becomes underdetermined. To align the whole mirror, at least one panel needs to be excluded. Finally, in practice, the alignment process is iterative, since the response matrices are measured only approximately.

Equation 6.14 can also be recast in terms of the minimization of the functional

$$\chi^2_p(\vec{L}^{\Omega}) = \sum_{\substack{s\in\Upsilon\cup\Psi\\c=1,2}} \sum_{\substack{p\in\Omega\\1\le i\le 6}} \left[R_{s_c,p_i}\delta L_{p_i} - (\widetilde{\sigma}_{s_c} - \sigma_{s_c})\right]^2. \tag{6.15}$$

This form is particularly useful when extending the alignment procedure to include readings from the global alignment system. While the current version of the software is fully capable

of obtaining all GAS readings, it does not incorporate them into the alignment procedure. However, an extension that includes GAS is straightforward, at least in principle, boiling down to minimizing the extended functional

$$\begin{aligned}\chi^2(\vec{L}^{\Omega}) = w_p\chi_p^2(\vec{L}^{\Omega}) + w_g\chi_g^2(\vec{L}^{\omega}) = w_p \sum_{\substack{s\in\Psi \\ c=1,2}} \sum_{\substack{p\,\in\,\Omega \\ 1\leq i\leq 6}} \left[R_{s_c,p_i}\delta L_{p_i} - (\widetilde{\sigma}_{s_c} - \sigma_{s_c})\right]^2 \\ + w_g \sum_{\substack{g\in\omega \\ 1\leq t\leq 6}} \sum_{\substack{g\,\in\,\omega \\ 1\leq j\leq 6}} \left[R_{g_t,g_j}\delta L_{g_j} - (\widetilde{\zeta}_{g_t} - \zeta_{g_t})\right]^2 .\end{aligned} \tag{6.16}$$

Here, the first term is identical to 6.15, with Ω comprising the set of all mirror panels and no external sensors, $\Upsilon = \varnothing$. The second term is analogous but goes over the readings and actuators of the GAS, with ω being the set of GAS panels and ζ_{g_t} the GAS readings. The matrix elements R_{g_t,g_j} are again the response of GAS readings to the motions of actuators on the relevant panels and form a block-diagonal matrix. The weights w_p and w_g are here to account for the difference in the effective numbers of degrees of freedom between the two subsystems. Correctly computing these weights, accounting for degeneracies in sensor readings, is very important, but extends beyond the scope of this work.

Bibliography

Abdo, A. A. et al. (2010). "Detection of gamma-ray emission from the starburst galaxies M82 and NGC 253 with the Large Area Telescope on Fermi." In: *The Astrophysical Journal* 709.2, pp. L152–L157. DOI: 10.1088/2041-8205/709/2/l152. URL: https://doi.org/10.1088%2F2041-8205%2F709%2F2%2Fl152.

Abraham, J. et al. (2010). "Measurement of the energy spectrum of cosmic rays above 1018 eV using the Pierre Auger Observatory." In: *Physics Letters B* 685.4, pp. 239 –246. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2010.02.013. URL: http://www.sciencedirect.com/science/article/pii/S0370269310001875.

Acero, F. et al. (Nov. 2009). "Detection of Gamma Rays from a Starburst Galaxy." In: *Science* 326, p. 1080. DOI: 10.1126/science.1178826. arXiv: 0909.4651 [astro-ph.HE].

Ackermann, M. et al. (Aug. 2012). "GeV Observations of Star-forming Galaxies with the Fermi Large Area Telescope." In: *The Astrophysical Journal* 755, 164, p. 164. DOI: 10.1088/0004-637X/755/2/164. arXiv: 1206.1346 [astro-ph.HE].

Aharonian, F. A. (2004). *Very high energy cosmic gamma radiation : a crucial window on the extreme Universe*. World Scientific Publishing Co. DOI: 10.1142/4657.

Allen, G. E. et al. (1997). "Evidence of X-Ray Synchrotron Emission from Electrons Accelerated to 40 T[CLC]e[/CLC]V in the Supernova Remnant Cassiopeia A." In: *The Astrophysical Journal* 487.1, pp. L97–L100. DOI: 10.1086/310878. URL: https://doi.org/10.1086%2F310878.

Ansoldi, S. et al. (2018). "The Blazar TXS 0506+056 Associated with a High-energy Neutrino: Insights into Extragalactic Jets and Cosmic-Ray Acceleration." In: *The Astrophysical Journal* 863.1, L10, p. L10. DOI: 10.3847/2041-8213/aad083. arXiv: 1807.04300 [astro-ph.HE].

Barrau, A. et al. (1997). "The CAT Imaging Telescope." In: *arXiv e-prints*, astro-ph/9705247, astro–ph/9705247. arXiv: astro-ph/9705247 [astro-ph].

Beatty, James J. and Stefan Westerhoff (2009). "The Highest-Energy Cosmic Rays." In: *Annual Review of Nuclear and Particle Science* 59.1, pp. 319–345. DOI: 10.1146/annurev.nucl.58.110707.171154. eprint: https://doi.org/10.1146/annurev.nucl.58.110707.171154. URL: https://doi.org/10.1146/annurev.nucl.58.110707.171154.

Bell, A. R. (Jan. 1978). "The acceleration of cosmic rays in shock fronts. I." In: *Monthly Notices of the Royal Astronomical Society* 182, pp. 147–156. DOI: 10.1093/mnras/182.2.147.

Berge, D., S. Funk, and J. Hinton (May 2007). "Background modelling in very-high-energy γ-ray astronomy." In: *Astronomy & Astrophysics* 466, pp. 1219–1229. DOI: 10.1051/0004-6361:20066674. arXiv: astro-ph/0610959 [astro-ph].

Bernlöhr, K. et al. (Mar. 2013). "Monte Carlo design studies for the Cherenkov Telescope Array." In: *Astroparticle Physics* 43, pp. 171–188. DOI: 10.1016/j.astropartphys.2012.10.002. arXiv: 1210.3503 [astro-ph.IM].

Bird, R. A. (Aug. 2015). "Probing Cosmic Rays with VERITAS: Observations of M 31 and the Positron Fraction." PhD thesis. University College Dublin.

Breiman, L. (Aug. 1996). "Bagging Predictors." In: *Mach. Learn.* 24.2, pp. 123–140. ISSN: 0885-6125. DOI: 10.1023/A:1018054314350. URL: http://dx.doi.org/10.1023/A:1018054314350.

Breitschwerdt, D., J. F. McKenzie, and H. J. Voelk (Mar. 1993). "Galactic winds. II - Role of the disk-halo interface in cosmic ray driven galactic winds." In: *Astronomy & Astrophysics* 269, pp. 54–66.

Brun, R. and F. Rademakers (1997). "ROOT — An object oriented data analysis framework." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1. New Computing Techniques in Physics Research V, pp. 81 –86. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(97)00048-X. URL: http://www.sciencedirect.com/science/article/pii/S016890029700048X.

Bykov, Andrei M. (2014). "Nonthermal particles and photons in starburst regions and superbubbles." In: *The Astronomy and Astrophysics Review* 22, 77, p. 77. DOI: 10.1007/s00159-014-0077-8. arXiv: 1511.04608 [astro-ph.HE].

CTA Consortium, The (Sept. 2017). "Science with the Cherenkov Telescope Array." In: *arXiv e-prints*, arXiv:1709.07997, arXiv:1709.07997. arXiv: 1709.07997 [astro-ph.IM].

Chevalier, R. A. and A. W. Clegg (Sept. 1985). "Wind from a starburst galaxy nucleus." In: *Nature* 317, p. 44. DOI: 10.1038/317044a0.

Christiansen, J. (2019). "Shower-Image Template Stereo Reconstruction Method." In: *VERITAS Internal Memo*.

Christiansen, J. and VERITAS Collaboration (2017). "Characterization of a Maximum Likelihood Gamma-Ray Reconstruction Algorithm for VERITAS." In: *International Cosmic Ray Conference* 301, 789, p. 789. arXiv: 1708.05684 [astro-ph.IM].

Cocconi, G. (Jan. 1960). "An air shower telescope and the detection of 10^{12} eV photon sources." In: *International Cosmic Ray Conference* 2, p. 309.

Cogan, P. (Dec. 2006). "Nanosecond Sampling of Atmospheric Cherenkov Radiation Applied to TeV Gamma-Ray Observations of Blazars with VERITAS." PhD thesis. University College Dublin.

— (2007). "VEGAS, the VERITAS Gamma-ray Analysis Suite." In: 3, p. 1385. arXiv: 0709.4233 [astro-ph].

Cortina, J. et al. (July 2009). "Technical Performance of the MAGIC Telescopes." In: *arXiv e-prints*, arXiv:0907.1211, arXiv:0907.1211. arXiv: 0907.1211 [astro-ph.IM].

Cronin, J. W. (2011). "The 1953 cosmic ray conference at Bagneres de Bigorre: The Birth of sub atomic physics." In: *Eur. Phys. J.* H36, pp. 183–201. DOI: 10.1140/epjh/e2011-20014-4. arXiv: 1111.5338 [physics.hist-ph].

Davies, J. M. and E. S. Cotton (Apr. 1957). "Design of the quartermaster solar furnace." In: *Solar Energy* 1, pp. 16–22. DOI: 10.1016/0038-092X(57)90116-0.

Fegan, D. J. (Sept. 1997). "TOPICAL REVIEW: γ/hadron separation at TeV energies." In: *Journal of Physics G Nuclear Physics* 23, pp. 1013–1060. DOI: 10.1088/0954-3899/23/9/004.

Fenech, D. M. et al. (Dec. 2008). "Deep MERLIN 5GHz radio imaging of supernova remnants in the M82 starburst." In: *Monthly Notices of the Royal Astronomical Society* 391, pp. 1384–1402. DOI: 10.1111/j.1365-2966.2008.13986.x. arXiv: 0810.0424.

Fermi, ENRICO (1949). "On the Origin of the Cosmic Radiation." In: *Phys. Rev.* 75 (8), pp. 1169–1174. DOI: 10.1103/PhysRev.75.1169. URL: https://link.aps.org/doi/10.1103/PhysRev.75.1169.

Frank, I. M. and I. E. Tamm (1937). "Coherent visible radiation of fast electrons passing through matter." In: *Compt. Rend. Acad. Sci. URSS* 14.3. [Usp. Fiz. Nauk93,no.2,388(1967)], pp. 109–114. DOI: 10.1007/978-3-642-74626-0_2,10.3367/UFNr.0093.196710o.0388.

Freund, Y. and R. E. Schapire (1997). "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting." In: *Journal of Computer and System Sciences* 55.1, pp. 119 –139. ISSN: 0022-0000. DOI: https://doi.org/10.1006/jcss.1997.1504. URL: http://www.sciencedirect.com/science/article/pii/S002200009791504X.

Galbraith, W. and J. V. Jelley (Feb. 1953). "Light Pulses from the Night Sky associated with Cosmic Rays." In: *Nature* 171, pp. 349–350. DOI: 10.1038/171349a0.

Gammell, S. (Oct. 2004). "A Search for Very High Energy Gamma-ray Emission from Active Galactic Nuclei using Multivariate Analysis." PhD thesis. University College Dublin.

Gao, Y. and P. M. Solomon (May 2004). "The Star Formation Rate and Dense Molecular Gas in Galaxies." In: *The Astrophysical Journal* 606, pp. 271–290. DOI: 10.1086/382999. eprint: astro-ph/0310339.

Greisen, Kenneth (1966). "End to the Cosmic-Ray Spectrum?" In: *Phys. Rev. Lett.* 16 (17), pp. 748–750. DOI: 10.1103/PhysRevLett.16.748. URL: https://link.aps.org/doi/10.1103/PhysRevLett.16.748.

H. E. S. S. Collaboration et al. (2018). "The starburst galaxy NGC 253 revisited by H.E.S.S. and Fermi-LAT." In: *Astronomy & Astrophysics* 617, A73, A73. DOI: 10.1051/0004-6361/201833202. arXiv: 1806.03866 [astro-ph.HE].

Hanna, D. et al. (Jan. 2010). "An LED-based flasher system for VERITAS." In: *Nuclear Instruments and Methods in Physics Research A* 612, pp. 278–287. DOI: 10.1016/j.nima.2009.10.107. arXiv: 0911.2015 [astro-ph.IM].

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The elements of statistical learning: data mining, inference and prediction*. 2nd ed. Springer. URL: http://www-stat.stanford.edu/~tibs/ElemStatLearn/.

Hatlo, M. et al. (2005). "Developments of mathematical software libraries for the LHC experiments." In: *IEEE Trans. Nucl. Sci.* 52, pp. 2818–2822. DOI: 10.1109/TNS.2005.860152.

Hays, Elizabeth (2007). "VERITAS Data Acquisition." In: *Proceedings, 30th International Cosmic Ray Conference (ICRC 2007): Merida, Yucatan, Mexico, July 3-11, 2007*. Vol. 3, pp. 1543–1546. arXiv: 0710.2288 [astro-ph]. URL: http://indico.nucleares.unam.mx/contributionDisplay.py?contribId=1166&confId=4.

Heck, D. et al. (1998). "CORSIKA: A Monte Carlo code to simulate extensive air showers." In:

Heitler, W. (1954). *Quantum theory of radiation*, p. 386.

Helder, E. A. and J. Vink (2008). "Characterizing the Nonthermal Emission of Cassiopeia A." In: *The Astrophysical Journal* 686.2, pp. 1094–1102. DOI: 10.1086/591242. URL: https://doi.org/10.1086%2F591242.

Hess, V. F. and C. D. Anderson (2013). "Physics 1936." In: *Physics 1922–1941*. Elsevier, pp. 351 –377. ISBN: 978-1-4831-9745-6. DOI: https://doi.org/10.1016/B978-1-4831-9745-6.50019-2. URL: http://www.sciencedirect.com/science/article/pii/B9781483197456500192.

Hillas, A. M. (1982). "Angular and energy distributions of charged particles in electron-photon cascades in air." In: *Journal of Physics G: Nuclear Physics* 8.10, pp. 1461–1473. DOI: 10.1088/0305-4616/8/10/016. URL: https://doi.org/10.1088%2F0305-4616%2F8%2F10%2F016.

Hillas, A. M. (1984). "The Origin of Ultra-High-Energy Cosmic Rays." In: *Annual Review of Astronomy and Astrophysics* 22, pp. 425–444. DOI: 10.1146/annurev.aa.22.090184.002233.

Hillas, A. M. (Aug. 1985). "Cerenkov Light Images of EAS Produced by Primary Gamma Rays and by Nuclei." In: *International Cosmic Ray Conference* 3, p. 445.

Hillas, A. M. (2013). "Evolution of ground-based gamma-ray astronomy from the early days to the Cherenkov Telescope Arrays." In: *Astropart. Phys.* 43, pp. 19–43. DOI: 10.1016/j.astropartphys.2012.06.002.

Hinton, J. A. and the HESS Collaboration (Apr. 2004). "The status of the HESS project." In: *New Astronomy Reviews* 48, pp. 331–337. DOI: 10.1016/j.newar.2003.12.004. arXiv: astro-ph/0403052 [astro-ph].

Hoecker, A. et al. (Mar. 2007). "TMVA - Toolkit for Multivariate Data Analysis." In: *arXiv e-prints*, physics/0703039, physics/0703039. arXiv: physics/0703039 [physics.data-an].

Holder, J. (Jan. 2005). "Exploiting VERITAS Timing Information." In: *International Cosmic Ray Conference* 5, p. 383. arXiv: astro-ph/0507450 [astro-ph].

Holder, J. et al. (Dec. 2008). "Status of the VERITAS Observatory." In: *American Institute of Physics Conference Series*. Ed. by Felix A. Aharonian, Werner Hofmann, and Frank Rieger. Vol. 1085. American Institute of Physics Conference Series, pp. 657–660. DOI: 10.1063/1.3076760. arXiv: 0810.0474 [astro-ph].

Hörandel, Jörg R. (2004). "Models of the knee in the energy spectrum of cosmic rays." In: *Astroparticle Physics* 21.3, pp. 241–265. DOI: 10.1016/j.astropartphys.2004.01.004. arXiv: astro-ph/0402356 [astro-ph].

IceCube Collaboration et al. (July 2018). "Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert." In: *Science* 361, pp. 147–151. DOI: 10.1126/science.aat2890. arXiv: 1807.08794 [astro-ph.HE].

James, G. et al. (2014). *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated. ISBN: 1461471370, 9781461471370.

Jelley, J. V. (Nov. 1963). "Čerenkov Radiation: its Origin, Properties and Applications." In: *The Physics Teacher* 1, pp. 203–209. DOI: 10.1119/1.2350661.

Kearns, M. and L. G. Valiant (1989). "Crytographic Limitations on Learning Boolean Formulae and Finite Automata." In: *Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing*. STOC '89. Seattle, Washington, USA: ACM, pp. 433–444. ISBN: 0-89791-307-8. DOI: 10.1145/73007.73049. URL: http://doi.acm.org/10.1145/73007.73049.

Kennicutt Robert C., Jr. (1998). "The Global Schmidt Law in Star-forming Galaxies." In: *The Astrophysical Journal* 498.2, pp. 541–552. DOI: 10.1086/305588. arXiv: astro-ph/9712213 [astro-ph].

Krause, M., E. Pueschel, and G. Maier (2017). "Improved γ/hadron separation for the detection of faint γ-ray sources using boosted decision trees." In: *Astroparticle Physics* 89, pp. 1–9. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2017.01.004. URL: http://www.sciencedirect.com/science/article/pii/S0927650517300166.

Kubo, H. et al. (Apr. 2004). "Status of the CANGAROO-III project." In: *New Astronomy Reviews* 48, pp. 323–329. DOI: 10.1016/j.newar.2003.12.002.

Lacki, Brian C., Todd A. Thompson, and Eliot Quataert (2010). "The Physics of the Far-infrared-Radio Correlation. I. Calorimetry, Conspiracy, and Implications." In: *The Astrophysical Journal* 717.1, pp. 1–28. DOI: 10.1088/0004-637X/717/1/1. arXiv: 0907.4161 [astro-ph.CO].

Lacki, Brian C. et al. (2011). "On the GeV and TeV detections of the starburst galaxies M82 and NGC 253." In: *The Astrophysical Journal* 734.2, p. 107. DOI: 10.1088/0004-637x/734/2/107. URL: https://doi.org/10.1088%2F0004-637x%2F734%2F2%2F107.

Le Bohec, S. et al. (Oct. 1998). "A new analysis method for very high definition imaging atmospheric Cherenkov telescopes as applied to the CAT telescope." In: *Nuclear Instruments and Methods in Physics Research A* 416, pp. 425–437. DOI: 10.1016/S0168-9002(98)00750-5. arXiv: astro-ph/9804133 [astro-ph].

Li, T. P. and Y. Q. Ma (1983). "Analysis methods for results in gamma-ray astronomy." In: *The Astrophysical Journal* 272, pp. 317–324. DOI: 10.1086/161295.

Longair, M. S. (Feb. 2011). *High Energy Astrophysics*.

Lundberg, J. et al. (Mar. 2010). "Limits, discovery and cut optimization for a Poisson process with uncertainty in background and signal efficiency: TRolke 2.0." In: *Computer Physics Communications* 181, pp. 683–686. DOI: 10.1016/j.cpc.2009.11.001. arXiv: 0907.3450 [physics.data-an].

Lundberg, Scott M., Gabriel G. Erion, and Su-In Lee (Feb. 2018). "Consistent Individualized Feature Attribution for Tree Ensembles." In: *arXiv e-prints*, arXiv:1802.03888, arXiv:1802.03888. arXiv: 1802.03888 [cs.LG].

Maier, G. and J. Knapp (2007). "Cosmic-ray events as background in imaging atmospheric Cherenkov telescopes." In: *Astroparticle Physics* 28.1, pp. 72 –81. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2007.04.009. URL: http://www.sciencedirect.com/science/article/pii/S0927650507000606.

Mirzoyan, R. (2013). "The History of Ground-Based Very High Energy Gamma-Ray Astrophysics with the Atmospheric Air Cherenkov Telescope Technique." In: *Nucl. Phys. Proc. Suppl.* 239-240, pp. 26–34. DOI: 10.1016/j.nuclphysbps.2013.05.004.

Mitchell, T. M. (1997). *Machine Learning*. 1st ed. New York, NY, USA: McGraw-Hill, Inc. ISBN: 0070428077, 9780070428072.

Mohanty, G. et al. (1998). "Measurement of TeV gamma-ray spectra with the Cherenkov imaging technique." In: *Astropart. Phys.* 9, pp. 15–43. DOI: 10.1016/S0927-6505(98)00005-X.

Müller, S. et al. (2017). "Single Photon Extraction for FACT's SiPMs allows for Novel IACT Event Representation." In: *International Cosmic Ray Conference* 301, 801, p. 801.

Naylor, B. J. et al. (Oct. 2010). "A Census of the High-density Molecular Gas in M82." In: *The Astrophysical Journal* 722, pp. 668–681. DOI: 10.1088/0004-637X/722/1/668. arXiv: 1006.1964.

Nieto, D. et al. (2015). "Construction of a medium-sized Schwarzschild-Couder telescope as a candidate for the Cherenkov Telescope Array: development of the optical alignment system." In: *arXiv e-prints*, arXiv:1509.02463, arXiv:1509.02463. arXiv: 1509.02463 [astro-ph.IM].

Nieto, D. et al. (2017). "Prototype 9.7 m Schwarzschild-Couder telescope for the Cherenkov Telescope Array: status of the optical system." In: *arXiv e-prints*, arXiv:1709.06324, arXiv:1709.06324. arXiv: 1709.06324 [astro-ph.IM].

Ohm, S., C. van Eldik, and K. Egberts (2009). "γ/hadron separation in very-high-energy γ-ray astronomy using a multivariate analysis method." In: *Astroparticle Physics* 31.5, pp. 383 –391. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2009.04.001. URL: http://www.sciencedirect.com/science/article/pii/S0927650509000589.

Paglione, Timothy A. D. and Ryan D. Abrahams (2012). "PROPERTIES OF NEARBY STARBURST GALAXIES BASED ON THEIR DIFFUSE GAMMA-RAY EMISSION." In: *The Astrophysical Journal* 755.2, p. 106. DOI: 10.1088/0004-637x/755/2/106. URL: https://doi.org/10.1088%2F0004-637x%2F755%2F2%2F106.

Parsons, R. D. and J. A. Hinton (2014). "A Monte Carlo Template based analysis for Air-Cherenkov Arrays." In: *Astropart. Phys.* 56, pp. 26–34. DOI: 10.1016/j.astropartphys.2014.03.002. arXiv: 1403.2993 [astro-ph.IM].

Patterson, J. R. and A. M. Hillas (Nov. 1983). "The relation of the lateral distribution of Cerenkov light from cosmic-ray showers to the distance of maximum development." In:

Journal of Physics G Nuclear Physics 9, pp. 1433–1452. DOI: 10.1088/0305-4616/9/11/015.

Persic, Massimo and Yoel Rephaeli (2014). "Estimates of relativistic electron and proton energy densities in starburst galactic nuclei from radio measurements." In: *Astronomy & Astrophysics* 567, A101, A101. DOI: 10.1051/0004-6361/201322664. arXiv: 1201.0369 [astro-ph.HE].

Pohl, M. (July 1994). "On the predictive power of the minimum energy condition. 2: Fractional calorimeter behaviour in the diffuse high energy gamma emission of spiral galaxies." In: *Astronomy & Astrophysics* 287, pp. 453–462.

Ratliff, G. (2015). "The Very Energetic Radiation Imaging Telescope Array System Observations of the Starburst Galaxy M82." PhD thesis. Illinois Institute of Technology.

Reynolds, P. T. et al. (Feb. 1993). "Survey of candidate gamma-ray sources at TeV energies using a high-resolution Cerenkov imaging system - 1988-1991." In: *The Astrophysical Journal* 404, pp. 206–218. DOI: 10.1086/172269.

Robitaille, Thomas P. and Barbara A. Whitney (2010). "The Present-Day Star Formation Rate of the Milky Way Determined from Spitzer-Detected Young Stellar Objects." In: *The Astrophysical Journal* 710.1, pp. L11–L15. DOI: 10.1088/2041-8205/710/1/L11. arXiv: 1001.3672 [astro-ph.GA].

Rojas-Bravo, César and Miguel Araya (Aug. 2016). "Search for gamma-ray emission from star-forming galaxies with Fermi LAT." In: *Monthly Notices of the Royal Astronomical Society* 463.1, pp. 1068–1073. ISSN: 0035-8711. DOI: 10.1093/mnras/stw2059. eprint: http://oup.prod.sis.lan/mnras/article-pdf/463/1/1068/18473067/stw2059.pdf. URL: https://doi.org/10.1093/mnras/stw2059.

Rousselle, J. et al. (2015). "Construction of a Schwarzschild-Couder telescope as a candidate for the Cherenkov Telescope Array: status of the optical system." In: *arXiv e-prints*, arXiv:1509.01143, arXiv:1509.01143. arXiv: 1509.01143 [astro-ph.IM].

Sakai, S. and B. F. Madore (Dec. 1999). "Detection of the Red Giant Branch Stars in M82 Using the Hubble Space Telescope." In: *The Astrophysical Journal* 526, pp. 599–606. DOI: 10.1086/308032. eprint: astro-ph/9906484.

Schapire, R. E. (1990). "The strength of weak learnability." In: *Machine Learning* 5.2, pp. 197–227. ISSN: 1573-0565. DOI: 10.1007/BF00116037. URL: https://doi.org/10.1007/BF00116037.

Schiel, R. W. and J. P. Ralston (2007). "The Greisen equation explained and improved." In: *Phys. Rev. D* 75 (1), p. 016005. DOI: 10.1103/PhysRevD.75.016005. URL: https://link.aps.org/doi/10.1103/PhysRevD.75.016005.

Shopbell, P. L. and J. Bland-Hawthorn (Jan. 1998). "The Asymmetric Wind in M82." In: *The Astrophysical Journal* 493, pp. 129–153. DOI: 10.1086/305108. eprint: astro-ph/9708038.

Sokolsky, P. (2011). "Final Results from the High solution Fly s Eye (HiRes) Experiment." In: *Nuclear Physics B - Proceedings Supplements* 212-213. Proceedings of the Cosmic Ray International Seminars (CRIS 2010), pp. 74 –78. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2011.03.010. URL: http://www.sciencedirect.com/science/article/pii/S0920563211000430.

Strickland, D. K. and T. M. Heckman (June 2009). "Supernova Feedback Efficiency and Mass Loading in the Starburst and Galactic Superwind Exemplar M82." In: *The Astrophysical Journal* 697, pp. 2030–2056. DOI: 10.1088/0004-637X/697/2/2030. arXiv: 0903.4175 [astro-ph.CO].

Strickland, D. K., T. J. Ponman, and I. R. Stevens (Apr. 1997). "ROSAT observations of the galactic wind in M 82." In: *Astronomy & Astrophysics* 320, pp. 378–394. eprint: astro-ph/9608064.

Tamm, I. E. (1939). "Radiation Emitted by Uniformly Moving Electrons." In: *Journal of Physics USSR* 1, pp. 439–454. DOI: 10.1007/978-3-642-74626-0_3.

Tammann, G. A., W. Loeffler, and A. Schroeder (June 1994). "The Galactic supernova rate." In: *The Astrophysical Journal Supplement Series* 92, pp. 487–493. DOI: 10.1086/192002.

VERITAS Collaboration et al. (Dec. 2009). "A connection between star formation activity and cosmic rays in the starburst galaxy M82." In: *Nature* 462, pp. 770–772. DOI: 10.1038/nature08557. arXiv: 0911.0873 [astro-ph.CO].

Vassiliev, V., S. Fegan, and P. Brousseau (Sept. 2007). "Wide field aplanatic two-mirror telescopes for ground-based γ-ray astronomy." In: *Astroparticle Physics* 28, pp. 10–27. DOI: 10.1016/j.astropartphys.2007.04.002. eprint: astro-ph/0612718.

Vassiliev, V. and CTA pSCT Project (2017). "Prototype 9.7m Schwarzschild-Couder telescope for the Cherenkov Telescope Array: Project Overview." In: *International Cosmic Ray Conference* 301, 838, p. 838.

Vincent, S. (2016). "A Monte Carlo template-based analysis for very high definition imaging atmospheric Cherenkov telescopes as applied to the VERITAS telescope array." In: *PoS* ICRC2015, p. 844. DOI: 10.22323/1.236.0844. arXiv: 1509.01980 [astro-ph.IM].

Weekes, T. C. (2005). "The Atmospheric Cerenkov imaging technique for very high energy gamma-ray astronomy." In: *International WE - Heraeus Summer School: Physics with Cosmic Accelerators Bad Honnef, Germany, July 5-16, 2004*. arXiv: astro-ph/0508253 [astro-ph].

Weekes, T. C. and K. E. Turver (July 1977). "Gamma-Ray Astronomy from 10-100 GeV: a New Approach." In: *Recent Advances in Gamma-Ray Astronomy*. Ed. by R. D. Wills and B. Battrick. Vol. 124. ESA Special Publication, p. 279.

Weekes, T. C. et al. (July 1989). "Observation of TeV Gamma Rays from the Crab Nebula Using the Atmospheric Cerenkov Imaging Technique." In: *The Astrophysical Journal* 342, p. 379. DOI: 10.1086/167599.

Weinstein, A. (2007). "The VERITAS Trigger System." In: 3, p. 1539. arXiv: 0709.4438 [astro-ph].

Wilson, J. G. and K. Greisen (1956). Vol. 3.

Wood, M. et al. (2016). "Monte Carlo studies of medium-size telescope designs for the Cherenkov Telescope Array." In: *Astroparticle Physics* 72, pp. 11–31. DOI: 10.1016/j.astropartphys.2015.04.008. arXiv: 1506.07476 [astro-ph.IM].

Yoast-Hull, Tova M. et al. (2013). "Winds, Clumps, and Interacting Cosmic Rays in M82." In: *Astrophysical Journa* 768.1, 53, p. 53. DOI: 10.1088/0004-637X/768/1/53. arXiv: 1303.4305 [astro-ph.HE].

Zatsepin, G. T. and V. A. Kuz'min (Aug. 1966). "Upper Limit of the Spectrum of Cosmic Rays." In: *Soviet Journal of Experimental and Theoretical Physics Letters* 4, p. 78.

Zitzer, B. for the VERITAS Collaboration (July 2013). "The VERITAS Upgraded Telescope-Level Trigger Systems: Technical Details and Performance Characterization." In: *arXiv e-prints*, arXiv:1307.8360, arXiv:1307.8360. arXiv: 1307.8360 [astro-ph.IM].

de Cea del Pozo, Elsa, Diego F. Torres, and Ana Y. Rodriguez Marrero (2009). "Multimessenger Model for the Starburst Galaxy M82." In: *The Astrophysical Journal* 698.2, pp. 1054–1060. DOI: 10.1088/0004-637X/698/2/1054. arXiv: 0901.2688 [astro-ph.GA].

de Naurois, M. and L. Rolland (Dec. 2009). "A high performance likelihood reconstruction of γ-rays for imaging atmospheric Cherenkov telescopes." In: *Astroparticle Physics* 32, pp. 231–252. DOI: 10.1016/j.astropartphys.2009.09.001. arXiv: 0907.2610 [astro-ph.IM].

Appendix

Optimized BDT Cuts

Table .1: BDT cuts for OA ATM21.

	Soft			Medium		
	Strong	Weak	Detection	Strong	Weak	Detection
BDTScoreLower00	0.926524	0.958173	0.962639	0.649633	0.695626	0.695619
BDTScoreLower01	0.965279	0.97834	0.98133	0.952584	0.973186	0.980221
BDTScoreLower02	0.968043	0.985154	0.995887	0.962756	0.983604	0.993538
BDTScoreLower03	0.968244	0.982672	0.991293	0.960831	0.981202	0.990464
BDTScoreLower10	0.953485	0.962043	0.963873	0.844047	0.936271	0.936203
BDTScoreLower11	0.975401	0.982673	0.984084	0.966615	0.98134	0.983718
BDTScoreLower12	0.968942	0.976087	0.980624	0.966876	0.974984	0.980294
BDTScoreLower13	0.900357	0.943999	0.968413	0.891111	0.936035	0.964891
BDTScoreLower20	0.942554	0.954525	0.955432	0.895345	0.9278	0.927754
BDTScoreLower21	0.978658	0.985238	0.987866	0.97511	0.983503	0.987295
BDTScoreLower22	0.920227	0.974317	0.985556	0.90934	0.972758	0.984746
BDTScoreLower30	0.969978	0.981283	0.981349	0.8722	0.891304	0.891261
BDTScoreLower31	0.974455	0.984292	0.98609	0.970764	0.979267	0.983318
BDTScoreLower32	0.975691	0.984928	0.989292	0.971962	0.983516	0.988636

Table .2: BDT cuts for OA ATM22.

	Soft			Medium		
	Strong	Weak	Detection	Strong	Weak	Detection
BDTScoreLower00	0.889815	0.923454	0.925592	0.780059	0.813718	0.813608
BDTScoreLower01	0.972245	0.980748	0.983159	0.944466	0.977424	0.982095
BDTScoreLower02	0.965939	0.977204	0.982593	0.961158	0.975349	0.982249
BDTScoreLower03	0.880716	0.938622	0.973968	0.876316	0.912506	0.973086
BDTScoreLower10	0.96134	0.975412	0.980415	0.897576	0.949583	0.949518
BDTScoreLower11	0.977594	0.986199	0.988143	0.962486	0.984071	0.987603
BDTScoreLower12	0.969714	0.980718	0.984426	0.966917	0.979534	0.984149
BDTScoreLower13	0.938208	0.970805	0.979524	0.905718	0.9693	0.978823
BDTScoreLower20	0.955277	0.96487	0.966369	0.898994	0.945767	0.945755
BDTScoreLower21	0.969612	0.980294	0.984912	0.96465	0.977101	0.984263
BDTScoreLower22	0.953588	0.971776	0.984174	0.949359	0.969249	0.983437
BDTScoreLower30	0.976909	0.985122	0.98514	0.888817	0.902276	0.902251
BDTScoreLower31	0.981281	0.986106	0.987465	0.974359	0.984307	0.985716
BDTScoreLower32	0.962641	0.973871	0.981172	0.956896	0.97275	0.978718

Table .3: BDT cuts for NA ATM21.

	Soft			Medium		
	Strong	Weak	Detection	Strong	Weak	Detection
BDTScoreLower00	0.900043	0.954292	0.968147	0.840165	0.893445	0.893445
BDTScoreLower01	0.94283	0.974915	0.981828	0.927093	0.96725	0.980895
BDTScoreLower02	0.962574	0.981595	0.985756	0.955651	0.980359	0.985362
BDTScoreLower03	0.967831	0.979801	0.985323	0.963349	0.978215	0.984964
BDTScoreLower04	0.904686	0.931751	0.964885	0.894671	0.928673	0.963526
BDTScoreLower10	0.753464	0.929598	0.933515	-0.73245	-0.421189	-0.421236
BDTScoreLower11	0.938155	0.962066	0.970217	0.894771	0.946223	0.965261
BDTScoreLower12	0.96171	0.981183	0.984214	0.957401	0.980088	0.983776
BDTScoreLower13	0.964399	0.980292	0.988316	0.956909	0.977109	0.987785
BDTScoreLower14	0.91666	0.966734	0.984018	0.904883	0.958967	0.982976
BDTScoreLower20	0.893198	0.944403	0.952786	0.837467	0.897038	0.940459
BDTScoreLower21	0.971612	0.982733	0.986665	0.964876	0.980206	0.985752
BDTScoreLower22	0.97416	0.986022	0.991684	0.972263	0.985169	0.99113
BDTScoreLower23	0.956808	0.974159	0.981129	0.953382	0.971508	0.980607
BDTScoreLower30	0.954394	0.976703	0.979978	0.926464	0.965238	0.970996
BDTScoreLower31	0.978984	0.985811	0.988375	0.970666	0.983802	0.987484
BDTScoreLower32	0.961529	0.979971	0.99109	0.959028	0.978099	0.990416

Table .4: BDT cuts for NA ATM22.

	Soft			Medium		
	Strong	Weak	Detection	Strong	Weak	Detection
BDTScoreLower00	0.83698	0.886641	0.920726	0.520345	0.819668	0.822098
BDTScoreLower01	0.937748	0.977455	0.982373	0.928457	0.96631	0.981241
BDTScoreLower02	0.959867	0.976274	0.981397	0.955153	0.974941	0.980872
BDTScoreLower03	0.924124	0.974016	0.981294	0.921423	0.970916	0.980721
BDTScoreLower04	0.938552	0.973614	0.986172	0.930692	0.967956	0.985526
BDTScoreLower10	0.756224	0.82112	0.822556	-0.390542	-0.388327	-0.388371
BDTScoreLower11	0.953753	0.977524	0.982476	0.914765	0.965696	0.97943
BDTScoreLower12	0.96717	0.978868	0.98336	0.961447	0.977276	0.982806
BDTScoreLower13	0.971765	0.982872	0.986886	0.968281	0.98132	0.986545
BDTScoreLower14	0.935682	0.979514	0.986677	0.920229	0.970991	0.986171
BDTScoreLower20	0.938815	0.958312	0.967512	0.893001	0.944605	0.955044
BDTScoreLower21	0.970519	0.986227	0.991008	0.961856	0.982403	0.98995
BDTScoreLower22	0.974179	0.982851	0.98765	0.972999	0.98199	0.987311
BDTScoreLower23	0.947946	0.979567	0.98996	0.938076	0.976754	0.98931
BDTScoreLower30	0.984509	0.987726	0.988483	0.980596	0.986184	0.986717
BDTScoreLower31	0.986202	0.990943	0.994717	0.983444	0.989171	0.992721
BDTScoreLower32	0.966562	0.985222	0.992748	0.964392	0.983707	0.992196

Table .5: BDT cuts for UA ATM21.

	Soft			Medium		
	Strong	Weak	Detection	Strong	Weak	Detection
BDTScoreLower00	0.838579	0.909248	0.921223	0.466544	0.62907	0.629004
BDTScoreLower01	0.924029	0.971196	0.977181	0.890666	0.9674	0.976141
BDTScoreLower02	0.951238	0.978385	0.987236	0.942359	0.976813	0.986637
BDTScoreLower03	0.897494	0.960551	0.982631	0.862083	0.957067	0.98151
BDTScoreLower04	0.867465	0.927581	0.952334	0.859614	0.916212	0.950698
BDTScoreLower10	0.825271	0.904226	0.904146	0	-0.0082813	-0.0082752
BDTScoreLower11	0.911211	0.963636	0.977412	0.869517	0.944818	0.974914
BDTScoreLower12	0.937816	0.967767	0.982491	0.931187	0.962658	0.981889
BDTScoreLower13	0.939839	0.974106	0.986666	0.929535	0.970724	0.986027
BDTScoreLower14	0.944304	0.966933	0.980053	0.939593	0.961091	0.979239
BDTScoreLower20	0.901068	0.951255	0.962723	0.335301	0.542931	0.54226
BDTScoreLower21	0.95549	0.980973	0.987186	0.940721	0.967202	0.984854
BDTScoreLower22	0.958067	0.98154	0.987776	0.950611	0.977789	0.987398
BDTScoreLower23	0.92634	0.959018	0.978898	0.901724	0.952716	0.97792
BDTScoreLower30	0.970792	0.981179	0.983922	0.833485	0.968929	0.968888
BDTScoreLower31	0.973791	0.984075	0.98792	0.965973	0.980766	0.987193
BDTScoreLower32	0.959102	0.978234	0.98288	0.955278	0.977553	0.982447

Table .6: BDT cuts for UA ATM22.

	Soft			Medium		
	Strong	Weak	Detection	Strong	Weak	Detection
BDTScoreLower00	0.848848	0.917469	0.930758	0.522352	0.522468	0.52237
BDTScoreLower01	0.930809	0.967907	0.980758	0.878428	0.958361	0.979572
BDTScoreLower02	0.940714	0.971551	0.979126	0.932726	0.969256	0.978587
BDTScoreLower03	0.92822	0.972939	0.991014	0.922047	0.962414	0.990255
BDTScoreLower04	0.896423	0.93901	0.960184	0.884418	0.936484	0.959098
BDTScoreLower10	0.771354	0.926868	0.926868	-0.0100622	-0.0100745	-0.0100751
BDTScoreLower11	0.926171	0.963544	0.978503	0.890055	0.954006	0.976411
BDTScoreLower12	0.938958	0.976149	0.985039	0.929422	0.973863	0.984572
BDTScoreLower13	0.949601	0.970563	0.989145	0.93762	0.968646	0.988499
BDTScoreLower14	0.895491	0.957655	0.978398	0.887369	0.953695	0.97692
BDTScoreLower20	0.92016	0.956064	0.967003	0.580576	0.580714	0.580677
BDTScoreLower21	0.965193	0.979417	0.98521	0.937905	0.970682	0.983045
BDTScoreLower22	0.968074	0.984549	0.988729	0.962841	0.983104	0.9884
BDTScoreLower23	0.919294	0.957704	0.985142	0.909107	0.951076	0.98312
BDTScoreLower30	0.964019	0.979495	0.982586	0.870034	0.954801	0.954689
BDTScoreLower31	0.97699	0.986047	0.989098	0.968913	0.984051	0.988393
BDTScoreLower32	0.959171	0.976372	0.984796	0.952956	0.974552	0.984231