A Mono Reconstruction Software for Phase II of the H.E.S.S. Experiment

MASTERARBEIT
zur Erlangung des akademischen Grades
Master of Science
im Fach Physik
eingereicht am
Institut für Physik
Mathematisch-Wissenschaftliche Fakultät I
Humboldt-Universität zu Berlin

von
Thomas Murach
geboren am 09. Januar 1988 in Rüdersdorf

Gutachter:
1. Prof. Dr. Thomas Lohse
2. Prof. Dr. Alexander Kappes

eingereicht am: 1.10.2012 überarbeitete Version
# Contents

1 Introduction 1

2 Gamma-Ray Astronomy 3
   2.1 Sources of Charged VHE Particles 3
      2.1.1 Pulsar Wind Nebulae 3
      2.1.2 Active Galactic Nuclei 3
      2.1.3 Supernova Remnants 4
   2.2 Gamma-Ray Production 4
      2.2.1 Leptonic Processes 4
      2.2.2 Hadronic Processes 5
   2.3 Propagation of Gamma Rays through the Universe 6
   2.4 Air Showers 6

3 The H.E.S.S. Experiment 9
   3.1 Experimental Setup 9
      3.1.1 Telescope Structure 9
      3.1.2 The Mirrors 10
      3.1.3 The Cameras 10
      3.1.4 The Trigger System 11
   3.2 Performance 12
   3.3 Event Simulation 13
   3.4 Image Processing and Parameterisation 14
      3.4.1 Image Cleaning 14
      3.4.2 Hillas Parameters 15

4 Multivariate Analysis Methods 17
   4.1 Toolkit for Multivariate Analysis 18
   4.2 Boosted Decision Trees 19
      4.2.1 Basic Concept 19
      4.2.2 Settings 21
   4.3 Multilayer Perceptrons 22
      4.3.1 Basic Concept 22
      4.3.2 Settings 24

5 Mono Reconstruction in Phase II 27
   5.1 Event Selection Cuts 27
   5.2 Calculation of the Simulated Time and Weights of Proton Data 32
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3 Direction Reconstruction</td>
<td>32</td>
</tr>
<tr>
<td>5.4 Energy Reconstruction</td>
<td>37</td>
</tr>
<tr>
<td>5.5 Particle Identification</td>
<td>40</td>
</tr>
<tr>
<td>6 Performance of the Mono Analysis</td>
<td>47</td>
</tr>
<tr>
<td>6.1 Angular Resolution</td>
<td>47</td>
</tr>
<tr>
<td>6.2 Energy Reconstruction Performance</td>
<td>50</td>
</tr>
<tr>
<td>6.3 Particle Identification Performance</td>
<td>53</td>
</tr>
<tr>
<td>6.4 Effective Areas and Event Rates</td>
<td>58</td>
</tr>
<tr>
<td>6.5 Sensitivities</td>
<td>61</td>
</tr>
<tr>
<td>7 Summary and Outlook</td>
<td>67</td>
</tr>
<tr>
<td>8 Zusammenfassung und Ausblick</td>
<td>69</td>
</tr>
<tr>
<td>A Appendix</td>
<td>71</td>
</tr>
<tr>
<td>A.1 Width and Length Lookup Tables</td>
<td>71</td>
</tr>
<tr>
<td>A.2 Simulation Time Calculation for Proton MC Events</td>
<td>71</td>
</tr>
<tr>
<td>A.3 MLP Input Variable Distributions</td>
<td>74</td>
</tr>
<tr>
<td>A.4 BDT Input Variable and Output Distributions</td>
<td>76</td>
</tr>
</tbody>
</table>
1 Introduction

Gamma ray astronomy is a modern scientific field of research combining particle physics, astronomy and astrophysics. Some main aims of gamma ray astronomy are the discovery of new gamma ray sources and the investigation of cosmic acceleration processes which are able to produce electrons and protons with energies of up to $O(100\text{ TeV})$, which is far beyond the energies accessible by particle accelerators on earth. According to current models, these processes happen in the most extreme regions in the universe, such as SuperNova Remnants (SNR), Active Galactic Nucleae (AGN) or pulsars. Gamma rays with energies comparable to the energies of the charged particles accelerated in such regions can be produced by these due to different processes described in more detail in Chap. 2. Such Very High Energy (VHE, i.e. with energies between 30 GeV and 30 TeV according to the conventional definition of Aharonian [1]) particles reach the earth’s atmosphere, where they interact with atmospheric particles, leading to the production of air showers. In the evolution of such showers, Cherenkov light is produced. By detecting and analysing this Cherenkov light, many important properties like the energy, the direction or the type of the incident particle can be reconstructed.

Modern gamma ray astronomy observatories like H.E.S.S. [2], MAGIC [3] or VERITAS [4] are able to perform such measurements by collecting Cherenkov light produced in air showers with mirrors, projecting the collected light onto a camera. Hence these experiments are commonly called Imaging Atmospheric Cherenkov Telescopes (IACTs). The parameters of the recorded images are the basis for many event analysis techniques.

This thesis deals with the H.E.S.S. telescope array in its phase II, which has started in summer 2012. The telescope array is situated in the Khomas Highland in Namibia. In phase I, this telescope array consisted of four telescopes called CT1–CT4, but recently it was enhanced by a fifth telescope referenced as CT5, initiating the experiment’s phase II. As CT5 has a very large mirror area of about 600 m², compared to about 100 m² for each of the telescopes of H.E.S.S. phase I, the performance of the H.E.S.S. experiment is going to be improved significantly. This holds true especially for events with low primary particle energies, as CT5 is going to lower the energy threshold from about 100 GeV to approximately 30 GeV. Due to the steep decline of the energy spectrum of cosmic ray particles about 60% of all events that can be detected with H.E.S.S. phase II will not be detected by CT1–CT4. Therefore it will be very important to possess a powerful analysis being capable of performing energy and shower direction reconstruction and particle identification. Especially the particle identification is a very important task, as only gamma rays serve as signal events, whereas protons provide the by far predominant background. Background events have to be identified and suppressed efficiently, which is a difficult task for low energy events. Concerning the energy resolution and the particle discrimination, this thesis is also meant as a crosscheck for the work of Masbou [5].
1 Introduction

In this thesis, studies concerning a reconstruction algorithm for CT5 mono events are presented, i.e. events which are only detected by CT5, whereas trigger information of CT1–CT4 is not considered at all. Some aims of the analysis presented in this thesis are to provide a good separation between incident photons and protons as primary particles, as well as providing a direction and an energy reconstruction using advanced methods like neural networks or boosted decision trees. In this analysis, only events simulated using Monte Carlo techniques are used, real data are not utilised.

In Chap. 2, the physical aspects of gamma ray astronomy are explained, including acceleration mechanisms of charged particles, VHE gamma ray creation, the physics of air showers and dominant decay and interaction processes.

The H.E.S.S. experiment is described in detail in Chap. 3. The telescope structures, mirrors, cameras and triggers are explained, also the performance of the experiment is described briefly. At the end of that chapter, the event simulation and the processing of images recorded by the H.E.S.S. telescopes are discussed.

In Chap. 4, the fundamentals of multivariate analysis methods are described, as such methods provide an important tool for the analysis presented here. The utilised methods called “boosted decision trees” and “multilayer perceptron” are introduced and explained.

Chap. 5 covers the conceptual approaches to the main goals of this analysis, including the key aspects like energy and direction reconstruction or particle identification. In Chap. 6, the performance of the methods introduced in the previous chapter is discussed.

Finally, in Chap. 7, the results are summarised and an outlook is given concerning possible improvements of the analysis. In Chap. 8, the same is done in German language.
2 Gamma-Ray Astronomy

In this chapter, the basic physical processes needed for the understanding of gamma ray astronomy are explained. This includes acceleration of charged particles, gamma ray production, propagation through the universe and the interaction with the atmosphere of the earth.

2.1 Sources of Charged VHE Particles

The acceleration of charged particles like electrons and positrons (from now on called electrons altogether, if not stated otherwise) or protons is, according to current models, a necessary step for the creation of VHE gamma rays. A selection of the most important acceleration sites and processes is described in this section.

2.1.1 Pulsar Wind Nebulae

Pulsar Wind Nebulae (PWN) are an important site of particle acceleration. The famous Crab Nebula is an example of such a source, being the first PWN detected in the TeV gamma ray regime [6]. Pulsars are fast rotating neutron stars with a mass bigger than the mass of the sun and a radius of some 10 km, thus being very compact objects. They lose their rotational energy mostly by accelerating a wind of charged particles up to relativistic velocities [1]. This wind is flowing away from the pulsar into the InterStellar Medium (ISM), creating a magnetic field. In case of the Crab Nebula, the mean field strength is on the order of $100 \mu$G. When the pulsar wind hits the ISM, the wind is shocked, meaning that it loses a significant fraction of its energy to the particles of the ISM, which are accelerated in the magnetic field created at the shock front up to energies of about $10^{16}$ eV. The largest part of the wind’s energy is carried away by electrons which then produce gamma rays via mechanisms explained in Sec. 2.2.

2.1.2 Active Galactic Nuclei

Another important source class are AGN, which are compact, extragalactic and very luminous active nuclei of galaxies [1]. They make up the largest fraction of sources detected in the TeV regime so far [6]. AGN are divided into several subclasses, like blazars or BL Lac objects. Mostly the only difference between these objects is the relative direction of the jet with respect to our observation direction [1], therefore the acceleration principle is described only once.

AGN are effective gamma ray emitters, delivering interesting temporal and spectral information. Depending on the direction of the jet, the gamma rays emitted from these
objects can be strongly Doppler-boosted, resulting in high gamma ray fluxes at very high energies. It is believed that electrons are dominantly responsible for creation of VHE gamma rays in AGN, although hadronic scenarios are possible as well. Especially gamma rays at extremely high energies between $10^{15}$ eV and $10^{19}$ eV can only be produced by hadronic processes.

There is another possible class of sites of gamma ray creation linked to AGN, called large scale AGN jets. These can be found in congregations of galaxies called galaxy clusters. The jets of single AGN are likely to point towards the inner space of the clusters, heating up the medium inside it. At the shock fronts of the jets, where the jets hit the medium, particles can be accelerated up to energies of $10^{20}$ eV [1].

### 2.1.3 Supernova Remnants

SNRs are an important source of cosmic rays with energies of up to $10^{15}$ eV [1], being also visible in gamma rays. SNR are extended, galactic sources with an interesting temporal development. Concerning VHE gamma rays, they reach their highest luminosity some 1000–10 000 years after the supernova explosion, when the number of relativistic particles reaches its maximum value. The charged particles are accelerated by diffuse shock wave, similar to the mechanism described above.

### 2.2 Gamma-Ray Production

Gamma rays are created in non-thermal processes, like scattering or interactions of charged particles with electromagnetic fields. Some dominant processes are explained in this section.

The gamma ray production processes are classified as leptonic and hadronic processes, depending on the type of incident VHE particle. The most important processes of both groups are explained briefly in the following.

#### 2.2.1 Leptonic Processes

The electrons which have been accelerated by processes described in the last section can produce VHE gamma rays in many ways. Inverse Compton scattering (IC) between these electrons and low energy photons is a very important leptonic process for VHE gamma ray production. The photons mostly stem from the Cosmic Microwave Background (CMB) (cf. Cui [6] and references therein). Such an IC process is illustrated in Fig. 2.1. Due to the kinematics of such scattering processes, the photons gain a significant fraction of the energy of the incident electron [1].

VHE gamma rays can also be produced by electron bremsstrahlung, electron synchrotron radiation and electron-positron annihilation, which are common mechanisms as well. Especially the last process can produce very high energy gamma rays if the particles are strongly boosted, while bremsstrahlung requires high gas densities to be efficient and extremely strong magnetic fields are required in order to produce synchrotron radi-
2.2 Gamma-Ray Production

\section*{2.2 Gamma-Ray Production}

\subsection*{2.2.2 Hadronic Processes}

There is a general mechanism of the production of VHE gamma rays by hadronic processes. A proton gains its energy in sources like the ones introduced in Sec. 2.1 and collides with matter or photons, which in turn leads to the production of neutral pions ($\pi^0$). With a likelihood of about 98.8\% [7], these pions decay into two photons with a mean life time of $8.3 \times 10^{-17}$ s in the rest frame of the $\pi^0$. Such a decay process is shown in Fig. 2.2.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2.2}
\caption{Feynman diagram of a $\pi^0$-decay, with the time axis in horizontal direction.}
\end{figure}

There is a wealth of possible interactions between protons and photons as well as for protons colliding with matter. An inelastic scattering process of a proton and a nucleus of a hydrogen atom of the ISM could look like the following process:

$$p + \gamma \rightarrow X + n \cdot \pi^0.$$  \hfill (2.1)

Here a particle or even a multitude of particles denoted as $X$ are produced along with one or more neutral pions, the decays of which lead to the production of gamma rays. The pions are not created at rest in our laboratory frame, hence photons with much higher
energies than half the $\pi^0$ mass are observed on earth. According to current models it is believed that photons from $\pi^0$ decays can have energies as high as $100$ TeV [1].

There are many other gamma-ray production mechanisms, like electron-positron pair-production in proton-photon collisions or proton synchrotron radiation, which are not going to be explained in detail. For more information, cf. for instance Aharonian [1].

2.3 Propagation of Gamma Rays through the Universe

After the production of VHE gamma rays at the source, the gamma rays observable on earth have to propagate through space for long distances. On their way there are many possibilities of absorption or scattering processes leading to a reduced number of VHE gamma rays which reach the earth. The most important one of these processes is electron pair production by scattering of the VHE gamma rays on low energy photons as shown in Fig. 2.3. Most of these are photons of the CMB, but especially near compact objects, the scattering on photons stemming from these objects dominates. In the energy range that is interesting for the H.E.S.S. experiment, the mean free path of photons decreases with energy, being in the order of $100$ Mpc for a photon with an energy of $1$ TeV [8].

2.4 Air Showers

VHE particles like photons, protons or electrons emitted as described above reach the earth’s atmosphere, which is opaque for particles in that energy domain. Due to inelastic scattering on nuclei of atoms in the air, electromagnetic or hadronic air showers are initiated [9].

If electrons or photons hit the atmosphere, electromagnetic air showers are initiated. The Heitler model [10] gives a simplistic view of this process. Neglecting the difference between the radiation length of electrons and the mean free path of gamma rays in the atmosphere, the model describes an electromagnetic shower as an alternating chain of
Figure 2.4: Simulated electromagnetic and hadronic air showers. The shower shown on the left was initiated by a 0.3 TeV photon, the shower on the right was produced by a 1 TeV proton \[12\].

Pair production and bremsstrahlung emissions, doubling the number of particles after each step. This process continues until the energy of the particles is too low to produce further electron-positron pairs. After reaching the shower maximum at a height of approx. 10 km above sea level in case of a 1 TeV photon \[11\], the number of particles decreases due to absorption processes.

If protons or atomic nuclei hit the atmosphere, they interact with atmospheric nuclei due to strong interactions, producing hadronic showers. In general pions are produced in such interactions, one third of them being neutral pions. These decay, as mentioned above, mostly into two photons, producing electromagnetic subshowers.

Due to the occurrence of multiple interactions, hadronic showers are much more irregular than electromagnetic showers, as shown exemplarily in Fig. 2.4. On the right side of the picture, a hadronic shower initiated by a proton with a primary energy of 1 TeV is shown, while a shower from a gamma ray with an energy of 0.3 TeV is shown on the left. The energies are chosen this way because on average, only one third of the energy of the proton is released in electromagnetic subshowers which are detectable by IACTs. Electromagnetic showers are more regular because most particles in them have low masses and are therefore strongly boosted along the track of the incident particle, and only Coulomb scattering is the major way of deflecting particles. In case of hadronic showers, the secondary particles are more massive and are created in inelastic scattering processes, and therefore gain a higher transverse momentum with respect to the direction of the primary particle \[9\].

The primary particles are very energetic, hence the secondary particles have very
high energies as well. Although the optical refraction index of air is very close to one ($n_{\text{air}} \approx 1.0003$ at normal air pressure of 1013.25 hPa and a temperature of $15{\degree}\text{C}$ [13]), the secondary particles are mostly faster than light is in air, hence charged particles produce Cherenkov light due to polarisation of air atoms and molecules. The Cherenkov radiation is emitted under an angle $\Theta_C$ with respect to the direction of flight of the relativistic particle. This angle is defined by the equation

$$\cos(\Theta_C) = \frac{c}{vn}. \tag{2.2}$$

As the refractive index of air $n$ increases as the air shower travels through the atmosphere down to lower altitudes and the velocity of the particles approximately equals the speed of light $c$, the $\Theta_C$ angle increases as the shower develops. Also the fact that the divergence of the tracks of the charged particles from the direction of the incident particle increases with decreasing altitude leads to higher Cherenkov light emission angles at lower altitudes.

The Cherenkov light can be detected by IACTs, as explained in more detail in Chap. 3. Due to the greater irregularity of hadronic showers, the Cherenkov light distribution at ground level is less homogeneous than in case of electromagnetic showers, as shown in Fig. 2.5. This information is going to be used in the analysis presented in this thesis.
3 The H.E.S.S. Experiment

In phase I, the H.E.S.S. experiment consisted of four Imaging Atmospheric Cherenkov Telescopes (IACTs) called CT 1–CT 4. They have been arranged as a square with a side length of 120 m [15]. The construction of the first of these telescopes was finished in June 2002 [16], the last of these telescopes was ready for data taking in December 2003 [17]. In summer 2012 the construction of a fifth telescope called CT 5 was finished. This new telescope is located in the centre of the square of the four already existing telescopes.

In this chapter, the experimental setup and the performance of the H.E.S.S. experiment are described. As this thesis is about a reconstruction software for CT 5 only, stronger emphasis is put on the description of the new telescope.

3.1 Experimental Setup

The H.E.S.S. experiment is located in the Khomas Highland in Namibia (23° South, 15° East [16]) at an altitude of approx. 1800 m [18]. This area has been chosen due to several reasons. First of all the galactic centre, which is a promising observation target, can be observed relatively close to zenith. Low zenith angle observations have the advantage that the Cherenkov light that is supposed to be detected is being less absorbed by the atmosphere between the air shower and the telescopes, leading to a better sensitivity for sources in this region of the sky.

The fact that a location at high altitude was chosen is also due to the minimisation of absorption. Another reason for choosing this region are the ideal climatic conditions. During most parts of the year the air is very dry. Therefore the cloud coverage is very low, as well as the likelihood of precipitation. Besides, the area around the site is poorly populated, which keeps the light pollution very low [5].

Besides the careful choice of the location of the site, the details of the construction of the telescopes are important for an optimal performance of the experiment. In the following sections, the experimental setup is described in detail.

3.1.1 Telescope Structure

The telescope mounts are important parts of the telescopes, granting high pointing precisions by providing the necessary stability as well as the ability to point the telescopes to arbitrary observation targets.

The support frames of the phase I telescopes as well as the mount of CT 5 are built in an alt-azimuth structure. The mounts themselves consist of two towers each, bearing the entire telescope dish structure and the camera. They are placed on circular rails [17], [19]. Four motors provide the ability to move the telescopes around the azimuthal
axis, two more motors provide the elevation of the optical axis of CT 5. The elevation axis is adjustable between $-35^\circ$ and $180^\circ$ ($175^\circ$ in case of CT 1–CT 4 [20]) with respect to the horizon. The slew rate around the azimuthal axis of CT 5 is approx. $200^\circ \text{min}^{-1}$, the slew rate around the elevation axis is $100^\circ \text{min}^{-1}$, and the pointing precision will be between 5 and $10''$, which is remarkable considering the overall weight of approx. 580 t [2], [16]. In case of the phase I telescopes, the slew rate is about $100^\circ \text{min}^{-1}$ around each axis, while the pointing precision after software correction is about $8''$.

The mount of CT 5 is connected to the dish at an altitude of 24 m. The dish is attached to a dish support structure at four points, which leads to a very low deformation of the dish during normal operation. The support structure alone has a weight of 55 t.

### 3.1.2 The Mirrors

The small telescopes are built in a Davies-Cotton style, with a mirror area of $107 \text{m}^2$ each [17]. The reflectors consist of 380 facets each, which are made of glass and a reflecting aluminum layer, protected by a layer of quartz [16]. Each of these facets is built as a circle with a diameter of 60 cm. The focal length of each of these telescopes is 15 m, hence the resulting focal length-to-diameter ratio is $f/d \approx 1.2$ [19].

The mirror area of the new telescope is $596 \text{m}^2$ and thus is significantly larger than the mirror areas of the small telescopes. The mirror is built up of 850 facets. These facets are of hexagonal shape with 90 cm distance between opposite, parallel sides. They are made of glass as well, with the same layered structure as the mirror facets of the phase I telescopes. The facets of all telescopes are held by actuators which are attached to the mirror support structure.

This dish is constructed in a way that the reflector obtains a parabolic shape. Its contour is rectangular with a size of $32 \times 24 \text{m}$. It is optimized such that it is very rigid and that its vibration eigenfrequencies are high so that the oscillations initiated by telescope movements or by wind have a small impact on the telescope performance.

Given the actual shape of the dish the telescope has a focal length of 36 m, resulting in the same $f/d$ ratio like the small telescopes [17].

### 3.1.3 The Cameras

The light that is reflected and focussed by the mirrors is collected by the cameras. In case of CT 1–CT 4 each of the cameras consists of 960 pixels. Each pixel is made of one Photo Multiplier Tube (PMT) and the respective read-out electronics. PMTs are chosen as detectors as they are able to detect very faint light signals and to count the number of photons which hit the respective pixel. Each pixel covers an opening angle of $0.16^\circ$. The values delivered by the PMTs after calibration and digitisation are the numbers of photoelectrons produced by the incident Cherenkov photons per pixel.

Apart from the PMTs, the camera contains electronics for digitisation of the signals and first trigger decisions. Each of the cameras weighs approx. 900 kg and would fit inside a cube with a side length of 1.5 m.
In case of CT5 the camera consists of 2048 pixels, each covering a solid angle of $0.07^\circ$ in diameter. The camera has a cylindrical shape with a height of 2.2 m and a diameter of 2.5 m [16]. Its weight adds up to approximately 3 t.

The camera is mounted in the focal plane of the mirror. When observing at different zenith angles the distance between the telescope and the shower maximum is changing. This leads to slightly different focal planes. In order to be able to compensate for these distance changes the camera can be moved by 10 cm along the optical axis of the telescope [21].

### 3.1.4 The Trigger System

In the H.E.S.S. experiment, there are several levels of triggers. There is one trigger per camera as well as one common system trigger [15].

The triggers on telescope level fire if a certain number of PMTs detected light within a time window of 1.3 ns [16]. Such a cut is well suited for suppression of background events like Night Sky Background (NSB) from starlight or random light pollution from terrestrial sources, as it is very unlikely for such background events to fulfil standard trigger conditions.

The system level trigger of phase I requires that at least two telescopes have triggered within a time window of 80 ns. For events fulfilling this requirement the system level trigger sends a signal to the respective telescopes, so that these can start to read out the event. This trigger chain allows for lowering the energy threshold by a factor of two compared to the case without a coincidence trigger. In H.E.S.S. phase I, this trigger on the system level allows for a good suppression of background events stemming from NSB and light pollution, but it is especially efficient for the suppression of muon events which have triggered only one telescope. It is very unlikely for such events to trigger more than one telescope due to their limited lateral Cherenkov light distribution.

The digitisation and read-out of the data lead to dead times of the telescope of 10 µs up to 10 ms, depending on the data acquisition system [15]. The overall trigger rate as well as the overall dead time is dominated by background events like proton-initiated air showers.

In phase II the trigger level setup is different, as the system trigger is not necessarily needed anymore. The trigger of CT5 is independent of the other triggers, but its information is considered by the system trigger as well [16]. Therefore several trigger patterns can be applied:

a) Only CT5 has to have triggered

b) CT5 plus one of the phase I telescopes have to have triggered

c) Stereo observation with at least two of the phase I telescopes; CT5 is also read out if it has triggered

d) Two out of all five telescopes have to have triggered
In case a), only information of CT5 is considered, the other telescopes are ignored. The trigger rate will be 3 kHz according to Vincent et al. [17]. This pattern is sensitive to events with an energy between 10 and 50 GeV. This trigger mode will be referenced as “Mono mode”. In the second case, which is sensitive to events between 50 and 100 GeV, the trigger rate will be approx. 1 kHz. Case c) is suited for events with energies over 100 GeV. Here the trigger rate will be equal to the H.E.S.S. phase I trigger rate of approx. 400 Hz [18]. Almost all of these events are going to trigger CT5 as well, which will result in a significant improvement of the energy and direction reconstruction as well as a better background suppression. Hence the sensitivity is going to be increased by a factor of at least two [16]. The same is valid for case d), which will result in a better performance compared to the phase 1 stereo mode requesting two out of four telescopes.

In the MC samples used in this thesis, two different trigger patterns are applied at the same time. The system trigger demands that either CT5 has triggered or that two out of all five telescopes have triggered. The simulation of the described trigger mechanisms is not part of the software presented in this thesis, thus only results of previous simulations are utilised.

### 3.2 Performance

The experimental setup described above allows for a very good performance, which is shown in this section. In phase I, an energy threshold of approx. 100 GeV for observations at zenith can be reached [18], while it is possible to reconstruct the energy of incident gamma rays with a resolution of approximately 15% [16].

The field of view is restricted to a solid angle with a diameter of 5°. The 80% containment point spread function of CT1–CT4 for a point-like light source is 1.4′. The energy-dependent angular resolution of up to 0.06° is small enough to investigate substructures of extended sources.

Another important parameter of the experiment is its sensitivity. In case of phase I a source with a flux comparable to the one of the Crab Nebula can be detected with a significance of 5 σ (five standard deviations) within about 30 s, while a source at low zenith angles with a flux of 1% of that flux can be detected within 25 h.

During the first years of operation, the H.E.S.S. experiment has been able to extend the catalogue of known VHE sources by a factor of ten by finding faint AGN signals [16] and by discovering more than 50 new sources inside of our galaxy during a systematic scan of the galactic plane performed between 2004 and 2007 [22], [23], [24], [25].

The field of view of CT5 is 3.5° and its angular resolution, depending on the event energy, is a factor of 5–10 worse than in H.E.S.S. phase I. The energy threshold is as low as 10 GeV [17], [18], which is unprecedentedly low compared to other IACTs. Nevertheless there are satellite-borne experiments like the Fermi experiment [26], which cover the energy range between 10 MeV and 100 GeV [16]. These experiments have the advantage of having a much larger field of view of almost 2π sr, compared to a few degrees in the case of IACTs, providing much higher exposure times. But due to the limited size of satellite-borne experiments the effective area of the main instrument onboard the Fermi
3.3 Event Simulation

In this thesis, only simulated events are analysed. There are several software packages available for the production of such events, the one used for the generation of the events analysed in this thesis is called CORSIKA [27]. It is capable of performing full simulations of extended air showers starting from incident photons, protons, heavy nuclei or other particles. The calculations of the processes happening in such a shower are performed using Monte Carlo (MC) techniques, in order to account for the statistical nature of the possible processes.

Every particle in the shower is simulated from its creation point along the way through the atmosphere to the point of its annihilation or absorption, or, if it is an unstable particle, to the point of its decay. If it is not annihilated and if it doesn’t decay, it is tracked until it reaches the ground. In case of a decay, possible decay channels with probabilities of more than 1% are considered.

On their way through the atmosphere, the particles can interact in many ways with the atoms or molecules of the medium. A wide range of possible interactions is considered. The production of Cherenkov light is included in the simulation as well. Also the deflection of charged particles in the magnetic field of the earth is calculated.

The simulation of the telescope response to the Cherenkov light is performed by a separate package called sim_telarray or, in case of the H.E.S.S.-specific package which is an adaption of the sim_telarray package, sim_hessarray [28]. This program calculates the paths of Cherenkov photons in a physically correct way, including reflections and absorptions. It also simulates the detection of the photons by the PMTs in the camera, the PMT response and the digitisation of the PMT signals for each of the five H.E.S.S. telescopes.

In the MC samples dealt with in this thesis, only the pixel information is available, details of the showers are not necessary and not included. The input MC samples were created by the Heidelberg working group [29], but several conversion steps were performed manually.
Figure 3.1: One of the first shower images taken with CT 5 [30]. It very likely shows an image of a gamma ray initiated air shower. The colour code represents the number of photoelectrons per pixel.

3.4 Image Processing and Parameterisation

With the experimental setup described above, images of Cherenkov light emitted in air showers as described in Sec. 2.4 can be recorded. An example of an image probably produced by a gamma ray initiated shower recorded as one of the first images by CT 5 is shown in Fig. 3.1. When observing a point-like gamma ray source, a frequently deployed observation strategy is to work in “wobble mode” [31], first introduced by Daum et al. [32]. This means that the telescope points next to the source with an offset of, for instance, 0.5° instead of pointing directly towards the source. The region around the source position is called ON region, the radius of which can be defined arbitrarily. Using the wobble offset strategy, it is possible to observe the source and at the same time measure the background in one or more OFF regions, which are camera regions with the same size and distance to the camera centre as the ON region. The specific direction between camera centre and source direction can be altered, making the background measurement depending less on systematic differences between specific sky regions.

In the following sections, the image processing and the image parameterisation are described.

3.4.1 Image Cleaning

A fundamental step in the image processing is the image cleaning. This is a procedure cleaning the images from noise. This is done by demanding that a certain number of photoelectrons is exceeded for each pixel as well as checking the photoelectron number of adjacent pixels. If the conditions are not fulfilled by a pixel, its number of photoelec-
trons is set to zero. In this analysis, two different image cleanings are used: A strict version demanding five photoelectrons (p.e.) per pixel as well as 10 p.e. in at least one neighbouring pixel and a less strict version demanding only one and three p.e., respectively. These cleanings are from now on called 0510 and 0103 cleaning, respectively. These image cleaning procedures are able to suppress NSB and other noise efficiently, as for such events it is very unlikely to exceed the threshold values in several adjacent pixels.

Apart from the noise removal capabilities, deploying two different image cleanings can deliver additional gamma / hadron discrimination opportunities, which is going to be shown in more detail in Sec. 5.5.

### 3.4.2 Hillas Parameters

In general, gamma ray shower images have an elongated, elliptical shape [33], as can be seen in Fig. 3.1. The properties of this ellipse can be parameterised using the set of so-called Hillas parameters, named after the inventor A. M. Hillas [34]. From these parameters, the properties of the air shower can be determined. For instance, the main axis of the ellipse represents the shower axis. The parameters are defined as moments of the photoelectron distributions in the camera [35]. Moments of first order describe the position of the image, the moments of second order describe the extent of the image. Moments of third order describe the overall shape of the image and can be used to indicate the asymmetry of the image.

Some of the most important Hillas parameters used in this thesis are shown in Fig. 3.2. Fundamental parameters are the width and length of the ellipse, the image size which is...
the sum of the pixel amplitudes and the centre of gravity (c.o.g.) of the ellipse. There are many variables which are calculated from these parameters, like the angle $\varphi$, which is the angle between the main axis of the ellipse and the $x$-axis of the camera coordinate system or the angle $\alpha$ between the ellipse’s main axis and the connection line between the camera centre and the c.o.g.. More variables derived from the Hillas parameters are explained in later chapters.
4 Multivariate Analysis Methods

In high energy physics, it is a common problem to identify a small fraction of signal events among a vast majority of background events. Such classification problems can be solved using standard, cut based analysis methods, which often is not satisfactory, because if the background events cover the entire possible phase space, such cuts cannot efficiently keep signal events without also accepting many background events. Furthermore, this method might be inefficient, as only a small fraction of the signal events is selected, as can be seen in Fig. 4.1a. Using the method as described, only one part of the signal region with $x_1 >$...
chosen in a way that the value of the signal distribution is larger than the value of the background distribution for an event which was classified with a $\zeta$ value of $\zeta > \zeta_0$. This way an event classification based on the $\zeta$ variable can be constructed.

In the analysis presented in this thesis, such methods are used for event classification and energy reconstruction. The utilised methods as well as the software toolkit used in the presented analysis are described in the following sections.

4.1 Toolkit for Multivariate Analysis

The MVA algorithms described later in this chapter are implemented in the “Toolkit for MultiVariate Analysis” (TMVA) [36]. This toolkit includes a variety of different machine learning algorithms like neural networks, likelihood estimators or Fisher discriminants. The general principle of working with these methods is very similar. The fundamentals are described in this section.

For any MVA method, a training phase has to be passed before the method can be used to classify events. In this training phase, events generated by Monte Carlo simulations are used to perform a variable ranking, which means that the variable with the highest discrimination power is ranked highest and therefore analysed first. Also the free parameters of the MVA algorithm are determined [22], [36]. This is possible as in case of simulated events the desired classification output is known, therefore the mentioned parameters can be adapted in such a way that the optimal classification performance is reached. It is also possible to select a specific subset of simulated events by applying preselection cuts which can be specified freely.
4.2 Boosted Decision Trees

After the training phase, MVA method tests are performed in order to calculate signal efficiency and background rejection values. The efficiency is defined as the fraction of all input signal events which pass the cut on the $\zeta$ value for an automatically determined cut value $\zeta_0$. The background rejection is defined as the fraction of background events which are rejected by that cut. Furthermore, a quantity called separation is calculated, which is defined as

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{\hat{\zeta}_S(\zeta) - \hat{\zeta}_B(\zeta)}{\hat{\zeta}_S(\zeta) + \hat{\zeta}_B(\zeta)}^2 d\zeta,$$  \hspace{1cm} (4.1)

Here, $\hat{\zeta}_S/B(\zeta)$ are the signal / background probability density functions of a classifier output $\zeta$ [36], [37]. The separation $\langle S^2 \rangle$ is a measure of the difference of the classifier shapes of signal and background events.

Also the discrimination significance $s$ is calculated, which is defined as

$$s = \frac{\langle \zeta_S \rangle - \langle \zeta_B \rangle}{(r_{\zeta_S} + r_{\zeta_B})^2},$$  \hspace{1cm} (4.2)

$\langle \zeta_S/B \rangle$ are the mean values of the classifier response for signal and background events, respectively, and $r_{\zeta_S/B}$ are the root mean square values of these distributions.

Finally, some overtraining checks are performed. Overtraining is a problem occurring if too many parameters of an algorithm have to be determined using too few training events. This results in an MVA method which is too well adapted to the training sample, hence being very sensitive to statistical fluctuations. In order to investigate whether overtraining occurs, the MC input sample is split up into a training sample and a test sample, which are independent from each other and equally sized [36]. During the training phase, only training sample events are used. In the test phase, quantities like $\langle S^2 \rangle$ or $s$ are calculated for both the training and the test sample, making it possible to compare the results. If overtraining occurs, the results differ much. After the MVA method has been trained and tested, the method can be used to calculate classifier responses $\zeta$ for arbitrary events.

The TMVA toolkit provides a simple interface for working with the MVA methods implemented in this toolkit. It provides tools that can be used to easily execute the steps necessary for training, testing and evaluating the MVA methods.

4.2 Boosted Decision Trees

4.2.1 Basic Concept

In the analysis presented in this thesis, Boosted Decision Trees (BDTs) are used for gamma / hadron separation. BDTs, as well as the MVA method presented in Sec. 4.3, are well-suited for classification problems as they consider non-linear correlations between input parameters, and parameters with low separation power are ignored efficiently [22]. The fundamentals of the BDT method as well as the tuning of the method’s parameters are explained in this section.

A fundamental part of the procedure of creating BDTs is the creation of binary
Decision Trees (bDTs). A sketch of a bDT is shown in Fig. 4.3. Each event \( \vec{x} \), consisting of \( N_{\text{var}} \) variables \( x_1, \ldots, x_{N_{\text{var}}} \), is analysed using such a bDT. At each node one of the variables \( x_i \) is analysed and tested if it exceeds a formerly determined cut value \( c_i \). Depending on the result of this test, the event is analysed further using the next of the remaining variables. The last-grown nodes are called “leaves”. According to the type of the majority of events in these leaves, each of them is classified as signal or background leaf, respectively [22]. Hence the phase space is divided into many small areas of signal and background leaves [36]. The details of the tree building are explained in the following.

The building of the tree starts at the first node, where the optimal cut value of the variable with the largest discrimination power is determined. This variable is found using a quantity called importance, defined as the rate of occurrence of a variable during the training, weighted with the number of events in the node and the squared separation gain obtained by applying a cut on the corresponding variable.

If an event fulfils the binary split criterion found in the aforementioned procedure, it is denoted as signal (S), and if it does not, it is denoted as background (B). In contrast to standard cut procedures, both sets of events are considered in the following steps of classification, which is worthwhile due to the incompleteness of the S/B classification after the first node [38]. Therefore in the next step the best suited variable and the corresponding cut value are determined among the remaining variables. This is repeated until all variables have been evaluated [22].

This procedure is prone to statistical fluctuations which occur due to the finiteness of the training sample [38], resulting in a tree that is too much adapted to this sample. To avoid problems resulting from such overtraining, the tree is boosted [22]. The fundamental principle of boosting algorithms is the application of boost weights to wrongly classified events before constructing a new tree. While for the first, unboosted tree the input samples were normalised and each event \( i \) was assigned a weight \( \omega_i = 1/N \), where \( N \) is the number of events in the training sample, the wrongly classified events in the boosted tree get a common boost weight \( \alpha > 1/N \). The variable \( \alpha \) is defined as follows:

\[
\alpha = \frac{1 - \varepsilon}{\varepsilon}. \tag{4.3}
\]
Here $\varepsilon$ is the fraction of misclassified events. After applying these weights to the respective events, the sample is renormalised. The new weights are used in the training phase of the next tree. Applying this algorithm recursively results in a forest containing many trees instead of a single tree.

The output of the BDT method is a single number indicating the probability of an event being a signal event. This number called $\zeta$ is obtained using the following formula:

$$\zeta(\vec{x}) = \frac{1}{N_{\text{Trees}}} \sum_{i} \ln(\alpha_i) h_i(\vec{x}).$$

(4.4)

This formula gives the BDT response $\zeta$ for an event $\vec{x}$. $N_{\text{Trees}}$ is the number of trees in the forest, and $h_i(\vec{x})$ is the classifier response for tree $i$, delivering the value +1 for signal-like and −1 for background-like events [36]. $\alpha_i$ is the boost weight as defined in Eq. (4.3) for the tree with the number $i$.

This procedure is called adaptive boost or “AdaBoost” [36], [39]. It applies for both signal and background samples, respectively, and it is the boosting algorithm used in this thesis.

### 4.2.2 Settings

Apart from understanding the basic principles of BDT methods it is important to know some of the BDT settings used in this thesis. Often the default settings implemented by the TMVA developer group [36] are suitable, as these guarantee fast training and stable responses [22]. In the following list the most important settings are briefly explained, especially if default values are not used:

- The number of trees $N_{\text{Trees}}$ is 200 by default, which is a compromise between classification speed and separation power [22]. In order to stabilize the BDT output this number has been changed to 400.

- The separation type used is the Gini index, which is the default type. It is used to find the optimal cut value for each variable in the tree by minimising the expression $p \cdot (1 - p)$, where the purity $p$ is defined as the ratio of signal events $N_S$ and the total number of events $N_{\text{tot}}$: $p = N_S / N_{\text{tot}}$ [36]. $N_{\text{tot}}$ is defined as the sum of signal and background events $N_{S/B}$.

- The node splitting stops when the number of events in a node falls below the default value of

$$\frac{N_S + N_B}{10 \cdot N_{\text{par}}^2},$$

where $N_{\text{par}}$ is the number of training parameters and $N_{S/B}$ is the number of signal / background events in the sample [22].

- In order to find the optimal cut value for each event variable $x_i$, the parameter space is scanned with a step size called $n_{\text{Cuts}}$. Its value is changed from the default value of 20 to −1 in order to achieve an optimal performance by testing
every possible cut value (with a certain granularity) and choosing the best one [36]. In this case, every single event is analysed.

- Tree pruning is an algorithm used to eliminate unimportant nodes from the tree in order to reduce the danger of overtraining. The default pruning method called CostComplexity is used. This method compares the cost and the additional yield in classification performance of further splitting below a certain node. A cost estimate is defined as $R = 1 - \max(p, 1 - p)$ describing the misclassification in a node with $p$ being the purity. This is used to calculate the cost complexity $\rho$ as

$$\rho = \frac{R(\text{node}) - R(\text{subtree below that node})}{N_{\text{nodes}}(\text{subtree below that node}) - 1}.$$  \hfill (4.5)

Here $N_{\text{nodes}}$ is the number of nodes. The node with the smallest value of $\rho$ is cut away. This is repeated as long as $\rho < \rho_0$ [36]. The default value for $\rho_0$ is used.

### 4.3 Multilayer Perceptrons

#### 4.3.1 Basic Concept

A different approach for using more advanced classification methods than standard cut series is the utilisation of "artificial Neural Networks" (NN). Different NN implementations exist, like the Clermont-Ferrand neural network or the MLP implementation (Multilayer Perceptron) [36]. These NNs are capable of building non-linear functional representations of the signal and background events [40]. In this thesis MLP is used as a tool for energy reconstruction in order to be compatible with the energy reconstruction mechanism used by Masbou [5], as the analysis presented in this thesis should serve as a crosscheck for this work. The working principle of MLP is described in this section, the energy reconstruction mechanism using MLP is described in Sec. 5.4.

The MLP method uses nodes called neurons which are grouped in layers for event classification. Every network constructed by MLP is a feed-forward model network [40] consisting of an input layer, an output layer and a number of hidden layers in between. The number of hidden layers can be defined freely, but at least one hidden layer has to be created. A sketch of such a network with one hidden layer can be seen in Fig. 4.4. The neurons in the input layer use the measured variables $x_i$ as function arguments in order to calculate a response value. The hidden layers build an internal representation of the data, and the output layer is used to read out the state of the network [41].

Each layer $L$ consists of $N^{(L)}$ neurons. The number of neurons in the hidden layers can be defined arbitrarily. MLP is capable of performing event classification as well as regression tasks. In case of classification, a single number is needed for indicating the type of an event, therefore the output layer consists of only one neuron. Furthermore, the input layer as well as every hidden layer contain an additional bias neuron, which is introduced for technical reasons [42]. A neuron should have a threshold value that has to be exceeded by its input in order for the neuron to deliver a non-zero output. The exact threshold value has to be determined during the training phase. As this is
4.3 Multilayer Perceptrons

**Figure 4.4:** Sketch of an MLP network with a single hidden layer containing five neurons and an output layer consisting of one neuron [36]. In this example, four input values \(x_1, \ldots, x_4\) and bias neurons as explained in the text are used.

Cumbersome to implement, another approach is chosen using bias neurons which always give an output value of one, and by adapting the connectivity weights during the training phase, threshold values become dispensable.

The neurons in one layer are connected to all neurons in adjacent layers. As can be seen in Fig. 4.4, the measured variables \(x_i\) defining an event are used as input for the neurons in the input layer. Each neuron represents a functional relation between an input and an output signal [36]. The output value \(y_j^{(L)}\) of neuron \(j\) in layer \(L\) is defined recursively by the following relation [40]:

\[
y_j^{(L)} = g \left( \frac{1}{T} \sum_{i=1}^{N^{(L-1)}} \omega_j^{(L)} y_i^{(L-1)} \right),\quad j = 1, \ldots, N^{(L)}, \quad L = 1, \ldots, M
\]

In this equation, the function \(g\) is the transfer function defining the functional relation between the neuron input and output. This function can be chosen freely and it is common to all hidden layer neurons. In the TMVA implementation, the transfer functions of the input and output layer neurons are simple linear functions. The \(\omega\) values are connectivity weights between the neurons in adjacent layers. \(M\) is the number of layers in the network. In general \(T\) is a free parameter of the algorithm, but in the TMVA implementation, \(T\) is fixed to one. Changing it could accelerate the training, but the \(T\) values can be fully absorbed in the weights. \(T\) is not shown in the following equations.

Given the relation between input values and neuron output, for a given set of measured
parameters $x_i$, the network reaches a defined state that can be read out via the response value of the output layer neuron. For a network containing only one hidden layer with $N^H$ neurons and with a linear output neuron transfer function, this mapping of a vector of input values to a number has the following, general form [36], [41]:

$$F_i(x_1, \ldots, x_N) = \sum_{j=1}^{N^H} \omega_{ij} \left( \sum_{k=1}^{N} \omega_{jk} x_k \right).$$  \hspace{1cm} (4.7)

In case of a single output neuron, $i$ is one. If the network should contain more hidden layers, the formula has to be adapted in a recursive way. $N$ is the number of input parameters, which is equal to the number of input layer neurons. The parameters $\omega_{ij}$ and $\omega_{jk}$ are the connectivity weights. These parameters have to be determined by fitting, for which an algorithm called back-propagation is used, which is executed during the training phase [36], [40]. It is used for minimising the quadratic error given as [41]

$$E = \frac{1}{2} \sum_{i=1}^{N^{(M)}} (y_i^{(M)} - t_i)^2.$$  \hspace{1cm} (4.8)

Here $y_i^{(M)}$ is the network output and $t_i$ is the “target” output value, which is zero for a background event and one for a signal event. $N^{(M)}$ is the number of neurons in the output layer.

The minimisation of $E$ is realized by using the gradient descent algorithm for updating the weights. Initially, the weights’ values are chosen randomly. In the implementation used, the weights are updated after presenting each event, which is called online learning. The update of the weights is done according to the following relation [40]:

$$\Delta\omega_{ji}^{(L)} = -\eta \delta_j^{(L)} y_i^{(L-1)} + \alpha \Delta\omega_{ji}^{(L)\text{old}}, \ i = 1, \ldots, N^{(L-1)}, \ j = 1, \ldots, N^{(L)}. \hspace{1cm} (4.9)$$

Here $\Delta\omega_{ji}^{(L)\text{old}}$ are the changes applied to the weights in the previous iteration. $\delta_j^{(L)}$ is defined recursively:

$$\delta_i^{(M)} = y_i^{(M)} \left(y_i^{(M)} - t_i\right) \left(1 - y_i^{(M)}\right)$$  \hspace{1cm} (4.10)

$$\delta_j^{(L)} = \sum_{k=1}^{N^{(L+1)}} \delta_k^{(L+1)} \omega_{kj}^{(L+1)} y_j^{(L)} \left(1 - y_j^{(L)}\right).$$  \hspace{1cm} (4.11)

$\eta$ and $\alpha$ are free parameters of the algorithm. The output values of the neurons belonging to adjacent layers and the weights of the connections between these neurons influence the weight updates. Also the update history is considered.

### 4.3.2 Settings

The TMVA implementation of the MLP method allows for the altering of many parameters. As already stated in Sec. 4.2.2, the default parameters found by the TMVA group

24
are well-suited for most analysis tasks. Still some parameters had to be changed. In the following list, the most important settings used in this thesis are given.

- The free parameters $\alpha$ and $\eta$ were not modified, hence the default values are used: $\alpha = 0.01, \eta = 0.02$.

- During the training phase, each event is evaluated $N_{\text{cycle}}$ times [40]. In this thesis, $N_{\text{cycle}}$ is raised from 500 to 600 in order to improve the classification performance.

- Data preprocessing can be performed. The input values can be transformed in many ways. In this thesis normalisation is applied, meaning that the interval of all variables is set to $[-1,1]$.

- Three hidden layers are utilized, consisting of $N$ neurons each, with $N$ being the number of input variables.

- The value of NeuronType defines the neuron transfer function $g(x)$. Systematic tests show that, in case of the present analysis, the predefined function called radial results in the best performance. It is defined as a Gaussian function:

\[ g(x) = e^{-\frac{1}{2}x^2}. \]

- An option was set in order to be able to obtain the probability of an event being a signal event using the probability density functions of the input variables.
5 Mono Reconstruction in Phase II

In this chapter, the software written for the event reconstruction in H.E.S.S. phase II in mono mode is presented. Mono mode means, as already mentioned in Sec. 3.1.4, that only information of CT5 is used. The trigger decisions as well as pixel information of any of the smaller telescopes (CT1–CT4) are not considered apart from the fact that the events have to have passed the system trigger.

In the following sections, the applied cuts are described as well as the basic conceptual approaches to tasks like direction reconstruction, energy reconstruction or particle identification. The performance of the introduced procedures and the key results of this analysis are described in the Chap. 6.

5.1 Event Selection Cuts

Applying cuts is necessary in order to reject events which do not fulfil certain quality criteria. This is necessary in order to only accept images containing meaningful information. Furthermore it is important to note the applied cuts because a comparison with other studies concerning mono reconstruction in H.E.S.S. phase II is only possible if the respective cuts are known. As this analysis serves as a crosscheck for the work of Masbou [5], the cuts are also compared to the cuts used in that analysis. Apart from describing the applied cuts, the necessary variables are introduced in the following.

First it is assured that CT5 is among the telescopes which have triggered. If this is the case, it is checked whether it is possible to calculate meaningful Hillas parameters for the corresponding event. This check is performed for both the 0103 and 0510 image cleanings. This is necessary as the analysis is based on the Hillas parameters calculated for both image cleanings. In more detail, the mentioned check assures that the length $L$ and the width $W$ are finite numbers and that a value called $s_{xy}$ exceeds a threshold value of $10^{-8}\text{rad}^2$. The value $s_{xy}$ is defined as $s_{xy} = \langle xy \rangle - \langle x \rangle \langle y \rangle$, with $x$ and $y$ being the pixel positions in the camera. The expectation values are calculated as pixel amplitude weighted mean. Thus this variable $s_{xy}$ is a measure of the width of the Hillas ellipse along the diagonal in the camera coordinate system. In the presented analysis, only the information whether the respective event fulfils the described cut is evaluated, the calculation of $s_{xy}$ as well as the definition of the cut value is done in the utilised software framework HAP [43] in version 11-02.

After making these preselection cuts, three more technical cuts are applied which are used to limit some of the input variables of the BDT algorithm to a reasonable range. This is necessary as otherwise the BDT implementation in the TMVA package would stop working due to large discrepancies between the variable ranges of gamma
and hadron distributions. Concretely, these variables are the scaled length \( \tilde{L} \), the scaled width \( \tilde{W} \) and the concentration \( C \). \( \tilde{L} \) and \( \tilde{W} \) are defined as

\[
\tilde{L} = \frac{L - \langle L \rangle}{\sigma_L} \quad \text{and} \quad \tilde{W} = \frac{W - \langle W \rangle}{\sigma_W},
\]

and \( C \) is defined as

\[
C = \frac{A}{\pi \cdot L \cdot W}.
\]

In these equations, \( W \) and \( L \) are the width and the length of the Hillas ellipse, respectively, and the term \( \pi \cdot L \cdot W \) represents the area of the ellipse. \( A \) is the size of the image. \( L \), \( W \) and \( A \) have already been introduced in Sec. 3.4.2. \( \langle W \rangle \) and \( \langle L \rangle \) are the mean width and length of gamma ray images with given values of \( A \) and a roundness \( \rho \), which in turn is defined as

\[
\rho = \frac{W}{L}.
\]

From \( A \) and \( \rho \), the values of \( \langle W \rangle \) and \( \langle L \rangle \) can be acquired by using lookup tables. In principle, the utilisation of lookup tables makes it possible to determine physical quantities which cannot be measured directly. The lookup tables used in this thesis are two-dimensional histograms representing the functional relation between measurable variables, for instance \( A \) and \( \rho \), and the mean value of the quantity one is interested in, which in this case would be \( L \) or \( W \). Using gamma ray MC data, a table for each \( L \) and \( W \) can be filled using the known values of \( A \), \( \rho \) and \( L \) and \( W \). The two last-mentioned variables are correlated with two input variables in a statistical sense, therefore a mean value of the possible \( L \) and \( W \) values for each combination of values of \( A \) and \( \rho \) is calculated and stored in the lookup tables. Also the mean values in the table are influenced by statistical fluctuations of the utilised MC data set. Therefore the table is smoothed, which means that each bin entry is adapted in order to fit better the bin entries of the surrounding bins. After the filling and smoothing of the lookup table, \( F \) can be determined using the measured values \( A \) and \( \rho \). The lookup tables for the width and the length are given in Fig. 5.1a and Fig. 5.1b.

The other variables occurring in Eq. (5.2), \( \sigma_W \) and \( \sigma_L \), are defined as

\[
\sigma_W = \sqrt{\langle W^2 \rangle - \langle W \rangle^2} \quad \text{and} \quad \sigma_L = \sqrt{\langle L^2 \rangle - \langle L \rangle^2},
\]

describing the widths of the \( \tilde{W} \) and \( \tilde{L} \) distributions. \( \langle L \rangle \) and \( \langle W \rangle \) have been introduced before, and \( \langle L^2 \rangle \) and \( \langle W^2 \rangle \) are defined analogously. The \( \sigma \) variables have to be inferred from lookup tables as well. These tables are shown in the appendix in Fig. A.1a and Fig. A.1b. By construction, the \( \tilde{W} \) and \( \tilde{L} \) distributions of gamma ray events are Gaussian distributions with a mean of approximately zero and a width of about one.

Considering the definitions of the three quantities \( C \), \( \tilde{W} \) and \( \tilde{L} \), it can be seen that it
5.1 Event Selection Cuts

is possible that some events have very high values in these variables due to the fact that the values in the denominators can be very small. The cuts preventing such behaviour by limiting the variable ranges are listed in Tab. 5.1.

It is plausible that the cuts on ˜W and ˜L have little impact on the overall number of events, as only very few events will exceed these cut values given the Gaussian distributions of the scaled length and width values described above. In case of the concentration, very high values are produced if the width W is very small, which can be the case if few camera pixels which are aligned in a row deliver non-zero photoelectron counts after image cleaning. In such cases, it cannot be guaranteed that the Hillas ellipse construction gives reliable results. Hence these events, which make up about 2.9% of the gamma ray events passing the previous cuts, are rejected due to this cut.

After performing these technical cuts, three more cuts which have a physics motivation are applied. They are listed in Tab. 5.1. The first one is a cut on the nominal distance n. The nominal distance is the distance between the c.o.g. of the reconstructed Hillas ellipse and the centre of the camera. As each point in the camera plane corresponds to a direction in three-dimensional space, distances are expressed as angular distance. Such a cut on n prevents the analysis of images which are located at the edge of the camera. The images of such events are very likely not contained entirely in the camera, as these images are often truncated and therefore cannot be parameterised in a standard way. Given the field of view of CT5 of 3.5°, the nominal distance is limited to 1° as a trade-off between the number of events passing this cut and the quality of the accepted events.

The second cut rejects events with non-zero photoelectron counts in less than four pixels in the camera of CT5 after a 0510 image cleaning. This is necessary in order to guarantee that the Hillas ellipses can be constructed in a well-defined way for events passing this cut, while keeping as many events as possible. Although it is in principle possible to reconstruct a Hillas ellipse for images with only two pixels with photoelectron

Figure 5.1: Lookup tables of the width W and the length L of gamma ray events. A and ρ are the size and the roundness of the Hillas ellipse. The colour scale represents the width and length values in rad.
## Mono Reconstruction in Phase II

<table>
<thead>
<tr>
<th>Cut Parameter</th>
<th>Cut Value</th>
<th>Rejected γ / p Events</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technical Cuts</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT5 has triggered</td>
<td></td>
<td>2.5% / 6.0%</td>
</tr>
<tr>
<td>Reasonable Hillas parameters</td>
<td></td>
<td>11.3% / 11.4%</td>
</tr>
</tbody>
</table>

| **BDT Variable Limitations**         |           |                       |
| Scaled width \( \bar{W} \)           | \( |\bar{W}| < 50 \) | 4.3% / 1.9%           |
| Scaled length \( \bar{L} \)           | \( |\bar{L}| < 50 \) | 0.00% / 0.02%         |
| Concentration \( C \)                 | \( C < \frac{10^9 \text{ p.e.}}{\text{rad}^2} \) | 2.9% / 2.8%           |

| **Physical Cuts**                     |           |                       |
| Nominal distance \( n \)              | \( n < 1^\circ \) | 38.7% / 59.6%         |
| Number of pixels \( N^{\text{pix}} \) | \( N^{\text{pix}} > 3 \) | 7.6% / 6.8%           |
| Size \( A \)                          | \( A > 40 \text{ p.e.} \) | 5.8% / 4.7%           |

| **Gamma/Hadron Separation Cuts**      |           |                       |
| BDT response \( \zeta \)              | \( \zeta > \zeta_0 \) | 4.8% / 35.7%          |
| Distance to ON region centre \( \Theta \) | \( \Theta < \Theta_0 \) | 24.9% / 46.4%        |
| Total rejected events                 |           | 69.3% / 90.2%         |

Table 5.1: List of the applied event selection cuts. The variables and the motivation for the choice of the cut values are explained in the text. The values given in the last column are calculated as the percentages with respect to the number of events passing the previous cuts. In case of proton events, these weighted number of events is used. The last two cuts are different for each energy band, so that numbers for these have to be given later (cf. Sec. 6.3).

Counts greater than zero (from now on only called “pixels”, if not stated otherwise), this would not be reasonable, as the number of degrees of freedom for the construction of the ellipse is too high compared to the number of pixels.

The last cut concerns the size of the image. Only events with more than 40 p.e. after a 0510 image cleaning are accepted. This value was chosen as a trade-off between the minimal intensity and the image quality. Especially events with low intensities are of interest for a mono analysis, because these events will not trigger one of the smaller telescopes. Still a minimal image size is required in order to guarantee the Hillas ellipses to be well-defined. In typical analyses performed in H.E.S.S. phase I, often 80 p.e. or more are required [44]. Due to the much larger mirror area of CT5, the two mentioned numbers are of comparable strength.

In Fig. 5.2, a cut flow chart is shown, depicting the influence of all of the cuts mentioned above. The content of the bins shown in that figure represents the percentage of the number of events passing the cut indicated in the respective bin labels and all previous cuts. In case of proton events, the reweighted number of events is used. The reweighting
5.1 Event Selection Cuts

Figure 5.2: Cut flow chart showing the percentage of events passing the cuts indicated in the bin labels. The distributions for gamma ray events and proton events are shown. The percentages are calculated with respect to the total number of events. The cuts are explained in the text.

It can be seen that most events are rejected due to the nominal distance cut, the first of the physical cuts. This cut especially reduces the number of proton events, as these are diffuse protons, meaning that the incident protons do not have a preferred direction. The source of the gamma ray particles is point-like, thus the incident gamma rays have a favoured direction. As the gamma ray point source is simulated with a wobble offset of 0.5°, less gamma ray events have a nominal distance of more than 1° compared to the number of corresponding proton events.

In the next to last bin shown in Fig. 5.2, labeled as “BDT cut”, the number of events passing the cut on the BDT response variable \( \zeta \) introduced in Chap. 4 is shown. This cut is going to be explained in more detail in Sec. 5.5.

The last bin shows the number of events passing a cut on the reconstructed shower direction. For the calculation of some quantities like effective areas (cf. Sec. 6.4), a cut on the reconstructed direction of the incident particle is performed. An ON region with an offset of 0.5° is defined as explained in Sec. 3.4. The calculation of the radius of the ON region is explained in Sec. 5.5. If this cut is applied, events whose reconstructed directions lie outside of the defined ON region are rejected.

Analysing Tab. 5.1, it becomes apparent that the cut with the highest impact on the number of events is the nominal distance cut, followed by the cut on the reconstructed shower direction and the cut on the BDT response \( \zeta \).

In the thesis of Masbou [5], only the nominal distance cut is applied, the other cuts
listed in Tab. 5.1 have not been applied. For particle identification, a cut based on the response value \( \zeta \) of an MLP neural network is used in the cited thesis, whereas in this analysis a BDT decision is used.

In the following, only events passing all of the cuts mentioned in this section besides the BDT cut and the ON region cut, if not stated differently, are discussed.

### 5.2 Calculation of the Simulated Time and Weights of Proton Data

The proton events contained in the corresponding MC samples have been simulated according to a power law spectrum with an index of \( \alpha_{MC} = 2 \), which is not the real spectrum as measured, for example, by the BESS experiment \[45\]. The experimental value of the differential, diffuse proton flux given by this experiment is

\[
\frac{d\Phi_B}{dE} = 0.099 \frac{1 \text{ TeV} \text{ m}^2 \text{ sr}}{(1 \text{ TeV})^{2.7}} \frac{1}{1 + \left( \frac{0.004 \text{ TeV}}{E} \right)^{1.75}}.
\]

In order to account for this spectrum which has an index of \( \alpha_B = 2.7 \), the MC events are reweighted using the expression for the weights \( w(E) \) (cf. Sec. A.2, Eq. (A.8)) of an event with an energy \( E \) given as

\[
w(E) = \left( \frac{E}{1 \text{ TeV}} \right)^{-\alpha_{MC}} \cdot 0.099 \cdot \left( \frac{E}{1 \text{ TeV}} \right)^{-\alpha_B} \cdot \frac{1}{1 + \left( \frac{0.004 \text{ TeV}}{E} \right)^{1.75}}.
\]

The distribution of the proton weights after all cuts except the BDT cut is shown in Fig. 5.3. An event with an energy higher than \( \sim 45 \text{ GeV} \) gets a weight bigger than one, events with a lower energy are assigned to weights smaller than one. The gamma ray events are simulated with a spectral index of \( \alpha_{MC\gamma} = 2 \) as well, which is a reasonable index for many gamma ray sources. Hence the gamma ray events have not been reweighted.

In order to calculate the background event rate, the simulated time \( T \) has to be calculated. For a given spectrum, \( T \) corresponds to the number of simulated events \( N_{Sim} \). Due to the reweighting of the proton events, the number of simulated events is altered, hence \( T \) has to be adapted to the BESS spectrum as well. The calculation is shown in the appendix in Sec. A.2. For a given number of proton MC events \( N_{Sim} = 1.1 \times 10^6 \), the correct time is \( T = 161.1 \text{ s} \).

### 5.3 Direction Reconstruction

An important part of the analysis is the reconstruction of the directions of the incident gamma rays and protons. This is necessary in order to be able to determine the exact
5.3 Direction Reconstruction

position of gamma ray sources and for the investigation of the question whether the observed source is point-like or extended. Particularly for point-like gamma ray sources it is possible to gain additional separation power between gamma ray and proton events due to a cut on the ON region. For this purpose it is important to have a good direction reconstruction.

There are two different direction reconstruction mechanisms implemented in the analysis presented here. The first one is based on the usage of Hillas parameters as well as on the utilisation of lookup tables. This method will be referred to as the $\delta$ method. For the second approach, the assumption is made that a point-like gamma ray source at a known position is searched for. The last-mentioned one will be referred to as the distance method.

The $\delta$ method approach is based on the fact that, ideally, the shower direction, which is represented by a point in the camera plane, is located somewhere on the straight line defined by the main axis of the constructed Hillas ellipse of an event. The shower direction can be reconstructed using the centre of gravity of the Hillas ellipse and a distance parameter called $\delta$ which is used for parameterising the distance between the c.o.g. and the true shower direction. The mentioned quantities are shown in Fig. 5.4. The determination of the direction of $\delta$ along the main axis is discussed later. The absolute value of the $\delta$ variable is a derived quantity and thus has to be acquired from other, known quantities. In this thesis, a lookup table is used. This table is filled using the known directions of simulated gamma ray showers. For such showers, the absolute distance between the c.o.g. and the shower direction is calculated and filled into a two-dimensional histogram as a function of the roundness $\rho$ and the natural logarithm of
Figure 5.4: Sketch of a Hillas ellipse including the quantities needed for the direction reconstruction. These are the centre of gravity (c.o.g.), $\delta$ and the angle $\beta$ between the main axis of the ellipse and the $x$ axis of the camera coordinate system.

the size $\ln(A)$. The lookup table is shown in Fig. 5.5. The value of $\delta$ can be looked up if the size $A$ and the roundness $\rho$, which are determined by the Hillas parameters, are known. As can be seen in the figure, large $\delta$ values occur if the size is very large and the roundness is small, hence for bright showers and elongated ellipses. The lowest $\delta$ values are obtained for very round shower images, which are likely produced by showers which occured very close to the telescope.

There is a degeneracy in the obtained $\delta$ values between events with high energies and high amounts of produced Cherenkov light which have a large distance to the telescope and events with low energies and low Cherenkov light intensities which occurred close to the telescope. This results in too low $\delta$ values for high energy events. In order to correct for this behaviour, the looked up $\delta$ values are scaled by a factor $G$, which is defined as given in the following formula:

$$G = \sqrt{\frac{g}{\langle g \rangle}}. \tag{5.9}$$

The above definition of $G$ as well as the method of using lookup tables for the reconstruction of $\delta$ is based on the work of Ehlert [46]. In Eq. (5.9), $g$ is a quantity defined as

$$g = \frac{L}{\ln A}. \tag{5.10}$$

$L$ and $A$ are the length and the size of the Hillas ellipse, as introduced before. This quantity is a function of the energy, and can therefore be used to correct the lookup values of $\delta$ for the energy. In order to obtain a variable centered at the value one, $g$ is divided by $\langle g \rangle = \langle g \rangle / (\delta)$, which is the corresponding mean value of $g$ for a given $\delta$. The square root of the resulting fraction is taken in order to yield a suitable scaling factor.
5.3 Direction Reconstruction

![Lookup table for the δ variable as a function of the roundness ρ and the natural logarithm of the size of the Hillas ellipse ln(A/[p.e.]). The colour scale represents the δ values in rad.](image)

Figure 5.5: Lookup table for the δ variable as a function of the roundness ρ and the natural logarithm of the size of the Hillas ellipse ln(A/[p.e.]). The colour scale represents the δ values in rad.

G. The corrected value $\delta'$ is defined as

$$\delta' = \delta \cdot G = \delta \cdot \sqrt{\frac{g}{\langle g \rangle}}. \tag{5.11}$$

From now on, $\delta'$ is meant when referring to “δ”.

After calculating the reconstructed δ value, the angular distance between the shower direction and the c.o.g. is known, but as indicated in Fig. 5.4, there are two possible shower directions differing only in the signs of the $\vec{\delta}$ components. In order to determine the correct sign, the Hillas parameter skewness $S$ is used. It describes the asymmetry of the intensity distribution inside the Hillas ellipse along the main axis and is calculated as a higher moment of the intensity distribution in the camera. This asymmetry, and thus also the skewness, is correlated with the signs of the components of $\vec{\delta}$. This can be explained by considering the Cherenkov angles under which the charged secondary particles in the air shower emit Cherenkov light. This angle depends on the refractive index of the air, so that at lower heights, the Cherenkov angle is larger than at higher altitudes, as explained in Sec. 2.4. This is indicated in the sketch shown in Fig. 5.6. The shower direction is to be found in the direction of the longer tail of the intensity distribution as shown in that figure. In order to calculate the correct shower direction, the angle $\beta$, as shown in Fig. 5.4, is used to find the correct pair of signs $\sigma_{x/y}$ of the $x$ and $y$ component of the connection vector $\vec{\delta}$ between the c.o.g. and the source position, as these signs depend on the position of the ellipse in the camera coordinate system.
Figure 5.6: Sketch of an air shower and CT5. The dashed lines indicate the tracks of the Cherenkov light photons. The intensity profile is used to calculate the skewness, which in this case would be negative. $\vec{x}_T$ denotes the position of the true source.

Here $\beta$ is used instead of $\varphi$ (cf. Fig. 3.2) because the $\beta$ value considers the skewness in order to find the correct orientation of the angle between the main axis and the $x$-axis of the camera system. Hence $\beta$ and $\varphi$ can differ by an angle of 180°.

From the c.o.g., $\beta$ and the signs of the components and the modulus $\delta$ of the vector $\delta$, the coordinates in the camera plane $x_r$ and $y_r$ representing the reconstructed shower direction can be calculated as

\begin{align}
  x_r &= x_{\text{c.o.g.}} + \sigma_x \cdot \delta \cdot |\cos \varphi|, \\
  y_r &= y_{\text{c.o.g.}} + \sigma_y \cdot \delta \cdot |\sin \varphi|.
\end{align}

Here $x_{\text{c.o.g.}}$ and $y_{\text{c.o.g.}}$ are the $x$- and $y$-components of the c.o.g.. The absolute values of the sine and cosine terms is taken because the angle $\varphi$ does not include information about the orientation of the ellipse’s main axis.

The distance method uses a different approach for the shower direction reconstruction. The fundamental assumption of this method is that the observed gamma ray source is point-like. The shower direction is defined to be the point on the main axis of the Hillas ellipse which is closest to the source position, as shown in Fig. 5.7. This is a much simpler approach than the previously described method, but the drawback of the distance method is that it is not only based on measurable quantities but also on the position of an assumed point-like source. Furthermore it is beneficial for this method to deal with sources with already known positions. When observing a source with an unknown position, an optimisation loop would have to be executed, finding the true source position as the point in the camera which minimises the mean angular distance of the reconstructed and the supposed true shower directions. Such an optimisation rou-
5.4 Energy Reconstruction

A very important part of this analysis is the implementation of an energy reconstruction mechanism, as it is crucial for the measurement of the energy spectrum of a source and also because in this analysis, the particle identification procedure is based on the reconstructed energy.

There are several possible approaches to the problem of energy reconstruction, like lookup tables or NN mechanisms. In this thesis, an MLP approach (cf. Sec. 4.3) has been chosen for several reasons, like the compatibility with the work of Masbou [5] and the good performance of such an approach. In principle, MC data are used in order to train the MLP neural network using several variables:

a) the width $W$,

b) the length $L$,

c) the size $A$,

d) the skewness $S$,

Figure 5.7: Sketch of a Hillas ellipse, with the indicated main axis on which the reconstructed shower direction $\vec{x}_R$ is located. The true direction is labeled as $\vec{x}_T$. 

tine for unknown source positions has not been implemented in the analysis presented here. Therefore, in the current state of the analysis, only sources with known positions can be analysed with the distance method. The distance method can also be seen as a method delivering the best possible direction reconstruction of an analysis based on Hillas parameters, as its performance only depends on the accuracy of the definition of the Hillas parameters.

Both of the direction reconstruction methods introduced above are applied in this thesis. Their performances are discussed in Sec. 6.1.

5.4 Energy Reconstruction
e) the kurtosis $K$,

f) $\Delta \varphi$.

The variables a)–d) have already been introduced. The kurtosis $K$ is another Hillas parameter, which describes the peak width of the intensity distribution along the main axis of the Hillas ellipse. It is defined as a higher moment of the intensity distribution in the camera in a similar way as the kurtosis known in statistics is defined.

The last of the listed variables, $\Delta \varphi$, is defined as

$$\Delta \varphi = \varphi_{0510} - \varphi_{0103}, \quad (5.14)$$

where $\varphi_{0510}$ is the already introduced angle $\varphi$ after a 0510 image cleaning and $\varphi_{0103}$ is the same angle after a 0103 image cleaning. $\varphi$ is computed in a range between 0 and 360°. It should, however, be independent of the orientation of the main axis of the Hillas ellipse, thus $\Delta \varphi$ is defined to be within the range between 0 and 90°.

Several other variables have been tested, like the nominal distance, the difference of the size values for the two image cleanings $\Delta A = A_{0103} - A_{0510}$ and the concentration, but they were abolished as they did not yield a further gain in the energy reconstruction performance.

In order to calculate the reconstructed energy $E_{\text{Reco}}$ using an MLP network approach, five energy bands were constructed per true energy ("$E_{\text{True}}$") decade. The energy band with the number $i$ contains events with an energy in the interval

$$10^{-2+0.2i} \leq \frac{E_{\text{True}}}{1 \text{TeV}} < 10^{-2+0.2(i+1)}, \quad (5.15)$$

starting with a lower energy boundary of approximately 0.016 TeV ($i = 1$, lower energy band boundary) and going up to an energy of 1 TeV ($i = 9$, upper energy band boundary). The simulated gamma ray event energies are in the range between 5 GeV and 10 TeV. The events with very low energies are mostly rejected by the cuts mentioned earlier, while the number of events with energies above 1 TeV is very low due to the simulated power law spectrum. For a mono analysis it is not necessary to provide a good energy reconstruction at energies above 1 TeV, as events with such high energies are very likely to trigger one of the four small telescopes.

The choice of the number of energy bands per decade can be motivated by considering the achieved energy resolution, a quantity that is going to be discussed in Sec. 6.2. The width of each band in the current setup corresponds to a relative deviation of about 40% of the central value of the respective band. This is approximately equivalent to the energy resolutions calculated in Sec. 6.2.

In each of the nine energy bands one MLP network is constructed which uses the variables listed above as input. Every MLP network is trained such that gamma MC events with a true energy contained in the corresponding energy band are considered as signal events, while gamma MC events from all other energy bands are treated as background events. In order to obtain $E_{\text{Reco}}$, the response values $\zeta_i$ of the MLP networks of all energy bins $i$ could be compared as it is done in Masbou [5], which has the
Figure 5.8: Example of a p-value distribution of a gamma MC event as a function of the mean energy of the corresponding energy band (see text for details). The points are fitted by a Gaussian distribution. Furthermore the positions of the reconstructed and the true energies of this event are indicated by the vertical lines.

disadvantage that the ζ scales are not necessarily comparable between different energy bands. Therefore a p-value \( p \) is used, giving the probability that an event belongs to the signal class of the corresponding MLP network. The p-values of different energy bands can be compared, because they represent the probability that the corresponding event belongs to the signal class of the corresponding energy band. Thus the absolute \( p \) scale is the same in all energy bands. The p-values can be plotted as a function of the mean energy of the corresponding energy band as shown in Fig. 5.8. On the abscissa of this figure, the logarithm of the energy \( E \) is shown, which is defined as

\[
E = 10^{0.5 \cdot (\log_{10}(E_1) + \log_{10}(E_2))}
\]

for an energy band with \( E_1 \) and \( E_2 \) as lower and upper energy boundaries, respectively. These p-values can be fitted by a Gaussian distribution. The reconstructed energy is identified with the position of the maximum of the fit function as indicated in the figure. In this example, the true energy of the event is \( E_{\text{True}} = 41 \) GeV, while the reconstructed energy is \( E_{\text{Reco}} = 44 \) GeV.

A problem arises if the maximal p-value occurs in either the first or the last energy band. In this case, the fit by a Gaussian distribution is not as well-suited as for cases in which the highest p-value is located at different energy bands. In order to cope with this problem, another point is added to the graph, either before or after the first or last energy band, respectively. The value of this point is fixed to 0.5. A more sophisticated solution has to be found in further studies.

In Masbou [5], some of the MLP input variables were chosen differently.
analysis, the length, the width, the skewness, the kurtosis are used similar to the analysis presented in this thesis. While in the analysis presented in this thesis the size $A$ is used, in the referenced thesis the logarithm of this quantity is used. Taking the logarithm does not result in a better energy reconstruction performance, as neural networks like MLP are able to construct complex nonlinear representations of the input variables. Lastly the nominal distance is used as MLP input in the cited thesis. As stated above, this variable does not provide a significant energy band classification improvement, so that it was chosen to be replaced by the variable $\Delta \varphi$.

Furthermore, gamma ray events with discrete energy values were used for the training of the MLP network in the cited thesis, while the protons were simulated with a continuous spectrum. In the analysis presented here, a simulated continuous energy spectrum is used for both gamma rays and protons.

5.5 Particle Identification

A central topic of this analysis is the identification of the type of the incident particles. This is a very important task, as in ground-based gamma ray astronomy, the vast majority of the observed air showers are proton-initiated, while the interesting events are initiated by gamma rays. There are many possible ways for accomplishing such a discrimination between signal and background events, like cut-based methods, NN or BDT methods as introduced in Chap. 4. In this analysis, an approach using BDTs was chosen due to several reasons. BDTs are able to provide very good discrimination power, even if the fraction of signal events among the entire number of events is low. Besides, the classification of events as well as the training is fast, resulting in a significantly higher computing performance of the analysis software compared to NN methods like MLP, which was the signal / background separation method chosen by Masbou [5]. The gamma / hadron separation performance of BDT and NN methods are comparable, so that both of these approaches are equally well-suited for the task discussed here.

The following variables have been chosen as BDT input:

a) the concentration $C$, 

b) the scaled width $\tilde{W}$, 

c) the scaled length $\tilde{L}$, 

d) $\Delta \varphi$, 

e) $\Delta \tilde{N}_{\text{pix}}$, 

f) $\Delta \tilde{A}$.

The variables a)–d) have already been introduced. The distributions of these four quantities are shown in Fig. 5.9 a–d.
5.5 Particle Identification

Figure 5.9: Distributions of the concentration $C$, $\Delta \varphi$, the scaled length $\tilde{L}$, the scaled width $\tilde{W}$, $\Delta \tilde{N}_{\text{pix}}$ and $\Delta \tilde{A}$ after all cuts except the cuts on the BDT response and on the reconstructed direction. The variables are explained in the text. The proton distributions have been scaled to the area of the gamma ray distributions.
The remaining variables e) and f), denoted as $\Delta \tilde{N}^{\text{pix}}$ and $\Delta \tilde{A}$, are defined according to the following equations:

$$\Delta \tilde{N}^{\text{pix}} = \frac{N_{0103}^{\text{pix}} - N_{0510}^{\text{pix}}}{N_{0103}^{\text{pix}}},$$

$$\Delta \tilde{A} = \frac{A_{0103} - A_{0510}}{A_{0103}}. \tag{5.17}$$

$N_{0103}^{\text{pix}}$ and $N_{0510}^{\text{pix}}$ are the number of pixels after 0103 and 0510 image cleaning, respectively. Analogously, the size values $A_{0103}$ and $A_{0510}$ for different image cleanings are defined. The distributions of the two variables are shown in Fig. 5.9e and f. Both of these variables are constructed such that their values are limited to a range between zero and one. The division by the corresponding values for 0103 image cleaning makes the $\Delta \tilde{N}^{\text{pix}}$ and $\Delta \tilde{A}$ values, which are built from the energy-dependent variables $A$ and $N^{\text{pix}}$, almost independent of the energy of the incident particle. Energy-dependent variables have the disadvantage that gamma ray events with high energies are likely to be removed by cuts on these variables, as the number of events decreases with increasing energy due to the spectrum following a power law. This leads to distributions of these variables which are dominantly influenced by low energy events, so that also the cut values are determined mostly by considering the properties of events with low energies. The particle identification is realised in energy bands, which is going to be explained later. This reduces the severity of this problem, and choosing energy-independent variables resolves it in each energy band.

The choice of most of the aforementioned variables can be motivated by considering the fact that proton-initiated showers are typically more irregular than photon-initiated showers, which has been explained in Sec. 2.4. Therefore the images of proton showers contain more pixels with, in general, lower intensities per pixel. This correlation between the type of the incident particle and the intensity density can be used for particle identification. Therefore the concentration $C$, which is defined as an intensity density, is chosen as a BDT input variable. It can be seen in Fig. 5.9a that the proton shower images have a lower density, which is the expected behaviour.

The fact that proton shower images are less dense can also be utilised in terms of the difference of variables when calculated after different image cleanings. In case of a proton shower, more pixels are removed from the image when applying a stricter image cleaning than in case of a gamma ray shower image, because in proton shower images, the individual pixels deliver lower photoelectron counts. This makes it more likely for these pixels to be removed when applying the stricter 0510 image cleaning. The variables $\Delta \varphi$, $\Delta \tilde{N}^{\text{pix}}$ and $\Delta \tilde{A}$ make use of this principle by calculating the difference of the respective variables after 0103 and 0510 image cleanings. In Fig. 5.9b, e and f, it can be seen that the main axis orientation, the number of pixels and the size, respectively, change more for proton showers than for gamma ray showers. This complies with the previously mentioned picture of less dense and more irregular proton shower images.

Finally, the two variables $\tilde{W}$ and $\tilde{L}$ are used as BDT input variables, the distributions of which are shown in Fig. 5.9c and d. As mentioned earlier, these quantities are
calculated under the assumption that the respective incident particle is a gamma ray, resulting in Gaussian $\tilde{W}$ and $\tilde{L}$ distributions with a width of one and a mean of zero for gamma rays. For proton events, the values of $\langle W \rangle$, $\langle L \rangle$, $\sigma_W$ and $\sigma_L$ for a given pair of $A$ and $\rho$ values are different from the corresponding mean values for gamma ray events. Therefore the $\tilde{W}$ and $\tilde{L}$ distributions of protons are different from the Gaussian gamma ray distributions by construction. Hence these variables can be used as BDT input.

It is also possible to construct $\tilde{W}$ and $\tilde{L}$ variables under the assumption that the corresponding event is a proton event, which could in principle deliver further discrimination power. These variables have been successfully used as input for BDTs in analyses in H.E.S.S. phase I [22]. Also different quantities can be defined analogously to $\tilde{N}^{\text{pix}}$, which has been attempted for the centre of gravity. As the additional separation power gained by using any of these variables is negligible, they are not used in this analysis.

Some of the variables used as BDT input were chosen differently than in the work of Masbou [5], where the length $L$, the width $W$, the natural logarithm of the size $\ln(A)$, the nominal distance $n$, the skewness $S$, the kurtosis $K$ and $\Delta \varphi$ were used as BDT input variables. The variables $L$ and $W$ were replaced by the corresponding scaled parameters, which has the advantage that they are independent of the event energy. The discrimination power of the scaled parameters is comparable to those of the two mentioned Hillas parameters. The size $A$ was not utilised because it does not deliver additional separation power. Also the nominal distance was not chosen due to the dependence of this variable on the position of the Hillas ellipse in the camera. As the protons are diffuse, the centres of gravity of protons have, in general, a larger distance to the centre of the camera system. Hence such a variable would classify all events which have a large distance to the telescope array as background, among them signal events, which is not desired. Furthermore, this cut results in a particle identification performance which is dependent on the chosen wobble offset, so that such a cut could limit the possible observation modes.

As explained in Sec. 5.4, nine energy bands have been constructed. For each energy band, one BDT is created and trained. This is done because the variables listed above are, in general, functions of the energy. Thus it is reasonable to only compare signal events and background events with comparable reconstructed energies.

After the energy reconstruction and assigning each event to an energy band, the BDT output $\zeta$ for each event is used for classifying it as signal or background event. For each energy band $i$, a separate cut value $\zeta^i_0$ has to be found. In this thesis, a method is used which finds the $\zeta^i_0$ values which maximise the significance of the observed signal excess for a source equivalent to the Crab Nebula for the corresponding energy bands. This is accomplished by increasing $\zeta^i_0$ from $-1$ to $1$ with a step size of $0.04$. The increase is stopped if the number of protons which are simulated with a shower direction located in the ON region and which pass the cut on $\zeta$ is smaller than 5. If such a constraint would not be applied, the sensitivity would be maximal for a value of $\zeta^i_0$ which rejects all proton events while keeping an arbitrary fraction of gamma ray events, the possibility of which is only an effect of the small number of simulated background events. In real measurements, the number of background events is usually much larger, making it impossible to reject all background events, as there is a certain fraction of proton events.
Figure 5.10: Sketch of the ON and OFF regions and the OFF region ring. Here the ON region is defined as a circle with an angular distance to the camera centre denoted as $d_{OFF}$ and a radius $\Theta_0$. There are several OFF regions shown as dashed circles and one OFF region ring containing the ON and OFF regions.

which are indistinguishable from gamma ray events. Furthermore it is beneficial if the number of proton events in the OFF region is large enough to be able to predict the amount of background events in the ON region in a statistically significant way. The aforementioned constraint accounts for this.

For each value $\zeta_0$, the significance $S_l$ is calculated using the formula

$$S_l = \sqrt{2} \cdot \sqrt{N_{ON} \cdot \ln \left[ \frac{1 + \alpha}{\alpha} \cdot \frac{N_{ON}}{N_{ON} + N_{OFF}} \right] + N_{OFF} \cdot \ln \left[ (1 + \alpha) \cdot \frac{N_{OFF}}{N_{ON} + N_{OFF}} \right]} \quad (5.18)$$

found by Li & Ma [47]. In this formula, $N_{ON}$ is the number of events whose reconstructed directions are located in the ON region and which pass the cut on $\zeta$. Both proton and gamma ray events are counted. $N_{OFF}$ is the number of events fulfilling the condition $\zeta \geq \zeta_0$ in the OFF regions. $\alpha$ is defined as

$$\alpha = \frac{1}{N}, \quad (5.19)$$

where $N$ is the number of OFF regions. $N$ is a function of the radius of the ON region, the latter one being calculated separately for each energy band. The radius of the ON region in the energy band $i$ is denoted as $\Theta_0^i$ (cf. Fig. 5.10). For each band, $\Theta_0^i$ is varied from 0.05° to 0.5° in steps of 0.05°. The $\Theta_0^i$ value which maximises the significance $S_l$
in the corresponding energy band is chosen. After fixing $\Theta_0^i$ to a certain value, $N$ can be calculated using the following formula, which can be derived using basic geometry relations:

$$N = \frac{\pi}{\arcsin \frac{\Theta_0^i}{d_{\text{OFF}}}} - 1. \quad (5.20)$$

Here, $d_{\text{OFF}}$ is the angular distance between the camera centre and the centre of the ON region. The term in squared brackets in general is not an integer value, hence it is rounded to the lower next integer, which is indicated by the brackets. The term in the brackets represents the total number of circles in the OFF region ring, which includes several circular OFF regions and one ON region. Thus the mentioned term has to be reduced by one in order to give the correct number of OFF regions.

The number of proton events in the circular OFF regions, $N_{\text{OFF}}$, is calculated using an OFF ring around the camera centre as shown in Fig. 5.10. The inner and outer radii $r_i^I$ and $r_i^O$ of the OFF region ring for energy band $i$ are calculated from the radius of the ON region and the angular distance between the ON region centre and the camera centre $d_{\text{OFF}}$:

$$r_i^I = d_{\text{OFF}} - \Theta_0^i \quad (5.21)$$

$$r_i^O = d_{\text{OFF}} + \Theta_0^i \quad (5.22)$$

In order to calculate the number of background events which are located in the ON region, the number of protons in the OFF region ring is counted and scaled to the area of the ON region circle. This is possible as the OFF region ring is radially symmetric with respect to the camera centre and because the incident protons do not have a preferred direction. The advantage of using the entire OFF region ring instead of using several circular OFF regions is that the full number of proton events with a respective radial distance to the camera centre can be used, which minimises the statistical error on the number of proton events in the ON and OFF regions. In order to calculate the number of proton events in the OFF regions $N_{\text{OFF}}$, which is needed for the calculation of $S_l$ according to Eq. (5.18), the number of proton events in the OFF region ring is scaled to the area of the ON region circle, the result of which is then multiplied by the number of OFF region circles $N$.

It should also be noted that there are also gamma ray events with poorly reconstructed shower directions, which thus are not reconstructed inside the ON region but in the OFF region ring. This fact is neglected here as the number of such events is low compared to the number of background events in the OFF region ring. The validity of this argument depends on the fact that the angular resolution is comparable to or smaller than the ON region radius. The values of these quantities are given in Sec. 6.1 and Sec. 6.3.

After finding the optimal combination of $\zeta_0^i$ and $\Theta_0^i$ values for each energy band, the particle identification can be performed by applying cuts on these values.
6 Performance of the Mono Analysis

In this chapter, the performance of the methods introduced in the previous chapter is discussed, including angular resolutions, the energy resolution or the performance of the gamma / hadron separation. Furthermore, quantities commonly used as comparison between different analyses, like effective areas or sensitivities, are introduced and discussed.

6.1 Angular Resolution

In this section, the performances of the two direction reconstruction methods introduced in Sec. 5.3 are discussed. A common measure of this performance is the angular resolution, given as the radius $R_{68}$ of the 68\% containment area. This area is a circular region in the camera around the true source position. The radius of this area depends on the quality of the direction reconstruction.

In this analysis, $R_{68}$ is calculated from the distributions shown in Fig. 6.1a and b. In Fig. 6.1a, the distributions of the squared angular distance between the reconstructed shower direction and the true point-source position for gamma ray and proton events after all cuts besides the cut on the shower direction for the distance method (left) and the $\delta$ method (right). The values of the radii of the 68\% containment circles $R_{68}$ for both kinds of events are given in the insets. In both figures, the area of the proton distributions have been scaled to match the area of the gamma ray distributions.

Figure 6.1: Squared angular distance $\Theta^2$ between the reconstructed shower direction and the true point-source position for gamma ray and proton events after all cuts besides the cut on the shower direction for the distance method (left) and the $\delta$ method (right). The values of the radii of the 68\% containment circles $R_{68}$ for both kinds of events are given in the insets. In both figures, the area of the proton distributions have been scaled to match the area of the gamma ray distributions.
6 Performance of the Mono Analysis

shower direction and the true position of the gamma ray source for both gamma ray and
proton events are shown. The reconstruction of the shower direction was done using the
distance method, hence the angular resolution resulting from this plot can be seen as a
lower limit on the angular resolution for events which passed the same cuts as applied
here. In this case, all cuts besides the one on the reconstructed shower direction have
been applied.

In order to calculate $R_{68}$, the integrals of the shown distributions are calculated. Then
the bin contents are summed up until $68\%$ of the total integral is reached. The upper
limit $\Theta_2^2$ of the last of these bins is used to calculate $R_{68}$:

$$R_{68} = \sqrt{\Theta_2^2}. \quad (6.1)$$

Due to the calculation of $R_{68}$, the given numbers are a conservative estimate of the true
angular resolution. The exact value of $R_{68}$ is slightly lower. From the numbers given in
the inset in Fig. 6.1a, it can be seen that the angular resolution is much better for gamma
ray events than for proton events, which is the expected behaviour, as the protons do
not have a preferred incident direction. Thus the main axes of the Hillas ellipses do not
have a preferred orientation, leading to larger mean distances of the true gamma ray
source position compared to the corresponding values of gamma ray events.

The distance method utilises Hillas parameters only. Therefore the angular resolution
could only be improved by finding a better reconstruction method of the Hillas ellipse,
which is out of scope of this thesis, or by choosing harder cuts on the shower images.
As this mono reconstruction analysis is intended for the reconstruction of events with
low energies, it is not reasonable to choose harder cuts. Therefore the distance method
direction reconstruction cannot be improved significantly.

In Fig. 6.1b, the distributions of the angular distance between the reconstructed di-
rection and the true gamma ray source position is shown for both gamma ray events and
proton events when using the $\delta$ method. In that figure, it can be seen that the average
gamma ray events have a much smaller angular distance to the true point-like source
position than the average proton events. There are several reasons for this behaviour.
One reason is that the lookup variables $\delta$ and $\langle g \rangle$ are constructed using gamma ray MC
samples. As the dependency between these two lookup variables and the corresponding
input parameters is different for proton and gamma ray events, the reconstructed proton
directions differ much from the corresponding true directions. Furthermore, the diffuse
protons do not have a preferred incident direction, therefore the reconstructed directions
are expected to not have a small angular distance to one single point in the camera even
if the $\delta$ and $\langle g \rangle$ values were reconstructed correctly.

In case of gamma ray events, there are two main reasons for the misreconstruction
of the shower direction when using the $\delta$ method, one of them being the calculation of
wrong signs of the components of $\vec{\delta}$, the other one being the assignment of a wrongly
reconstructed absolute $\delta$-value to an event. The total amount of events which have an
angular deviation of more than $\sim 0.3^\circ$ between the reconstructed and the true shower
direction is 49\%. The misreconstruction of the signs of the $\delta$ components accounts for
68\% of these events, while the misreconstruction of the absolute value of $\delta$ is responsible
6.1 Angular Resolution

Figure 6.2: Angular resolution, expressed in terms of $R_{68}$, as a function of the reconstructed energy. Only the plots for gamma ray events are shown, for both the distance and the $\delta$ method as indicated in the legend.

for the remaining 32%. The error resulting from a wrong reconstruction of the ellipse’s main axis is negligible. It is difficult to improve the performance of the reconstruction of the signs of the $\delta$ components. A possible solution would be the training of a BDT including several input variables. The performance of the reconstruction of the absolute $\delta$ value could be improved by finding better-suited lookup variables than $\rho$ and $\ln(A)$. Using multivariate techniques could also improve the $\delta$ reconstruction performance, but further studies have to be done in order to quantify these statements.

Comparing the two figures discussed above, it can be seen that the distance method provides significantly smaller $R_{68}$ values for both gamma ray events and proton events, but as the distance method has disadvantages like being applicable only for point-like gamma ray sources, the analysis is performed for both direction reconstruction methods.

The angular resolution can also be calculated as a function of the energy. In the work of Masbou [5], the angular resolution is given as $R_{68} = 0.15^\circ$ at an energy of 500 GeV and $R_{68} = 0.35^\circ$ at an energy of 20 GeV. The corresponding curves for both direction reconstruction methods resulting from the analysis presented in this thesis are shown in Fig. 6.2. In case of the distance method, the lowest value of $R_{68}$ is 0.07°, which is reached at an energy of $\sim$ 200 GeV, while the maximum value of 0.15° occurs at very low energies of $\sim$ 20 GeV. The same general behaviour can be observed for the $\delta$ method as well. Here the worst angular resolution is 0.72° at 20 GeV and the best angular resolution is 0.22° at 500 GeV. The general decrease of $R_{68}$ with increasing energy is due to the fact that the errors on the Hillas parameters are larger if the event energy is low. At such energies, there are only few pixels, the Cherenkov light intensities are low and the shower
6 Performance of the Mono Analysis

Figure 6.3: (a) Energy reconstruction bias $b_E$ for gamma ray events before and after a cut on the BDT response $\zeta$ (indicated by the superscript) as a function of the true energy.
   (b) Energy resolution $r_E$ as a function of the true energy before and after the BDT cut.

images often have a large roundness, hence the definition of the direction of the ellipse’s main axis is not precise. With increasing energy, the shower images get a more distinct shape, so that the Hillas parameters are better defined. As both direction reconstruction methods depend directly on the Hillas parameters, the angular resolution improves with increasing energy.

6.2 Energy Reconstruction Performance

The performance of the energy reconstruction methods introduced in Sec. 5.4 is discussed in this section. There are two important quantities which quantify the energy reconstruction performance. The first one is the bias, the second one is the energy resolution.

The bias $b_E$ is defined as the expectation value of the distribution of $\sigma_E$ of reconstructed energies for a given true energy. $\sigma_E$ and thus $b_E$ are defined as

$$\sigma_E = \frac{E_{\text{Reco}}(E_{\text{True}}) - E_{\text{True}}}{E_{\text{True}}}$$

$$\rightarrow b_E = \langle \sigma_E \rangle = \frac{\langle E_{\text{Reco}} \rangle (E_{\text{True}}) - E_{\text{True}}}{E_{\text{True}}}.$$  

Here $\langle E_{\text{Reco}} \rangle (E_{\text{True}})$ is the expectation value of the reconstructed energy for a given true energy $E_{\text{True}}$. The bias is calculated in bins of the true energy. In Fig. 6.3a, the energy reconstruction bias is shown for gamma ray events before and after cuts on the BDT response, respectively. It can be seen that for true energies smaller than approximately $50 \text{ GeV} \approx 10^{-1.3} \text{ TeV}$, the reconstructed energy is too high regardless of the cut on $\zeta$, 

...
6.2 Energy Reconstruction Performance

whereas for energies higher than $500\,\text{GeV} \approx 10^{-0.3}\,\text{TeV}$, the bias is negative, meaning that the reconstructed energy is too low.

At low energies, the shapes of the signal and background distributions of the MLP input variables are very similar, hence it is difficult to correctly assign gamma ray events to one of the first three energy bands. An example of the input distributions for the second energy band is shown in Fig. 6.4. This energy band covers the range from 25 to 40 GeV. The distributions for three more energy bands are given in the appendix in Sec. A.3. Apart from the problematic identification of the correct energy band at such low energies, another effect is responsible for the high bias at very low energies. Due to the power law spectrum, there are many events with true energies close to the lower limit of the first energy band. If due to statistical fluctuations in the shower development the amount of produced Cherenkov light is too low, these events are likely to not trigger CT5. Hence such showers are only accepted if they produce comparably much Cherenkov light, leading to a higher reconstructed energy and thus to positive biases. There is a compensating effect that events with higher true energies produce less light than the average gamma ray shower at such energies, but due to the mentioned power law spectrum, this case is less common than the upward fluctuations.

In the consecutive true energy range up to approximately 500 GeV, the bias of events before the BDT cuts is lower than 10\%, indicating that the reconstructed energy on average is in good agreement with the true energy. For events with true energies between 500 GeV and 1 TeV, the bias is negative with a modulus increasing up to 40\%. The reasoning is analogue to the case of very low energies, with the difference that in this...

Figure 6.4: Distributions of the MLP input variables for the second energy band, which includes energies between 25 GeV and 40 GeV. The variables have been introduced in the text.
case only events with lower reconstructed energies are kept. Hence the reconstructed energies of both very low and very high energy events tend to be closer to the centre of the entire energy range than the corresponding true energy values. For events with an applied BDT cut, the general behaviour is similar. At energies below $\sim 35\,\text{GeV}$, the bias is positive and very high. The bias of events with higher true energies is negative, reaching a minimum of $\sim -25\%$ at $70\,\text{GeV}$, then increasing up to a bias value of $5\%$ at $450\,\text{GeV}$ and then dropping again reaching $-40\%$ at $1\,\text{TeV}$.

The difference of the bias curves before and after BDT cuts can be explained by considering the influence of the $\zeta$ cut on the type of events. Both at low and very high energies, there are many events with a poorly reconstructed energy. The images of these events are likely to have a less regular shape than images of events with a better agreement between the true and reconstructed energy. Such irregular events are more likely to be rejected by the cut on the BDT response $\zeta$. Thus, on average, events with a better reconstructed energy pass this cut, leading to smaller deviations from a bias value of zero at both very low and high energies.

At mean true energies, between $\sim 40\,\text{GeV}$ and $\sim 630\,\text{GeV}$, the mean reconstructed energies of events before the BDT cut are in good agreement with the corresponding true energies. The cut on $\zeta$ again removes events which are more irregular, selecting the subset of events which have the properties of regular gamma ray events. As the training of the MLP networks is performed using events with both regular and irregular shapes, the removal of the subset of events with irregular properties introduces the observed energy bias. This effect is also present at lower and higher energies, but in these energy regions it is dominated by the aforementioned effects of systematically rejecting events with too low or too high energies, respectively, and by the poor discrimination power of the MLP input variables at low energies.

Another quantity to be discussed is the energy resolution $r_E$. It is defined as the width of the distribution of $\sigma_E$, which has been defined in Eq. (6.2). As in case of the energy bias, it is evaluated per true energy bin. The distributions shown in Fig. 6.3b are the energy resolutions for gamma ray events before and after the BDT cut. It can be seen that the resolutions are decreasing with increasing energy. This is due to the fact that with increasing event energies, the gamma-ray shower images become more regular, which leads to a more distinctive response of the MLP networks. This in turn results in a value of $E_{\text{Reco}}$ which is closer to the true value, because the distribution of the p-values from the MLP networks has a lower width and thus the mean value is defined more clearly.

The cut on $\zeta$ improves the energy resolution, particularly at low energies. The reasoning done for explaining the impact of the BDT cut on the energy reconstruction bias can be applied here, too. The BDT cut selects a subset of events which have more regular properties, hence the deviation of the reconstructed energies around the corresponding mean value is smaller than in the case when all events are used regardless of the BDT classification. As the energy resolution is a measure of this deviation, its values are lowered if the BDT cut is applied. While at low energies this effect results in a factor of two between the two shown distributions, at energies higher than $200\,\text{GeV}$ the impact of the cut on $\zeta$ is negligible. This is again due to the fact that events with high energies
6.3 Particle Identification Performance

A crucial part of the particle identification is the event classification with respect to the BDT response $\zeta$. As explained in Sec. 5.5, a separate cut value $\zeta^0_i$ is determined for each energy band $i$. The cut values are shown in Fig. 6.5a. For the $\delta$ method, the $\zeta^0_i$ distribu-
6 Performance of the Mono Analysis

tion shows a maximum at mean energies between 100 GeV and 250 GeV, corresponding to the fifth and sixth energy band. The very first band is strongly influenced by statistical fluctuations, as few events are reconstructed in that energy band due to the large energy reconstruction bias at low energies. Furthermore, low energy events are likely to be rejected by the cuts introduced above. For the used number of simulated events, only \( \sim 12,000 \) events are reconstructed in the first energy band and pass the cut on the BDT response, while \( \sim 120,000 \) events pass the \( \zeta \) cut in the second energy band. According to the simulated power law spectrum and the given widths of the energy bands, the number of simulated events in the first bin is higher than the respective number in the second bin, so that the low number in the first bin is an effect of the energy reconstruction bias together with the high likelihood of rejecting these events due to the applied cuts.

In the next energy bands up to the maximum bins, the cut value is increasing with increasing energy. This is due to the fact that the separation at low energies is very difficult, as the input variable distributions for both gamma ray and proton events are very similar. The distributions for the second energy band, which covers energies from 25 GeV up to 40 GeV, are shown in Fig. 6.6a. In Fig. 6.6b, the resulting distribution of the \( \zeta \) values for both gamma ray and proton events are shown. In the appendix in Sec. A.4, the distributions for three more energy bands are given. From the figures shown in the appendix, it can be concluded that the particle identification improves significantly with the energy.

Due to the poor discrimination performance at energies as low as those in the second energy band, many signal events are rejected due to the \( \zeta \) cut, so that the cut optimisation leads to comparably small \( \zeta_0 \) values.

At higher energies, the small number of proton events limits the choice of the cut values. The condition explained in Sec. 5.5 that the number of proton events in the ON region is required to be larger than five together with the good gamma / hadron separation at high energies leads to smaller cut values compared to the energy bands at mean energies. At mean energies of about 100 GeV, the number of proton events is still large enough that the constraint on the number of protons has low impact, while the gamma / hadron separation is already good, so that the \( \zeta \) cut values can be chosen higher than in other energy regimes.

The influence of the chosen direction reconstruction method on the choice of the \( \zeta_0 \) cut values is significant. In most energy bands, the cut values resulting from the \( \delta \) method are higher than the values from the distance method. This is a result of the fact that due to the cut on the reconstructed shower direction, the number of proton events is reduced more when applying the distance method instead of the \( \delta \) method. Therefore it is more likely for the \( \delta \) method than for the distance method to fulfil the condition that the number of proton events has to exceed the value five.

Along with the optimisation of the \( \zeta_0 \) cut values, the radii of the ON regions \( \Theta_0 \) have been calculated. The distribution of the resulting values is shown in Fig. 6.5b. According to Fig. 6.2, the angular resolution is worse for low energies (\( E \lesssim 80 \) GeV) than for higher energies, therefore more signal events are reconstructed with a larger angular distance to the true source position. In order to keep more of these signal events while rejecting as many background events as possible, the radius of the circular ON
6.3 Particle Identification Performance

Figure 6.6: Top: BDT input variables for the second energy band, which covers energies from 25 up to 40 GeV. The variables have already been introduced in Sec. 5.5. Bottom: Resulting $\zeta$ distributions (“BDT response”) for this energy band. The “Signal” distribution corresponds to gamma ray events, “Background” stands for proton events.
region has to be chosen higher for low energies, although the effect is small according to the shown distributions. With increasing energy, the $\Theta_0$ values are increasing as well, despite the fact that the angular resolution is decreasing with increasing energy. This can be explained by the aforementioned condition that the number of proton events inside the ON region and with an energy corresponding to the respective energy band has to be higher than five. At higher energies, the proton flux is lower than at lower energies, therefore the radius of the ON region has to be chosen larger.

The ON region radii differ significantly between the two direction reconstruction methods, as the angular resolution is much better in case of the distance method. Nevertheless, the general development of the $\Theta_0$ with increasing energy values is very similar for both methods, only the absolute scaling is different. The explanations made above can be applied to both methods.

According to Tab. 5.1, the number of gamma ray events is reduced by only 4.8% by the cut on the BDT response with respect to the number of events after the size cut, while the number of proton events is reduced by 35.7%. These numbers relate to the distance method. Hence this cut improves the ratio of signal events to background events significantly.

In all energy bands besides the first one, the BDT input variables providing the highest discrimination power are the concentration and $\Delta \tilde{A}$, while $\Delta \varphi$ is the least important variable. In the work of Masbou [5], the last-mentioned variable is found to be the most important discrimination variable. The $\Delta \varphi$ distributions from this thesis, averaged over all energy bands, as well as the distribution from the cited thesis, are shown in Fig. 6.7. It can be seen that the distributions shown in the two figures at the top differ much, which might be due to the application of different event selection cuts. In the cited thesis, only a cut on the nominal distance ($n < 1^\circ$) and a cut on the shower direction were applied. In order to investigate the proposition that the different cuts are the reason for the difference between the two distributions, the distributions of the $\Delta \varphi$ variable after applying all cuts up to the technical cuts (cf. Tab. 5.1), a cut on the nominal distance and a cut on the shower direction is shown in Fig. 6.7c. It can be seen that the impact of the change of cuts on the distributions is low.

The impact of the change of cuts as well as the energy dependence of the width of the $\Delta \varphi$ distributions can be quantified by analysing the distributions shown in Fig. 6.7d. In this figure, the widths of the mentioned distributions are shown as functions of the true energy for gamma ray and proton events. The width $\sigma_{\Delta \varphi}$ of the $\Delta \varphi$ distribution is defined as the position on the abscissa at which the $\Delta \varphi$ distribution reaches half the maximum value, which in turn is identified with the number of entries in the first bin. For each type of incident particles, the width distributions are shown for two different cut schemes. The first one is the standard cut chain except the BDT cut and the cut on the shower direction, the second one conforms to the same cuts as applied in the work of Masbou [5]. In the first energy bands, the $\sigma_{\Delta \varphi}$ values have been set to 90° due to two reasons. In the very first bin, the number of events is so low that it is unreasonable to define a width of the distribution. In the next bins, the number of events is large enough, but the $\Delta \varphi$ distributions in these energy bands are very flat, so that at no point on the abscissa the half maximum value is reached. In the remaining bins, the widths are
6.3 Particle Identification Performance

Figure 6.7: $\Delta \phi$ distributions resulting from this analysis after all cuts except the BDT cut and the cut on the reconstructed direction (top left) and after applying only a nominal distance cut and a cut on the reconstructed shower direction (bottom left). The top right figure shows the distributions from Masbou [5]. Here the gamma ray distribution is denoted as “Signal” and the proton distribution is labeled as “Background”. Also the widths of the $\Delta \phi$ distributions as functions of the true event energy are given for both gamma rays and protons for two different cut patterns (bottom right). The first cut pattern requires all cuts except the BDT cut and the cut on the reconstructed shower direction, the second cut pattern, which is indicated by the subscript “C”, requires the same cuts as applied in the plots shown in the bottom left figure.
6 Performance of the Mono Analysis

decreasing with increasing energy, which is the expected behaviour, as the Hillas ellipses are better defined if the Cherenkov light intensities are increasing. Therefore the main axis directions change less after different image cleanings for both gamma ray and proton events. In case of gamma ray events, the choice of the cut pattern has negligible influence on the $\Delta \varphi$ distribution, whereas the proton distribution changes significantly in at least one energy band. Here the widths of the distributions are smaller when the cuts similar to the ones applied in the work of Masbou [5] are used. Regardless of the influence of the cuts, the widths of the gamma ray distributions exceed the width of the distribution shown in Fig. 6.7b of approximately $10^\circ$ in the last energy band only, which corresponds to energies between 600 GeV and 1 TeV. At lower energies, the widths of the respective gamma ray distributions are significantly larger. Thus further investigations are needed in order to find an explanation for the inconsistencies between the distributions resulting from this thesis and from the cited one.

6.4 Effective Areas and Event Rates

A quantity needed for the measurement of fluxes which is also frequently used for the quantification of the performance of analyses like the one presented here is the so-called effective area of gamma ray events $A_{\gamma}^{\text{eff}}$. It is defined as the circular area $A_\gamma$ around the telescope perpendicular to the pointing direction of the telescope, multiplied by the efficiency $\varepsilon$ with which the events pass the cut chain:

$$A_{\gamma}^{\text{eff}} = A_\gamma \cdot \varepsilon.$$  \hfill (6.4)

For diffuse protons, the effective area is defined in a different way. Here $A_\gamma$ is the area around the telescope multiplied by the solid angle corresponding to the cone in which the incident directions of the simulated protons are contained. This cone is defined by its opening angle $\theta_c$. The multiplication with a solid angle is not necessary for gamma rays originating from a point-like source.

In case of gamma ray events, the radius of the area $A_\gamma$ around the telescope is $R_{\gamma \text{MC}} = 600$ m, the corresponding value for proton events is $R_{p \text{MC}} = 1000$ m. The opening angle is $\theta_c = 5^\circ$. All of these values are defined in the MC data samples. The corresponding areas $A_{\gamma/p}$ are then given as

$$A_\gamma = \pi (R_{\gamma \text{MC}})^2 = 1.13 \times 10^6 \text{ m}^2$$ \hfill (6.5)

$$A_p = 2\pi (1 - \cos\theta_c) \cdot \pi (R_{p \text{MC}})^2 = 7.51 \times 10^4 \text{ m}^2 \text{ sr.}$$ \hfill (6.6)

These constants are multiplied by the energy-dependent efficiency values $\varepsilon_{\gamma/p}(E)$, which are defined as the number of events after all cuts, including the cuts on the shower direction and on the BDT response $\zeta$, divided by the total number of simulated events. In case of proton events, the weighted number of events is used. The calculation of the $\varepsilon$ values is done for each energy band, and also for each of the two direction reconstruction methods. The direction reconstruction has an influence on the efficiencies due to two reasons. The cut on the reconstructed direction depends on the distance between the true
Figure 6.8: Efficiencies $\varepsilon_{\gamma/p}$ of both gamma ray events (top left) and proton events (top right) as functions of the true event energy. The effective areas $A_{\gamma/p}^{\text{eff}}$ as functions of the energy are shown at the bottom. The distributions of the effective areas for gamma ray events when using each of the direction reconstruction methods are shown in the bottom left figure, the proton distribution resulting from the distance method is shown in the bottom right figure.

...position of the gamma ray source and the reconstructed direction, so that events with a large angular distance to the true source positions are rejected. Also the definition of the ON region for each energy band depends on the direction reconstruction (cf. Sec. 5.5). Apart from the influence of the cut on the shower direction, the cut values for the BDT response depend on the choice of the direction reconstruction procedure, which in turn leads to different $\varepsilon$ values.

The distributions of the $\varepsilon$ values for gamma ray events as well as for proton events as a function of the true energy are shown in Fig. 6.8a and b. It can be seen that the acceptance of events depends strongly on the energy. Especially low energy events are rejected, as the Cherenkov light intensities are low for such events. Therefore these events are more likely to be rejected by the size cut and the cut on the number of...
pixels. Furthermore, the BDT input variable distributions for both gamma ray and proton events are more similar for small event energies, making it much more difficult to discriminate between signal and background events. Hence the cut on the BDT response is less efficient for low energy gamma ray events than for events with higher energies.

The proton efficiencies are much smaller than the gamma ray efficiencies. This is due to the definition of the efficiencies. The total number of simulated events is much larger than the number of events which have triggered. Especially events with very low energies are unlikely to trigger at all, and the Cherenkov light gain is reduced by a factor of approximately three compared to gamma ray events with the same energy (cf. Sec. 2.4), therefore a significantly larger fraction of simulated proton events do not trigger compared to the fraction of gamma ray events.

The resulting effective area distributions for gamma ray events is shown in Fig. 6.8c and d. At low energies the values resulting from the $\delta$ method are larger than the values from the distance method. This can be explained by considering the fact that in case of distance method the particle identification relies more on the cut on the shower direction than on the cut on $\zeta$. As the ON regions are smaller than the angular resolution, more gamma ray events are rejected even though the overall significance reached at these low energies is high. In case of the $\delta$ method more gamma ray events are accepted, while also accepting more protons. Hence the efficiencies and the effective areas are higher in case of the utilisation of the $\delta$ method. At high energies, the angular resolutions are comparable to the radii of the ON regions, so that relatively many gamma ray events pass the cut on the shower direction, while the angular resolution of proton events remains unchanged. Thus the cut on the direction reduces the efficiency of gamma ray events less significantly than at low energies, so that the cut on $\zeta$ becomes more important for the gamma ray efficiencies also in case of the distance method. Hence here the values from the $\delta$ method are lower than the ones from the distance method.

At energies above 1 TeV, the reconstructed energy deviates very much from the true energy, as the energy bands in which the MLP networks have been trained only cover energies up to this threshold. Also the BDTs are not trained for events at such high energies, so that these events are likely to be rejected by the BDT cut. Thus the effective area in this region decreases with increasing energy.

The effective areas of proton events as a function of the energy when using the distance method is shown in Fig. 6.8d. The general form is similar to the shape of the gamma ray distributions, only the absolute scaling is different. The reasoning applied to gamma ray events can be applied here, too.

In Tab. 6.1, the effective areas calculated in this thesis for two energy values and the respective values given in Masbou [5] are listed. The effective areas in the cited thesis are larger than the values resulting from the analysis discussed here. This is due to the fact that the performance of the gamma / hadron separation presented in the referenced thesis is higher than the performance achieved in this analysis. Therefore the cut on $\zeta$ has less influence on the number of gamma ray events compared to this analysis, resulting in higher efficiencies and thus in larger effective areas.

Apart from the effective areas, the background event rate is an interesting measure of the performance of the analysis. In Fig. 6.9, the integral rate of proton events is
6.5 Sensitivities

<table>
<thead>
<tr>
<th>$E_{\text{True}}$</th>
<th>$A^\gamma_{\text{eff}}$ using distance / $\delta$ method</th>
<th>$A^\gamma_{\text{eff}}$ given in Masbou [5]</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 GeV</td>
<td>6200 m$^2$ / 12 000 m$^2$</td>
<td>40 000 m$^2$</td>
</tr>
<tr>
<td>80 GeV</td>
<td>47 000 m$^2$ / 69 000 m$^2$</td>
<td>90 000 m$^2$</td>
</tr>
</tbody>
</table>

Table 6.1: Effective areas for two values of the true event energy. The values calculated using the two different direction reconstruction methods and the values given in Masbou [5] are listed.

Figure 6.9: Integral rate of proton events $R_p$ as a function of the reconstructed energy $E_{\text{Reco}}$.

shown. This quantity is defined as the rate of background events with a reconstructed energy larger than the energy the function is evaluated at. For energies smaller than approximately 500 GeV, the proton rates are smaller when using the distance method instead of the $\delta$ method for the direction reconstruction. Thus at low energies, both the effective area of gamma ray events and also the proton rate is higher in case of the $\delta$ method, while both of these quantities are lower for this method at high energies. The reasoning about the relation of the proton rates for the different direction reconstruction methods is similar to the case of the effective areas.

6.5 Sensitivities

Two of the most important performance measures of an analysis are the differential and the integral sensitivity curves. Both of these quantities are explained and discussed in this section.
6 Performance of the Mono Analysis

The differential sensitivity is defined as the minimal flux in a certain energy range which can be detected with a significance of $5\sigma$ in an observation time of $T_{\text{obs}} = 50$ h. The value of this minimal flux is calculated for each energy band individually. The spectral index $\alpha_C = 2.39$ of the flux is chosen to be identical with the spectral index of the Crab Nebula flux in order to have the possibility to denote the calculated minimal flux as a percentage of the latter one. According to Aharonian et al. [48], the differential flux $\frac{d\Phi_C}{dE}$ of the Crab Nebula is given as

$$\frac{d\Phi_C}{dE} = 3.76 \times 10^{-7} \left(\frac{E}{1 \text{ TeV}}\right)^{-2.39}.$$

In the cited paper from which this formula is taken a cutoff energy of $E^C = 14.3$ TeV is given. As mono analyses like the one presented in this thesis are going to be used for the analysis of events with low energies of $E \lesssim 100$ GeV, the exponential cutoff of the spectrum at higher energies can be neglected. The impact of neglecting the cutoff factor at such low energies is less than 1%. Nevertheless it should be noted that the discussion about the spectrum has not settled yet, because alternatives as given in Albert et al. [49] have to be discussed.

From the differential flux given in Eq. (6.7) and the effective area introduced in Sec. 6.4, the number of signal events $N_i^\gamma$ which pass the cut chain including the cut on the shower direction in each energy band can be calculated as

$$N_i^\gamma = T_{\text{obs}} \int_{E_1}^{E_2} \frac{d\Phi_C}{dE}(E)A_{\text{eff}}(E) \, dE$$

$$\approx T_{\text{obs}} \cdot A_{\text{eff}}(\langle E \rangle) \int_{E_1}^{E_2} \frac{d\Phi_C}{dE}(E) \, dE.$$  

(6.8)

(6.9)

Here $E_1^{1/2}$ are the lower and upper limits of the corresponding energy band $i$, and $A_{\text{eff}}(\langle E \rangle)$ is the effective area value at the mean energy of the respective energy band. The approximation made in Eq. (6.9) is possible if the widths of the energy bands are sufficiently small, so that the effective area can be considered as a constant in the energy band. This is approximately true for energies higher than 40 GeV. Below this threshold, the results produced by this approximation have to be regarded as a rough estimate.

In order to calculate the significance according to Eq. (5.18), the number of background events $N_i^p$ which pass the entire cut chain per energy band $i$ has to be known. This is done using the proton event rate per energy band $R_i^p$, which is defined as

$$R_i^p = \frac{\sum_{j=1}^{n_i^p} w_j(E_j)}{T}.$$  

(6.10)
6.5 Sensitivities

Figure 6.10: Differential and integrated sensitivity curves plotted together with functions representing the respective values of a source with a flux of one percent of the Crab Nebula’s flux. The sensitivities calculated using both the distance method and the $\delta$ method for the direction reconstruction are shown in each figure. The values resulting from the utilisation of the $\delta$ method are indicated by the “$\delta$” superscript.

In this equation, $T$ is the simulated time and $w_j(E_j)$ is the weight of event $j$ with a true energy $E_j$. The calculation of both $T$ and $w_j$ is shown in Sec. A.2. $n^p_i$ is the number of proton events reconstructed in the the respective energy band after all cuts. From the proton rate, the number of proton events in that energy band for the given BESS spectrum can be calculated as

$$N^p_i = T_{\text{obs}} \cdot R^p_i.$$  

(6.11)

From the values $N^p_i$ and $N^\gamma_i$, the differential sensitivity can be calculated by scaling $N^\gamma_i$ by a factor $\eta$ in such a way that the significance $S_i = 5\sigma$. The differential sensitivity $d\Phi_{\text{min}}/dE$ is then defined as the differential flux of the Crab Nebula at the mean energy of the corresponding energy band, multiplied by $\eta$:

$$d\Phi_{\text{min}}/dE(E) = \eta(E) \cdot d\Phi_C/dE(E).$$  

(6.12)

This is the minimal flux which is detectable in 50h with a significance of $5\sigma$. The resulting distributions of this quantity as a function of the true energy, calculated for both direction reconstruction methods, are shown in Fig. 6.10a, together with the curve representing one percent of the differential Crab flux. The numbers for three exemplarily chosen energies are given in Tab. 6.2. The sensitivity curves shown in the figure are decreasing with increasing energy. One reason for this behaviour is that the spectral index of the proton events is bigger than the index of the gamma ray events. Thus the number of proton events decreases faster with increasing energy than the number of gamma ray events. Hence the gamma ray flux can be multiplied by a smaller number. Furthermore, the background suppression by the BDT cut is more efficient at higher
energies. Thus not only the absolute flux values, but also the flux measured in units of the Crab flux decreases. For energies higher than 100 GeV, the percentage of the differential Crab flux increases again. This can be explained by considering the distribution of the cut values \( \zeta_0 \), which decrease with increasing energy due to the previously mentioned constraint that the number of proton events in the ON region has to exceed the value five. Thus the number of proton events remains constant, while the values of the flux integrals decrease with increasing energy number. Hence the \( \eta \) values have to be chosen larger in order to reach the same significance of 5\( \sigma \). The general behaviour is independent of the utilised direction reconstruction method.

At energies below approximately 60 GeV \( \approx 10^{-1.2} \) TeV, the differential sensitivities calculated when using the \( \delta \) method are smaller than the corresponding values for the distance method. This is due to the fact that the efficiency for gamma ray events is higher for the \( \delta \) method, as explained in Sec. 6.4. The proton rate is higher as well, but for the calculation of the sensitivities, the resulting significance per energy band is the important quantity. The significance is approximately proportional to \( \frac{N_{\gamma}^{\text{ON}}}{\sqrt{N_{\gamma}^{\text{ON}} + N_{p}^{\text{ON}}}} \), where \( N_{\gamma}^{\text{ON}} \) and \( N_{p}^{\text{ON}} \) are the number of events in the ON region for gamma rays and protons, respectively. The significance at low energies is higher in case of the \( \delta \) method, so that the scaling factor \( \eta \) can be chosen smaller. At higher energies, the distance method delivers higher values. The reasoning is similar to the case of small energies. The two direction reconstruction methods differ by a factor of less than two.

In the standard procedure of the differential sensitivity calculation, usually two conditions have to be fulfilled. The first one demands that the number of events after all cuts is bigger than ten. This condition is applied for each energy band. The second condition requires that the ratio of the number of signal events and proton events after all cuts is larger than 5\%. This condition has not been implemented in this analysis, as it cannot be fulfilled in all but two energy bands. The reason for the fact that the second condition cannot be fulfilled is inherently a part of the calculation of the minimal flux which can be detected in 50 h, as the gamma ray flux is scaled down, starting from the flux of the Crab Nebula, until a significance of 5\( \sigma \) is reached. Hence the number of proton events and gamma ray events is fixed, and the ratio of the resulting numbers is almost always smaller than the required 5\%.

A quantity closely related to the differential sensitivity is the integral sensitivity. It is

<table>
<thead>
<tr>
<th>( E_{\text{True}} )</th>
<th>( \frac{d\Phi_{\text{min}}}{dE} / \frac{d\Phi_{C}}{dE} ) for the distance / ( \delta ) method</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 GeV</td>
<td>21.6% / 11.3%</td>
</tr>
<tr>
<td>100 GeV</td>
<td>7.1% / 8.9%</td>
</tr>
<tr>
<td>250 GeV</td>
<td>10.5% / 19.7%</td>
</tr>
</tbody>
</table>

Table 6.2: Differential sensitivity values for three different energies. The values obtained from using both the distance method and the \( \delta \) method for the direction reconstruction are shown. The sensitivities are given in percent of the differential Crab flux \( \frac{d\Phi_{C}}{dE} \).
6.5 Sensitivities

defined as the minimal integrated flux $\Phi_{\text{min}}(E)$ which can be detected with a significance of $S_l = 5\sigma$ in an observation time of $T_{\text{obs}} = 50$ h. The procedure of calculating the significance is similar to the one described above, with the only difference being the definition of the number of signal and background events. In this case, the number of signal events $N^\gamma(E)$ is defined as

$$N^\gamma(E) = T_{\text{obs}} \cdot \int_E^\infty \frac{d\Phi_C(E')}{dE'} A_{\text{eff}}^\gamma(E') dE'.$$

(6.13)

The number of proton events is calculated analogously to Eq. (6.11), but the proton rate is defined differently:

$$R^p(E) = \frac{1}{T} \int_E^\infty \frac{d\omega}{dE'} (E') dE'.$$

(6.14)

Instead of summing up the weights per energy band, the differential weights are integrated over the energy range from the starting energy $E$ up to infinity. From this rate, the number of proton events is calculated via

$$N^p(E) = T_{\text{obs}} \cdot R^p(E).$$

(6.15)

From $N^p$ and $N^\gamma$, the significance $S_l$ can be calculated. The number of signal events $N^\gamma$ is scaled with a factor $\eta'$ such that $S_l = 5\sigma$. $\eta'$ is then used as a scaling factor for the integrated Crab flux:

$$\Phi_{\text{min}}(E) = \eta'(E) \int_E^\infty \frac{d\Phi_C(E')}{dE'} dE'.$$

(6.16)

The fluxes calculated this way are shown in Fig. 6.10b. The distributions of the integral sensitivities calculated for both direction reconstruction methods are shown as functions of the true energy. It can be seen that, analogously to the differential sensitivities, the values resulting from the two direction reconstruction methods are very similar. Due to the higher differential sensitivity values resulting from the $\delta$ methods at high energies, the integral sensitivity is almost always higher for this method, except for very low energies. The integral sensitivities of the two methods again differ by a factor of not more than two.

In Tab. 6.3, the integral sensitivity values for the same three energies as in Tab. 6.2 are given, along with the respective values taken from the integral sensitivity plot given in Masbou [5]. The values taken from the cited thesis are mostly smaller than the sensitivities calculated for both the distance and the $\delta$ method. This is due to the better gamma / hadron separation of the cited analysis, which is mostly a result of the powerful discrimination variable $\Delta\varphi$.

A quantity related to the sensitivities is the observation time $T_{5\sigma}$ which is necessary in order to detect a source similar to the Crab Nebula. It is calculated using the entire
6 Performance of the Mono Analysis

<table>
<thead>
<tr>
<th>$E_{\text{True}}$</th>
<th>$\Phi/\Phi_C$ for the distance / $\delta$ method</th>
<th>$\Phi$ given in Masbou [5]</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 GeV</td>
<td>5.2% / 5.0%</td>
<td>10%</td>
</tr>
<tr>
<td>100 GeV</td>
<td>6.9% / 10.2%</td>
<td>2%</td>
</tr>
<tr>
<td>250 GeV</td>
<td>13.9% / 22.7%</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

Table 6.3: Integral sensitivity values for three different energies. The values obtained from using both the distance and the $\delta$ method for the direction reconstruction are shown, as well as the values taken from Masbou [5]. The sensitivities are given in percent of the integral Crab flux $\Phi_C$.

energy range covered in the analysis. As for the sensitivity values, the number of signal and background events in the ON and OFF regions is used as input for the significance formula given in Eq. (5.18). The value $\alpha$, which is the inverse of the number of circular OFF regions (cf. Sec. 5.5), has to be calculated as an average value, as the number of OFF regions varies with the energy band number. The average value of $\alpha$ is calculated as the weighted mean of the $\alpha_i$ values for each energy band according to the equation

$$\bar{\alpha} = \frac{\sum_{i=1}^{9} \alpha_i N_{\text{OFF}}^i}{\sum_{i=1}^{9} N_{\text{OFF}}^i}.$$  \hspace{1cm} (6.17)

The $\alpha_i$ values of each energy band are weighted with the number of events in the OFF regions $N_{\text{OFF}}^i$ for each energy band. The mean $\alpha$ value is then used in the significance formula in order to calculate $T_{5\sigma}$. In this thesis, the value resulting from the distance direction reconstruction method is

$$T_{5\sigma} = 250 \text{ s} \approx 4.2 \text{ min}.$$ \hspace{1cm} (6.18)

In the given accuracy, the corresponding value for the $\delta$ method is equal to the one given in Eq. (6.18). The mean value of $\alpha$ has been defined analogously to Eq. (6.17) for the $\delta$ method as well, replacing the definition of the ON and OFF regions according to the values calculated for the $\delta$ method.

The corresponding value of $T_{5\sigma}$ given in the work of Masbou [5] is about 1 min, which is significantly lower than the values resulting from this analysis. This is again a result of the better particle identification performed in the cited analysis.
7 Summary and Outlook

Recently a new telescope called CT5 has been added to the array of the four existing telescopes of H.E.S.S. phase I, marking the beginning of phase II. The new telescope will enhance the performance of the experiment significantly as it has a much larger mirror area than each of the four small telescopes. This thesis presents a full reconstruction of events which have triggered at least the new telescope. Only simulated gamma rays and protons, the latter ones providing the predominant background, are regarded, while heavier atomic nuclei or electrons are neglected. The calculation of the properties of the incident particles is based on the Hillas parameters reconstructed from the intensity distribution in the camera of CT5. The analysis covers the reconstruction of the direction and the energy as well as the identification of the type of the incident particles.

The direction reconstruction is accomplished in two different ways. The $\delta$ method uses the Hillas parameters and a displacement value $\delta$, which is obtained from a lookup table. The distance method utilises the Hillas parameters and assumes that a point-like source with a known position is searched for. This method provides better angular resolutions than the $\delta$ method, but in the current state of the analysis it is only applicable for the observation of point-like gamma ray sources. The $\delta$ method, though providing a worse angular resolution, is implemented as an alternative method as in case of this method no assumptions have to be made about the properties of the gamma ray source. The mean angular resolution for gamma ray events provided by the distance method is 0.11°, the corresponding value for the $\delta$ method is 0.50°. The angular resolutions are improving with increasing energy for both direction reconstruction methods.

Another important topic of this analysis is the reconstruction of the event energy using multilayer perceptrons (MLP). The input variables of these neural networks are the width $W$, length $L$, size $A$, skewness $S$ and kurtosis $K$ of the reconstructed Hillas ellipses and a variable called $\Delta\varphi$, which represents the difference between the two orientations of the main axis of the Hillas ellipse after two different image cleanings. In this analysis, the 0103 and the 0510 image cleanings are used. The energy resolution at energies below 100 GeV is 60–70% if a cut on the BDT response is not applied and 35–55% if it is applied. At high energies, there is no difference between the resolution before and after the BDT cut. Here a resolution of about 10% at 1 TeV is reached. The energy reconstruction bias is 20–80% at energies below 25 GeV, but it decreases rapidly with increasing energy. In the energy range between 25 and 800 GeV, the bias is lower than 25% both before and after the cut on the BDT response $\zeta$.

The reconstructed energy is used to assign each event to one out of nine energy bands. In each of these bands, a boosted decision tree (BDT) is created and trained in order to separate gamma ray events from proton events with similar reconstructed energies. The input variables for the BDTs are the scaled width $\tilde{W}$ and length $\tilde{L}$, the concentration $C$, and so on.
7 Summary and Outlook

\(\Delta \varphi\), and the two quantities \(\Delta \tilde{N}_{\text{pix}}\) and \(\Delta \tilde{A}\). The last-mentioned variable is calculated as the difference between the image size \(A\) after the two image cleanings, divided by the size after the 0103 image cleaning. \(\Delta \tilde{N}_{\text{pix}}\) is defined analogously. The performance of the particle identification using cuts on the BDT response and on the reconstructed shower direction can be expressed in terms of the effective area and sensitivities. The effective area for gamma ray events with a true energy of about 100 GeV is 57 000 m\(^2\) in case of the distance method and 76 000 m\(^2\) for the \(\delta\) method. The differential sensitivities at 100 GeV resulting from the distance and the \(\delta\) method are 7.1\% and 8.9\% of the differential flux of the Crab Nebula, respectively. The integral sensitivities for the two direction reconstruction methods at the same energy are 6.9\% and 10.2\% of the integral flux of the Crab Nebula, respectively. These values are slightly higher, but still comparable to those given in the work of Masbou [5]. In that thesis, a mono reconstruction software is presented which makes use of the \(\Delta \varphi\) variable as the most important variable for the gamma / hadron separation. In the analysis presented in this thesis, the outstanding separation power of this variable could not be confirmed. Here the most important discrimination variables are the concentration and \(\Delta \tilde{A}\), while \(\Delta \varphi\) is found to be the least important of the utilised BDT input variables.

Further studies could be done in order to improve the \(\delta\) method, especially the determination of the signs of the components of \(\vec{\delta}\). A feasible attempt could be the utilisation of additional BDTs constructed for the purpose of the determination of the mentioned signs or even for the reconstruction of \(\vec{\delta}\) vector including the signs of the components. In case of the energy reconstruction, the function by which the MLP p-values are fitted could be chosen differently, as it is not proven that a Gaussian distribution is the optimal choice. Apart from the choice of the fit function, better input variables to the MLP networks could, in principle, be found. The same applies to the BDT input variables. Furthermore, it is possible to use the timing information, which is available for the new telescope. This means that the information about time at which each camera pixel has been illuminated is available, so that the development of the shower image in the camera becomes traceable. This could in principle help to improve the determination of the correct signs of \(\vec{\delta}\) in the \(\delta\) method.
8 Zusammenfassung und Ausblick


Die Richtungsrekonstruktion wird in dieser Analyse auf zwei verschiedene Arten durchgeführt. Die δ-Methode verwendet die Hillasparameter und eine Verschiebungsgröße δ, die aus Lookup-Tabellen gewonnen wird. Die Distance-Methode arbeitet ebenfalls mit Hillasparametern, es muss im Rahmen dieser Methode jedoch angenommen werden, dass nach Punktschwellen mit bekannter Richtung gesucht wird. Diese Methode ergibt bessere Winkelauflosungen als die δ-Methode, ist aber im gegenwärtigen Zustand der Analyse nur bei der Beobachtung punktartiger Quellen von Gammastrahlen einsetzbar. Daher wurde auch die δ-Methode, obwohl sie eine schlechtere Winkelauflosung bietet, als Alternative implementiert, da bei dieser Methode keinerlei Annahmen über die Eigenschaften der Gammastrahlenquelle getroffen werden müssen. Die mittlere Winkelauflosung für Gammastrahleneereignisse beträgt 0.11° für die Distance-Methode und 0.50° im Falle der δ-Methode. Die Auflösungen beider Methoden verbessern sich mit steigender Energie.

Eine weitere, wichtige Aufgabe dieser Analyse ist die Rekonstruktion der Energie der einfallenden Teilchen, wofür mehrlagige Perzeptrons (MLP) verwendet werden. Die Eingabeveränderlichen dieser neuronalen Netzwerke sind die Breite $W$, die Länge $L$, die Größe $A$, die Skewness $S$ und die Kurtosis $K$ der rekonstruierten Hillas-Ellipsen. Weiterhin wird eine Variable $\Delta \Phi$ verwendet, welche als Differenz zwischen den Richtungen der Hauptachsen der beiden Hillasellipsen nach zwei unterschiedlichen Bildbereinigungsprozeduren („Cleanings“) berechnet wird. In der vorliegenden Analyse werden das 0103- und das 0510-Cleaning verwendet. Die Energieauflösung bei Energien kleiner als 100 GeV beträgt 60–70 % für den Fall dass kein Cut auf die $\zeta$-Variable durchgeführt wird und 35–55 % für den Fall dass dieser Cut angewandt wird. Bei hohen Energien ist der Unterschied zwischen diesen beiden Fällen gering, wobei eine Auflösung von etwa 10 % bei 1 TeV erreicht wird. Der Bias der Energierkonstruktion bei sehr niedrigen Energien ist
sehr hoch, nimmt aber schnell mit steigender Energie ab. Im Bereich zwischen 30 und 800 GeV ist er sowohl für Ereignisse mit und ohne angewandtem Cut auf die $\zeta$-Variable kleiner als etwa 25%.

Die rekonstruierte Energie wird verwendet, um die Ereignisse in eines der neun möglichen stehenden Energiebänder einzuteilen. In jedem Band wird ein „Boosted Decision Tree“ (BDT) erstellt und trainiert, um Gammastrahlereignisse von Protonereignissen mit ähnlichen rekonstruierten Energien separieren zu können. Die Eingabevariablen der BDTs sind die skalierte Breite $\tilde{W}$ und Länge $\tilde{L}$, die Concentration $C$, $\Delta \phi$ sowie die beiden Größen $\Delta N_{\text{pix}}$ and $\Delta A$. Die letztgenannte Variable wird als die Differenz zwischen den Bildgrößen $A$ nach den beiden verschiedenen Cleanings berechnet, wobei das Resultat durch den entsprechenden Wert nach dem 0103-Cleaning dividiert wird. $\Delta N_{\text{pix}}$ ist analog definiert. Die Effizienz der Teilchenidentifikation kann mit Hilfe von effektiven Flächen und von Sensitivitäten ausgedrückt werden. Die effektive Fläche für Gammastrahlereignisse mit einer wahren Energie von etwa 100 GeV beträgt 57 000 m² für die Distance-Methode und 76 000 m² für die $\delta$-Methode. Die differentiellen Sensitivitäten bei 100 GeV betragen 7.1 % bzw. 8.9 % des differentiellen Flusses des Krebsnebels. Die integrierten Sensitivitäten bei der erwähnten Energie betragen 6.9 % bzw. 10.2 % des integrierten Flusses des Krebsnebels. Diese Werte sind etwas größer als die in der Arbeit von Masbou [5] genannten Werte, befinden sich aber in der gleichen Größenordnung. In der zitierten Arbeit wird eine Mono-Rekonstruktionssoftware vorgestellt, die besonders auf der Verwendung von $\Delta \phi$ als wichtigster Diskriminierungsvariable beruht. Das herausragende Unterscheidungspotential dieser Variable konnte nicht bestätigt werden. In der vorliegenden Arbeit wurde festgestellt, dass die Concentration und $\Delta \tilde{A}$ die wichtigsten Eingabevariablen für die BDTs sind, während $\Delta \phi$ die Variable mit der geringsten Relevanz ist.

In weiterführenden Studien könnte die $\delta$-Methode verbessert werden, insbesondere die Bestimmung der Vorzeichen der $\vec{\delta}$-Komponenten. Eine mögliche Methode hierzu ist die Verwendung von BDTs. Solche Methoden könnten auch verwendet werden, um die vorzeichenbehalteten $\vec{\delta}$-Komponenten zu rekonstruieren. Im Falle der Energierekonstruktion könnte eine besser geeignete Fitfunktion als die verwendete Gauf-Verteilung gefunden werden, da nicht bewiesen ist, dass diese die beste Wahl darstellt. Abgesehen von der Wahl der Fitfunktion könnten besser geeignete MLP-Eingabevariablen gefunden werden. Dies gilt ebenso für die BDT-Variablen. Weiterhin wäre es möglich, Timing-Informationen zu verwenden, die im Falle des neuen Teleskops verfügbar sind. Dies bedeutet, dass die Information über den Zeitpunkt, an dem jeder Pixel Licht detektiert hat, bekannt ist. Somit ließe sich die Entwicklung der Schauerbilder in der Kamera nachvollziehen, was prinzipiell dazu verwendbar ist, die Bestimmung der Vorzeichen von $\vec{\delta}$ zu verbessern.
A Appendix

A.1 Width and Length Lookup Tables

In this section, the lookup tables of the widths of the \( W \) and \( L \) distributions, \( \sigma_W \) and \( \sigma_L \), are shown. The meaning of these variables is discussed in Sec. 5.1.

![σW lookup table](a)

![σL lookup table](b)

Figure A.1: Lookup tables of the widths of the \( W \) and \( L \) distributions, \( \sigma_W \) and \( \sigma_L \). \( A \) and \( \rho \) are the size and the roundness of the ellipse, explained in more detail in Chap. 5.1. The colour scale represents the \( \sigma \) values in rad.

A.2 Simulation Time Calculation for Proton MC Events

In order to calculate the time corresponding to the number of events in a proton MC sample for a given proton flux spectrum, several quantities have to be known. These are the number of events in the sample \( N \), the solid angle \( \Delta \Omega \) corresponding to the possible incident directions of the protons and the spectrum \( \frac{d\Phi}{dE} \) in the simulated energy range \( E_1 - E_2 \). Also the radius \( R \) of the circular area perpendicular to the pointing direction of the telescope in which the incident protons would pass the plane of this area has to be known. The simulated time \( T \) can be calculated using the following general relation:

\[
N = T \cdot \Delta \Omega \cdot \pi R^2 \cdot \int_{E_1}^{E_2} \frac{d\Phi}{dE} dE. \quad (A.1)
\]
In this case, the view cone corresponding to $\Delta \Omega$ has a radius of 5°, $R = 1000$ m and the energy range is given as $E_1 = 10 \text{ GeV}$ and $E_2 = 50 \text{ TeV}$. In the MC samples used in this analysis, the simulated proton spectrum was set to a pure power law with an index of $\alpha = 2$, while the true spectrum $\frac{d\Phi_B}{dE}$, as measured by the BESS collaboration [45], [50], is given as

$$\frac{d\Phi_B}{dE} = \frac{0.099}{1 \text{ TeV m}^2 \text{s sr}} \cdot \frac{\left( \frac{E}{1 \text{ TeV}} \right)^{-2.7}}{1 + \left( \frac{0.004 \text{ TeV}}{E} \right)^{1.75}}. \quad (A.2)$$

The number of events in the MC sample has to be calculated as

$$N_{MC} = T \cdot \Delta \Omega \cdot \pi R^2 \cdot \Phi_{MC} \cdot \int_{E_1}^{E_2} \left( \frac{E}{1 \text{ TeV}} \right)^{-\alpha} dE, \quad (A.3)$$

while the appropriate true number of events is, according to Eq. (A.1),

$$N_T = T \cdot \Delta \Omega \cdot \pi R^2 \cdot \int_{E_1}^{E_2} \frac{d\Phi_B}{dE} dE. \quad (A.4)$$

In order to calculate the correct time, $N_T$ can be expressed by the given quantity $N_{MC}$, and the simulated spectrum has to be weighted by a weight function $w(E)$ in order to give the BESS spectrum:

$$N_T = T \cdot \Delta \Omega \cdot \pi R^2 \cdot \Phi_{MC} \cdot \int_{E_1}^{E_2} \left( \frac{E}{1 \text{ TeV}} \right)^{-\alpha} w(E) dE \quad (A.5)$$

$$= \frac{N_{MC}}{\int_{E_1}^{E_2} \left( \frac{E}{1 \text{ TeV}} \right)^{-\alpha} dE} \cdot \int_{E_1}^{E_2} \left( \frac{E}{1 \text{ TeV}} \right)^{-\alpha} w(E) dE. \quad (A.6)$$

$w(E)$ then has to be chosen as

$$w(E) = \frac{d\Phi_B}{dE} \cdot \left( \frac{E}{1 \text{ TeV}} \right)^{\alpha \cdot 1 \text{ m}^2 \text{s sr TeV}^{-\alpha \cdot \beta}} \quad (A.7)$$

$$= \left( \frac{E}{1 \text{ TeV}} \right)^{-\alpha_{MC}} \cdot \frac{0.099 \cdot \left( \frac{E}{1 \text{ TeV}} \right)^{-2.7}}{1 + \left( \frac{0.004 \text{ TeV}}{E} \right)^{1.75}}. \quad (A.8)$$
A.2 Simulation Time Calculation for Proton MC Events

The factor $\beta$ has to be introduced in order to yield the right units. Now $N_T$ becomes the following:

\[
N_T = \frac{N_{MC}}{E_1} \left( \frac{E}{1\text{TeV}} \right)^{-\alpha} \int_{E_1}^{E_2} \frac{E^\alpha}{E^{1\text{TeV}}} dE \beta \cdot \int_{E_1}^{E_2} \frac{E^\alpha}{E^{1\text{TeV}}} dE \cdot \beta \\
= \frac{N_{MC}}{E_1} \left( \frac{E}{1\text{TeV}} \right)^{-\alpha} \int_{E_1}^{E_2} \frac{d\Phi_B}{dE} dE.
\]

(A.9) \hspace{1cm} (A.10)

Identifying the two expressions for $N_T$ (Eq. (A.4) and Eq. (A.10)) with each other, the correct time $T$ can be calculated:

\[
T = \frac{N_{MC} \cdot \beta}{\Delta \Omega \cdot \pi R^2 \cdot \int_{E_1}^{E_2} \left( \frac{E}{1\text{TeV}} \right)^{-\alpha} dE}.
\]

(A.11)

For $N_{MC} = 1.1 \times 10^6$, the corresponding time $T$ is

\[T \approx 161.1 \text{s}.
\]

(A.12)

This value is used for the calculation of the rate of proton events.
A Appendix

A.3 MLP Input Variable Distributions

In the following three figures, the distributions of the MLP input variables for three energy bands are shown. It can be seen that the discrimination power increases with increasing energy, as the inequality between the signal and background distributions increases with the energy.

Figure A.2: Distributions of the MLP input variables for the fourth energy band.
Figure A.3: Distributions of the MLP input variables for the sixth energy band.

Figure A.4: Distributions of the MLP input variables for the eighth energy band.
In the following three figures, the distributions of the BDT input variables for three energy bands are shown. It can be seen that the discrimination power increases with increasing energy, as the inequality between the signal and background distributions increases with the energy.

The resulting BDT response plots for these three energy bands are given in Fig. A.8–Fig. A.10. It can be seen that the gamma / hadron separation improves significantly with increasing energy.

Figure A.5: Distributions of the BDT input variables for the fourth energy band.
A.4 BDT Input Variable and Output Distributions

Figure A.6: Distributions of the BDT input variables for the sixth energy band.

Figure A.7: Distributions of the BDT input variables for the eighth energy band.
Figure A.8: Distributions of the BDT response values $\zeta$ for the fourth energy band.

Figure A.9: Distributions of the BDT response values $\zeta$ for the sixth energy band.
Figure A.10: Distributions of the BDT response values $\zeta$ for the eighth energy band.
Bibliography


Bibliography


Bibliography


List of Figures

2.1 Feynman diagram of an IC scattering process, with the time axis in horizontal direction. .............................................. 5
2.2 Feynman diagram of a π0-decay, with the time axis in horizontal direction. ........................................... 5
2.3 Feynman diagram of electron pair production by scattering of a VHE photon γ_{VHE} on a low energy photon γ_{LE}, with the time axis in horizontal direction. .............................................. 6
2.4 Simulated electromagnetic and hadronic air showers. The shower shown on the left was initiated by a 0.3 TeV photon, the shower on the right was produced by a 1 TeV proton [12]. .............................................. 7
2.5 Simulated Cherenkov light distribution on ground level in case of 0.3 TeV photon (left) and 1 TeV proton (right) initiated showers [14]. The side length of each of these images is 400 m. .............................................. 8
3.1 One of the first shower images taken with CT5 [30]. It very likely shows an image of a gamma ray initiated air shower. The colour code represents the number of photoelectrons per pixel. .............................................. 14
3.2 Sketch of an image parameterisation using Hillas parameters. The image, which has an elliptical shape, is represented by the green area. The coordinate axes represent the camera coordinate system. The meaning of the parameters is explained in the text. .............................................. 15
4.1 Left: Sketch of two-dimensional signal (S) and background (BG) event phase spaces. x_1 and x_2 are event characterising variables, x_{c1,2} are the respective cut values. Right: Concept of MVA methods, with x_1 \ldots x_N being variables characterising an event and ζ being the output variable of the MVA method. .............................................. 17
4.2 Example of idealised signal and background ζ distributions. These distributions can be used to find a cut value ζ_0 as described in the text. .............................................. 18
4.3 Sketch of a bDT, inspired by Ohm et al. [22]. At each node, a cut criterion x_i < c_i is evaluated. According to the result of each check, the event x follows a path along the tree to finally reach the last-grown leaves, which are classified as signal (S) or background (B). .............................................. 20
4.4 Sketch of an MLP network with a single hidden layer containing five neurons and an output layer consisting of one neuron [36]. In this example, four input values x_1, \ldots, x_4 and bias neurons as explained in the text are used. .............................................. 23
List of Figures

5.1 Lookup tables of the width $W$ and the length $L$ of gamma ray events. $A$ and $\rho$ are the size and the roundness of the Hillas ellipse. The colour scale represents the width and length values in rad. ................................. 29

5.2 Cut flow chart showing the percentage of events passing the cuts indicated in the bin labels. The distributions for gamma ray events and proton events are shown. The percentages are calculated with respect to the total number of events. The cuts are explained in the text. ......................... 31

5.3 Distribution of proton event weights after all cuts except the BDT cut and the ON region cut. ................................................................. 33

5.4 Sketch of a Hillas ellipse including the quantities needed for the direction reconstruction. These are the centre of gravity (c.o.g.), $\delta$ and the angle $\beta$ between the main axis of the ellipse and the $x$ axis of the camera coordinate system. ......................................................... 34

5.5 Lookup table for the $\delta$ variable as a function of the roundness $\rho$ and the natural logarithm of the size of the Hillas ellipse $\ln(A/[p.e.])$. The colour scale represents the $\delta$ values in rad. .............................................. 35

5.6 Sketch of an air shower and CT5. The dashed lines indicate the tracks of the Cherenkov light photons. The intensity profile is used to calculate the skewness, which in this case would be negative. $\vec{x}_T$ denotes the position of the true source. .............................................................. 36

5.7 Sketch of a Hillas ellipse, with the indicated main axis on which the reconstructed shower direction $\vec{x}_R$ is located. The true direction is labeled as $\vec{x}_T$. ................................................................. 37

5.8 Example of a p-value distribution of a gamma MC event as a function of the mean energy of the corresponding energy band (see text for details). The points are fitted by a Gaussian distribution. Furthermore the positions of the reconstructed and the true energies of this event are indicated by the vertical lines. ................................................................. 39

5.9 Distributions of the concentration $C$, $\Delta \phi$, the scaled length $\tilde{L}$, the scaled width $\tilde{W}$, $\Delta N_{\text{pix}}$ and $\Delta A$ after all cuts except the cuts on the BDT response and on the reconstructed direction. The variables are explained in the text. The proton distributions have been scaled to the area of the gamma ray distributions. ........................................... 41

5.10 Sketch of the ON and OFF regions and the OFF region ring. Here the ON region is defined as a circle with an angular distance to the camera centre denoted as $d_{\text{OFF}}$ and a radius $\Theta_0$. There are several OFF regions shown as dashed circles and one OFF region ring containing the ON and OFF regions. ................................................. 44
6.1 Squared angular distance $\Theta^2$ between the reconstructed shower direction and the true point-source position for gamma ray and proton events after all cuts besides the cut on the shower direction for the distance method (left) and the $\delta$ method (right). The values of the radii of the 68\% containment circles $R_{68}$ for both kinds of events are given in the insets. In both figures, the area of the proton distributions have been scaled to match the area of the gamma ray distributions. ................. 47

6.2 Angular resolution, expressed in terms of $R_{68}$, as a function of the reconstructed energy. Only the plots for gamma ray events are shown, for both the distance and the $\delta$ method as indicated in the legend. ................. 49

6.3 Left: Energy reconstruction bias $b_E$ for gamma ray events before and after a cut on the BDT response $\zeta$ (indicated by the superscript) as a function of the true energy. Right: Energy resolution $r_E$ as a function of the true energy before and after the BDT cut. ................. 50

6.4 Distributions of the MLP input variables for the second energy band, which includes energies between 25 GeV and 40 GeV. The variables have been introduced in the text. ................................. 51

6.5 Optimised cut values $\zeta^i_0$ (left) and $\Theta^i_0$ (right) as a function of the energy, which is given in terms of the energy band number $i$. The distributions resulting from the two direction reconstruction methods are shown. The “$\delta$” superscripts in the legends indicate the distributions obtained from the $\delta$ method. ................................. 53

6.6 Top: BDT input variables for the second energy band, which covers energies from 25 up to 40 GeV. The variables have already been introduced in Sec. 5.5. Bottom: Resulting $\zeta$ distributions (“BDT response”) for this energy band. The “Signal” distribution corresponds to gamma ray events, “Background” stands for proton events. ................................. 55

6.7 $\Delta \phi$ distributions resulting from this analysis after all cuts except the BDT cut and the cut on the reconstructed direction (top left) and after applying only a nominal distance cut and a cut on the reconstructed shower direction (bottom left). The top right figure shows the distributions from Masbou [5]. Here the gamma ray distribution is denoted as “Signal” and the proton distribution is labeled as “Background”. Also the widths of the $\Delta \phi$ distributions as functions of the true event energy are given for both gamma rays and protons for two different cut patterns (bottom right). The first cut pattern requires all cuts except the BDT cut and the cut on the reconstructed shower direction, the second cut pattern, which is indicated by the subscript “C”, requires the same cuts as applied in the plots shown in the bottom left figure. ................................. 57
List of Figures

6.8 Efficiencies $\varepsilon_{\gamma/p}$ of both gamma ray events (top left) and proton events (top right) as functions of the true event energy. The effective areas $A_{\gamma/p}^{\text{eff}}$ as functions of the energy are shown at the bottom. The distributions of the effective areas for gamma ray events when using each of the direction reconstruction methods are shown in the bottom left figure, the proton distribution resulting from the distance method is shown in the bottom right figure. .......................................................... 59

6.9 Integral rate of proton events $R_p$ as a function of the reconstructed energy $E_{\text{Reco}}$. .......................................................... 61

6.10 Differential and integrated sensitivity curves plotted together with functions representing the respective values of a source with a flux of one percent of the Crab Nebula’s flux. The sensitivities calculated using both the distance method and the $\delta$ method for the direction reconstruction are shown in each figure. The values resulting from the utilisation of the $\delta$ method are indicated by the “$\delta$” superscript. .......................................................... 63

A.1 Lookup tables of the widths of the $W$ and $L$ distributions, $\sigma_W$ and $\sigma_L$. $A$ and $\rho$ are the size and the roundness of the ellipse, explained in more detail in Chap. 5.1. The colour scale represents the $\sigma$ values in rad. . . . . 71

A.2 Distributions of the MLP input variables for the fourth energy band. . . . . 74

A.3 Distributions of the MLP input variables for the sixth energy band. . . . . 75

A.4 Distributions of the MLP input variables for the eighth energy band. . . . . 75

A.5 Distributions of the BDT input variables for the fourth energy band. . . . . 76

A.6 Distributions of the BDT input variables for the sixth energy band. . . . . 77

A.7 Distributions of the BDT input variables for the eighth energy band. . . . . 77

A.8 Distributions of the BDT response values $\zeta$ for the fourth energy band. . . 78

A.9 Distributions of the BDT response values $\zeta$ for the sixth energy band. . . 78

A.10 Distributions of the BDT response values $\zeta$ for the eighth energy band. . . 79
## List of Tables

5.1 List of the applied event selection cuts. The variables and the motivation for the choice of the cut values are explained in the text. The values given in the last column are calculated as the percentages with respect to the number of events passing the previous cuts. In case of proton events, these weighted number of events is used. The last two cuts are different for each energy band, so that numbers for these have to be given later (cf. Sec. 6.3).

6.1 Effective areas for two values of the true event energy. The values calculated using the two different direction reconstruction methods and the values given in Masbou [5] are listed.

6.2 Differential sensitivity values for three different energies. The values obtained from using both the distance method and the δ method for the direction reconstruction are shown. The sensitivities are given in percent of the differential Crab flux $\frac{d\Phi_C}{dE}$.

6.3 Integral sensitivity values for three different energies. The values obtained from using both the distance and the δ method for the direction reconstruction are shown, as well as the values taken from Masbou [5]. The sensitivities are given in percent of the integral Crab flux $\Phi_C$. 

---

30

61

64

66
Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 01.10.2012                                           Thomas Murach