

# Multiwavelength Analysis of the TeV-Radio Galaxy 3C 84/NGC 1275

Lena Marie Linhoff  
2021

A document submitted in partial fulfillment of the requirements for the degree of  
*Doctor rerum naturalium*  
at  
Fakultät Physik, Technische Universität Dortmund

Supervised by  
Prof. Dr. Dr. Wolfgang Rhode and Prof. Dr. Kevin Kröninger





## Abstract

The radio galaxy 3C 84 is a well-studied source of radio emission and was detected as NGC 1275 also in the MeV/TeV regime by gamma-ray detectors like MAGIC and *Fermi*-LAT. It is still unclear where and how the gamma-ray emission is produced. In this thesis, I will confine possible emission sites and exclude the region near the black hole as the origin of the gamma-ray production.

For this aim, I investigate the optical depth of the broad-line region using data published by MAGIC and *Fermi*-LAT. Furthermore, a cross-correlation study is performed to find a possible correlation between the light curves of the two radio components in 3C 84 detected by the VLBA and the gamma-ray light curve measured by *Fermi*-LAT. A significant correlation between the core component and the gamma-ray emission is found, which is in line with the results I derive from analyzing the optical depth of the broad-line region. For the first time, I perform a long-term analysis of NGC 1275 for four years of MAGIC data, which reveals a short flare at the beginning of 2017 and a very low state of activity since then.

To perform this long-term analysis, the software framework `autoMAGIC` was developed in the course of this thesis. `autoMAGIC` enables fully automatic and reproducible analyses of long-term data and can be used for the automatic processing of MAGIC data in the future.

## Zusammenfassung

Der aktive galaktische Kern 3C 84, bekannt als Quelle von Radiostrahlung, wurde in den letzten Jahren auch von Gammastrahlungsdetektoren wie MAGIC und *Fermi*-LAT detektiert. Um die Entstehung der Gammastrahlung zu verstehen, grenze ich die Region ein, in der diese Strahlung entstehen kann, indem ich die Absorption der Gammastrahlung in der Broad-Line-Region berechne. Aufgrund dieser Studie kann die direkte Umgebung des schwarzen Lochs als Entstehungsregion ausgeschlossen werden. Außerdem führe ich eine Korrelationsanalyse der Lichtkurven der zwei bekannten radio-hellen Regionen mit der Lichtkurve im MeV/TeV-Bereich durch. Diese Analyse findet eine signifikante Korrelation zwischen der Kernkomponente der Quelle und der *Fermi*-LAT Lichtkurve. Im Rahmen dieser Arbeit wurde außerdem die erste Langzeitstudie der von MAGIC gemessenen hochenergetischen Gammastrahlung von NGC 1275 durchgeführt. Die Ergebnisse dieser Analyse zeigen ein kurzes Aufleuchten der Quelle Anfang 2017 und danach lediglich einen sehr schwachen Fluss.

Um diese Langzeitanalyse durchzuführen, wurde `autoMAGIC` entwickelt, ein Programm zur automatischen und reproduzierbaren Analyse der von MAGIC gemessenen Daten. `autoMAGIC` kann über diese Arbeit hinaus für alle Standardanalysen genutzt werden und wird maßgeblich dazu beitragen die Daten der MAGIC Kollaboration für die kommenden Jahrzehnte zu archivieren.



# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>I</b>	<b>Multiwavelength Astronomy: Theoretical Groundwork</b>	<b>3</b>
<b>2</b>	<b>There is more. What We See and Don't and Why</b>	<b>5</b>
2.1	Atmospheric Windows . . . . .	5
2.2	Extragalactic Sources . . . . .	6
2.3	Active Galactic Nuclei . . . . .	9
2.4	In Particular: Perseus A / NGC 1275 / 3C 84 . . . . .	13
2.5	Extragalactic Background Light . . . . .	17
2.6	Cosmic Microwave Background . . . . .	17
2.7	Galactic Sources . . . . .	18
<b>3</b>	<b>Observer's View</b>	<b>21</b>
3.1	Spectral Energy Distribution . . . . .	21
3.2	Light Curve . . . . .	24
<b>II</b>	<b>MAGIC</b>	<b>25</b>
<b>4</b>	<b>Ground-Based Gamma-Ray Detectors</b>	<b>27</b>
4.1	Cherenkov Effect . . . . .	27
4.2	Imaging Air Cherenkov Telescopes . . . . .	30
4.3	Water Cherenkov Telescopes . . . . .	31
<b>5</b>	<b>Data Analysis of the MAGIC Telescopes</b>	<b>33</b>
5.1	Data Taking . . . . .	33
5.2	Hillas Parameters . . . . .	36
5.3	Gamma/Hadron Separation and Event Reconstruction . . . . .	38
5.4	Instrument Resonse Functions . . . . .	42
5.5	The Inverse Problem . . . . .	46
5.6	On/Off Regions . . . . .	49
5.7	Implementation of the Analysis Chain . . . . .	52
<b>6</b>	<b>Data Formats for Gamma-Ray Astronomy</b>	<b>57</b>
6.1	MAGIC Legacy . . . . .	58
<b>7</b>	<b>Gammapy</b>	<b>59</b>

<b>8</b>	<b>autoMAGIC</b>	<b>61</b>
8.1	History and Goal . . . . .	61
8.2	Analysis Flow . . . . .	62
8.3	Technical Prerequisites . . . . .	65
8.4	Long-Term Analysis of the Crab Nebula . . . . .	69
8.5	Open Issues and Next Steps . . . . .	76
<b>9</b>	<b>Gamma-Ray Emission from NGC 1275</b>	<b>79</b>
9.1	Observational History in High Energies . . . . .	79
9.2	Long-Term Observations of NGC 1275 . . . . .	80
<b>III</b>	<b>Fermi-LAT</b>	<b>87</b>
<b>10</b>	<b>Space-Based Gamma-Ray Detector: Fermi-LAT</b>	<b>89</b>
<b>11</b>	<b>Fermi-LAT Data Analysis</b>	<b>91</b>
11.1	Data Flow and Levels . . . . .	91
11.2	Fermi-LAT Data Analysis Tools . . . . .	92
<b>12</b>	<b>Fermi-LAT Results for NGC 1275</b>	<b>93</b>
<b>IV</b>	<b>VLBA</b>	<b>95</b>
<b>13</b>	<b>Radio Astronomy</b>	<b>97</b>
13.1	Radio Interferometry . . . . .	98
13.2	Very Long Baseline Array . . . . .	104
<b>14</b>	<b>Radio Data Analysis</b>	<b>107</b>
14.1	Imaging . . . . .	107
14.2	Image Segmentation: Random Walker . . . . .	109
<b>15</b>	<b>Flux Variations in 3C 84's Radio Components</b>	<b>113</b>
15.1	Identification of the Radio Components in VLBA Radio Images . . . . .	113
15.2	Total Flux Calibration . . . . .	114
<b>V</b>	<b>Multiwavelength Analysis</b>	<b>119</b>
<b>16</b>	<b>Constraining the Gamma-Ray Emission Region in NGC 1275</b>	<b>121</b>
16.1	Theoretical Groundwork . . . . .	121
16.2	Modeling the Photo-Absorbed SED . . . . .	125
16.3	Results . . . . .	128
16.4	Acceleration Mechanisms . . . . .	129

## Contents

<b>17 Multi-Wavelength Cross-Correlation Analysis</b>	<b>131</b>
17.1 Discrete Cross-Correlation Functions . . . . .	131
17.2 Light Curve Simulation . . . . .	132
17.3 Cross-Correlation Analysis for <i>Fermi</i> -LAT and VLBA Light Curves . . . .	133
17.4 Possible Gamma-Ray Suppression and Neutrino Production? . . . . .	136
 <b>VI Conclusion &amp; Outlook</b>	 <b>139</b>
<b>18 Conclusions and Outlook</b>	<b>141</b>
18.1 New Insights into 3C 84/NGC 1275 . . . . .	141
18.2 Towards an Automated and Reproducible MAGIC Data Analysis . . . . .	143
 <b>VII Appendix</b>	 <b>145</b>
<b>A The autoMAGIC Database Structure</b>	<b>147</b>
<b>B Job Submission with HTCondor</b>	<b>153</b>
<b>C MARS vs. Gammapy: Study on <math>\theta^2</math> Cuts</b>	<b>155</b>
<b>D MARS vs. Gammapy: Comparison by Observation Conditions</b>	<b>161</b>
<b>E Light Curve Simulation</b>	<b>165</b>
<b>F Funding Acknowledgement and Data Usage Statement</b>	<b>175</b>
<b>Bibliography</b>	<b>177</b>
<b>Glossary</b>	<b>195</b>
<b>Affidavit (Eidesstattliche Versicherung)</b>	<b>199</b>
<b>Acknowledgements</b>	<b>200</b>



# Introduction

---

As simple as it is, a look up high to the night sky is and has always been a source of inspiration and the starting point for questions that drove the scientific evolution of humankind. During the past several hundred years, scientists went to the boundaries of what was possible at their time to gain knowledge and study the things nobody had understood before. This way, the understanding of our direct environment but also about the universe around us grew with every technical invention and every genius thought that was contributed by outstanding women and men.

After people had studied the objects in the sky with the naked eye over thousands of years, Galileo Galilei and Johannes Kepler invented and built the first optical telescopes in 1608 and 1611, and the technological development of astronomy began. Since the 17th century, there has been a giant evolution of telescopes in two dimensions: On the one hand, their size increased and their resolution was improved so that we can observe celestial objects at a very great distance. On the other hand, the range of the accessible wavelengths was enhanced. Today we can observe not just visible light but also radiation in the infrared, ultraviolet, radio, X-ray, and gamma-ray wavebands. In the last decades, we also gained access to the universe by neutrinos and gravitational waves.

Along the way, the instruments got bigger, the amount of people involved got larger, and so did the expense that is needed to reveal the secrets of our existence. Despite all efforts, the series of questions never stopped. But neither did the ambition of scientists all over the world to solve them.

Since the electromagnetic spectrum is almost fully covered with detectors, we can explore other dimensions to gain new insights: First, the possibilities that come with the computational capabilities that expanded exponentially in the last years. Second, the combination of observations made by multiple detectors at different ranges of the electromagnetic spectrum.

The key to new discoveries is not hidden in one image or a single time series anymore, but it comes from analyzing hundreds of images or millions of data samples. This amount of data can not be analyzed by a single scientist. To achieve anything with our incredibly precise, large and expensive detectors, we need the best possible data analysis methods and software.

For this thesis, I choose to use the power of modern astronomy to have a deeper look into a very special galaxy named 3C 84 or NGC 1275 located in the constellation of Perseus. 3C 84 is a well-known source for radio astronomers and has been observed over decades. In the gamma-ray regime, we got contact only for the last 13 years.

I combine the data of three instruments to get new insights into the behavior of 3C 84 over several years: In the radio regime, I use data provided by the Very Long Baseline Array (VLBA). The gamma-ray data comes from the Large Area Telescope (LAT) on board

## 1 Introduction

the *Fermi* satellite and from the Major Atmospheric Gamma Imaging Cherenkov (MAGIC) telescopes that are located on the Canary Island of La Palma.

Combining the data of these different instruments comes with some challenges. Especially for MAGIC, the prerequisites were not given to allow for an analysis of data measured over several years, which is a usual time range for studies made with radio data. This is why, besides the scientific questions regarding 3C 84, a major focus of this thesis lies in the development of an automated analysis for the MAGIC data, which enables the analysis of huge data sets, in the first place.

Within the frame of this work, I developed `autoMAGIC`, an analysis framework that delivers reproducible high-level results from data taken by MAGIC in an automated and scalable manner. `autoMAGIC` is not exclusively used for this work, but will also manage the massive data processing of all MAGIC data ever recorded and preserve the legacy of MAGIC for future generations of scientists.

This thesis is structured as follows: The first part is dedicated to the theoretical foundations that are needed to understand all aspects of the multi-wavelength analysis performed later. In the following parts, the three instruments MAGIC, *Fermi*-LAT and the VLBA are described, together with their data analysis approaches and the scientific results for 3C 84. In the last part, the high-level results from the different experiments are combined in a multi-wavelength study that addresses two questions: Where does the gamma-ray emission in NGC 1275 come from and is the evolution of the source in the radio regime connected to its behavior in the gamma-ray range?

## **Part I**

# **Multiwavelength Astronomy: Theoretical Groundwork**



## There is more. What We See and Don't and Why

---

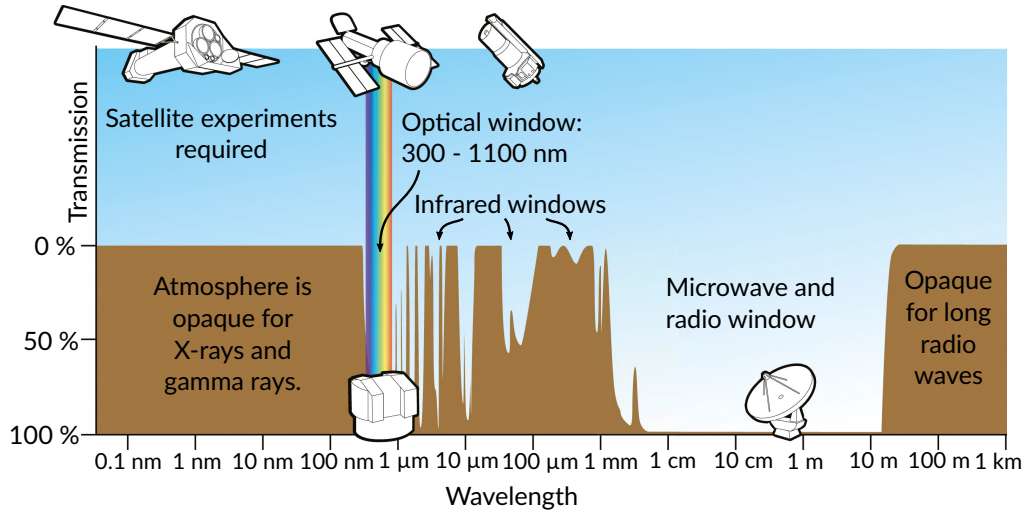
### 2.1 Atmospheric Windows

The reason why stars were the first objects noticed by men is pretty obvious: People can see them. Or Earth's atmosphere is transparent for the optical light the human eye can detect. So, optical astronomy was born, and the bright dots at the night sky were observed, studied, and categorized over centuries with telescopes of increasing size and resolution. From these observations, planetary movements and the concept of galaxies were derived.

If we transform the term *light* to *radiation* or *emission*, the optical emission is just one form of radiation a source can send to us. All this refers to the energy that is transferred from a source to an observer or detector and the name we call it depends only on the amount of energy. At the very low end of this energy range, called the electromagnetic spectrum, the energy is transferred by radio waves. Radio waves have a long wavelength of micrometers up to kilometers. In the middle part of this range, Earth's atmosphere is transparent for radio waves, which makes this kind of radiation directly accessible from Earth's surface, see [Figure 2.1](#). This is why radio astronomy was the next field that captured the market of astronomic science (see [chapter 13](#)). Going to higher energies and shorter wavelengths, the next window allows for ground-based detections in the nanometer range: the optical window for the already mentioned visible light and its non-visible margins, the infrared and the ultraviolet light, which are already partly absorbed by the atmosphere. Above these energies, Earth's atmosphere becomes completely opaque to radiation, which would otherwise be dangerous for living organisms. To be able to study the higher energetic X- and gamma rays, large space-based detectors become inevitable, which obviously requires space flights, aerospace technology, and a huge amount of money (see [chapter 10](#)). For the highest energies, astronomers exploit the secondary processes that happen when gamma rays hit the atmosphere and create big light cones (Cherenkov effect), detectable again with ground-based telescopes (see [section 4.2](#)).

Apart from the electromagnetic spectrum, which is the subject of a *multi-wavelength study*, we can also detect neutrinos or gravitational waves from astronomical sources. Combining observations of electromagnetic radiation with detections of neutrinos or gravitational waves would result in a *multi-messenger study*. In the scope of this work, I focus on emission from the electromagnetic spectrum.

Being able to detect all these different kinds of radiation brings us to the next question: Where does all this radiation come from?



**Figure 2.1:** Atmospheric opacity depending on the wavelength of the radiation. The atmosphere is transparent mainly in two energy ranges, at the optical window around 400 nm to 900 nm and at the radio window at wavelengths between 1 cm and 10 m. At all other wavelengths, the atmosphere is opaque to electromagnetic radiation. This is why satellites are used to detect gamma rays and X-rays. Image credit: ESA/Hubble (F. Granato)

## 2.2 Extragalactic Sources

Next to the objects we can see in the sky, like the Moon, the Sun, and the stars, there are, as said above, far more objects we don't see with our eyes. Our solar system, hosting Earth and seven further planets (eight, if you count Pluto, but this is another issue) with their moons, is just a small crumb in a much bigger system, our galaxy called Milky Way. The Milky Way is a barred spiral galaxy that hosts, besides our Sun, 100 to 400 billion stars. All this matter surrounds a supermassive black hole of  $3$  to  $4 \times 10^6 M_{\odot}$  in its center (e.g. [81, 66, 144]). The center of the galaxy, the black hole and its surroundings, is known as a strong radio emitter and called Sagittarius A\*. Zooming out further, the Milky Way with a radius of roughly 290 kpc ( $= 9 \times 10^{21}$  m) [53] is, again, just one amongst billions.

There are a lot of other galaxies, some of them discovered, named, and cataloged in various ranges of the electromagnetic spectrum, and a much bigger part still undiscovered. All of these galaxies have something in common: the spinning black hole at the center, surrounded by a disk of matter and a dust torus. The black hole accretes matter from the innermost regions of the disk and is able to re-eject a part of this accreted matter as radiation.

Galaxies can now be classified by the intensity and the kind of this radiation.

### 2.2.1 The Unified Scheme

Extremely bright central regions of galaxies are called Active Galactic Nuclei (AGN) and feature emission lines in the optical, X-ray, ultraviolet, infrared, and gamma-ray regime.

Not all types of radiation are present for all sources, which further separates different types of AGN. Often, this radiation is emitted as a pair of relativistic jets, a collimated stream of outflowing ionized matter perpendicular to the accretion disk along the rotational axis of the black hole.

Depending on the viewing angle, what we see of an AGN from Earth can vary widely. Looking directly into the jet or seeing just the dust torus makes a huge difference in the appearance of the same object. This is why, historically, AGN were called a number of names before scientists found these alleged different sources to fit in one unified scheme, see [Figure 2.2](#).

**Blazars** are sources directly pointing their jet towards Earth. Because the jet is mostly responsible for high-energy radiation, blazars are detected by ground- and space-based gamma-ray detectors. Blazars can be further divided into subclasses of BL Lac objects and Flat Spectrum Radio Quasars (FSRQs).

**BL Lac** objects, named after the prototype BL Lacertae in the constellation Lacerta, are very luminous and highly variable objects that feature, unlike other AGN types, no emission lines<sup>1</sup> in their energy spectrum. The lack of these lines could be caused by the luminous jet whose nonthermal continuous emission covers the emission from other parts of the galaxy and the emission lines occurring there. Furthermore, BL Lac emission is characterized by strong optical polarization, which is also variable.

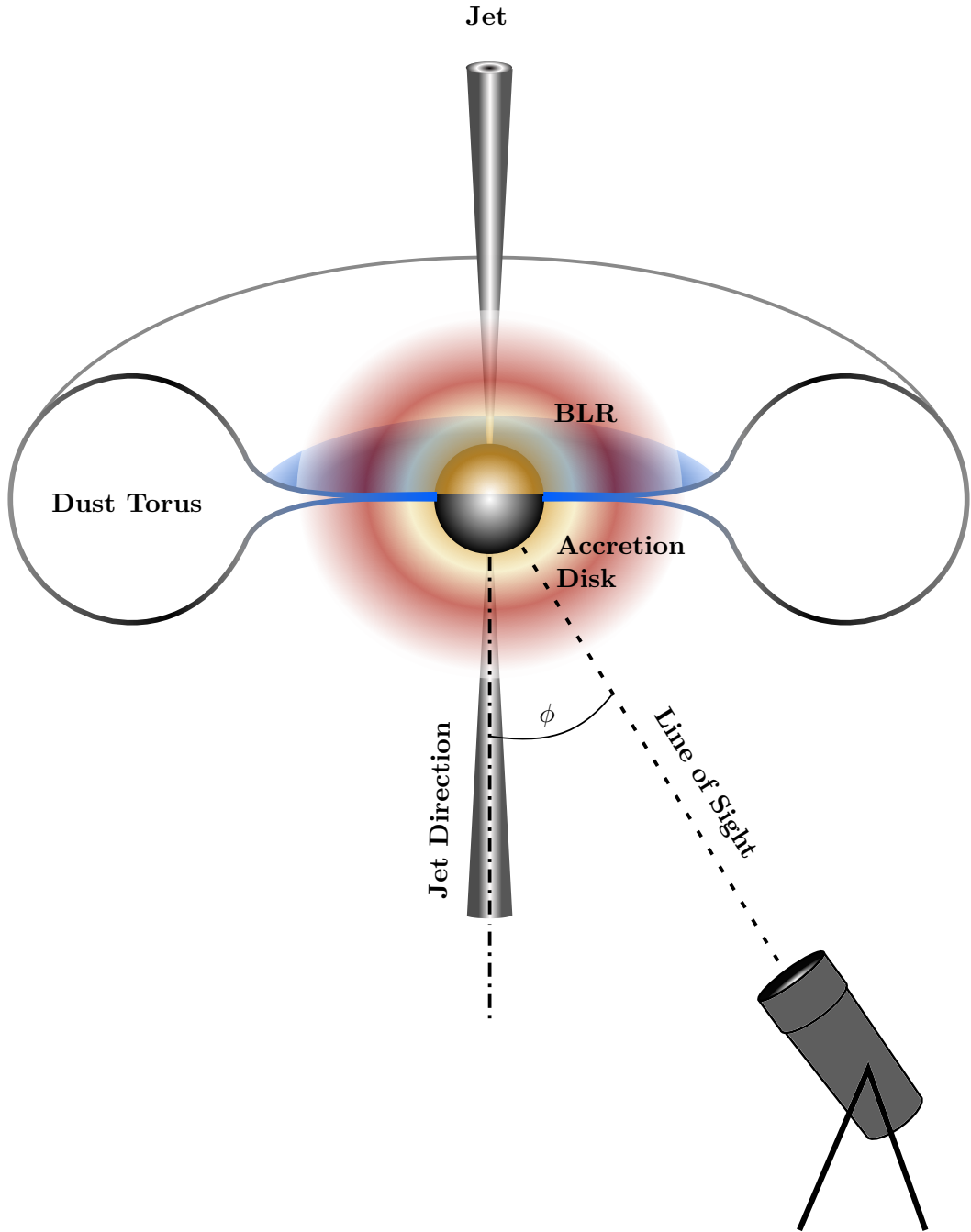
**FSRQs** and BL Lac have the flux variability in common, but in turn, FSRQs feature strong and broad optical emission lines and a higher luminosity. There is an overlap between both blazar subclasses, especially for objects with weak emission lines that sometimes can not be classified explicitly.

Objects possessing a jet that is tilted against the line of sight are classified as **Fanaroff-Riley (FR)** galaxies. Such objects show significant luminosity at radio wavelengths, whereas FR type I galaxies have their maximum luminosity at the upper parts of the jet, close to the center, while FR type II galaxies show bright lobes at the ends of the jets, further away from the core. The classification of Fanaroff and Riley is based on the first systematic sample of radio sources which was provided by the Third Cambridge Catalogue of Radio Sources (3C) (436 sources) [65] and its revised edition [30]. All objects from that sample are cataloged as 3C xxx (e.g. 3C 84), and this enumeration survives until today. Nonetheless, several sources appeared independently in other surveys over the years and got additional names.

In contrast to radio-loud FR galaxies, **Seyfert** galaxies are radio-quiet objects, which feature no or non-relativistic jets. Due to the low radio luminosity of the core, the host galaxy becomes visible. Seyfert galaxies are furthermore characterized by their emission lines, which are produced by ionized gas clouds or parts of the rotating accretion disk. The characteristics of these emission lines – if narrow or broad – further separate Seyfert galaxies into two subclasses. Seyfert type I galaxies show broad emission lines coming from fast-rotating gas close to the central black hole. Seyfert type II galaxies have in turn narrow emission lines, suggesting that the emitting material is moving slower and, therefore, has

---

<sup>1</sup>Optical emission at a specific wavelength is called “line emission” and originates mostly from gas clouds in the central region of an AGN. Width and emitted wavelength of these lines provide some information about the composition of the gas clouds and their position and rotation speed within the AGN.



**Figure 2.2:** Scheme of an AGN. The dot-dashed line marks the jet direction, whereas the dashed line represents the line of sight. The angle  $\phi$  between these lines is called *viewing angle*.



a greater distance to the central engine. This separation between broad and narrow line emitting objects turns out to be not that strict in reality. Seyfert galaxies were found featuring narrow as well as broad emission lines, which is why further subclasses were introduced, but which will not be discussed in detail, here.

## 2.3 Active Galactic Nuclei

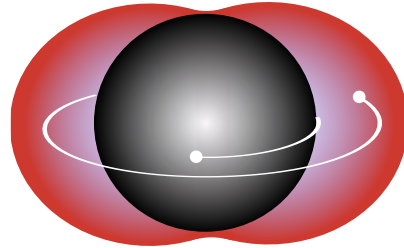
In the case of an Active Galactic Nucleus (AGN), the central Black Hole (BH) and its surroundings eject one or two cones of emission outflow, the so-called jets. These jets are oriented along the rotational axis of the BH and transfer huge amounts of energy away from the core. Despite the existence of some models explaining the formation of a jet and the processes within, jets keep a lot of secrets for themselves until today. The first models that build the foundation of our understanding of jets and black holes today were created in the 1960s and 1970s. In 1963, Roy Kerr first described uncharged rotating BHs by solving Einstein's field equations [100]. In 1971, Remo Ruffini and John Wheeler predicted the existence of an ergosphere around a rotating BH [142], which was used by Roger Penrose between 1969 and 1971 to create a model explaining the extraction of energy from a BH [130, 131]. In 1977, Roger Blandford and Roman Znajek introduced a mechanism similar to Penrose's, which is now called *Blandford-Znajek mechanism*, describing the formation of jets (although the term *jet* is never used in their publication) [33]. Until today, these ideas remain and were questioned, extended, or modified by many others.

### 2.3.1 Black Holes

The very basic idea of jet formation stated by these men are the following: At first glance, a BH's gravity is supposed to swallow all matter that comes close to it. This is a strong contradiction with the appearance of jets which seem to have their origin very close to the BH and are clearly transporting matter away from the BH.

If the BH is rotating, it possesses an ergosphere, in which an entering particle is accelerated rapidly by the radial pull of the spinning BH so that an observer can not determine its location on its trajectory, anymore (see Figure 2.3). The particle is accelerated to the speed of light and becomes non-stationary. The outer surface of the ergosphere is given by the envelope of all locations where an approaching particle would reach the speed of light. This is why the ergosphere is not a perfectly spherical shell: its radius has a maximum at the BH's equator and decreases until it reaches the poles where no ergosphere exists, see Figure 2.3. The inner boundary surface is given by the event horizon. A particle falling behind this event horizon is captured by the BH and will never be seen again.

But, as stated by Penrose, particles are able to escape the ergosphere, if they have enough

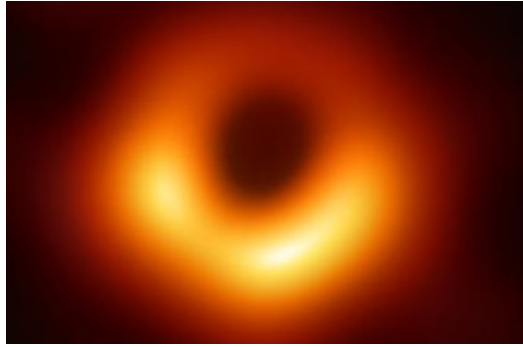


**Figure 2.3:** Ergosphere around a BH.

energy. Caused by the acceleration, a particle might split into two sub-particles. One of these particles has negative mass-energy and falls behind the event horizon. The other one catches a very small amount of the BH's angular momentum and can escape the ergosphere, with – and that is the crucial point, here – more energy than the original particle. The energy difference is provided by the BH itself. Penrose calculated that a particle can gain 20.7% of its original energy in that process. Theoretically, a rotating BH could lose its total angular momentum this way and stop rotating, eventually.

For his “discovery that black hole formation is a robust prediction of the general theory of relativity”, Roger Penrose won a Nobel Prize in 2020 [161].

In 2019 the Event Horizon Telescope (EHT) published the first “picture of a black hole” of the radio galaxy M87 [46], [Figure 2.4](#). This image is by no means a photograph of a BH in the conventional sense. We see radiation at 1.3 mm from gas close to the ergosphere in a distance approximately 4.5 to 5 times the event horizon radius away from the center of the BH. The gas is emitting via synchrotron emission at radio wavelengths while it is accelerated and forced to change its moving direction by the rotating BH. We see a steady, bright but unresolved blob in radio maps, where we expect the center of a galaxy and the BH. This happens for all AGN, but resolving this process spatially is not possible with today's detectors, but for the prominent exception of the EHT.



**Figure 2.4:** Image of a BH at the center of the galaxy M87. Published by the EHT collaboration in 2019 [46].

### 2.3.2 Jet Formation

To explain jet formation with the existence of a spinning BH, Blandford and Znajek take the accretion disk and its electromagnetic field into account. Like everything else, the electromagnetic field is twisted by the rotation of the BH and accelerates particles from the accretion disc. These accelerated particles radiate and produce further particles in cascades. The particle cascades then leave the ergosphere along the twisted field lines, extracting energy from the black hole. The twisted magnetic field lines collimate the particle outflow to a great stream of plasma: the jet. This is a very simplified picture of the theory of jet formation; in the real world, there is a whole field of science dedicated to this topic, which I can not discuss here, further. For further review, see e.g. [122].

Despite the underlying concepts seeming to be well understood since the 1980s, there are a lot of aspects unsolved in the understanding of jet formation. Among them are the size and position of the jet base. It is still unclear by which parts of an AGN's core – the ergosphere or the inner regions of the accretion disk – a jet is driven. The resolution of currently operating telescopes is still too poor to resolve the very beginning of the jet. Highly resolved images of M 87 and 3C 84 show a rapidly expanding, conical profile of the jet base [120, 169, 87]. But especially for 3C 84, the observed profile of the jet base could not be reproduced in simulations based on current theories of jet formation.

These images reveal another feature of AGN jets: bright outer edges and a dimmer inner part. This model is referred to as the *spine-sheath structure*, where the jet consists of a slow-moving sheath and a faster spine. These structures are expected to be produced in ergosphere-driven jets [113] and are further described in the next section.

### 2.3.3 Jet Components

A jet starts as continuous outflow of plasma, modulated by the accretion rate (the amount of matter from the accretion disk that is absorbed by the BH). Depending on the galaxy type, the jet is very bright at the beginning (1 kpc to 10 kpc) and just fades out further away or terminates in bright radio hot spots, called *lobes*. In the latter case, the moving plasma interacts with external material in the galaxy and is slowed down rapidly. Through this shock, synchrotron emission is emitted and the lobes can be detected as bright regions by radio telescopes.

Apart from these lobes and the central region of a radio-loud galaxy, other bright features show up in radio observations. In some cases, these features are stationary, in others they are moving along the jet and are variable in velocity, size, and brightness, see e. g. Figure 2.8.

These features are often called *knots* or *radio components* and appear sometimes in a spatially regular structure along the jet, for example in M 87 [106, 128].

Radio-bright knots are considered to result from shock waves within the jet, when the plasma emits synchrotron radiation as it is decelerated. The smooth circular shape of the knots we see in radio maps not necessarily represents the true shape of these illuminated shock fronts, but is a result of the limited resolution of the telescope and its beam. How these shocks are initiated and how they evolve along the jet is still a topic of debate [52, 158]. In this research area, magnetohydrodynamic (MHD) simulations are used and continuously improved to examine and develop theoretical models to understand the images we detect at Earth.

Shocks can happen when the jet outflow collides with the matter of the surrounding galactic material, but this scenario does not explain moving or evenly spaced knots. A widely accepted concept is based on the pressure difference between the jet and the external medium, resulting in so-called *re-collimation shocks*. The pressure of the external galactic medium decreases faster along the path of the jet than the pressure in the highly relativistic jet plasma. This leads to strong local acceleration of the jet plasma and causes rarefaction waves within the jet.

To explain moving knots and jet features varying in size and brightness, the mentioned spine-sheath model or *transverse-stratified jets model* [93] comes into play. As described in

subsection 2.3.2, the fast-moving inner part of the jet is isolated from the galactic medium by a slower outer layer. Note that “slow” here still means “mildly relativistic speed”, but slower than the velocity of the highly-relativistic spine. The different pressure and speed of these layers are driven by the accretion rate and other characteristics of the accretion disk. Depending on the configuration of the two layers, the shock waves form and evolve in different ways, see [93]. Observational evidence for this jet model was already found in Very Long Baseline Interferometry (VLBI) radio images, also for 3C 84, see section 2.4.

Moving jet components gain access to several characteristics of the jet and the center of the AGN. Interpolating their motion back to the origin, the time of formation can be calculated. Comparing data obtained at other wavelengths but similar time ranges can supply valuable insights in the mechanisms happening near the BH. In case of the emerging radio component, strong gamma-ray flares are observed at the same time for some sources.

### 2.3.4 Broad-Line Region

So far, the central parts of an AGN were mentioned: The black hole, the accretion disk in the equatorial plane, and the jets emerging from the BH's poles. Still, some important parts are missing to understand the fundamental processes happening within these objects.

The accretion disk is surrounded and fed by an optically thick dust torus that covers the central regions of an AGN if it is seen edge-on (perpendicular to the jet axis). The dust torus absorbs radiation from the inner regions and re-radiates the absorbed energy in the infrared regime. Since the dust torus obscures the luminous central parts, an AGN may look different depending on the viewing angle, as described in subsection 2.2.1.

Another key feature of an AGN is the so-called Broad-Line Region (BLR), an accumulation of gas clouds accompanying the accretion disk in a distance of 0.01 pc to 1 pc from the BH. The resolution of currently operating telescopes is too poor to resolve the BLR. Hence, its size, shape, and dynamics have to be assessed by theoretical assumptions or indirect measurements and are subjects of current research. The shape of the BLR is described as “bird-nestlike” [111]. For mathematical assumptions, the BLR's geometry is often estimated as geometrically thick *spherical shell* surrounding the black hole in a certain distance, e. g. [84, 41]. The pole regions, where the jet might pierce the shell, are cut out from the shell. Therefore, the accretion disk and the jet's cross-section are not covered by an evenly distributed amount of gas. Another approximation is the *ring geometry* where the gas clouds are arranged in a ring around the jet axis, e. g. [155]. Both geometries have their weaknesses, whereas the shell geometry remains the more common one. The BLR geometries are further discussed in section 16.1.

Additionally to the shape, the evolution of the BLR is another matter of scientific discussion. In the 1980s and 1990s, mainly two models were proposed, the *discrete clouds model* and the *disk wind model*.

In the case of the discrete cloud model, the BLR consists of – as the name suggests – discrete clouds. The clouds form by evaporating material from the dust torus that reaches the sublimation radius, the transition layer between the dust torus and the gas of the BLR. Like every matter close to the BH, the gas clouds are accelerated by its rotation and partly accreted, eventually.

## 2.4 In Particular: Perseus A / NGC 1275 / 3C 84

The disk wind model (e.g. [117]) holds magnetic winds driven by the accretion disk responsible for the BLR’s evolution. The origin of the accretion disk wind is still unclear. Mostly radiation pressure, thermal pressure [174], or magnetocentrifugal pressure [32] are considered as driving mechanisms.

The Broad-Line Region earned its name by the emission of broad lines in the optical and ultraviolet (UV) range, first discovered by Carl Seyfert [146] who used them to define types of galaxies (see subsection 2.2.1). Over the years, emission lines from hydrogen, (e.g. Lyman- $\alpha$  line, Balmer lines), helium, oxygen, nitrogen, carbon, magnesium, calcium, and iron were found. After this discovery, it soon became clear that the de-facto monochromatic emission lines of the ionized gas appear broadened for a distant observer because the emitting medium is moving. But the nature of this movement was a crucial issue since it was not clear if the gas is moving radially (inflow/outflow) or on Keplerian orbits. Among others, Gaskell [78] found that the rotation of the ionized gas around the black hole is mainly responsible for the emission line broadening, together with a smaller turbulence component, proposed by Osterbrock [127]. The possibility of outflowing material was completely overruled by their models, whereas a hint of outflow has been detected in some sources (e.g. NGC 3227 [57]).

For further reading regarding the BLR specifics, see the reviews by Gaskell [79], Peterson [132] and Yong et al. [178].

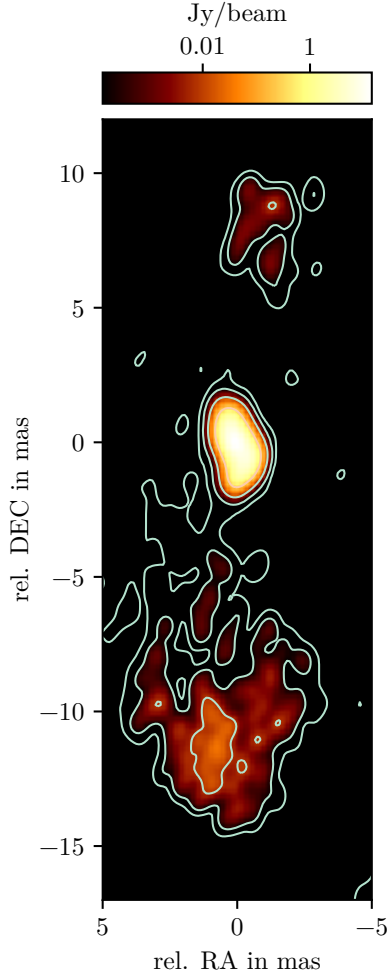
In the context of this work, the BLR becomes important when discussing the origin of the high-energy gamma-ray emission. Depending on the optical depth of the BLR and the energy of the gamma ray, the emission region can be placed closer to or further away from the BH. This is discussed in detail in chapter 16.

## 2.4 In Particular: Perseus A / NGC 1275 / 3C 84

This work has a closer look at the radio galaxy 3C 84, which is identified with the gamma-ray source NGC 1275 and was detected as a gamma-ray emitter in 2008 by the *Fermi*-LAT [3]. NGC 1275 is the central galaxy of a galaxy cluster located at the center of the constellation Perseus, which is why it is also referred to as Perseus A. In optical images, we see NGC 1275 as a giant elliptical galaxy with a bright core and a lacy filamentary structure in its outer regions. Radio images provide a closer look at the bright core at the center and reveal a two-sided parsec-scale jet, see Figure 2.5. The best-resolved images at 22.2 GHz obtained by the RadioAstron space-VLBI mission even show a limb-brightened jet base at sub-parsec scales [87].

Since the Perseus cluster is a bright X-ray source, NGC 1275 was also observed with X-ray telescopes like Chandra [70].

As a very bright and “nearby” radio source with a distance of 62.5 Mpc [165], 3C 84 is one of the best-observed AGN, with observations reaching back to the 1950s, where the first radio telescopes were built. With the upcoming radio interferometry technology in the 1960s, 3C 84 became one of the most observed and best-studied radio galaxies [27, 58]. For example, the Metsähovi Radio Observatory is monitoring 3C 84 since the 1970s, see Figure 2.7.

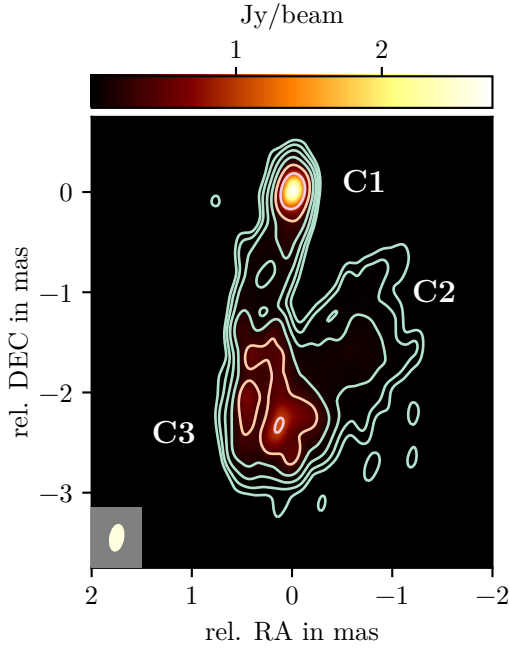


**Figure 2.5:** Monitoring of Jets in Active Galactic Nuclei with VLBA Experiments (MOJAVE) 15 GHz radio map of 3C 84 from 1997-06-01. We see the radio core and diffuse emission originating from the jet. Contour levels at 0.004, 0.02, 0.12, 0.60 and 3.16 Jy/beam. Flux is plotted in log-scale to visualize the faint diffuse emission of the jets.

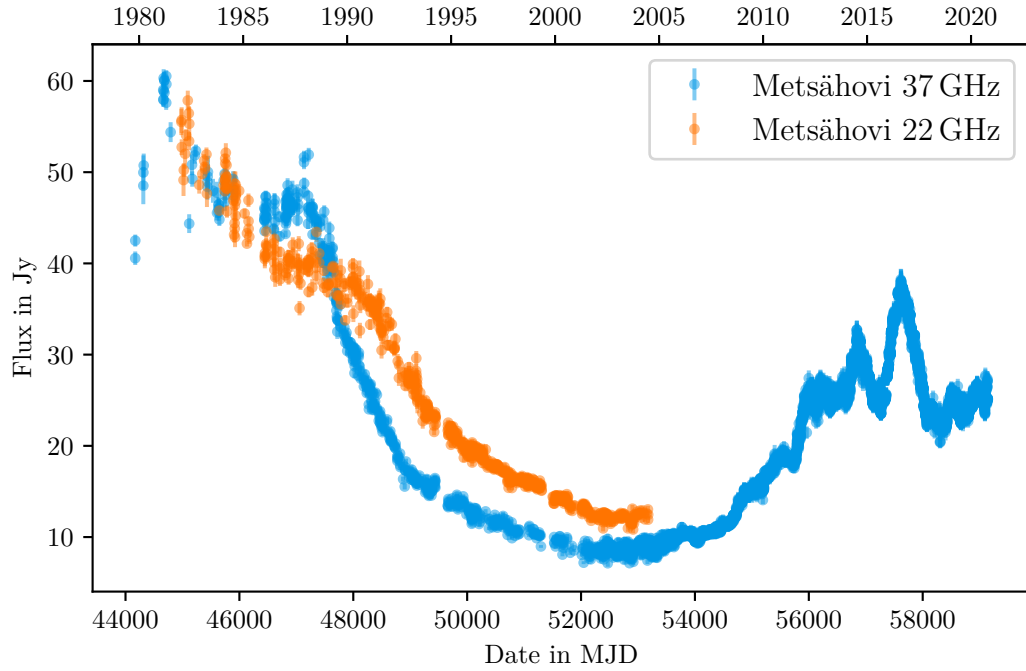
Especially in the radio regime, some remarkable features were observed. Close to the center, radio knots were ejected from the core, moving downstream the jet. One of these components (called C2) was ejected around 1980 [167] and is barely detectable these days. In 2005, a new radio component emerged from the core, called C3 [119]. This component moved along the jet and had its brightness maximum in 2016. After that, the component broadened and faded and is not detectable as a compact feature anymore. The components are shown in Figure 2.6.

Despite being well studied and observed over decades at different energies, some key aspects of 3C 84 remain unknown and are still subject of ongoing research. The existence and the comparatively low temperature of the mentioned filamentary structures can not be explained, since they are surrounded by gas at much higher temperatures. Since *Fermi*-LAT detected gamma-ray emission from NGC 1275, the origin of this radiation and the mechanisms responsible for it are still unclear. Also, key parameters of an AGN like the mass of the BH or the doppler factor remain unknown.

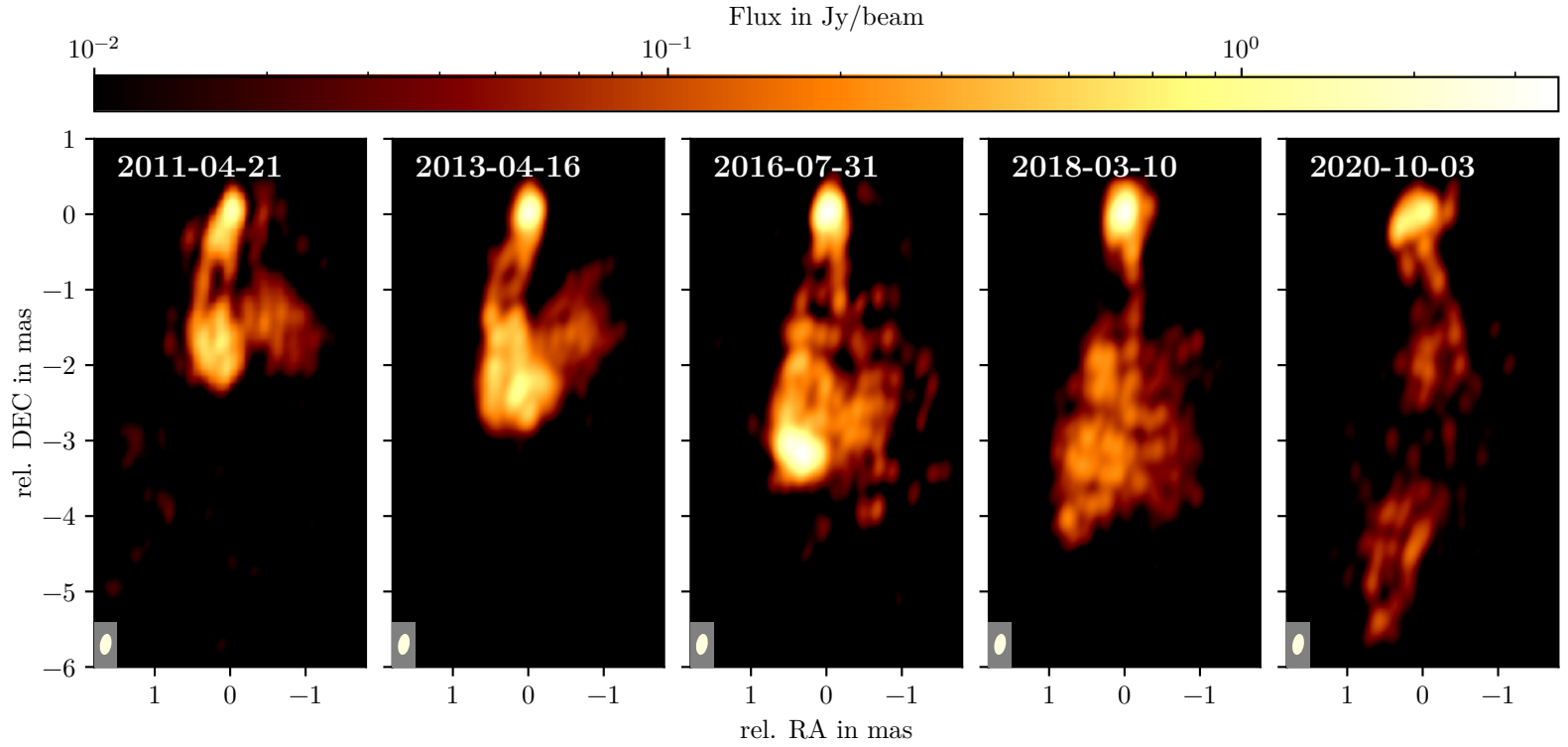
Furthermore, the calculations regarding the inclination angle, the angle between the jet direction and the line of sight, are very ambiguous. There are strong hints that the jet is precessing, which would explain the different values, see Figure 2.8. A dedicated study regarding the precessing jet was performed as a Master's thesis by Rune Dominik, supervised by me [62]. We show that the motion of the jet's direction from south-west to south-north in the radio images is consistent with a precession of 28.8 years period (21.4 years for a model with additional nutation). The results are published in Dominik, Linhoff, Elsässer, and Rhode [61]



**Figure 2.6:** Radio map of 3C 84 on 2013-06-16 observed by the VLBA-BU-BLAZAR program at 43 GHz. Contourlines are shown at 0.009, 0.02, 0.06, 0.15, 0.40, 1.05 and 2.76 Jy/beam. The beam size is depicted in the bottom left corner. C1 marks the radio core of the source and is probably hosting the BH. C2 is an old and faint radio component and not detectable anymore. C3 is a moving component which emerged from the core around 2005.



**Figure 2.7:** Radio flux of 3C 84 at 22 GHz and 37 GHz measured by the Metsähovi radio telescope [159].



**Figure 2.8:** Radio maps of 3C 84 observed by the the VLBA-Boston University Blazar Monitoring Program (BU-BLAZAR program) program at 43 GHz over several years. For better visualization, the colorbar is scaled logarithmically. The radio component C3 is moving in south direction, along the jet and away from the radio core (C1). It has its brightness maximum in 2016 at a distance of roughly 3 mas to the core and fades out afterwards. In the epoch of 2020-10-03 we see the curved shape of the jet, which is probably a result of the precessing jet base.



## 2.5 Extragalactic Background Light

The evolution of galaxies, the formation of stars and the radiation of AGN has filled the universe with diffuse radiation distributed over the entire energy spectrum, called the Extragalactic Background Light (EBL). This radiation still exists, although red-shifted to longer wavelengths at the UV, optical and infrared (IR) range ( $100\text{ }\mu\text{m} - 1\text{ mm}$ ). A Very-High Energy (VHE) photon coming from a source of interest can interact via pair production with this background radiation, which reduces its mean free path and affects the amount of radiation that can be measured at Earth. The opacity of the EBL strongly depends on the particle's energy and the distance of the source. For example, for sources at large distances (redshift  $z \approx 4$ ), the EBL can be completely opaque to radiation above 20 GeV [86]. The impact of the EBL must be considered, when the Spectral Energy Distribution (SED) of a source is studied, especially in the VHE range.

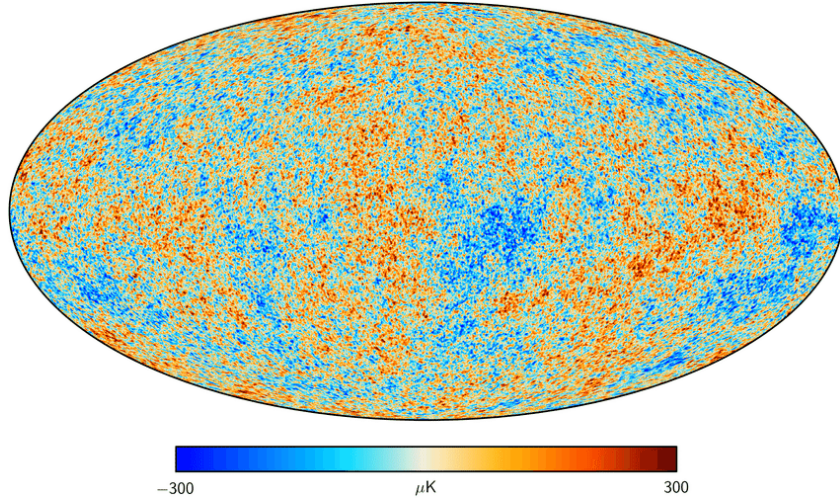
The EBL is not directly measurable from Earth, because it is much fainter than the zodiacal light. This is why complex modeling of the radiation coming from star formation and AGN is necessary to estimate the attenuation by the EBL. This effort was undertaken among others by Dominguez et al. [60], Gilmore, Prada, and Primack [86] and Franceschini, Rodighiero, and Vaccari [76], whose EBL models are widely accepted.

## 2.6 Cosmic Microwave Background

In the microwave range, the diffuse radiation in the universe is dominated by the Cosmic Microwave Background (CMB). The CMB is a remnant of the universe's recombination epoch, 380 000 years after the Big Bang. As the universe expanded and cooled during its evolution, the originally short wavelengths were stretched to millimeter wavelengths of microwaves. Today, the CMB can be measured as isotropic radiation with an almost perfect black-body spectrum at 2.73 K [75].

The CMB was theoretically predicted by Alpher and Herman in 1948 [20]. Since the late 1980s, it is studied systematically by space missions, beginning with the Cosmic Background Explorer (COBE), launched in 1989 [148]. The current mission lead by the European Space Agency (ESA), was started by launching the Planck satellite in 2009 [7].

Together with the EBL, the CMB also corresponds to the attenuation of the VHE gamma-ray spectrum and has to be modeled for very distant sources.



**Figure 2.9:** The Cosmic Microwave Background observed by the Planck satellite [8] showing the oldest light in the universe emitted 380 000 years after the Big Bang. The temperature fluctuations mark regions with higher density that are considered to be the origin of structures like galaxies and stars. Image credit: ESA and the Planck Collaboration.

## 2.7 Galactic Sources

So far, we had a closer look at extragalactic sources, which are, as the name says, located outside of our galaxy, the Milky Way. But also our direct neighborhood inside the Milky Way hosts a variety of objects that radiate at different wavelengths and can be detected at Earth. Among them are supernova remnants, Pulsar Wind Nebula (PWN), or binary objects, which I will introduce briefly.

### 2.7.1 Supernova Remnants

When a star has reached the end of its lifecycle, it collapses to a neutron star or a BH, depending on its mass. The collapse causes strong shock waves that propagate rapidly through the surrounding material and burst the shell away from the collapsed star. This event is called *supernova*. Such a collapse can happen if the nuclear fusion that keeps a star alive comes to a halt and the equilibrium between radiation pressure and gravitation is gone. Another mechanism causing a supernova is when a star accretes more and more matter from a nearby object up to a critical mass. If this mass is reached, the star is destroyed in a thermonuclear explosion.

The shell of gas and plasma that expands into space is called *remnant*. The shock waves propagating through the plasma create radiation at multiple wavelengths, from radio to gamma-ray emission.

### 2.7.2 Pulsar Wind Nebulae

If a star collapsed to a neutron star, it remains as a magnetized and fast rotating object, a so-called *pulsar*<sup>2</sup>. The rotating star ejects beams from its magnetic poles that can be detected as a precise periodical signal at Earth.

By its rotation, the pulsar's magnetic field pulls on charged particles and gas clouds inside the remnant and accelerates them to relativistic speed. This mechanism is called Pulsar Wind Nebula. PWN were detected across the spectrum in the radio, optical, X-ray, and gamma-ray range.

### 2.7.3 Binary Systems

A binary system is composed of two objects (stars, pulsars, BHs) rotating around their common center of gravity. If they come close enough, the more compact object with greater mass can accrete matter from the lighter object or absorb it completely. The gravitational potential energy of the transferred matter is then re-radiated from the compact object in the X-ray and gamma-ray band.

### 2.7.4 Crab Nebula: The Standard Candle for Gamma-Ray Astronomy

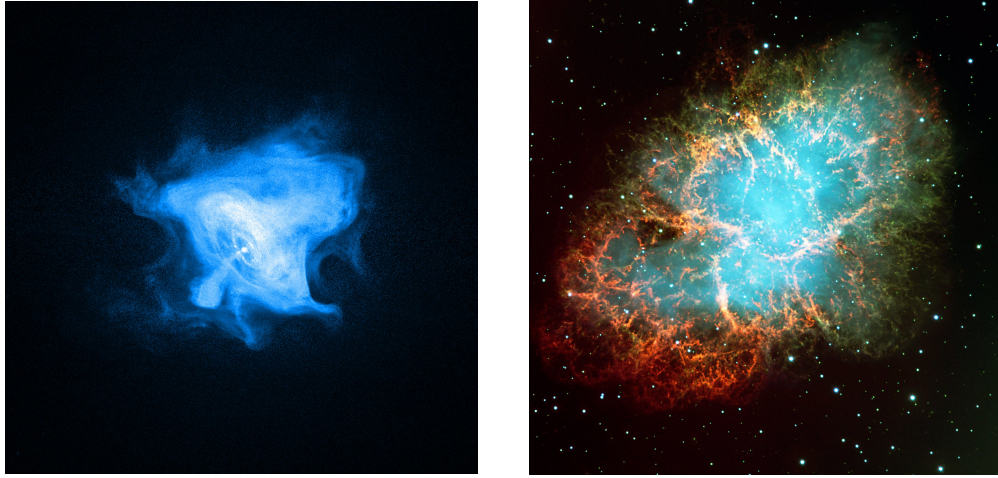
The most famous supernova remnant is the Crab Nebula, created by the gravitational collapse of a star at the end of its lifetime. In 1054, this supernova was reported by Chinese astronomers as very bright object at the sky, which was even visible at daylight [151].

Driven by the collapse of the star, a shock wave expands around the star and creates a shell of gas and dust. The collapsed star builds the Crab Nebula pulsar, a rotating neutron star, which is responsible for a steady and strong magnetic field within the nebula. The combination of these ingredients (charged particles, a strong magnetic field, and a shock wave propagating through the ejected matter) leads to a strong and steady emission which covers a broad energy spectrum from synchrotron emission at radio wavelengths, over X-ray emission to gamma rays at very-high energies. Due to its bright and non-variable flux at GeV to TeV energies, the Crab Nebula is one of the most observed and best studied gamma-ray sources. As so-called “standard candle” of gamma-ray astronomy, it is often used for testing new hardware or analysis approaches.

The Crab Nebula was detected as the first source at TeV energies by the Whipple telescope in 1989 [172]. Its emission was assigned to the remnants' shell, whereas dedicated studies about the pulsar emission followed years later, when further Cherenkov telescopes like the MAGIC were built [2, 12]. At radio wavelengths, the Crab Nebula pulsar was detected much earlier in 1968 [49] and named PSR 0531.

---

<sup>2</sup>short for: pulsating radio source. As the name indicates, pulsars were first discovered at radio wavelengths.



**Figure 2.10: Left:** Crab Nebula pulsar and its surrounding PWN as seen at X-rays by the Chandra X-ray observatory between 2001 and 2004 [145]. The white dot in its center marks the rotating neutron star that remains after the star collapse. Image credit: NASA/CXC/SAO F. Seward et al. **Right:** Optical image of the Crab Nebula revealing the structures of the exploded gas shell. The blueish structures at the center belong to the PWN inside the shell. Image credit: ESO.

## Observer's View

---

As stated in the last chapter, an AGN emits radiation at a great range of energies, and very different detectors and techniques are necessary to capture this emission at all energies (or: frequencies). The measured data and the high-level results obtained by analyzing these data are very different. Radio and optical astronomy can provide images of the observed object. With these images, the source morphology, its expansion, and its brightness at different wavelengths can be studied. Some parts of an AGN, like the core and the jet, can be resolved by radio maps, whereas optical images offer insights into a galaxy's disk and formations far away from the core.

### 3.1 Spectral Energy Distribution

Since not all detectors are capable of resolving the inner components of a source and actually “take a picture”, an astronomer is also interested in the amount of emission that is produced at a certain energy. This amount of emission, or: *flux*, at a certain energy can change over time, sometimes very rapidly. The shape of the so-called Spectral Energy Distribution (SED), which is flux per energy, time, and area, provides insights into the mechanisms that take place in the AGN itself. Theoretical predictions, resulting from models describing the processes within a source, can be checked against the measured SEDs. The overall spectral energy distribution of an AGN is a result of multiple processes which take place at different regions within the source and involve several particle populations. On top, the radiation emitted from a source far away can be attenuated or completely absorbed by the interstellar medium and the magnetic fields within. An SED is, therefore, always a superposition of models describing these various components.

Since this work deals with a specific AGN, we will have a closer look at the SED of NGC 1275. As one can imagine, collecting data over a great energy range from various facilities and research groups is challenging, even more, if a source is known to be variable in flux, which requires simultaneously measured data. Otherwise, an SED would contain information about different states of the source and different acceleration processes, which would make the results less conclusive. One publication, that has undertaken that effort is from Abdo et al., 2009 [3], see Figure 3.1. Not all of the data shown in Figure 3.1 was collected simultaneously, which is why *Fermi*-LAT measured flux points far above the upper limit that was calculated by the Energetic Gamma Ray Experiment Telescope (EGRET).

Like all standard SEDs of this source type, the SED of NGC 1275 consists of two bumps: the synchrotron bump on the lower-energy range of the spectrum (left) and the Inverse Compton (IC) bump on the higher energies (right).

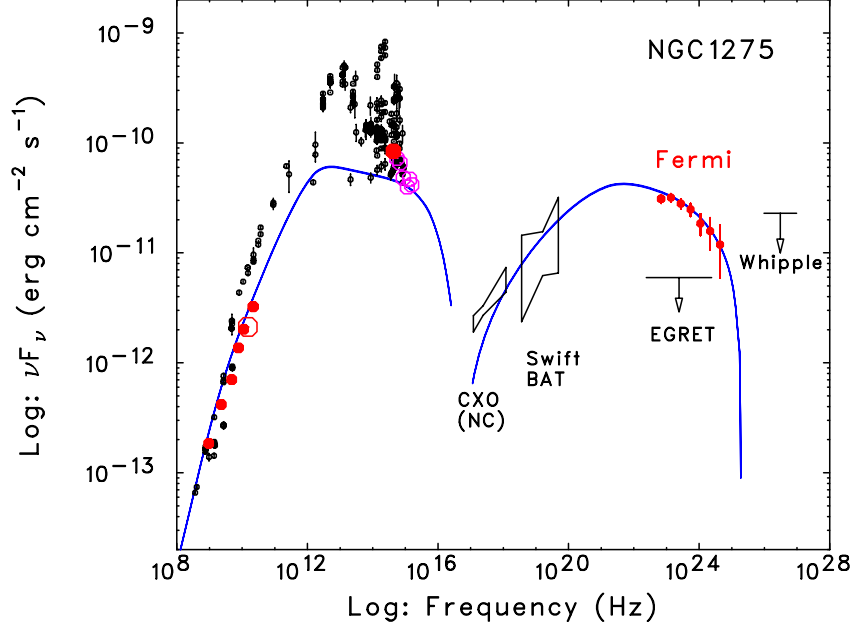
### 3.1.1 SED Modelling

The synchrotron bump, extending from radio to X-ray, results from free ultra-relativistic electrons emitting synchrotron radiation while forced on bent paths through magnetic fields. Synchrotron radiation has its upper limit somewhere at 50 GeV [104], which means that another process comes into play for the higher energies.

The IC scattering process is responsible for the bump at higher energies, from MeV to TeV, which represents the gamma-ray regime. In this process, low-energy photons are scattered to very-high energies by ultrarelativistic electrons, so that VHE gamma-ray emission can be produced. The so-called seed-photon field, the photons that can be scattered to such high energies, can originate from the synchrotron radiation (Synchrotron Self-Compton (SSC) process) or external fields, for example, photons from the BLR (External Compton (EC) scattering).

One may have noticed that both mechanisms only rely on leptonic particles. No hadrons are required to explain the two-bump structure. That is why they are referred to as *leptonic models*. Most blazar SEDs are well-described by leptonic models, but in some cases, they fail to adequately explain the high-energy bump, and other particle populations are considered. In these so-called *hadronic models*, high-energy emission is produced by synchrotron radiation of protons or secondary particles produced in proton-proton interactions or pion decays, see e. g. [42].

In the case of NGC 1275, the SED presented by Abdo et al. in Figure 3.1 was modeled using two leptonic models, see [3, 74, 80] for further reading.



**Figure 3.1:** SED of 3C 84/NGC 1275 published by Abdo et al. in 2009 [3] using radio data from radio, optical, X-ray and gamma-ray telescopes. The SED is fitted with a leptonic one-zone model (blue dashed curve) and a decelerating flow model by Georganopoulos and Kazanas [80]. The data combined in this plot was not taken simultaneously. In the MeV range, *Fermi*-LAT detected the source in 2008 above EGRET’s upper limit, which was measured in the 1990s. This is a hint that NGC 1275 was in a higher state of activity after 2000.

## 3.2 Light Curve

Another key to decode the characteristics of an AGN is its light curve, which means the amount of flux in a specific energy range over time. AGN are highly variable objects, especially in the gamma-ray regime. What causes these variations is still unclear in some cases. For multi-wavelength studies, light curves obtained at various energies are compared to identify parts of the source or particle populations that were involved in these flux variations. There are mainly two types of variation: long-term trends and rapid flux changes, named *flares*. Changes in an AGN's light curve in the radio regime can be connected to flux variation in the gamma-ray regime. Recalling the leptonic model, this is the case when increasing synchrotron emission creates more seed photons for SSC scattering and the gamma-ray flux increases also.

But not for every gamma-ray flare activity at lower energies increases. For these so-called *orphan flares*, no counterpart at other energies can be detected. In January 2017, an orphan flare of NGC 1275 was detected by MAGIC, where the gamma-ray flux was 50 times higher for two days than in previous observation campaigns [22]. At the same time, no radio observatories reported any special behavior, and no new radio component was detected from 3C 84 since then.



## **Part II**

### **MAGIC**



## Ground-Based Gamma-Ray Detectors

---

At the highest energies, our Earth’s atmosphere is not transparent for incoming particles. Instead of propagating straight through the atmosphere and hitting the surface, they initiate particle cascades and the energy of the primary particle is distributed over a bunch of secondary particles, see [Figure 4.1](#). This is a big advantage for all organic life on the planet, since being exposed to gamma radiation all the time would mean a severe health risk for all living things. In turn, for a gamma-ray astronomer interested in these exact particles, this circumstance causes some extra efforts.

There are three ways to account for this “problem”: First, measuring the gamma radiation above Earth’s atmosphere using satellites equipped with gamma-ray detectors. Second, measuring not the gamma particle itself but the light cone produced by its secondary particles when it enters the atmosphere. And third, detect the secondary particles reaching the ground with water Cherenkov detectors. Space-based detection of gamma rays is possible with the *Fermi*-LAT, which is described further in [chapter 10](#). Most techniques used by ground-based detectors are built on the *Cherenkov Effect*.

### 4.1 Cherenkov Effect

The Cherenkov effect, named after Pavel Cherenkov, who first described this effect in 1934 [\[43\]](#), occurs if a particle moves faster in a medium than light in that particular medium. This is the case for all the secondary particles in a cascade started by a high-energetic particle that enters the atmosphere. Which kinds of secondary particles are produced depends on the primary particle. There are mainly two types of particles that are of interest for gamma-ray astronomy: photons and hadrons. Photons are the messengers of gamma radiation, whereas hadrons are the main background and have to be eliminated later. There are more particles entering our Earth’s atmosphere like electrons, positrons, or compounds like helium or iron nuclei, but for Cherenkov astronomy, we focus on photons and hadrons. Although labeled as background in the gamma-ray community, the group of hadronic particles is called Cosmic Rays and the subject of a dedicated field of research.

The particle showers initiated by hadrons and photons are produced several tens of kilometers above Earth’s surface and are composed of different secondaries, see [Figure 4.1](#). Nevertheless, the shapes of hadron- and photon-induced showers appear very similar with only minor differences, which makes it hard but possible to separate them.

An incoming high-energetic photon will produce an electron-positron pair by colliding with an atom in Earth’s atmosphere. Electron and positron produce new photons via bremsstrahlung in the field of other air molecules, which again can trigger pair production if carrying enough energy. This cascade goes on until all secondary particles have lost

#### 4 *Ground-Based Gamma-Ray Detectors*

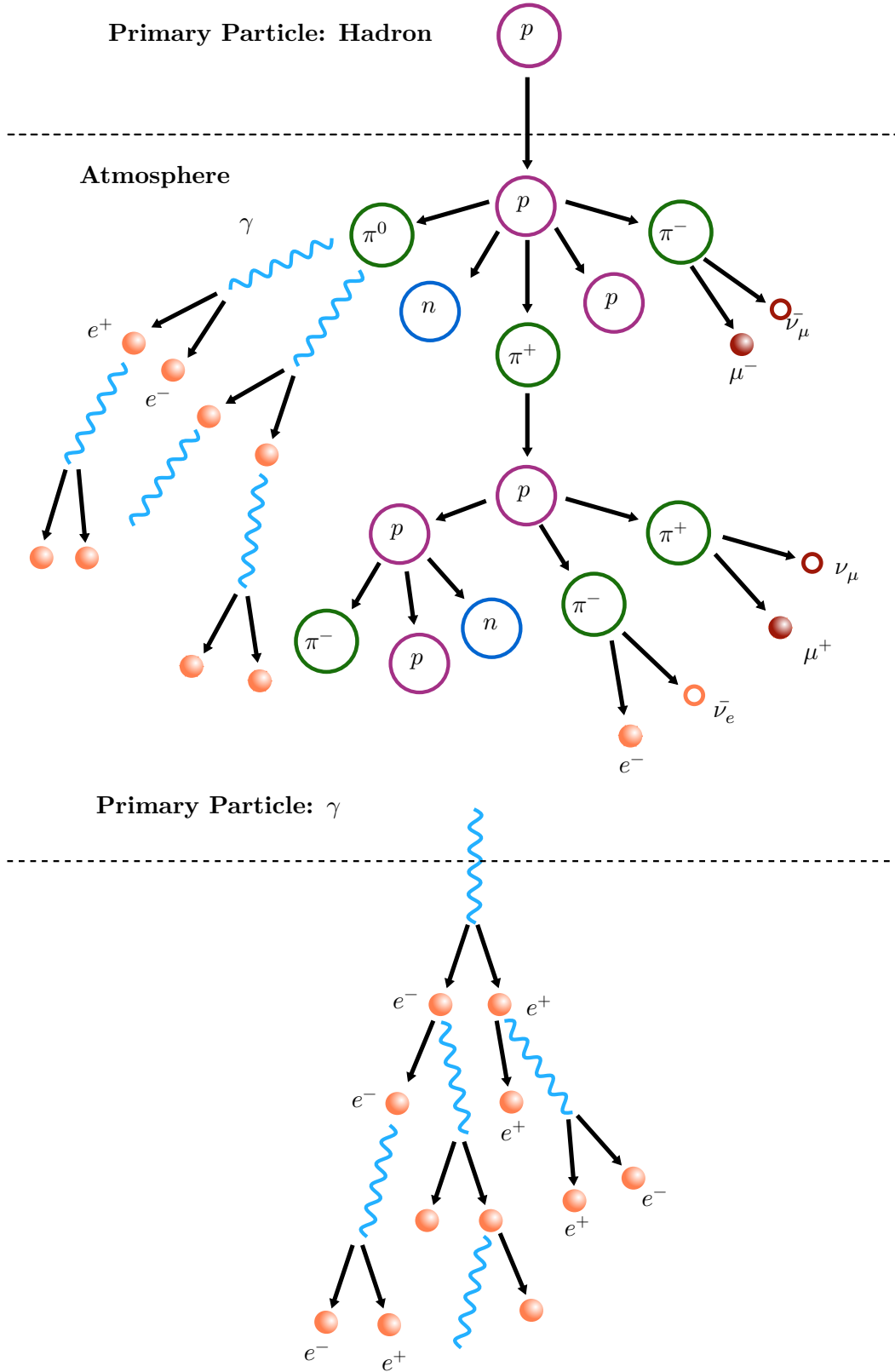
their energy and no more photons or electrons are produced. For hadrons, more particle types come into play. Among others, hadrons produce pions when they interact with the atmosphere, which then decay into two photons (neutral pion) or electrons, muons and neutrinos (charged pions). These particles interact again with the surrounding matter, producing a similar cascade as the gamma-ray secondaries.

All of these cascade particles move through the atmosphere at a very high velocity, higher than the speed of light in this medium. The charged particles cause a short moment of polarization of the surrounding molecules at every point on their track. At the moment when the polarized molecules relax back to their ground state, visible and UV light at a continuous spectrum peaking at 420 nm is emitted, the so-called *Cherenkov Photons* [107]. Due to its high velocity, the particle is always ahead of this wavefront of the emitted light. All the wavefronts interfere at a certain angle  $\theta$  to the particle's propagation direction and build a light cone. The angle depends on the refractive index of the medium and can be calculated with:

$$\cos \theta = \frac{c}{nv_p}$$

with  $c$  the speed of light in vacuum,  $n$  the refractive index, and  $v_p$  the particle's speed.

The light cone produced this way by the particle cascade is visible for a very short moment, lasting several nanoseconds.



**Figure 4.1:** Particle cascades for hadron- and photon-induced air showers. Hadronic particles involve, among others, pion production and decay, while gamma-ray showers are purely electromagnetic.

## 4.2 Imaging Air Cherenkov Telescopes

Ultra-sensitive telescopes at Earth’s surface, so-called Imaging Air Cherenkov Telescopes (IACTs), can take a very short video of the air showers. Algorithms and software tools can then extract information about the showers from these videos, e.g. its size and the direction of its origin. The extracted features enable us to reconstruct the characteristics of the primary particle, mostly its type, energy, and source position. To be able to catch these short and faint events, IACTs need extremely sensitive cameras, a very fast electronic trigger, a clear sky, and as little background light as possible. This is why IACTs are operated only at night, high altitudes, and far away from any civilization.

IACTs consist of a mount that can be rotated to every position in the sky, a parabolic-shaped mirror with several meters in diameter collecting the light from air showers, and a camera taking short videos of the reflected light. Since it is very expensive to build mirrors of the required size, the reflector is composed of smaller mirrors, quadratic, hexagonal, or round. Most IACTs operate in systems of two to five telescopes to improve the reconstruction of a particle’s features, but IACT systems with more telescopes operating simultaneously are under construction ([subsection 4.2.2](#)). The cameras used for these telescopes consist of hundreds of pixels, sensitive enough to detect single photons. There are mainly two technologies used to produce a pixel: Photomultiplier Tubes (PMTs) and Silicon Photomultipliers (SiPMs).

PMTs are evacuated tubes containing dynodes and anodes and a photocathode at the front side. Photons from the Cherenkov light hit the photocathode and produce electrons via the photoelectric effect. These electrons are multiplied by being accelerated to the following dynodes. At the last step, the bunch of electrons reaches the anode and leaves the tube creating a voltage pulse, which is used to analyze the incoming primary photon. PMTs have to be operated at a very high voltage and can easily be damaged, for example by moon- or daylight. Nevertheless, most of the currently operating Cherenkov Telescopes use PMTs.

SiPMs are groups of photo-diodes that are operated with reverse bias. A single Cherenkov photon initiates an avalanche of charge carriers (electron-hole pairs) that can be measured as an avalanche current. This technology has some advantages compared to PMTs: SiPMs can be operated at a much lower voltage, which reduces wear. Furthermore, SiPMs are much more robust against background light as PMTs, which makes operation even at full moonlight technically possible. (For further information, see [\[21, 40, 177\]](#).) For this reason, most of the upcoming telescope cameras are built with SiPMs.

### 4.2.1 MAGIC

The Major Atmospheric Gamma Imaging Cherenkov telescopes are a stereoscopic system of two Cherenkov telescopes located at the Roque de Los Muchachos on the Canary Island of La Palma, Spain, at an altitude of about 2200 m. The first telescope, MAGIC I, was inaugurated in 2004 and the second telescope saw its first light in 2009. From 2011 to 2012, a major upgrade was performed, changing some hardware and improving the overall performance, see [\[15, 16\]](#). Both telescopes have a parabolic mirror with 17 m in diameter.

MAGIC was built and is used to detect gamma-rays in the VHE gamma-ray regime above 30 GeV up to 30 TeV. Further technical details are described in [chapter 5](#).

#### 4.2.2 CTA

The Cherenkov Telescope Array (CTA) represents the next generation of gamma-ray detectors. It will be built in two locations where multiple telescopes of different sizes and characteristics will be placed. Fully built, CTA will be composed of 70 Small-Sized Telescopes (SSTs) most sensitive in an energy range above 10 TeV, 40 Mid-Sized Telescopes (MSTs) covering the core energy range of 100 GeV to 10 TeV, and eight Large-Sized Telescopes (LSTs) responsible for the low-energy particles below 100 GeV. The first telescopes are currently built next to the MAGIC telescopes at the Canary island La Palma (Northern hemisphere). As the second site, Paranal in Chile (Southern hemisphere) has been chosen. This unique combination of two sites at both Northern and Southern hemispheres enables observing the full sky. The first LST was already built at La Palma and started data taking in 2018. The full array will not be finished before 2025. With its full-sky coverage, the broad energy range, and the ability for long-term observations, CTA will deliver very valuable observations, longed for especially in the multi-wavelength community. For detailed information, see [\[5\]](#).

### 4.3 Water Cherenkov Telescopes

Besides measuring the Cherenkov light of the particle cascade emitted in the air, another opportunity exists to exploit the Cherenkov effect for particle detection. For very high-energetic cascades, some of the secondary particles can reach the ground. These particles are detected in large water tanks where they also emit Cherenkov light. Unlike IACTs, the water Cherenkov detectors can operate nearly 24/7 and cover a much larger Field of View (FoV). On the downside, they have a much higher energy threshold because they require secondary particles to reach the ground. This happens only for high-energetic particles above 10 TeV in sufficient quantity.

Currently operating facilities using this technique are, for example, the Pierre Auger Observatory in Argentina [\[4, 48\]](#) and the High-Altitude Water Cherenkov Gamma-Ray Observatory (HAWK) in Mexico [\[19, 10\]](#).





## Data Analysis of the MAGIC Telescopes

---

As already mentioned in [section 4.2](#), the MAGIC telescopes record very short videos of air showers induced by charged particles in Earth's atmosphere. To derive scientific insights into the secrets of a source from these videos, a complex data analysis has to be performed. This is done by an analysis chain customized for the detector hardware, the observed source, weather and Moon conditions at the time the data was taken. In the next sections, I will describe how an air shower is recorded ([section 5.1](#)), how the properties of the primary particle are estimated ([section 5.3](#)), and how the observation of thousands of these showers finally leads to scientific results ([section 5.5](#)).

### 5.1 Data Taking

A typical observation by the MAGIC telescopes starts with calibrating procedures before the actual target object is observed. One observation unit, called *run*, takes about 20 minutes and is divided into several *subruns*. Both telescopes measure simultaneously and point to the exact same position in the sky. Depending on the source position in the sky and the importance of the source, the observation time varies, but is usually in an order of one to several hours.

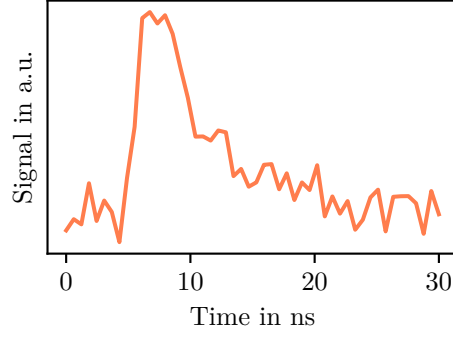
Both telescopes are equipped with identical cameras consisting of 1039 hexagonal pixels. Each pixel is made of a PMT with a hexagonally-shaped Winston cone mounted on its entrance window. The cone is used as a light guide with one side adapted to the PMT and the other side of hexagonal shape with a diameter of 3 cm (flat-to-flat), to fill the camera's surface completely. Each cone has a FoV of  $0.1^\circ$  which results in a total FoV of  $3.5^\circ$  for each telescope [\[16\]](#).

The PMT is connected to a Domino Ring Sampler version 4 (DRS4) chip which reads out the signal produced by the incoming photons with  $1.64 \times 10^9$  samples per second. If the signal exceeds a certain trigger threshold, a time series of 50 values ( $\approx 30$  ns) is stored for all pixels in the camera. In stereo mode, which is the usual mode for MAGIC observations, the cameras of both telescopes are read out at the same time.

Whatever triggered the electronics is called an *event*. If the event was induced by an air shower, the time series follows a typical pulse shape: a steep increase followed by a flat decay, see [Figure 5.1](#). The area under this curve scales with the number of photons arriving at the PMT. But, naturally, not only photons from air showers trigger the readout, but also light from human activity near the telescopes, or just electronic artifacts can cause events.

As one can imagine, this highly-sensitive system does not deliver smooth and meaningful photon-pulses out of the box. Careful calibration and pre-processing are necessary to obtain

**Figure 5.1:** Sketch of a pulse shape as measured by a camera pixel. Real data can contain more noise and artifacts that have to be removed during the calibration process. Photon charge and arrival time are calculated by finding the maximal integral of 6 consecutive time slices.

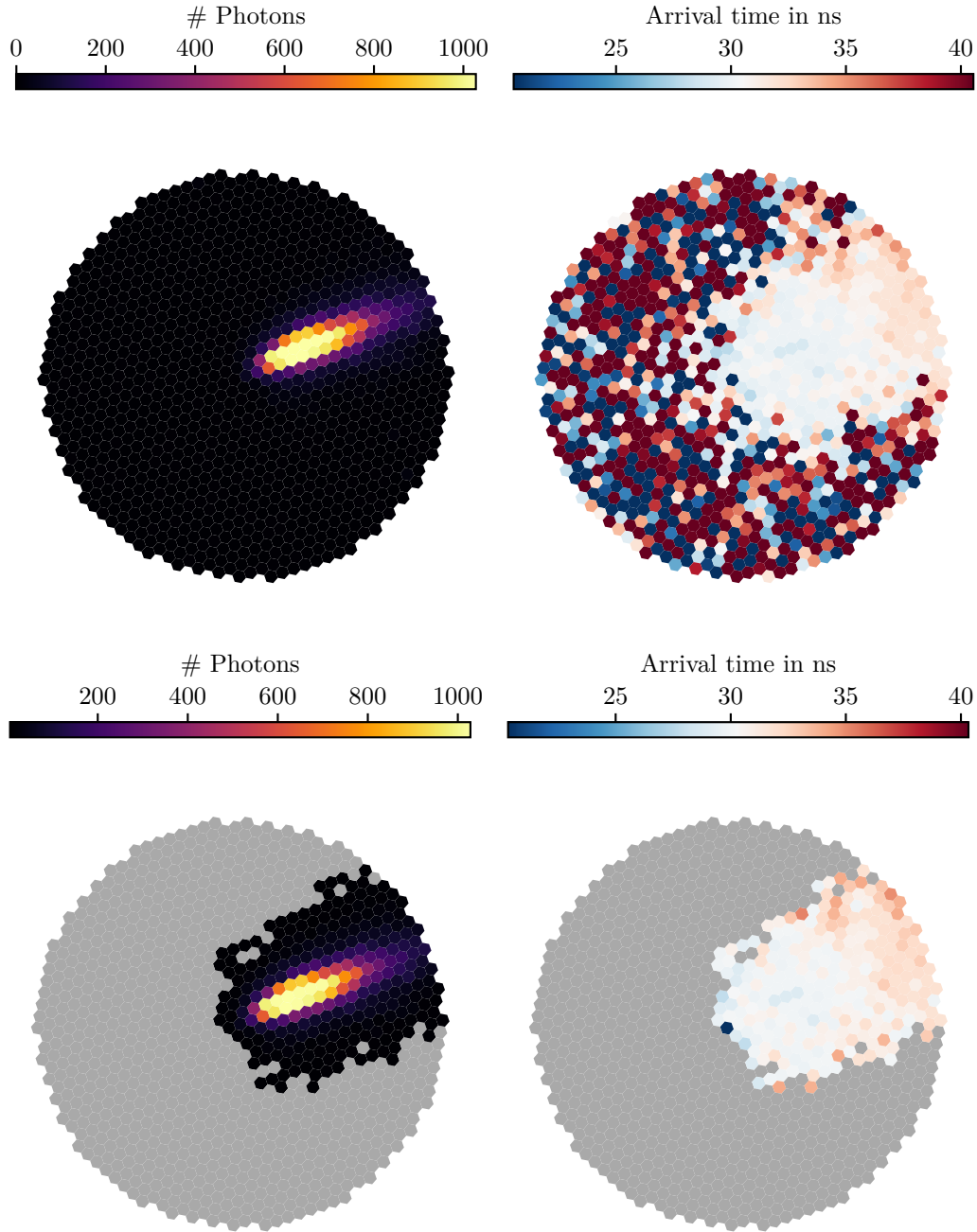


data that can be used for further reduction. The calibration technique is further described in [15, 16].

Once the system is set up, the values for *arrival time* and the number of photons (*photon charge*) can be extracted from the pulses in every pixel. This is done by a “sliding window” algorithm that finds the six consecutive time slices with the maximum integral. The arrival time is then calculated as the average of time stamps belonging to these time slices weighted with their charge. The photon charge is computed from the charge in the selected time slices by applying the F-factor method, described in [115]. For further details regarding the extraction of photon charge and arrival time, see [16, 12].

Having the number of photons in a certain time interval per pixel results in an image containing a bright area of a typically elliptical shape, see Figure 5.2 upper panel.

Since the whole camera is read out after a trigger, the camera image shows pixels that contain shower photons, as well as pixels containing just noise, for example from the Night Sky Background (NSB), moonlight, or electronic artifacts. These noisy pixels have to be removed from the images and the pixels containing the shower photons have to be selected, see Figure 5.2 bottom panel. This task is called *cleaning* and is another important step in the data analysis chain. The cleaning suppresses isolated pixels and selects groups of pixels that by their arrival time and photon charge are assumed to belong to the shower, see, e.g. [179].



**Figure 5.2:** Top: Full MAGIC camera image showing the photon charge (left) and the arrival times of the photons (right) per pixel. Bottom: Selected pixel after the image is cleaned. Based on the remaining pixels, the image features are calculated. The arrival times reveal the direction of the shower development from early (blue) to later (red) arriving photons. Images are created using [103].

## 5.2 Hillas Parameters

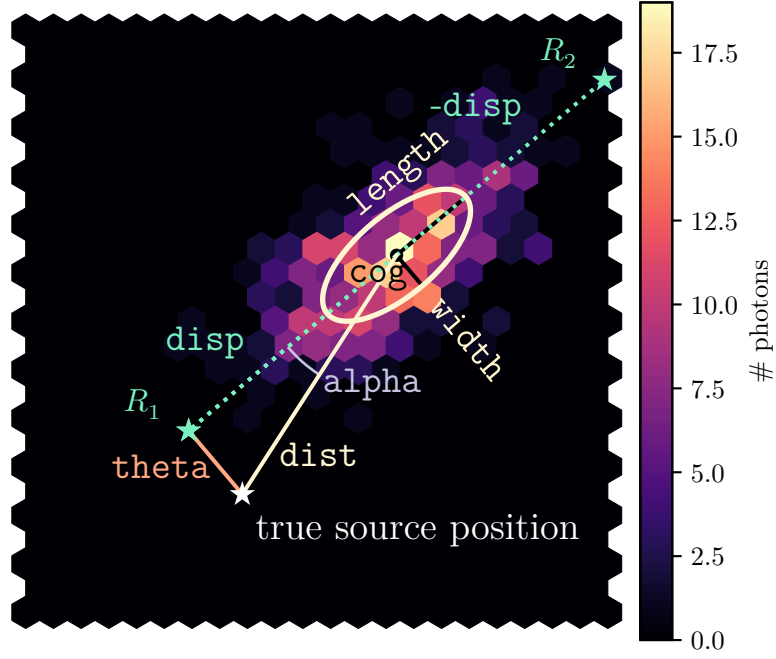
To use the cleaned shower images for further analysis steps, like gamma/hadron separation or energy reconstruction, they have to be parameterized, first. Although there has been major progress in technical equipment and modern software frameworks, the way IACT images are parameterized has stood the test of time and is still based on the “Hillas parameters”. Michael Hillas introduced his idea of describing a shower in 1985 [94] and established the basis for today’s data analysis. To estimate the two axes of a shower image, a Principal Components Analysis (PCA) is performed with the pixels that were selected as shower pixels by the cleaning.

The original *Hillas Parameters* are (see also Figure 5.3):

- **length**: Root Mean Square (RMS) distance of the light distribution along the shower axis, calculated from a principal component analysis
- **width**: RMS distance of the light distribution perpendicular to the shower axis, calculated from a principal component analysis
- **frac(2)**: The fraction of the amount of photons in the two brightest pixels of the shower and the sum of photons over all other pixels
- **cog**: The Center of Gravity (CoG) weighted by the brightness of the pixels
- **dist**: The distance from the centroid to the known source position
- **alpha**: Angle between the centroid-source connection and the shower axis

Over the years, additional image features were defined to complement the Hillas parameters. For the MAGIC data analysis, the arrival times are taken into account, as well as the amount of edge pixels hit by the shower (*leakage*), or higher momenta of the brightness distribution. Furthermore, information gained from the stereo-observation is added to the event. Once a shower is parameterized this way, the data volume is heavily reduced from two images of 1039 pixels to a list of parameter values per shower event.

Now, for every detected air shower, one data sample per telescope exists, consisting of a list of features that can be used by the Machine Learning (ML) algorithms that are described in the next section. But we can learn very little from a single shower. In fact, we have to collect hundreds of events to gain relevant information about an observed object. This is called *event-based* analysis, in contrast to, for example, radio astronomy, which is *image-based*.



**Figure 5.3:** Hillas parameters calculated for a shower image. `length` and `width` denote the two components of the principal component analysis, `cog` is the Center of Gravity of the pixels, weighted by their number of photons. Along the main axis of the resulting ellipse, two source positions  $R_1$  and  $R_2$  can be reconstructed in a distance `disp` from the `cog`. `disp` is calculated using a Random Forest (RF). The angular distance between the true and the reconstructed source position is called `theta`, the corresponding angle is `alpha`. Figure adapted from [125].

### 5.3 Gamma/Hadron Separation and Event Reconstruction

To obtain scientifically valuable results from the parameter sets we obtained in the last section, such as light curves and spectra (see [chapter 3](#)), three tasks have to be managed beforehand by the data analysis:

- Estimate the energy of the primary particle
- Reconstruct its direction
- Separate the gamma-ray events from the background events.

An IACT detects all sorts of air showers, no matter which particle induced them, see [section 5.1](#). For gamma-ray astronomy, as the name says, exclusively the gamma rays coming from a particular source (*signal*) are of interest. Showers induced by all other particles, mainly hadrons, are defined as *background*. Looking at the numbers, a background event happens roughly  $10^3$  to  $10^6$  times more often than a signal event, depending on the source flux. For weak sources, the background rate is at the higher range. Observing gamma rays becomes a search for the photon-needle in a haystack of hadrons.

All these tasks are performed based on features extracted from the shower image. These features serve as input for Machine Learning algorithms that classify a shower and estimate the energy of the primary particle and its direction. In the next subsections I will describe the features that are extracted from the shower images followed by the ML algorithm that is used to estimate particle type, source position and energy in [subsection 5.3.1](#).

#### 5.3.1 Random Forest

After extracting the image parameters for every shower event, the gamma-ray-induced showers have to be separated from hadron-induced showers. Furthermore, their energy and the distance of the reconstructed source position from the *cog* of the shower have to be estimated. For MAGIC, this is done by a ML algorithm, the Random Forest [[36](#), [35](#)].

A RF is a supervised algorithm, meaning that the classes which have to be separated are known in advance (for example *gamma* and *hadron* in our case). The algorithm “learns” to distinguish pre-labeled gamma and hadron events by their parameters (also called *features*). The data set with pre-labeled events is called the *training set*.

For hadron-labeled data, observations with no strong gamma-ray source in the FoV are used. For gamma events, no pure data set exists, which is why these events are artificially created with Monte Carlo (MC) simulations (see next section).

The RF is an ensemble of decision trees, which are trained independently from each other. The decision tree is described in the next subsection. Since a single decision tree would tend to fit only the training set perfectly (so-called *overtraining*), an ensemble of  $n_{\text{tree}}$  decision trees is trained.

If the trained RF is applied to unlabeled data, every decision tree classifies every event as gamma or hadron. The average over all the classifications results in a parameter with multiple names (*prediction value*, *confidence*, *gammaness*, *hadronness*), which depends mostly on the community slang. According to MAGIC, I go with *hadronness* from now on (which is the same as  $1 - \text{gammaness}$ ).

### 5.3 Gamma/Hadron Separation and Event Reconstruction

The hadronness, by nature, has a value between 0 and 1, which is calculated for every event by the RF. Based on this value, a cut can be applied, for example

$$\text{event type} = \begin{cases} \text{gamma}, & \text{hadronness} \leq 0.28 \\ \text{hadron}, & \text{hadronness} > 0.28 \end{cases} \quad (5.1)$$

for a cut at  $\text{hadronness} = 0.28$ .

#### Decision Tree

To build a decision tree, the pre-labeled training data of both classes is split based on one feature. The subsets are then split again, based on another or the same feature, and so on. Every split has the aim to enlarge the homogeneity in the subsets, which means containing more and more events of one class. The splitting stops if a subsample contains only events of one class or after a pre-defined number of splits. Every split of a decision tree is called *node*; a node that is not split further is called *leaf*.

At every node, a feature is chosen based on which the sample is split. In case of RFs, this feature is selected as the best-splitting feature from a randomly sampled subset of all features. If a feature is non-binary, which means that more possibilities exist to split the data sample, multiple splits are tested and the best split is taken.

After several splits, the data is separated into leaves with only one class of events or, depending on the user specifications, leaves containing a majority of one class.

For calculating “how good” a split performs in order to separate the target classes, the *Gini-Index* is used [11]. The Gini-Index is one common value to calculate the goodness of a split. Another common measure is the *Information Gain* or *Entropy*.

The Gini index for one node can be calculated as

$$Q_{\text{Gini}} = 4 \frac{N_{\gamma}}{N} \frac{N_{\text{h}}}{N} \in [0, 1] \quad (5.2)$$

with  $N$  the total number of events in a subnode and  $N_{\gamma}$  and  $N_{\text{h}}$  the number of events belonging to one class.

For a certain split  $\Gamma$  into two subnodes, denoted *left* and *right*, the Gini indices of the subnodes are added and scaled to  $[0,1]$ :

$$G_{\text{Gini}}(\Gamma) = 2 \left( \frac{N_{\gamma, \text{left}}}{N_{\text{left}}} \frac{N_{\text{h}, \text{left}}}{N_{\text{left}}} + \frac{N_{\gamma, \text{right}}}{N_{\text{right}}} \frac{N_{\text{h}, \text{right}}}{N_{\text{right}}} \right) \in [0, 1] . \quad (5.3)$$

For a perfect split with completely pure subsamples, the Gini index would be zero. Calculating the “best split” now means minimizing  $G_{\text{Gini}}$ :

$$\Gamma^* = \text{argmin}_{\Gamma} G_{\text{Gini}}(\Gamma) . \quad (5.4)$$

A RF can also be used for regression tasks. In that case, the events are split by minimizing the variance of the subsamples’ energies. The energy variance in a subnode is given as:

$$\sigma_{\text{subnode}}^2(E) = \frac{1}{n-1} \sum_{i=1}^N (E_i - \bar{E})^2 \quad (5.5)$$

## 5 Data Analysis of the MAGIC Telescopes

with  $N$  the number of events in the subnode, their energy  $E_i$  and the mean energy  $\bar{E}$ .

The term that has to be minimized is then the sum of the variances in both subnodes weighted by their population:

$$\sigma^2(E) = \frac{1}{N_L + N_R} (N_L \sigma_L^2(E) + N_R \sigma_R^2(E)) . \quad (5.6)$$

This approach is used for the estimation of a particle's energy and the `disp` (subsection 5.3.2). For further details of the implementation of RFs in the MAGIC analysis, see Albert et al. [11].

### 5.3.2 Reconstructing the Source Position

Besides the particle type, another important feature of the primary particle has to be derived from the shower images: the source position. In the case of the MAGIC analysis, the distance of the estimated source position from the CoG of the shower along the shower axis, the so-called `disp`, is estimated by a RF. This approach results in two possible positions, also known as head-tail ambiguity (c.f.  $R_1$  and  $R_2$  in Figure 5.3).

The decision, which position to take, is made by joining the images of the shower from both telescopes, if they are available. The estimated source position is now defined as average of the two closest positions, if they have a distance smaller than  $0.22^\circ$ , see Figure 5.4. If no pair of positions can be found fulfilling this requirement, the corresponding event is assumed to be hadron-induced and rejected [16].

### 5.3.3 Hadron Training Datasets

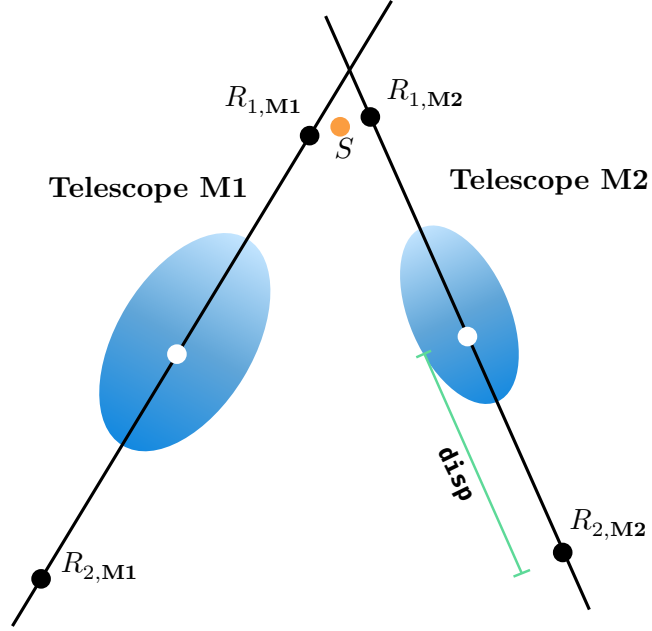
To train a RF, labeled training data for all classes is needed. In the case of the gamma-hadron separation, data sets for photon-induced air showers and hadron-induced showers are required. For the MAGIC telescopes, the hadron-induced training data is obtained from real measurements. The telescopes are pointed to a position at the sky with no or a very weak gamma-ray source. The measured data is assumed to contain only hadronic showers. Therefore, all showers from these observations are labeled as “hadron”. Events observed and labeled this way are also called *Off data* or *Off events*.

Not all IACTs use real observations as data for hadron events. It is also possible to simulate hadronic interactions in the atmosphere and their following cascades. In contrast to gamma-ray induced air showers, hadronic air showers involve various kinds of particles, which makes the simulation much more complex and expensive. Simulating suitable hadron MCs that match the real observations is a very difficult task. Therefore, MAGIC decided to use real measurements as training data. This data comes for free if weak sources are observed and can not be detected. Even if the data might contain no or very few gamma-ray events from the target source, enough events from the hadronic background are recorded.

### 5.3.4 Gamma-Ray Training Datasets: Monte Carlo Simulation

In contrast to the training data set for hadronic events, a pure gamma-ray data set can not be obtained by observations. Gamma-ray-induced shower images are, therefore, created





**Figure 5.4:** Stereo reconstruction of the source position from shower images of the two MAGIC telescopes. The distance `disp` is estimated by a RF, which results in two possible source positions per shower along the shower axis  $R_{1,2}$ . The two closest positions are averaged weighted with the number of pixels in the shower image to calculate the estimated source position  $S$  (orange dot). If the smallest distance between two positions is  $> 0.22^\circ$  the event is rejected. White dots mark the cog of the shower image.

artificially by propagating a particle of known origin and energy through the atmosphere with all its interactions and secondaries. The Cherenkov light emitted by these secondaries can be ray-traced to the known geometry of the telescope's mirror and into the camera, where it can exceed a simulated trigger threshold and result in a simulated image. This is done several million times for different particle energies and directions, resulting in a set of gamma-ray events with known origin. These images can now be cleaned and parameterized the same way as observed data, and the calculated image parameters serve as input for the RF.

The simulation of gamma-ray events is implemented in the software package Cosmic Ray Simulations for Kascade (CORSIKA) [92], which was originally invented to simulate air showers for the Karlsruhe Shower Core and Array Detector (KASCADE) experiment [59] in the 1980s. CORSIKA's core is written in FORTRAN 77, but several extensions were developed over the years, written in C. The MC simulations used in this work were produced with the MAGIC Monte Carlo Software (MMCS), a customized CORSIKA version for MAGIC [149].

The detector response from reflector and camera of the MAGIC telescopes is implemented within the MAGIC Analysis and Reconstruction Software (MARS), which is described further in [section 5.7](#).

Simulating air showers is by no means an easy task. Extremely precise knowledge about the particle interactions and cross-sections, especially in the VHE regime, is required as well as a complete understanding of the detector with all its hardware features (and bugs). This is why CORSIKA and MARS are constantly modified and improved. Only if the simulations fit exactly to the measured data, a RF is able to distinguish the events based on their primary particle and not between measured and simulated data.

For the same reason, a single set of simulated gamma-induced showers is not suitable to analyze a large amount of measured data. The more exactly the parameters of the simulation match the conditions of the measured data, the better the RF can reject the background and estimate the energy and the `disp` parameter. This is why a variety of MC sets exist for different software configurations, observation modes, and zenith angles, for example. Furthermore, changes in the telescope's hardware, through weather, Calima<sup>1</sup>, wear and upgrades have to be considered in MARS. A MC set is therefore valid only for a specific time range and specific observation conditions. This circumstance may appear as a side note here but increases the effort for an analyzer immensely, as we will see later in [section 8.4](#).

Simulated gamma-ray events are not only used to train a RF, but also to calculate the detector response, as we will see in [section 5.4](#).

### 5.4 Instrument Resonse Functions

After all events are equipped with a hadronness value and a cut was applied, the data sample is in most cases still dominated by isotropically distributed background events. Until

---

<sup>1</sup>Typical weather phenomenon on the Canary Islands: Southeasterly, sand-laden wind that carries dust from North-African Sahara desert to the islands. The dust covers everything, including the telescope mirrors, which results in reduced reflectivity and affects the measured data.

now, only individual air showers were reconstructed and classified. To gain scientific results, all observed events have to be combined and global features, like their energy distribution, need to be investigated. If this is accomplished, we can additionally study the variations of these characteristics over time.

As astronomers, we are interested in the true astrophysical flux  $\Phi(E, t)$  of a single source at the sky. The term “flux” here means number of events per energy, area, and time, and is given in  $\text{TeV}^{-1} \text{cm}^{-2} \text{s}^{-1}$  for point sources and  $\text{TeV}^{-1} \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$  otherwise.

This true flux is hidden in the data we measure and can not be observed directly just from counting the events we classified as gamma ray before. Instead, it is convoluted with the detector response  $R$  of the telescope, meaning that a particle with a *true* energy  $E$  and position  $\mathbf{p}$  will be detected at another *estimated* position  $\hat{\mathbf{p}}$  and energy  $\hat{E}$  or might be not detected at all. The probability distributions for this estimated position and energy depend on the true features of a particle and detector characteristics.

The signal we observe  $N$  is given by

$$N(\hat{E}, \hat{\mathbf{p}}) = \int_{t_0}^{t_1} dt \int_{\Omega} \int_E \frac{d\Phi(E, \mathbf{p}, t)}{dE d\mathbf{p}} R(\hat{E}, \hat{\mathbf{p}}|E, \mathbf{p}, t) dE d\mathbf{p} + b(\hat{E}, \hat{\mathbf{p}}) \quad (5.7)$$

with the observed counts rate  $N$ , the true flux  $\Phi$ , the detector response  $R$  and the background  $b$ , that is added to the signal of every source.

The detector response  $R(\hat{E}, \hat{\mathbf{p}}|E, \mathbf{p}, t)$  serves as convolution kernel and maps the true energy and position to the estimated ones. In general, the estimators for energy and position could be correlated. But we ignore this fact when we split  $R$  into an effective area, an energy migration and a Point Spread Function (PSF) as described in the following:

- **Point Spread Function (PSF):**  $\text{PSF}(\hat{\mathbf{p}}|E, \mathbf{p})$  is the spatial probability distribution of the reconstructed source positions for events emitted by the same point source. For example, a Gaussian distribution can be used. The standard deviation  $\sigma_{\text{PSF}}$  of this distribution is also called *angular resolution*. As the name says, a PSF is by definition a function. For an IACT analysis, the function is integrated in spatial bins to define the probability to reconstruct an event with an offset  $r$  from the true source position. To visualize the PSF, the containment radius  $r_{\text{contain}}$  is used. It denotes a maximum distance between the true and the reconstructed source position containing a certain percentage of the simulated events, see [Figure 5.5](#).
- **Effective Area:**  $A_{\text{eff}}(E, \mathbf{p})$  accounts for the ratio of events that survive the analysis chain. This ratio varies with the event’s true energy, its offset from the FoV’s center and the pointing of the telescope (zenith distance<sup>2</sup>, azimuth). It can be translated into an *effective collection area*, which is energy- and offset-dependent, see [Figure 5.6](#). The effective area is given for every bin  $i$  in energy and offset by

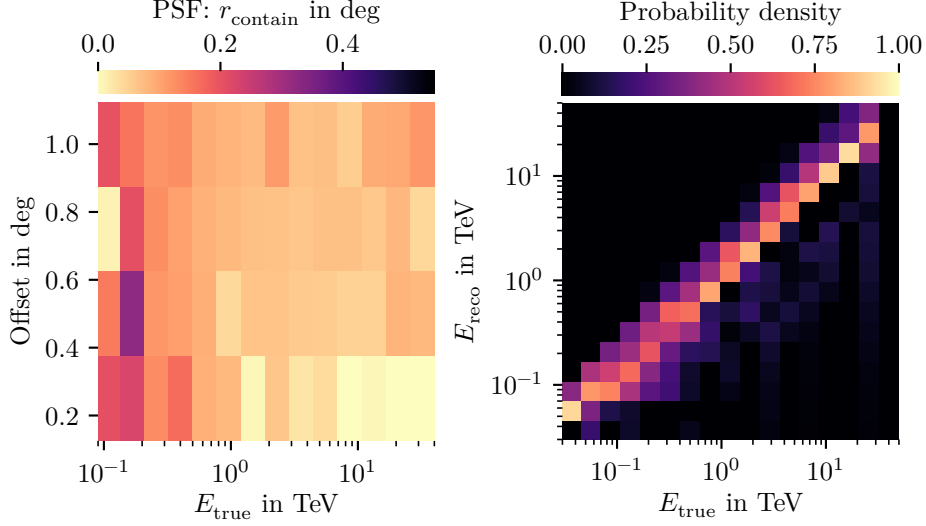
$$A_{\text{eff}, i} = \frac{N_{\text{detected}, i}}{N_{\text{simulated}, i}} \cdot 2\pi R_{\text{max}} \quad (5.8)$$

---

<sup>2</sup>The zenith distance is measured in degrees with respect to the point directly above the telescope (0°). The closer the pointing position comes to the horizon, the higher the zenith value will be. Typically, MAGIC observes at zenith ranges between 5° and 50°. Zenith values above 50° are technically possible, but the measurement quality and data reconstruction starts to get worse.

with the maximum impact radius  $R_{\max}$  of the simulated events.  $A_{\text{eff}}$  is always binned in the known simulated (*true*) energy.

- **Energy Dispersion:**  $E_{\text{disp}}(\hat{E}|E, \mathbf{p})$  describes the probability of a particle with true energy  $E$  to be reconstructed at  $\hat{E}$  (Figure 5.5). It is expressed as one-dimensional probability distribution. Similar to the PSF, the function is integrated in bins of the true energy.



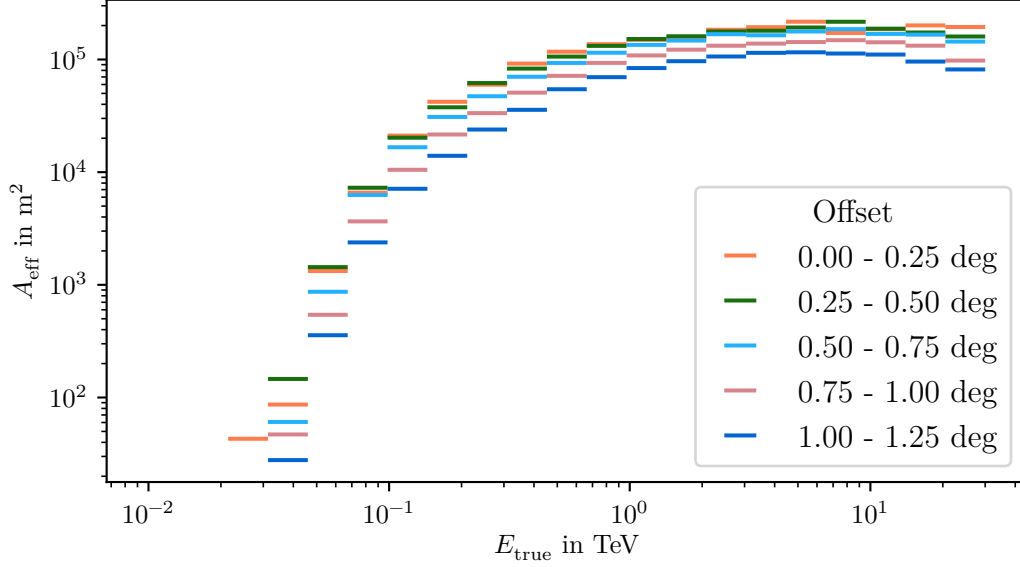
**Figure 5.5:** **Left:** PSF expressed by the 68% containment radius  $r_{\text{contain}}$ , which is the maximum distance between the true and the reconstructed source position which includes 68% of the reconstructed events.  $r_{\text{contain}}$  is given for five offset (distance from the FoV's center) bins and 20 bins in energy. **Right:** Energy migration matrix showing the probability of an event simulated with  $E_{\text{true}}$  to be reconstructed at  $E_{\text{reco}}$ . Both Instrument Response Functions (IRFs) are shown for MC simulations from production ST0311.

As we see from Equation (5.7), the detector's response changes over time. To deal with this dependence, the IRF calculation is done for small time intervals where the pointing does not change much and the detector response is assumed to be stable. For MAGIC this is the duration of one run, which is typically 5 to 20 minutes long. This is why the IRFs described above are not time-dependent.

Having all these functions defined, the detector response becomes

$$R(\hat{\mathbf{p}}, \hat{E}|\mathbf{p}, E) = A_{\text{eff}}(E, \mathbf{p}) \cdot \text{PSF}(\hat{\mathbf{p}}|E, \mathbf{p}) \cdot E_{\text{disp}}(\hat{E}|E, \mathbf{p}) . \quad (5.9)$$

Equation (5.9) describes the detector response at any point in the FoV, so it is valid for observing a possibly extended source at an unknown position. The data set used for this thesis only contains observations of point sources at a known position with a fixed offset from the center of the FoV, which is common for observing extra-galactic sources. If the



**Figure 5.6:** Energy-dependent  $A_{\text{eff}}$  for five different offsets. At low energies, the  $A_{\text{eff}}$  is restricted by the capability of the telescope’s hard- and software to trigger and reconstruct very faint and small air showers properly. At high energies, the air showers are very large and often not fully contained in the shower image, which also leads to a bad reconstruction.

source position and, therefore, its offset is known and the source is point-like, the spatial dependence  $\mathbf{p}$  can be dropped. In this case, we define an On region in the FoV around the known source position of certain size  $\theta$  and count the events within this region as  $N_{\text{on}}$ , further described in [section 5.6](#). The effective area must then be calculated for the same sized region. [Equation \(5.9\)](#) simplifies to

$$R(\hat{E}|E) = A_{\text{eff}}(E) \cdot E_{\text{disp}}(\hat{E}|E) . \quad (5.10)$$

[Equation \(5.7\)](#) finally becomes:

$$N(\hat{E}) = t_{\text{obs}} \int_E \frac{d\Phi(E)}{dE} \cdot A_{\text{eff}}(E) \cdot E_{\text{disp}}(\hat{E}|E) dE + b(\hat{E}) . \quad (5.11)$$

The only way to calculate the detector response  $R(\hat{E}|E)$  is from simulated data, where the true characteristics of the events are known. The pure IRF calculation therefore needs no measured data, but the MC simulations used to calculate the IRFs must have passed through the same analysis steps as the measured data. Furthermore, the MC simulations must fit the target data in terms of zenith angle, analysis period/MC production, and the amount of moonlight.

Since the simulated data used for the IRFs is classified by a RF which was also trained with simulated data, the same data set must not be used for both tasks! This is why the original MC simulations are split into two subsets: one for training the RFs and one for the

IRF calculation. From the statements above it is obvious that the quality of an IACT's data analysis and the reliability of the scientific results depend crucially on the quality of the MC simulations.

So far, all equations are formulated as analytical expressions. For a counting instrument like an IACT, the true flux, the detector response and the background are discretized and expressed as matrices with counts per interval in energy. Equation (5.11) becomes:

$$\mathbf{N} = \mathbf{\Phi} \cdot \mathbf{R} + \mathbf{b} . \quad (5.12)$$

## 5.5 The Inverse Problem

The naive approach to obtain the true flux directly by inverting Equation (5.12) unfortunately fails since inverting the detector response  $\mathbf{R}$  yields large statistical errors, which causes an ill-posed problem. There are, in general, many ways to overcome such an inverse problem. Which solution is appropriate depends on the individual circumstances. A common approach in gamma-ray astronomy is a likelihood maximization using a pre-defined model or a step function without any model assumptions.

To use these approaches, we first have to define a likelihood function. Since we deal with event-based measurements, the data comes as counts per time interval, area, and energy bin. We expect these counts to be Poisson-distributed around an unknown mean value  $\mu$ :

$$\mathcal{P}_\mu(k|\mu) = \frac{\mu^k}{k!} e^{-\mu} \quad (5.13)$$

with  $k$  the measured counts.

The likelihood is then defined as the product over all bins of the Poisson values given a certain model assumption. Since we defined our IRFs to be valid for the duration of one run, we have to calculate the Poisson value for every energy bin and run individually.

$$\mathcal{L}(N, b|\boldsymbol{\lambda}) = \prod_{j=1}^{n_{\text{runs}}} \prod_{k=1}^{n_{\hat{E}}} \mathcal{P}(N_{jk}|\mu_{s,jk} + \mu_{b,jk}) \times \mathcal{P}(b_{jk}|\mu_{b,jk}/\alpha) \quad (5.14)$$

with  $N_{jk}, b_{jk}$  the measured counts in the On and Off regions for the  $j$ -th run in the  $k$ -th bin of reconstructed energy,  $\mu_s, \mu_b$  the unknown true number of events for signal and background and  $\alpha$  the ratio of the observation time and area between the On and Off region. How these counts for the On and Off region are obtained is described in section 5.6.

In Equation (5.14), the *expected* background events are denoted as  $\mu_{b,jk}$  and the *expected* signal events depending on the model parameters  $\boldsymbol{\lambda}$  are given as:

$$\mu_{s,jk}(\hat{E}|\boldsymbol{\lambda}) = t_{\text{obs},j} \int_{\Delta\hat{E},k} d\hat{E} \int_E \frac{d\Phi(E|\boldsymbol{\lambda})}{dE} R(\hat{E}|E) dE \quad (5.15)$$

with the known detector response  $R(\hat{E}|E)$  from Equation (5.10).

To deal with  $\mu_b$ , the simplest approach is also assuming a poisson distribution which is done in the approach described above. Nevertheless, more approaches exist to model the background, see for example [141, 51].

The aim is now to maximize Equation (5.14) while varying the model parameters  $\lambda$ . This is equivalent to finding the model parameters that most likely produce the observed output given the detector response calculated from MC simulations. The absolute value of  $\mathcal{L}(N, b|\lambda)$  is of no interest, which is why the logarithm of Equation (5.14) can be used to simplify the equation and speed up the computation. Instead of maximizing the likelihood, the negative log-likelihood has to be minimized, which is a task that can be solved properly.

There are now multiple ways to define the model parameters  $\lambda$ . Here, we focus on two approaches:

First, *forward folding*, which optimizes the parameters of a given spectral model (e.g. power law, log-parabola). This model can be derived from known acceleration mechanisms (see subsection 3.1.1) or just be assumed by the experience of the analyzer, which can introduce a severe bias, of course! In turn, the number of free parameters is small.

Second, *unfolding*, which is free from any model assumptions. In this case, we treat  $\lambda$  as parameters of a step function and hence directly as the unknown true flux per bin of true energy, the so-called *flux points*. This approach comes with a higher number of free parameters, depending on the bin size, and might not be suitable if not enough events are available. Furthermore, the results might contain un-physical oscillations of the bins. To avoid such behavior, a so-called *regularization* term can be added to the likelihood, penalizing a large second derivative and smoothing the spectrum, see [162] for further reading.

The mentioned methods are by far not the only approaches to overcome the inverse problem. Depending on the conditions, various deconvolution theories and regularization methods exist.

### 5.5.1 Spectral Flux Modeling

In case of the forward folding, a common a-priori assumption for the spectral shape is the power law. The power-law shape is motivated by acceleration models, for example, by the First Order Fermi Mechanism [71, 29], which correctly predicts the shape of the Cosmic Rays (CR) spectrum. Nonetheless, there are various models to account for the behavior of different source types. All of them are based on a power-law function. In the following, I will introduce some common models that are relevant for this work.

The **power-law** has the form

$$\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^\gamma, \quad \lambda = (N_0, \gamma), E_0 = \text{fixed} \quad (5.16)$$

with the amplitude  $N_0$ , the reference energy  $E_0$ , and the index  $\gamma$ . The reference energy is fixed and not part of the fit parameters.

The **log-parabola** spectrum is derived from this simple power-law, which is typically used for modelling blazars:

$$\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^{-\left(\alpha + \beta \log\left(\frac{E}{E_0}\right)\right)}, \quad \lambda = (N_0, \alpha, \beta), E_0 = \text{fixed} . \quad (5.17)$$

Taking absorption, e. g. by the EBL, into account, an **exponential cutoff** above a certain cutoff energy  $E_c$  is modeled by

$$\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^\gamma e^{-\left(\frac{E}{E_c}\right)}, \quad \lambda = (N_0, \gamma, E_c), E_0 = \text{fixed} . \quad (5.18)$$

In some cases, a stronger cutoff is modeled by a **super-exponential cutoff** with

$$\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^\gamma e^{-\sqrt{E/E_c}}, \quad \lambda = (N_0, \gamma, E_c), E_0 = \text{fixed} . \quad (5.19)$$

Additionally to the mentioned ones, more models exist, but for this work only the log-parabola and the exponential cutoff model are relevant.

### 5.5.2 Calculating Flux Points and Light Curves

Additionally to the parameters of a spectral energy distribution, also flux points are a common high-level analysis result in gamma-ray astronomy. Flux points denote the flux in a certain energy bin and can be calculated either as *integral* flux points or *differential* flux points. In the first case, the differential flux (photons per energy, area and time) is calculated by fitting a model in a given energy bin. In the latter case, the flux point is calculated as integral over the differential flux in the edges of the energy bin (photons per area and time). To obtain the energy density, the differential flux is often multiplied with  $E^2$ .

Furthermore, the flux variations over time are of importance to understand the behavior of a source. In contrast to the calculations of flux points, the dataset is now split by observation time, for example, run-wise, night-wise or monthly binned. The total flux for every bin is computed by integrating the SED above a threshold energy  $E_{\text{threshold}}$ , which has to be identical for every time bin. With sufficient statistics in every time bin, the spectral parameters can be fitted separately for every bin. If the spectrum of the source is well-known or the number of events per time bin is too small to obtain a valid fit, a fixed spectrum with just the normalization as a free parameter is fitted to the data in that time bin.

The right choice of the time bin size depends on the general flux level of the source. For strong sources with high flux, shorter time bins are sufficient to obtain enough events to be able to get reliable fit results. For weak sources, more observation time is needed to collect enough events.

The minimization of the log-likelihood function results in an estimation for  $\mu_s$  per bin. “How good” this estimation is, depends on the amount of events that can be used to calculate it. For low statistics, when just a few events were counted in a bin, we might get a value from the minimization that is still not a proper estimation for  $\mu_s$ . To communicate these uncertain estimations, usually *upper limits* are plotted for these bins instead of flux points.



To decide which flux points should be represented by their upper limits, the Test Statistics (TS) value is used. The TS value describes the absolute difference between the fit statistic value for two hypotheses  $H_0$  and  $H_1$ . In our case,  $H_1$  means that we measured  $\mu_s > 0$  signal counts,  $H_0$  stands for no signal counts. Hence, the negative log-likelihood function is evaluated at the minimum for  $H_1$ , and at the point, which results in zero signal events. If this difference is smaller than a certain (user-defined) threshold, an upper limit is evaluated instead of the flux point that would result from the estimation for  $\mu_s$ . The TS value is connected to the Li&Ma significance via its square root. For example, if we want to use upper limits for flux points with a Li&Ma significance less than  $3\sigma$ , we choose all flux points with  $\sqrt{\text{TS}} < 3$ .

After the decision is made *if* an upper limit is plotted, we have to calculate its absolute value. To derive this value, we have to choose the point left (above) the minimum position where the negative log-likelihood increased by a value of TS. Which TS value to choose is again a decision of the analyzer. A common value is  $2\sigma$ , which means an absolute increase of 4.

The uncertainties for a flux point are calculated in the same way. In that case, a lower limit is calculated by searching for the value *below* the minimum position where the fit statistic increased by a certain value. For flux point uncertainties, usually a TS value of 1 is used, which relates to a Li&Ma significance of  $1\sigma$ . For further reading, see e.g. [141].

## 5.6 On/Off Regions

To solve the inverse problem described in the last section, we need measured counts for  $\mathbf{N}$  and  $\mathbf{b}$ . Recall that we end up after section 5.3 with a list of gamma-like events with reconstructed energy and origin. These events are spread over the FoV for strong sources with a peak around the source position. Most of the observed area contains just background events initiated by CRs, diffuse gamma-ray emission, electrons, or just false reconstruction. This continuous background has to be taken into account when claiming any statement about the observed source.

Around the known source position an *On* region is defined of a size that is optimized in a way that most of the signal is included but as little background events as possible. The radius of the region is called  $\theta$ . Analog to the On region, *Off* regions of the same size and distance to the center of the FoV are defined. The amount of events in the On region is a superposition of the source flux and the continuous background in the Off region:

$$N_{\text{on}} = N_{\text{signal}} + N_{\text{off}} \quad (5.20)$$

where  $N_{\text{signal}} = N_{jk}$  and  $N_{\text{off}} = b_{jk}$  from Equation (5.14). Usually, more than one region is used to average the  $N_{\text{off}}$  and reduce statistical uncertainties.

### 5.6.1 Li&Ma Significance

The first thing, an astronomer would want to know before doing any advanced studies, is if the source could be detected at the assumed position at all. Since an observation would always contain background events, we have to quantify if the amount of On events is

significantly higher than the amount of Off events. This is calculated by the Li and Ma [108] likelihood-ratio test:

$$S = \sqrt{2} \left[ N_{\text{on}} \cdot \ln \left( \frac{(1 + \alpha)N_{\text{on}}}{\alpha(N_{\text{on}} + N_{\text{off}})} \right) + N_{\text{off}} \cdot \left( \frac{(1 + \alpha)N_{\text{off}}}{N_{\text{on}} + N_{\text{off}}} \right) \right]^{\frac{1}{2}} \quad (5.21)$$

where  $N_{\text{on}}$  denotes all events measured in the On region,  $N_{\text{off}}$  denotes all events measured in the Off region, and  $\alpha$  denotes the fraction of the observing times or the ratio of the numbers of On and Off regions. In case of one On and one Off region which are of the same size and were observed over the same time,  $\alpha$  would be 1.

$S$  is the number of standard deviations denoting the deviation from the null-hypothesis, which is “all events are background events”. In the gamma-ray community, an observation resulting in  $S \geq 3\sigma$  is denoted as “hint”, whereas  $S \geq 5\sigma$  is denoted as “detection” of a source.

### 5.6.2 $\theta^2$ Cut

To define the size of the On/Off regions, usually a radius  $\theta$  is chosen that maximizes the significance states in Equation (5.21). Figure 5.7 shows the gamma-like events binned by their squared distance from the On and Off regions’ centers. Many events at small distances to the source position indicate that the source was detected at the expected position.

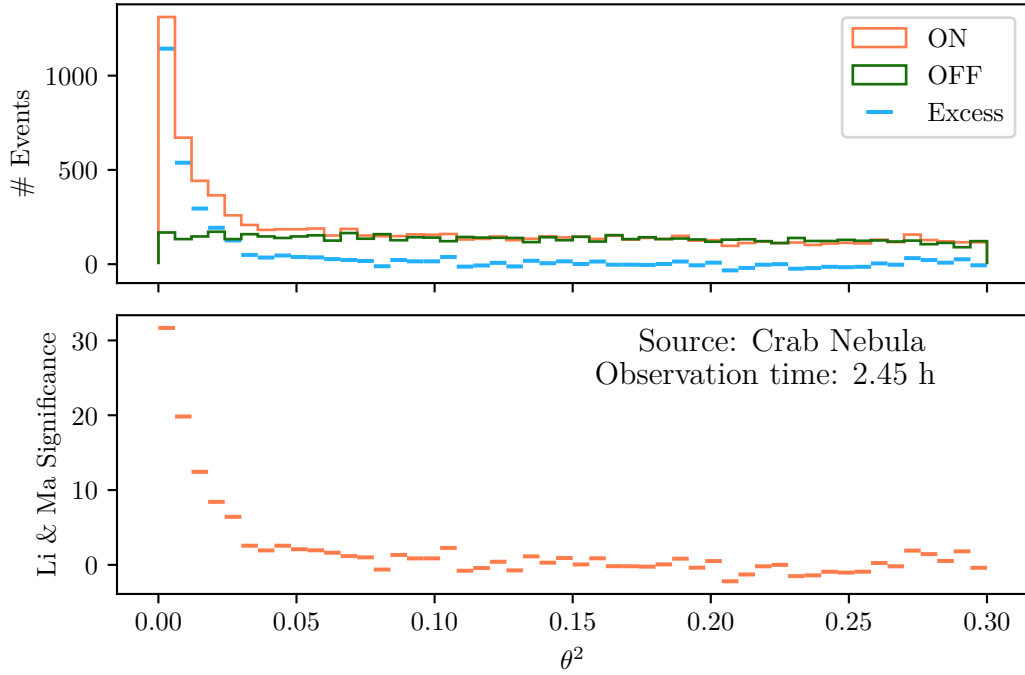
Applying a cut in  $\theta^2$  and rejecting every event outside the defined region drops the spatial dependency in Equation (5.7). The same cut must then be applied in the calculation of the effective area to obtain comparable counts.

To improve the sensitivity of the analysis, the size of the On region can be calculated separately for different bins of reconstructed energy  $\hat{E}$ , because the  $\theta^2$  distribution varies at different energies. For very low energies, only a few events were detected and survived the hadronness cut. A cut at small  $\theta^2$  values would exclude most of the events from the On region. Whereas, at higher energies, a “loose cut” at higher values would make the On region unnecessarily large and include too many background events.

### 5.6.3 Observation Mode

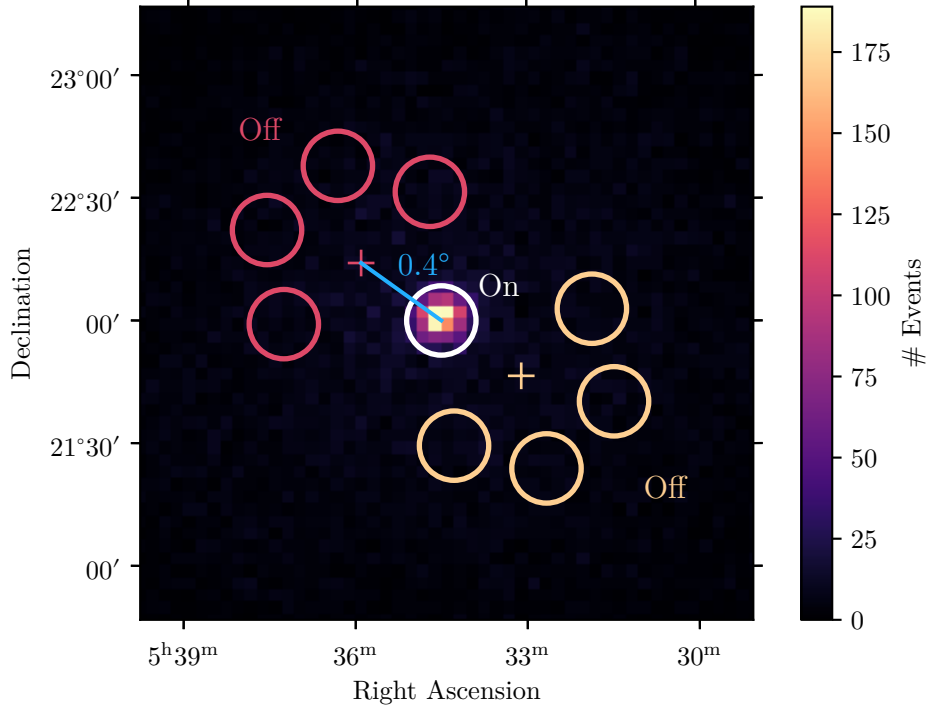
There are two techniques to observe and count events in the On and Off region. The first and intuitive one is pointing the telescope to the source region for a specific time to count the On events and after that, pointing to a region without any known source to count the Off events. Two problems arise for this method: First, a part of the already very limited observation time is spent on Off observations and can not be used for the actual target source. Second, two different time slots can result in different observing conditions, for example, through changing weather or rising Moon. The counts for  $N_{\text{on}}$  and  $N_{\text{off}}$  would not be comparable anymore.

The second observation technique is called *wobble mode* and allows for measuring  $N_{\text{on}}$  and  $N_{\text{off}}$  at the same time, which solves both problems. To achieve this, the telescopes are pointed not directly at the source but at a position with typically  $0.4^\circ$  offset, see Figure 5.8. This way, there are several other areas in the FoV with the same distance to its center and



**Figure 5.7:** Upper panel:  $\theta^2$  plot for 2.45 h of Crab Nebula observation on 2016-10-12. “ON” denotes the number of events with a squared distance of  $\leq \theta^2$  from the expected source position, “OFF” denotes the number of events with a distance of  $\leq \theta^2$  from the Off position, which is placed at the same distance to the FoV’s center where no known sources are located. If more than one Off position is used, the OFF value is weighted with  $\alpha = \frac{1}{\# \text{ Off positions}}$ . “Excess” is the difference between  $N_{\text{on}}$  and  $\alpha N_{\text{off}}$ . Lower panel: Li & Ma significance depending on the squared On and Off region radius  $\theta^2$ .

the same detector acceptance as the On position. These areas can be used as Off positions. Usually, a source is observed at two to four different wobble positions with the same offset. All data used for this work was measured in wobble mode.



**Figure 5.8:** Counts map of Crab Nebula observations from 2016-10-12 with overlaid On and Off regions from two runs. The crosses mark the pointing positions for two wobble runs. On and Off regions have the same distance of  $0.4^\circ$  from the pointing position. For bright sources as the Crab Nebula, Off regions close to On region are avoided because they can contain badly reconstructed gamma-ray events from the source.

## 5.7 Implementation of the Analysis Chain

The analysis chain described above is realized in MARS, a software package created for the data analysis for MAGIC [16, 179]. MARS is written in C++ using ROOT [38, 39] and takes care of all analysis tasks from the raw data calibration up to producing the final scientific results like spectra and light curves. All analysis results of the different analysis steps are stored in the tree-like file structure of the ROOT framework, a C++-based open-source software package developed by CERN [38]. An overview of the most important MARS packages and their tasks and relations is given in Figure 5.9 and the individual executables are described in the following.

- **Sorcerer:** Calibration, quality checks, and spike removal for raw DRS4 data of each

telescope. The output are *calibrated files*, which are needed as input for **Star**. **Sorcerer** can be used for measured and simulated data. For measured data, the raw files have to be combined with the corresponding report files first, which is done using **Merpp**. Report files contain e.g. weather data and hardware reports which do not exist for simulated data, of course. **Sorcerer**'s in- and output are provided at subrun-level, which means one file per subrun and telescope.

- **Star**: Image cleaning and Hillas parameterization (see [section 5.2](#)) for each telescope and subrun. For data measured under moonlight, the training data for the RF is modified, since a RF trained with data taken at dark nights would not perform well on data recorded under moonlight. Therefore, additional noise has to be added to every pixel in the camera image to simulate the moonlight before the image is cleaned and the Hillas parameters are calculated. According to the moonlight, the cleaning thresholds have to be adjusted.
- **Superstar**: Run-wise merging and stereoscopic reconstruction. The individual parameter sets received from the two MAGIC telescopes are merged. Every event is now mapped to a single parameter set. To generate a more comfortable data set, the single subrun files are merged into one file per run.
- **Selectmc**: For MC data, no subruns exist and therefore no merging is performed for simulated data by **Superstar**. This task is completed by **Selectmc**, which merges all (usually roughly 5000) files and splits them into two subsets for training the RF and calculating the IRFs. In general, **Selectmc** is able to perform several tasks to fit the MC sets to the data, but for this work, only the splitting and merging is used.
- **Coach**: Trains the RFs with **Superstar** output files from hadronic (Off-) data and simulated gamma-ray data. RFs are trained for gamma/hadron separation and **disp**-estimation (see [section 5.2](#)). The energy reconstruction is done via Look-Up Tables (LUTs), which are generated by binning the simulated events in the parameter **size** and in  $R_{M1}/r_{\text{cher}}$ , with the impact parameter for the telescope M1  $R_{M1}$  and the Cherenkov radius  $r_{\text{Cher}}$ . The estimated energy can then be “looked-up” for both (M1 & M2) parameter sets of an unlabeled event. The final  $E_{\text{est}}$  is calculated as the weighted average of both values. Optionally, a RF can be used in **Coach** for the energy estimation, too. For the gamma/hadron separation, Off data is needed while the **disp** and energy estimation are performed on the simulated data only.
- **Melibe**a: Applying the RF and LUT for unlabeled data (*On data*) to create a hadronness value, energy estimation and **disp** for every event. Based on the hadronness value, an event can be classified as gamma- or hadron-induced. Furthermore, **Melibe**a calculates the reconstructed source position as described in [subsection 5.3.2](#). All the following executables use **Melibe**a output files as input.
- **Odie**: Producing  $\theta^2$ -Plots (see [subsection 5.6.2](#)) and define PSF. From the  $\theta^2$ -Plots, the Li&Ma significance is calculated (cf. [Equation \(5.21\)](#)).
- **Caspar**: Creating skymaps to search for a source and define its expansion. Additionally, TS value distributions are calculated to indicate how likely the observed flux is different

from the Null hypothesis of only background events.

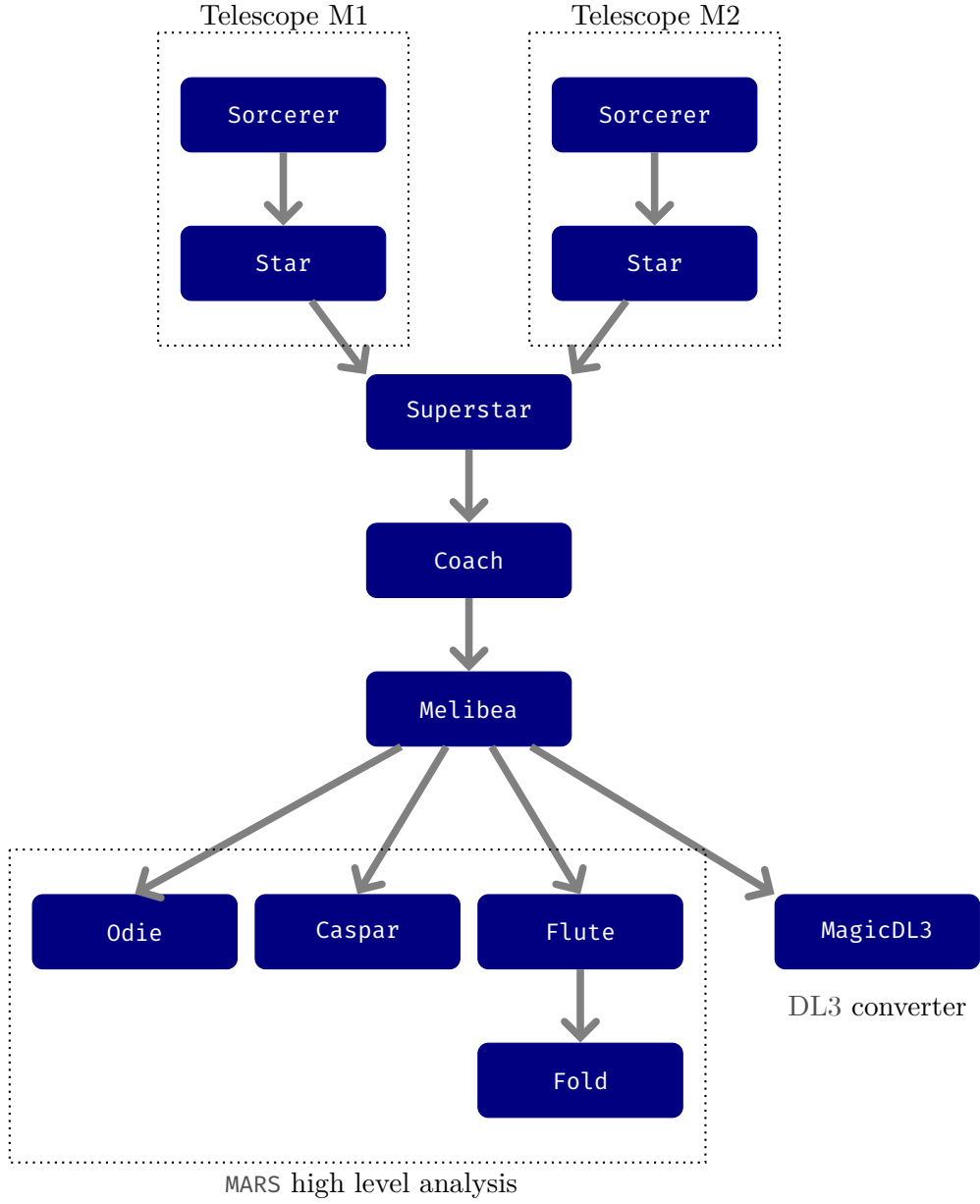
- **Flute**: Performs the background estimation from On and Off regions and calculates IRFs as explained in [section 5.4](#). Furthermore, **Flute** produces flux points and a light curve using a pre-defined energy spectrum.
- **Fold**: Performs the likelihood optimization (forward-folding, see [section 5.5](#)) and creates the spectral energy distribution depending on the true energy.
- **MagicDL3**: The newest member of the MARS family: Converts **Melibeia** output files into Data Level 3 (DL3) files, a new standard format for gamma-ray astronomy, see [chapter 6](#), and computes the IRFs in the same manner as **Flute**.

All MARS executables are used as command-line tools. For every analysis step, a user has to set some parameters (e. g. paths to the data files or cuts on the used events) via the command line or a configuration file and call the executable. After one step is finished, the next program can be executed with the output from the last step as input. The whole workflow is optimized for processing a small data set and leaves a lot of options for quality checks by the user, which is mainly done by inspecting plots generated by **ROOT**. For analyzing data of several nights without moonlight, this approach is suitable since the competent MARS user has control over every step and can modify the analysis anytime.

For a less competent user or an observing time range of several years, this approach has its challenges<sup>3</sup>. This is the reason, why the **autoMAGIC** project was initiated, which is described further in [chapter 8](#).

---

<sup>3</sup>Which is actually an understatement.



**Figure 5.9:** Overview of the MARS executables and their connection. Up to **Melibea**, measured data and MC simulations have to go through the same analysis chain. For this work, **Selectmc** is only used for simulated data to merge and split all files or a more convenient data handling. For measured data, this task is done by **Superstar**.





## Data Formats for Gamma-Ray Astronomy

---

The implementation of the analysis chain described in [section 5.7](#) is custom-designed for the MAGIC telescopes. The software packages cannot be used out of the box for data analysis for other experiments. The data products, respectively the `MeLibea` output files, can not be used as input for any other software to perform the high-level analysis steps.

The MARS software is not an exception: all operating gamma-ray telescopes have their tailor-made analysis software and data formats. This circumstance makes it rather impossible to combine data from different facilities. First, a scientist needs access to both data and telescopes. Second, a scientific user has to be capable of understanding and applying the software products of the respective collaboration, which makes combined and reproducible analyses even more difficult.

Because of the effort that has to be undertaken to obtain high-level data like light curves and energy spectra, combined studies or studies of very large data sets are still very rare. Another issue is the common habit of gamma-ray observatories to restrict the access to the data to members of the respective collaboration, which has political and historical reasons.

The upcoming CTA will be the first observatory that will provide open access to their data. For this reason, a data format is needed which on the one hand provides reduced and easy-to-handle data, but on the other hand, provide all necessary information for users to customize them for a special research project. Such a general data format additionally enables scientists from different groups to combine and compare their data and results.

In 2015 the Gamma Astro Data Format (GADF) project was started with the scope to develop a data format which can be used by all gamma-ray facilities [\[55\]](#).



**Documentation: Data Formats for Gamma-Ray Astronomy**  
<https://gamma-astro-data-formats.readthedocs.io/en/latest/index.html>

This project today defines a standard on how to store event-based data at the so-called DL3<sup>1</sup> and standardize the handling of time and coordinates. The DL3 represents lists of events with their reconstructed origin, energy, and arrival time.

Besides the DL3 event lists, the Data Format for Gamma-Ray Astronomy contains the Good Time Intervals (GTI)<sup>2</sup>, the pointing information and all IRFs to enable a user to calculate spectra and light curves as described in [section 5.4](#). The IRFs that can be stored are:

---

<sup>1</sup>The declaration of different data levels is mainly driven by *Fermi*-LAT and CTA. The higher the data level, the more reconstructed and aggregated the data.

<sup>2</sup>The time interval when the telescope is observing, denoted as the time between TSTART and TSTOP. The dead time of the detector is not considered here.

- Effective area  $A_{\text{eff}}$
- Energy dispersion (migration)
- PSF
- Background models<sup>3</sup>

Equipped with this information, a scientist can make their own high-level analysis, no matter which telescope recorded the data. All the data is stored in Flexible Image Transport System (FITS) files, a simple, very common, open data format, which was developed by National Aeronautics and Space Administration (NASA) in the 1980s. More information on FITS can be found in the [FITS standard](#).

### 6.1 MAGIC Legacy

All data MAGIC recorded since the beginning was processed with the proprietary software MARS and stored as ROOT files, see [section 5.7](#). To preserve the data for future generations of scientists and to move forward in the spirit of open gamma-ray astronomy, MAGIC decided to transfer all its data at DL3 into the open GADF described above. As a first act, a joint analysis of the Crab Nebula, a standard source regularly observed by all gamma-ray telescopes, was performed with data from *Fermi*-LAT, MAGIC, the Very Energetic Radiation Imaging Telescope Array System (VERITAS), the First G-APD Cherenkov Telescope (FACT) and the High Energy Stereoscopic System (H.E.S.S.) [[123](#)]. For further reading, see also [[124](#), [37](#), [125](#)].

In the next years, massive data processing will be necessary to transform all MAGIC data into the GADF. This is the second reason why the data analysis pipeline `autoMAGIC` is developed as part of this work. It enables us to automatically perform the MAGIC data analysis chain up to `Melibe`a and convert the `Melibe`a output (ROOT files) into FITS files. The `autoMAGIC` project is described in [chapter 8](#).

---

<sup>3</sup>Background models are usually not used for the standard MAGIC analysis since the background is estimated as described in [section 5.6](#) via On and Off regions.




## Gammapy

---

Together with the GADF, software is needed that can be used to work with this data format. In our case, this task is accomplished by Gammapy [55, 54, 123], which is an actively developed and widely accepted tool for modern data analysis in gamma-ray astronomy. In 2021, Gammapy was also introduced as the official science tool for the upcoming CTA and H.E.S.S.

Gammapy is written in Python and mainly built on numpy [91], scipy [168] and astropy [24, 25]. Starting around 2015, it was intended to be used for the H.E.S.S. galactic plane survey [1], but quickly became a more general tool for other experiments like MAGIC, CTA, *Fermi*-LAT and VERITAS. The software package is community-developed with the aim to involve active scientists from different working groups and provide user-friendly high-level analysis tools for all gamma-ray astronomers. Together with the GADF, Gammapy is a strong leverage factor for open and reproducible gamma-ray astronomy.

Gammapy works with DL3 data as input and therefore provides high-level analysis tasks like likelihood-fitting, modeling, time analysis, and (IRF-)simulation. In contrast to the MARS framework, it is open-source, easier to handle, and scriptable for large analyses while keeping the benefits of plotting-tools and an easy-to-access and easy-to-modify analysis chain.

 Gammapy Documentation: <https://gammapy.org>  
 Gammapy Repository: <https://github.com/gammapy/gammapy>  
And most importantly:  
 Gammapy Song: [https://gammapy.org/gammapy\\_song.mp3](https://gammapy.org/gammapy_song.mp3)



## 8.1 History and Goal

The autoMAGIC project was initiated in 2020 by Simone Mender, Cosimo Nigro, and me for the noble goal to nothing less than automatize the whole MAGIC analysis chain. Well, the original goal was more driven by pure laziness after I estimated the effort of manually creating a light curve of NGC 1275 with all available MAGIC data. The data of NGC 1275 was taken over several years, zenith ranges, and partly during moonlight. As mentioned in [subsection 5.3.4](#) and [section 5.4](#), it is necessary to train an individual random forest and compute IRFs for each combination of the different parameters, since the quality of the gamma-hadron separation crucially depends on the strong similarity of the training data with the target data. Doing all these analysis steps by hand would have meant a huge effort and a lot of redundant tasks. All in all, I assume analyzing all this data manually would have taken about one year of tedious and hardly reproducible work. Therefore, we started to automatize some tasks.

Additionally, we learned about the “MAGIC Legacy” project, which gave our initial idea a greater meaning. The aim is now to implement an automated analysis, suitable for all kinds of sources and projects, which reduces the needed user interaction to a minimum and generates reproducible results. The future goal is to produce DL3 FITS files for all MAGIC data fully automatically. For a dedicated analysis, a scientist would just send a data request for given meta parameters like source name and time interval of interest and directly receives the DL3 data in the open-source FITS format. This procedure would completely release a scientist from the burden of low-level data analysis with MARS and enable for the first time easy access to long-term studies.

There is another trivial but very time-consuming task autoMAGIC is going to solve: The Off-data search. As written in [subsection 5.3.1](#) and [5.3.3](#), a data set containing only hadronic showers is needed to train a random forest. For the MAGIC data analysis, real observations without any gamma-ray signal are used. The data used as training set must contain only hadronic showers but, furthermore, be recorded under good weather conditions. Bad weather like a cloudy sky and strong wind affects the data and reduces the similarity to the simulated gamma-ray data, which contain no weather effects. A RF would rather separate the simulated good weather from the measured bad weather than separate gamma- and hadron-induced showers.

The selection of suitable Off data is barely automatized, which is another crucial task for autoMAGIC. There is no reason for doing this task by hand since the metadata containing weather information is stored in every file. Additionally, a first estimate for the significance of the gamma-ray signal for every observation is calculated by the MAGIC Online Analysis

(MOLA), a preliminary analysis that is done on the island directly after the data was recorded. Based on these values, an observation is automatically classified as potential Off data, and no need for human interaction is given in most cases<sup>1</sup>.

Summing up, *autoMAGIC* has the following goals:

- Automatize redundant tasks and Off data search
- Reduce human interaction
- Generate reproducible results
- Enable long-term studies
- Make data taken under moonlight conditions accessible
- Lower the skill-level necessary to do science with MAGIC data
- Deliver the framework for the MAGIC Legacy project
- Take the step to open and reproducible gamma-ray astronomy with MAGIC
- Process the data of NGC 1275 for the science case of this thesis

## 8.2 Analysis Flow

*autoMAGIC* is an automated analysis pipeline to process from *calibrated* data (output of *sourcerer*) to DL3 FITS files. The aim is to perform the low-level standard analysis and provide a user with an output they can use for further high-level analysis. To be able to compare the new approach of DL3 files and their results with the traditional MARS approach, *autoMAGIC* currently produces two output formats: GADF DL3 files for further analysis with *Gammapy* and *Melibe* output files for high-level analysis with *Caspar*, *Odie* and *Flute*.

*autoMAGIC* takes care of the Off data search, Off- and target data processing as well as MC processing. All steps that have to be performed on the target data must be applied for the simulated data in the same manner to be able to produce IRFs, as described in [section 5.4](#). If I describe a specific analysis step in the following, keep in mind that this step is done for measured and simulated data, mostly.

Starting with the calibrated data, *Star* is executed to apply the image cleaning and Hillas parameterization. In case of a moon data analysis, randomly generated noise has to be added to the Off and simulated data to account for the higher NSB level in the target data.

The *Star* output for both telescopes is merged in *Superstar* into a single parameter set per event. For the MC data, an additional step is necessary to merge all *Superstar* output files together and split them into a training set for the RF and a second set to compute the IRFs. This is done with the MARS executable *Selectmc*.

After Off and MC data are prepared, the RFs can be trained and LUT can be created with *Coach*. The trained RFs and LUT are then used to classify the target events, estimate their

---

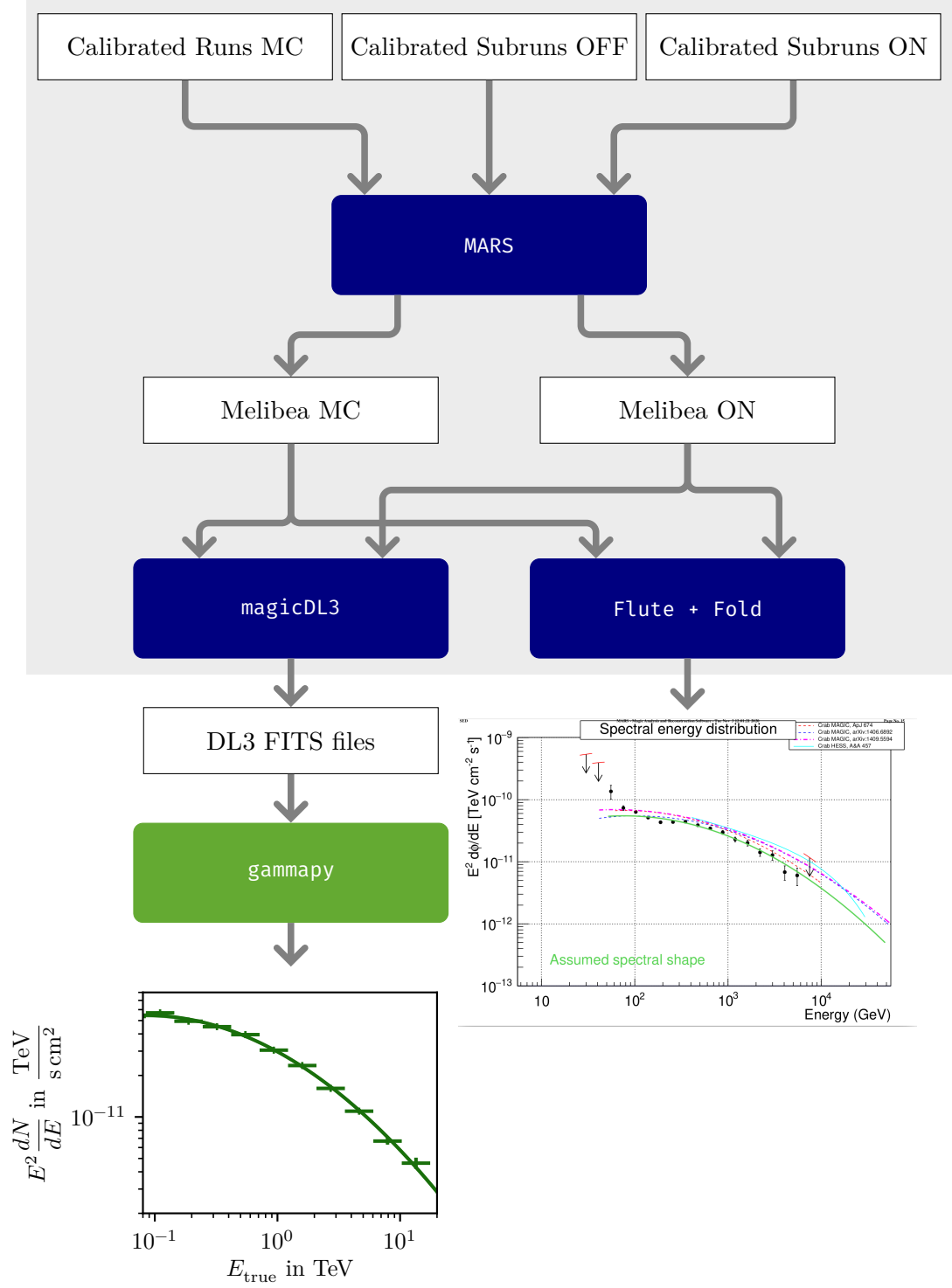
<sup>1</sup>It may occur for some special cases that weather conditions and significances are not properly stored in the date because of temporarily broken hardware or software problems. If a situation comes up where not enough Off data can be selected automatically, a human analyzer needs to have a look at the data.

energy and reconstruct the source position by using `Melibe`. Additionally, the second MC set is classified as well to be prepared for the calculation of the IRFs, see [section 5.4](#).

A user is now able to perform a high-level analysis with the `Melibe` output and the proprietary `MARS` executables. If the data is to be analyzed with `Gamm`, the conversion into the GADF is necessary. This is done by the DL3 converter `magicDL3`. `magicDL3` uses `Melibe` output files and the second MC set (that was not used to train the RF), to compute event lists and IRFs as required by the GADF standard. With these files, a user is equipped to create energy spectra and light curves with `Gamm`.

Especially for the task of comparing `MARS` and `Gamm` in [section 8.4](#), also `Flute` and `Fold` are implemented into `autoMAGIC`. `Flute` takes care of producing a light curve while `Fold` is responsible for fitting the SED, see [section 5.7](#).

This workflow is shown in [Figure 8.1](#).



**Figure 8.1:** Workflow of *autoMAGIC* (grey shaded area) and *Gammapy*. Calibrated files are used as input for the MARS block, which performs *Star*, *Superstar*, *Coach* and *Melibea*. The *Melibea* output can be used as input for the high-level analysis tools provided by MARS or they are converted to GADF files for high-level analysis



## 8.3 Technical Prerequisites

The calibrated data are provided by the MAGIC data center, which is located at the Port d'Informació Científica (PIC) in Barcelona. The PIC is also the location of the computer cluster which is used to run the autoMAGIC pipeline and store the database. The code base is stored in a GitLab repository also placed at the PIC.

autoMAGIC does not implement the analysis chain itself, for example, the Hillas calculation, but uses Python wrappers to call the individual executables. Calling an executable, for example *Star*, with a specific configuration and input data set, can be understood as one small work package as part of the whole analysis. All these work packages are executed separately and partly in parallel if the individual tasks have no interdependencies. This is done via jobs, which run independently and in parallel on multiple nodes of a cluster, one job for each work package.

The files that are needed as input and are generated as output are stored on the PIC file system and copied to the node (input) or to the file system (output) by the job. Furthermore, each job produces a logfile, which is also stored in the file system. The logfile provides insights into the process executed by a job and, in case of an error, gives hints about what went wrong.

Since access to the file system at the PIC and the cluster is needed to perform an analysis with autoMAGIC, a user can not perform an automated analysis on their local computer without access to the PIC. In the future, it is planned that a scientist can give their request to autoMAGIC (e.g. via a web form) and gets the DL3 output without the need for a PIC account or having any knowledge about autoMAGIC or MARS.

In the following subsections, I will give a brief introduction about the general idea how to use a relational database for automagic. Further information regarding the structure of this database and the job submission with HTCondor [160] are given in appendix A

### 8.3.1 Job Management

For a complete analysis, several ten-thousand jobs have to be processed, depending on how many different observation conditions are covered by the target data. One MC set consists of 5000 files, stored in one ROOT file each.

So many jobs and their interdependencies can not be coordinated manually. This is where a relational database comes into play. For all job types (e.g. *Coach*, *Melibeia*, *magicDL3*), a table exists, and each job is mapped to one row in that table. This way, all the necessary parameters, the job's status and the dependencies can be monitored. For example, if all *Superstar* jobs are finished, the *Coach* job can start. Furthermore, if one job fails and ends in state *error*, all jobs depending on its output are also set so state *error* without being submitted. For example, no *Melibeia* job will be submitted, if the job generating the input *Superstar* file ended in state *error*. The database keeps all this information and enables us to control and manage all small tasks that have to be done for a whole analysis.

### 8.3.2 Data Management

Besides monitoring the jobs, the database is needed to store all information about the measured and simulated data and make them accessible. For example, we want to select all observations from a specific source in a given time range, which is, in principle, the first question the database has to answer. Without a database, this question can only be answered by looking into the file system and rely on the file names and paths, which should include date and source name. The MAGIC file system is built that way, so it is generally possible (and the common way for manual analyses) to select files by their name. But if a user wants to access information about the observation condition, for example, the zenith range or the amount of moonlight in the data, every single ROOT file has to be opened and checked for the information stored there. In *autoMAGIC* all subruns are opened one time initially, and all values that could be useful for any analysis task are stored in the database. The same happens for MC runs.

### 8.3.3 Analysis-Related Information

Jobs and observations build two main components of the database, the third is given by all analysis-related information that has to be stored somewhere. This is, for example, the beginning and end of the analysis periods, which is the time range where a specific MC production should be used to train a RF. Another important analysis-related information is the threshold and cleaning level for the Moon data analysis or the used MARS version. Based on this analysis-related information, which is mostly stored in the MAGIC internal Wiki or some peoples' heads, a competent human analyzer would customize the analysis. But exactly these highly customized analyses are hard to reproduce in most cases. For an automated analysis, all this information has to be stored in the database to ensure a customized analysis *and* reproducible results.

### 8.3.4 Special Remarks: Coach

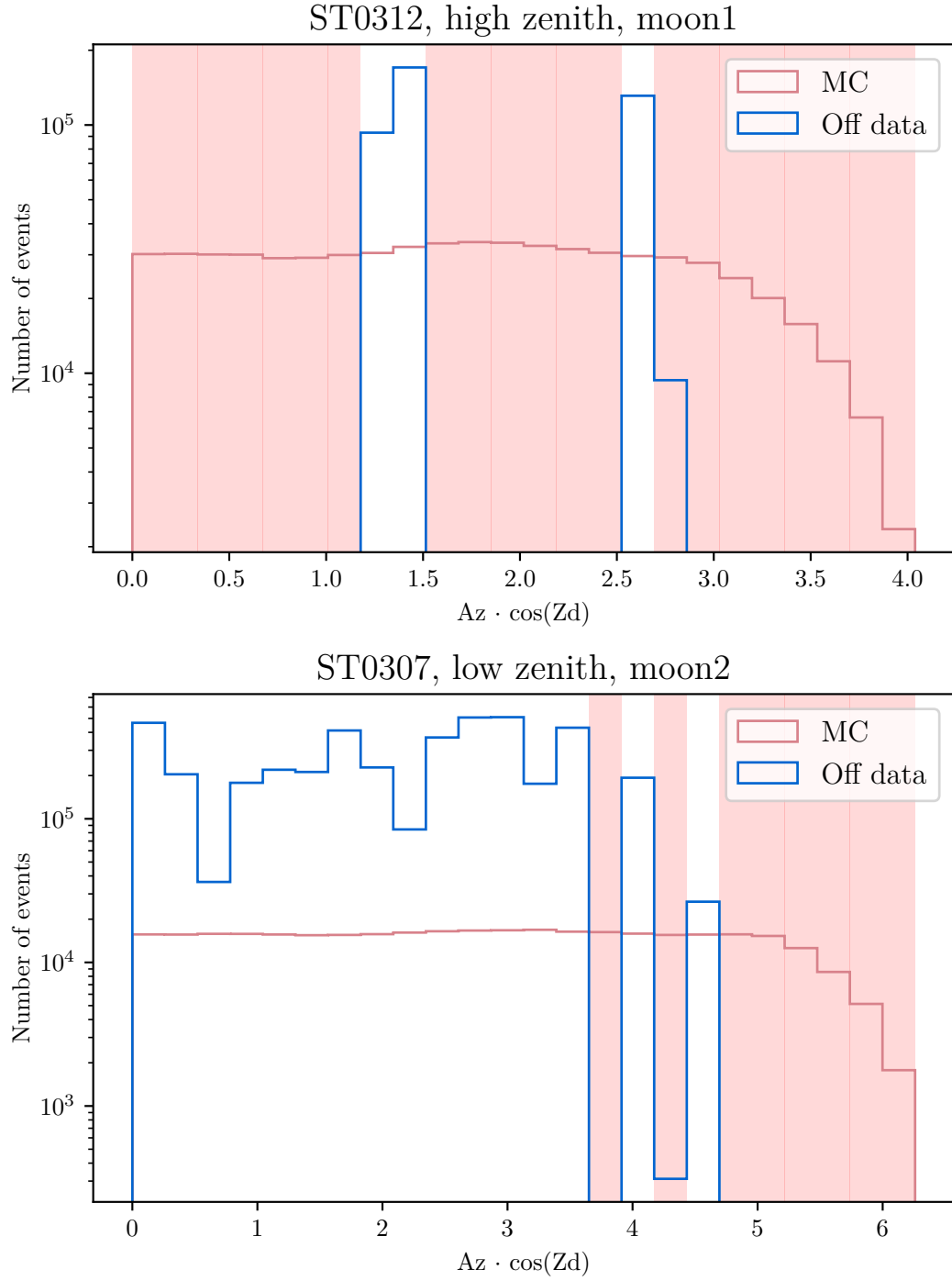
The Coach step (training the RFs and creating the LUTs) is the core piece of the analysis and the most complex part. Selecting the training data so that they cover a certain parameter space which fits to the target data is crucial for a good gamma-hadron separation. Moreover, the amount of gamma-ray and hadron-induced showers should be roughly equal or with a benefit for the Off data. As mentioned, zenith angle, moon range (which is quantified by the Direct Current (DC) value of the M1 telescope) and MC production of the training have to fit the target data.

Additionally, the impact of the azimuth angle must not be neglected. Depending on the azimuth angle, Earth's magnetic field changes, which has an impact on the air shower which is generated by charged particles (c.f. [section 4.1](#)). Imagine a sphere with the telescope at its center. The greater the zenith angle (measured from the pole directly above the telescope), the greater the area that is covered by the azimuth angle, which is why the azimuth angle becomes more important for large zenith angles. To take this dependency into account, the training is chosen so that  $Az \cdot \cos(Zd)$  is covered uniformly by Off data and MC simulations.

For the simulated gamma rays, this aspect is fulfilled by design, whereas for the Off data, the user (or `autoMAGIC` in this case) has to take care of this.

As one can imagine, for data measured at low to medium zenith, this constrain is relatively easy to fulfill, whereas for high zenith data, the  $Az \cdot \cos(Zd)$  range is often covered very poorly by the Off data, which affects the performance of the RF. For `autoMAGIC`, I implemented the Off data selection in a way that, first, a good  $Az \cdot \cos(Zd)$  coverage is pursued and, second, the total number of Off data events is equal or greater than the number simulated events. If these constraints are not fulfilled, the user can decide whether the RF should still be processed. Unfortunately, underpopulated Off data sets are quite common for the high zenith range (above  $50^\circ$ ) because there is not enough measured Off data available. For a quick quality check by the user, `autoMAGIC` can generate plots with the binned training data, see [Figure 8.2](#).

Another important task is to ensure that the target source data is never used to train the RF. Since we want to re-use the RFs that are already produced for other analyses, we have to track the sources that are used as Off data. To solve this issue, always two RFs are trained for the same parameter set with disjunct source sets. If another source is to be analyzed with the same observation conditions, the dedicated RF is used, which does not contain the target data in its source set. This approach saves computation time because `Coach` has to be executed just two times for given observation conditions and not for each new source individually. In turn, the amount of Off data, which is partly very limited anyhow, is reduced further. To distribute the amount of data roughly equally, the data sets for the respective sources are ordered by their number of events and are allocated in alternating order.



**Figure 8.2:** Quality check plot for Coach. The number of Off events should be equal to or greater than the number of simulated events in every  $Az \cdot \cos(Zd)$  bin. Especially for high zenith this is not fulfilled in most cases. Underpopulated bins are shaded in red. “moon1” denotes DC values of the telescope M1 in the range of 2.2 mA to 3.3 mA, “moon2” stands for DC values of 3.3 mA to 5.5 mA. Both measured at the M1 telescope.

## 8.4 Long-Term Analysis of the Crab Nebula

To demonstrate the full capabilities of `autoMAGIC`, I performed a data reduction for all data of the Crab Nebula, which was recorded by `MAGIC` between September 2016 and February 2020<sup>2</sup>. The data set covers five MC productions, three zenith ranges, and five moon ranges. This means that 51 separate analyses would have to be performed by an analyzer to create the light curve for the Crab Nebula from this data set.

To validate the light curve I calculate with `Gammapy` from the DL3 data produced with `autoMAGIC`, I compare them to the light curve obtained from `Flute` for the same input data. The SED is compared to the Crab spectrum published by the `MAGIC` collaboration [13].

### 8.4.1 Analysis Setup: Data Selection and $\theta^2$ and Hadronness Cuts

For the Crab Nebula data, a minimum transmission at 9 km of 0.55 is required, the maximum DC at telescope M1 is set to 13.2 mA and the maximum zenith angle to  $62^\circ$ . For the Off data, the minimum transmission at 9 km is set to 0.8 to keep the training data for the RF as similar as possible to the simulated data. Most of the used RFs are trained with MC events simulated at a distance range of  $0^\circ$  to  $2.5^\circ$  from the center of the FoV (view cone). Where simulations with this view cone are not available, diffuse MCs sets with a view cone of  $1.5^\circ$  are chosen.

`MARS` offers the possibility to apply an energy-dependent  $\theta^2$  and hadronness cut (see subsection 5.6.2). This increases the sensitivity for events at lower energies, since the analysis performs a worse reconstruction at these energies and the corresponding events have a higher possibility to be reconstructed as proton and/or further away from the true source position. The  $\theta^2$  cut at lower energies can therefore be applied at greater distances to include more events coming from the target source into the analysis. At high energies, where events are reconstructed closer to the true source position, such a loose cut would include too much background events. It is therefore recommended to apply an energy-dependent  $\theta^2$  and hadronness cut.

Unfortunately, the energy-dependent  $\theta^2$  cut is not yet implemented into `Gammapy`, and therefore I choose a fixed global  $\theta^2$  cut at  $0.02 \text{ deg}^2$  for the analysis with `Flute` and `magicDL3+Gammapy` to keep the analysis approaches comparable. The hadronness cut is also fixed at 0.28 to ensure that the analyses are performed as similarly as possible. Nevertheless, an energy-dependent approach for the hadronness cut is implemented for `Flute` and `magicDL3` and would not conflict with `Gammapy`. To perform “the best” scientific analysis with `Flute` and `magicDL3+Gammapy`, at least the hadronness cut should always be applied energy-dependent; for `Flute` also the  $\theta^2$  cut should be energy-dependent.

Another important ingredient is the spectrum that is used to calculate a light curve with a forward-folding approach. Usually, a spectrum is fitted to the data before or during the light curve calculation. For the Crab Nebula, the SED in the gamma-ray regime is well

---

<sup>2</sup>`MAGIC` observed the Crab Nebula since 2011, but the calibrated data of these observations are only available on tape. Restoring this data would have been a massive workload, since we need not only the Crab Nebula data but suitable Off data, as well. Therefore, the long-term Crab Nebula light curve is calculated from all data directly available at PIC. In general, `autoMAGIC` is capable of analyzing data before 2016.

known, which is why I used a fixed log-parabola spectrum. The normalization parameter remains free to be fitted. Again, this is a technical choice to keep the analysis setups as similar as possible in order to compare the two software packages. For a scientific analysis, the spectrum must be estimated, first.

All other analysis parameters, as for example, the number of Off positions, the energy range, and the energy binning in true and estimated energy are set to the same values, where possible. Nevertheless, small discrepancies between the two analysis approaches remain, which are mainly caused by slightly different flux point calculation methods and minimization algorithms.

A dedicated analysis regarding the performance of `Flute` and `magicDL3+Gammapy` using energy-dependent of fixed cuts can be found in appendix C.

### Crab Nebula Light Curve 2016 – 2020

Having all the parameters in line, the light curve of the Crab Nebula for 4.5 years can be calculated with `Flute` and `Gammapy`. Figure 8.3 shows the run-wise binned long-term light curves for the Crab Nebula from both software frameworks.

All in all, the flux points from `Flute` and `Gammapy` are in very good agreement with each other and with the reference flux from Aleksić et al. [13]. Some mismatches between the `MARS` and `Gammapy` fluxpoints are still visible. They mainly result from the reasons mentioned in the last section, namely slightly different calculations for the likelihood maximization. Some runs were processed only by `Flute` or `Gammapy` which could be caused by a failed minimization.

Figure 8.4 shows the histogram of the flux values for all runs computed by `Flute` and `Gammapy` and their mean values. Both mean fluxes are close to the reference flux from Aleksić et al., 2015 [13] ( $1.2 \times 10^{-10} \text{ cm}^{-2} \text{ s}^{-1}$ ), whereas `Gammapy` ends up below the reference flux and `Flute` above. For `Flute`, I obtain a mean flux value of  $1.26 \pm 0.32 \times 10^{-10} \text{ cm}^{-2} \text{ s}^{-1}$ , for `Gammapy` the mean flux is computed to  $1.07 \pm 0.27 \times 10^{-10} \text{ cm}^{-2} \text{ s}^{-1}$ . Further comparison plots are shown in the appendix D.

Note that the reference flux from Aleksić et al. was calculated also with `Flute` and with a smaller data set that is not include in the data set used here. It is therefore a good benchmark to check the performance of `autoMAGIC`, but not the holy truth that I aim to reproduce exactly with this analysis.

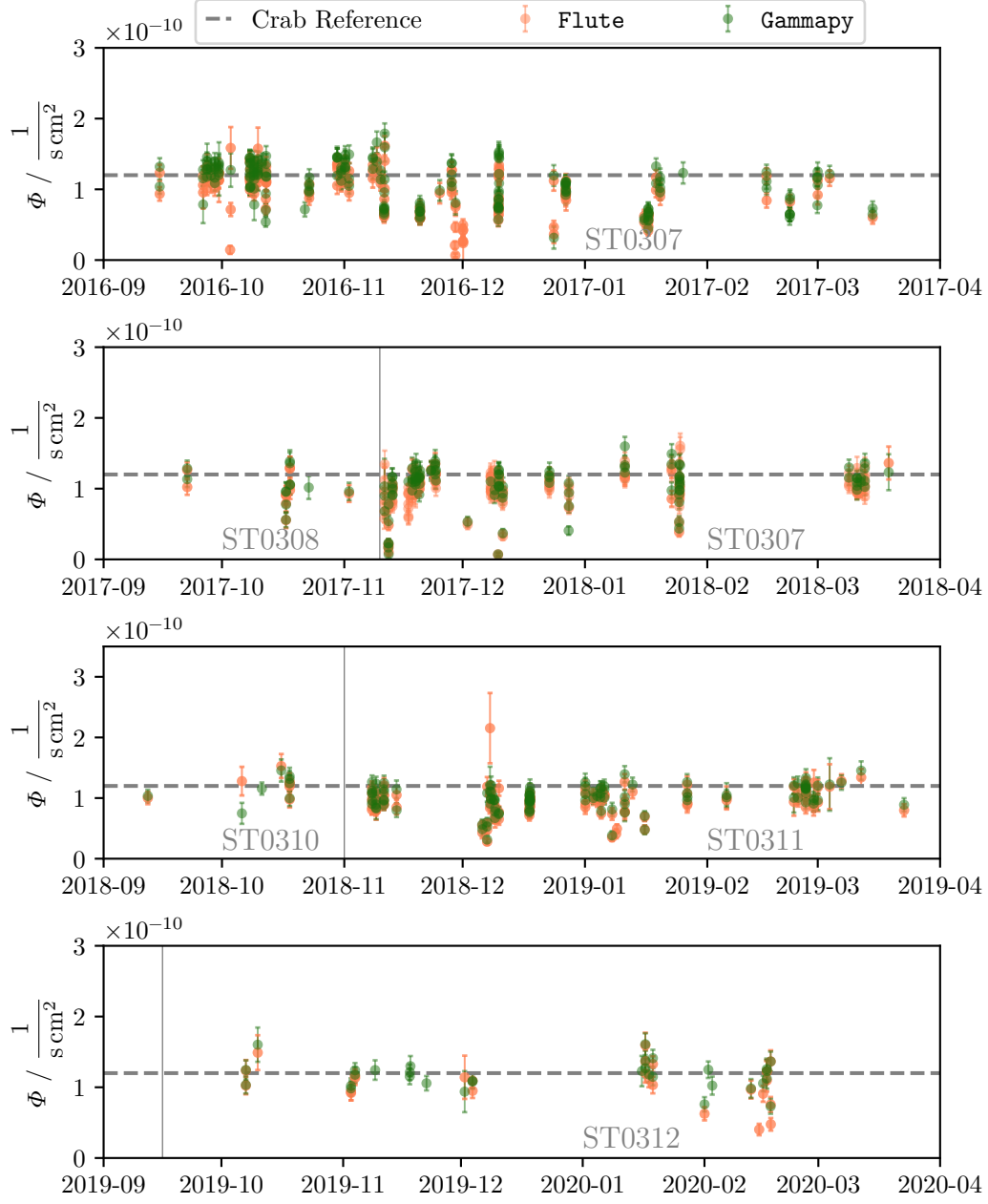
### Crab Nebula Spectral Energy Distribution

Figure 8.5 shows the SED for the long-term Crab Nebula data set calculated with `Gammapy` and the reference model from [13]. The spectrum is fitted with a log-parabola model with free parameter  $N_0, \alpha, \beta$  and a fixed reference energy  $E_0 = 300 \text{ GeV}$ . The fit results in

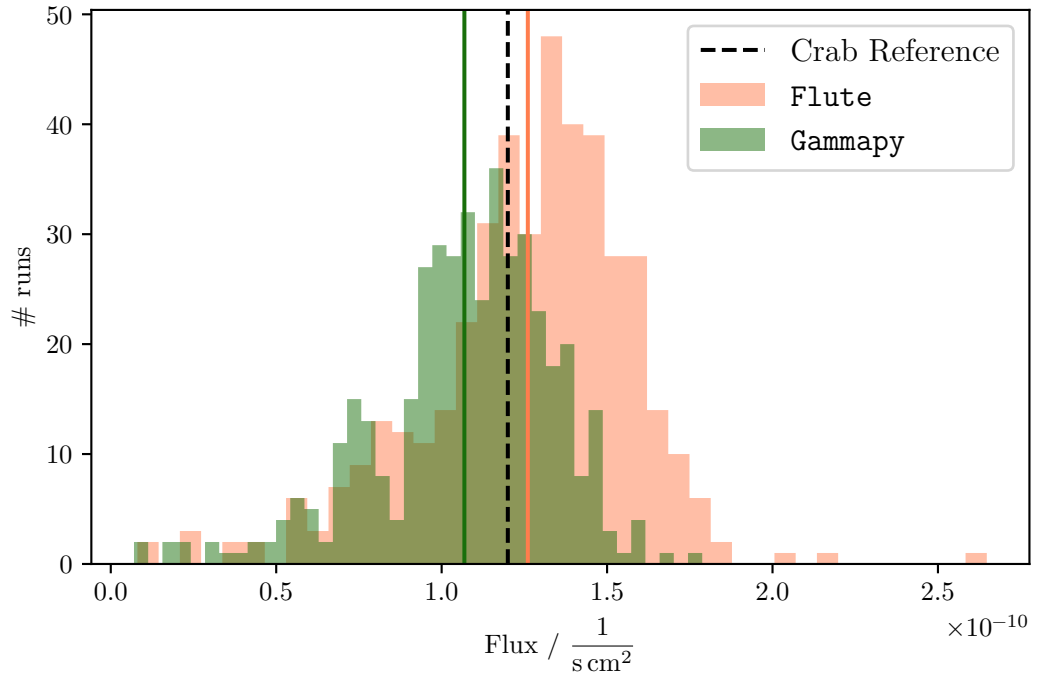
$$\frac{dN}{dE} = (2.97 \pm 0.02) \times 10^{-11} \frac{1}{\text{TeV cm}^2 \text{ s}} \quad (8.1)$$

$$\times \left( \frac{E}{300 \text{ GeV}} \right)^{-(2.49 \pm 0.01) - (0.098 \pm 0.003) \cdot \log(E/1 \text{ TeV})} \quad (8.2)$$

and is also in very good agreement with the reference spectrum, c. f. Figure 8.5.

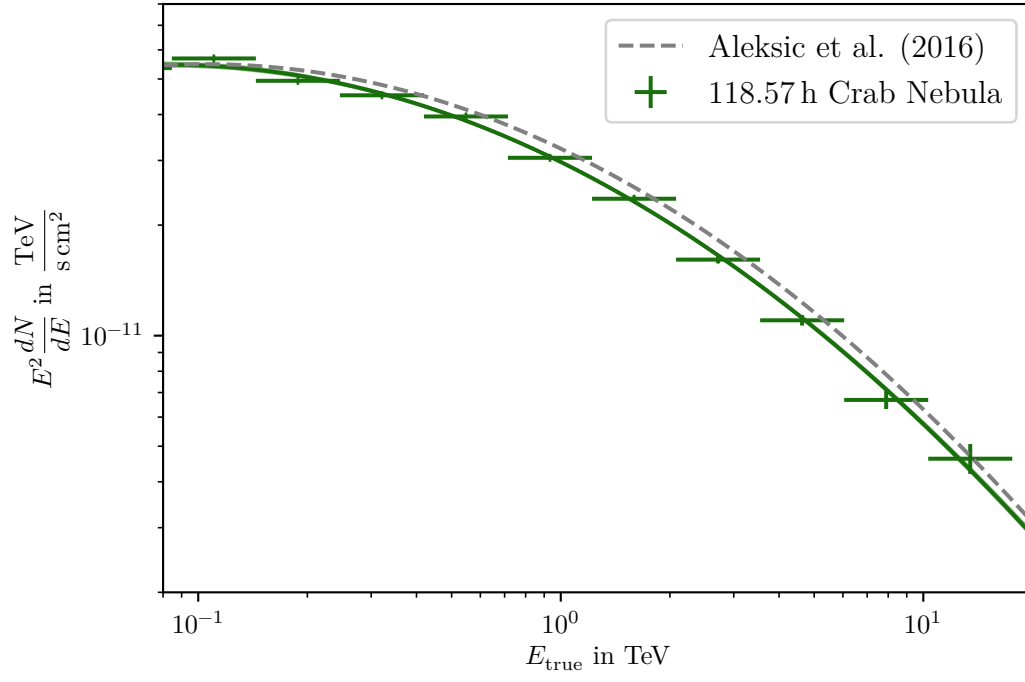


**Figure 8.3:** Long-term light curve of the Crab Nebula produced with Flute and Gammapy in run-wise binning using all Crab Nebula data measured by MAGIC from 2016-04 until 2020-02. Each panel represents one winter observation period from September to April. During the summer, the Crab Nebula is no visible from MAGIC's location. Both analyses were performed with a fixed global  $\theta^2$  cut  $< 0.02^\circ$  and a hadronness cut  $< 0.28$ .



**Figure 8.4:** Histogram of flux values computes run-wise with `Flute` and `Gammapy` for the Crab Nebula. The black dashed line marks the reference flux from [13]. The solid lines denote the mean fluxes of the corresponding histogram.





**Figure 8.5:** SED of the Crab Nebula modeled with data measured from 2016 to 2020. Fit results are computed using `Gammapy`.

### Analysis with autoMAGIC

Analyzing the long-term Crab Nebula dataset manually would have been a huge amount of work, most of it of repetitive nature. As an educated guess I would estimate several months of work for a human analyzer, depending on the skills and the hardware equipment. Performing the whole analysis with autoMAGIC takes only the time to fill out a configuration file and start the analysis. To process all the Crab Nebula data, the Off data and the corresponding MC simulations from calibrated level to DL3, 407 142 jobs are submitted to the cluster and need about 36 hours to finish. All RFs that were produced during this analysis can be re-used for data sets from other target sources taken under the same observation conditions, which extremely reduces the amount of computation time. Table 8.1 shows the number of jobs that are performed for every analysis step.

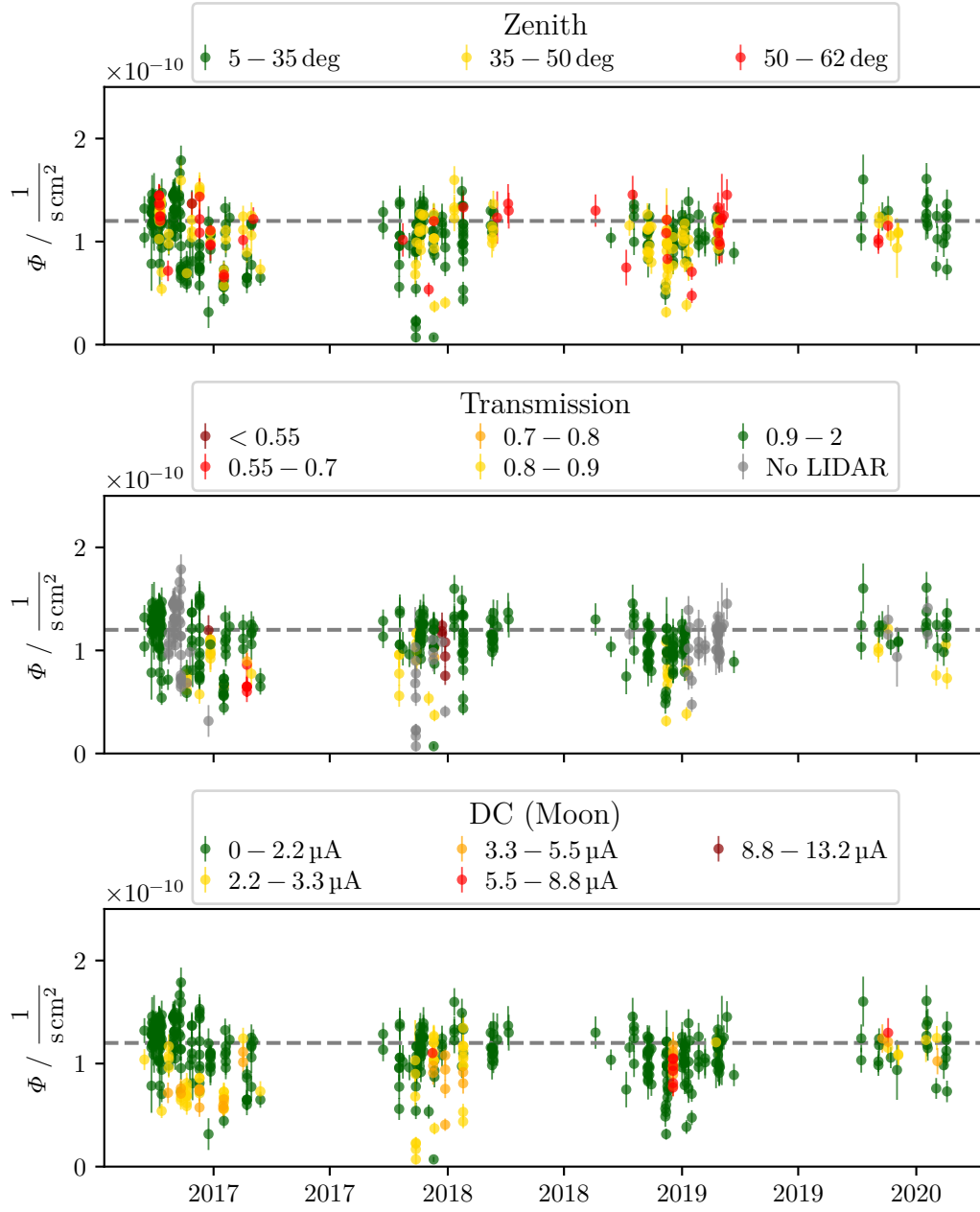
**Table 8.1:** Number of jobs for all analysis steps performed to produce the long-term Crab Nebula light curve (2016-2020) and the corresponding SED for all available data.

	Data	MC
Star + Superstar On	758	403934
Star + Superstar Off	1371	
Selectmc		84
Coach		51
Melibeia	758	84
magicDL3		51
Flute + Fold		51
Total		407142

#### 8.4.2 Observation Conditions

Using the autoMAGIC database together with Gammapy offers the possibility to connect the observation conditions for every single run to its flux point at the end of the analysis chain. Figure 8.6 shows the run-wise binned Crab Nebula light curve with color-coded zenith range, transmission (as an indicator for the weather), and DC of telescope M1 (as an indicator for the amount of moonlight). For all observation conditions, we see no systematic bias at specific ranges. Also, the flux points calculated from data measured at moonlight are in good agreement with the reference flux and show no obvious bias. Therefore, the Moon analysis is successfully integrated into the automated analysis, which was one of the key requirements of the autoMAGIC project.

Appendix D shows comparison plots for all flux points computed by Gammapy and Flute for all observation conditions and analysis periods.



**Figure 8.6:** Run-wise binned light curve of the Crab Nebula over several years. Observation conditions are color-coded for zenith, transmission (weather) and DC (moon). We see no obvious bias for individual observation conditions.

## 8.5 Open Issues and Next Steps

*autoMAGIC* is a software framework for an operating experiment based on software that is actively developed and continuously improved. By nature, *autoMAGIC* will never be finished as long as the MAGIC telescopes are operating and data has to be processed. Currently, it is in good shape to perform the first automated analyses and compare them to results that were obtained by human analyzers. These tests are ongoing and will reveal many things that can be improved.

As yet, we are aware of some issues that are to be addressed in the near future.

### 8.5.1 The High-Zenith Issue

Separating the target data by its observation conditions and train dedicated RFs for every combination of observation conditions is crucial to obtain a good separation and energy estimation. This is why the data is usually split in three zenith ranges: low ( $5^\circ$  to  $35^\circ$ ), medium ( $35^\circ$  to  $50^\circ$ ) and high ( $50^\circ$  to  $62^\circ$ ). The standard MC sets are also simulated for these zenith ranges. To create a well-trained RF, Off data is needed that covers the same  $Az \cdot \cos(Zd)$  range as the simulated data. The amount of Off data is often quite small for the high zenith range, because MAGIC observes predominantly at low zenith angles because of the better performance of the detector at lower angles. A source is observed at a high zenith angle only if there is no other choice to obtain measurements from this source or there are no known source positions at a lower zenith angle, which is not very common. Therefore, compared to the low and medium zenith, only a small amount of Off data exists for the high zenith range.

To account for this problem, we have to find a tradeoff between the lack of RF performance we cultivate by a too-small training data sample or by using improper data. For example, we could soften the time intervals given by the analysis periods so that some observations are valid as Off data for two periods. Another possibility is to enlarge the threshold for the significance (currently  $1\sigma$ ) or loosen the weather condition cut.

How such changes affect the quality of the RF and how they should be implemented into *autoMAGIC* must be investigated in dedicated studies.

### 8.5.2 The $\theta^2$ Cut Issue

As described in [subsection 5.6.2](#), the  $\theta^2$  cut is usually made in bins of estimated energy to account for the different amount and distribution of events at different energies. This energy-dependent  $\theta^2$  cut is implemented in MARS' *Flute*, where the cut is applied in usually 30 bins independently. Applying this energy-dependent cut accounts especially for the underpopulated bins at low energies and improves the analysis' sensitivity.

For the DL3 data, this approach is theoretically possible but currently not implemented in the *Gammapy* package, which is used here for the high-level analysis of DL3 data. Therefore, results may differ from the results obtained with *Flute*.

This issue will be fixed in the near future by implementing this technique into *Gammapy*'s code base.

### 8.5.3 Upcoming Projects

- Background models: Several advanced analyses techniques, for example, the creation of sky maps or the analysis of extended sources, require dedicated spatial background models. These models can be obtained by Off observations or events that are classified as hadron-induced.
- Massive data processing: Up to now, only analyses of individual sources were performed. For a first step to the MAGIC legacy, *all data* taken in a specific MC period should be processed to DL3 level.
- Integration of the sum trigger analysis: To detect gamma rays at very low energies ( $< 30$  GeV), MAGIC performs observations using a special trigger, the so-called *Sum Trigger*. To analyze the data taken in sum trigger mode, special treatment of the MC simulations is needed and the data processing has to start at a lower data level. These extensions are currently developed and will be integrated into the autoMAGIC workflow.
- Flare advocate: Urgent analyses, for example, after a Gamma-Ray Burst (GRB) or a flare, are performed by so-called “flare advocates” at the moment. These are analyzers on call prepared with pre-trained RFs to perform and deliver analysis results as fast as possible. This task requires a lot of personal effort for the analyzers and leads to preliminary results that can be used for a first estimation, but a full custom-designed analysis has to be performed later, anyway. Ideally, autoMAGIC will be used to produce reliable results directly after the data is accessible and without much human interaction.



## Gamma-Ray Emission from NGC 1275

### 9.1 Observational History in High Energies

As mentioned in [section 2.4](#), the radio galaxy 3C 84 was detected as high-energy gamma-ray source by *Fermi*-LAT in 2008. In gamma-ray astronomy, 3C 84 is identified with NGC 1275, an object from the New General Catalogue of Nebulae and Clusters of Stars (NGC), which was published at the end of the 19th century by Johan Ludvig Emil Dreyer and is still one of the most common used catalogues in astronomy [[63](#)].

The first notice of NGC 1275 as gamma-ray emitter was made by the COS-B satellite between 1975 and 1979 at energies from 70 MeV to 5 GeV [[154](#)]. At that time, it was unclear if NGC 1275 is the only source responsible for gamma-ray emission in the Perseus cluster. COS-B operated from 1975 to 1982 and provided the first detailed view of the gamma-ray universe. The next satellite experiment, the EGRET, which operated from 1991 to 2000, did not detect NGC 1275, although its energy range and threshold were sufficient for the flux detected ten years before by COS-B [[139](#)]. From 2008, *Fermi*-LAT received a well-observable gamma-ray signal from NGC 1275 directly after its launch and was able to monitor the source since then. This leads to the conclusion that the source must have been in a lower flux state in the 1990s and restarted its activity after 2000. Unless, it has to be considered that EGRET had a FoV of  $\approx 20^\circ$  which is much smaller than the sky coverage of *Fermi*-LAT which is about 20 % of the sky.

In the very-high energy regime, the MAGIC telescopes first detected NGC 1275 between 2009 and 2011 in stereoscopic mode [[18](#)]. Aleksić et al. report a mean flux of  $(1.6 \pm 0.3_{\text{stat}} \pm 0.3_{\text{sys}}) \times 10^{-11} \text{ cm}^{-2} \text{ s}^{-1}$  above 100 GeV from October 2009 to February 2010 and a very similar mean flux of  $(1.3 \pm 0.3_{\text{stat}} \pm 0.3_{\text{sys}}) \times 10^{-11} \text{ cm}^{-2} \text{ s}^{-1}$  between August 2010 and February 2011 [[17](#)].

Six years later, MAGIC observed NGC 1275 from September 2016 to February 2017 and detected a remarkable gamma-ray flare on New Year's Eve 2016/2017 [[22](#)]. With a peak value of  $9.5 \times 10^{-10} \text{ cm}^{-2} \text{ s}^{-1}$  the flux was around fifty times higher than in the first campaign from 2009 to 2011.

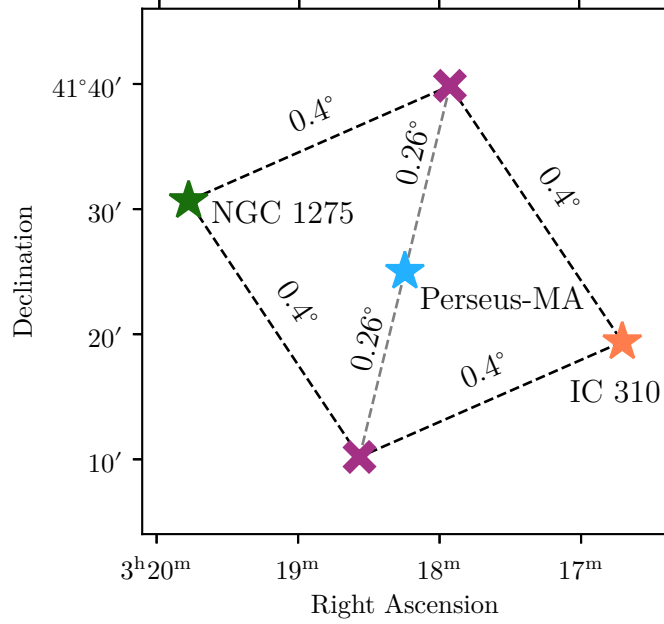
MAGIC observed NGC 1275 on a regular basis since then, but no further flare was observed, and no additional data was systematically analyzed or published. The aim of this thesis is to investigate the data since 2017, which was also the original motivation for the autoMAGIC project. Since 2016, NGC 1275 was observed under varying observation conditions (moon, zenith, MC period), which would sum up to 23 individual analyses.

The analysis results presented in the following were created by reducing the calibrated data to DL3 files with autoMAGIC and performing the high-level analysis with Gammapy.

## 9.2 Long-Term Observations of NGC 1275

Since its detection, NGC 1275 is observed by MAGIC on a regular basis. For this work, all data taken since 2016-01-01 is used<sup>1</sup>. Additional to NGC 1275, another gamma-ray source is located in the Perseus cluster: IC 310. With a distance of  $0.6^\circ$ , IC 310 is visible in the FoV of MAGIC ( $3.5^\circ$ ) if the telescopes are pointed at the wobble positions of NGC 1275. Therefore, the region around the position of IC 310 has to be excluded for any further analysis.

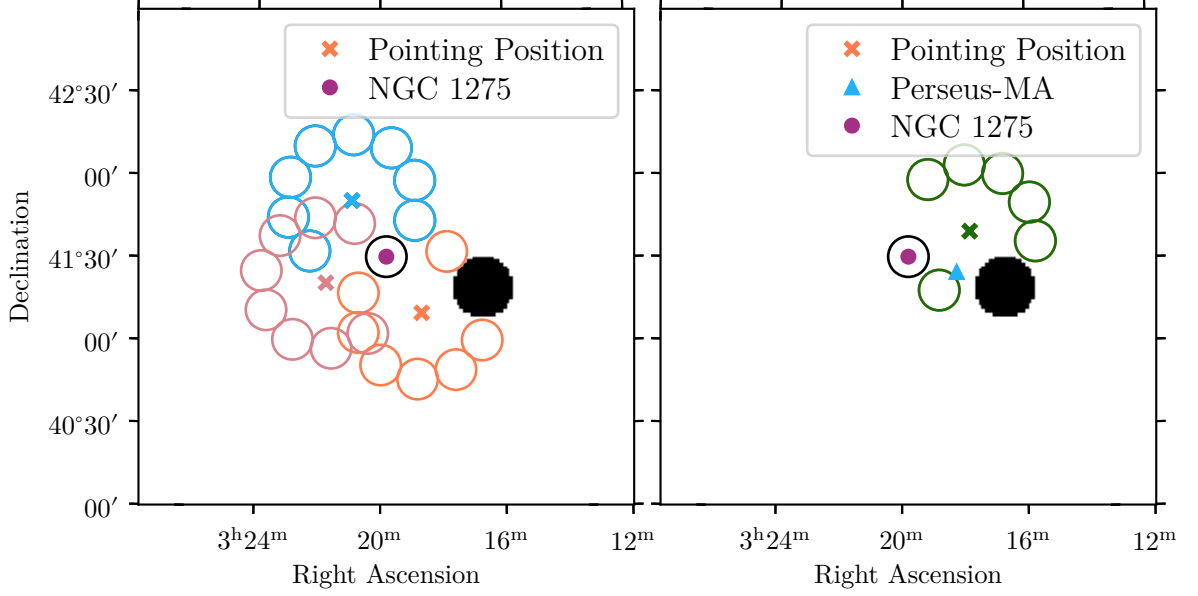
On the other hand, we use the advantage of having two such nearby sources by pointing to two positions in a distance of  $0.4^\circ$  from both sources. These observations are suitable for analyzing NGC 1275 and IC 310 and are categorized as “Perseus-MA” data, which is the position half-way between the two constructed wobble positions where no source is located. This data can be used for analyzing both sources with the standard MC simulations suitable for an offset of  $0.4^\circ$ . Figure 9.1 shows the source positions, Perseus-MA, and the respective wobble positions. For the long-term analysis of NGC 1275, the datasets observed by pointing at wobble positions around Perseus-MA and NGC 1275 are combined. The number of possible Off regions in the field of view is constrained by the exclusion of a region around IC 310 and the size of the On region ( $\theta^2$  cut), see Figure 9.2.



**Figure 9.1:** Pointing positions in the Perseus cluster. The small distance of NGC 1275 and IC 310 causes some special treatment during the analysis but enables us to observe them both at the same time. The wobble positions of Perseus-MA (marked with an x) are designed to be in a distance of  $0.4^\circ$  of both sources.

<sup>1</sup> All data recorded before this date was removed from the file system at the PIC and is stored on tape. This makes the data access quite difficult. For the longterm-goal of reprocessing all MAGIC data to DL3, this data will be recovered and processed with autoMAGIC in the next years.





**Figure 9.2:** On and Off positions for observations of NGC 1275 (left) and Perseus-MA (right). The wobble positions are marked with an x. The black region is the excluded region around the source IC 310 (radius of  $0.18^\circ$ ). The On and Off regions have a radius of  $\sqrt{0.015 \text{ deg}^2} = 1.22^\circ$ .

### 9.2.1 Two States of Activity

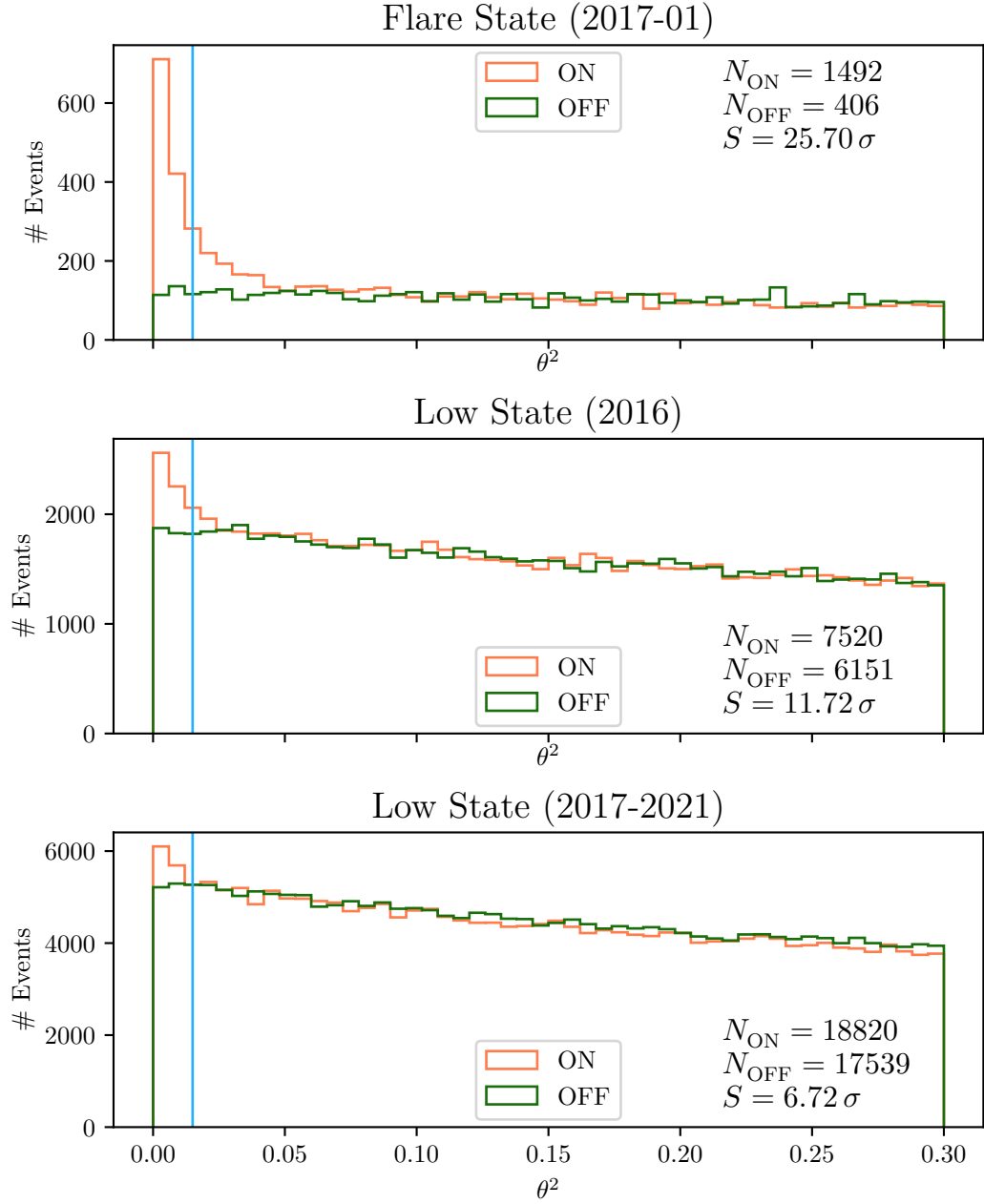
As mentioned before, MAGIC detected a gamma-ray flare of NGC 1275 in January 2017. Since this very short flare would dominate the whole analysis, the MAGIC dataset is divided into three subsets:

- Low state before the flare: 2016-06-02 to 2017-12-30 (MJD<sup>2</sup>: 57424.85 to 57752.5)
- Flaring state: 2016-12-30 to 2017-01-03 (MJD: 57752.5 57756.5)
- Low state after the flare: 2017-01-03 to 2021-03-16 (MJD: 57756.5 to 59289.88)

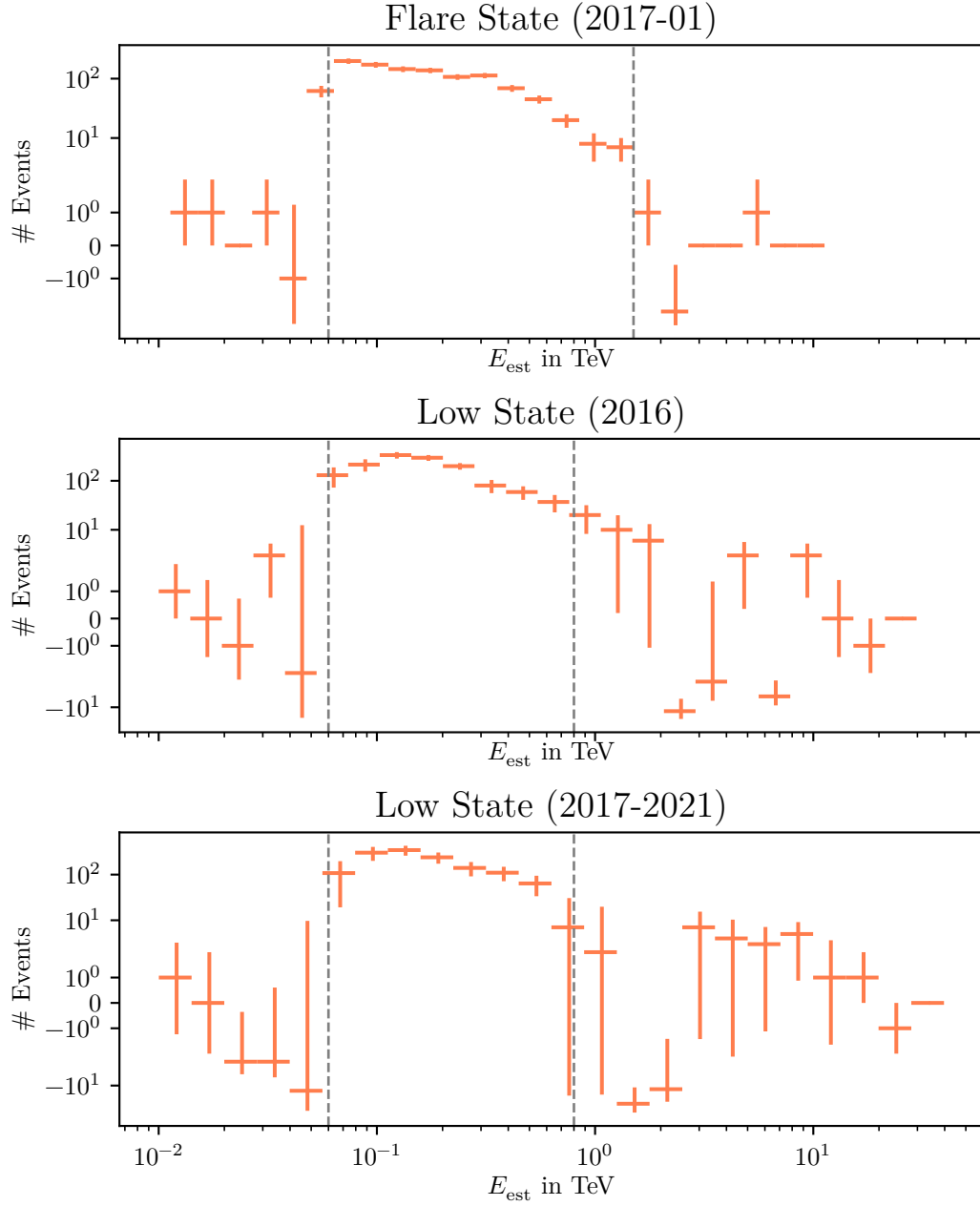
The size of the On region is chosen as  $\theta = 0.122^\circ$  for all data sets. Figure 9.3 shows the  $\theta^2$  plots, containing the number of events within the squared distance to the On and Off positions and the chosen cut.

To set the energy boundaries, we have to examine the energy-dependent excess events ( $= N_{\text{ON}} - \alpha N_{\text{OFF}}$ ), see Figure 9.4. For the flaring state, the energy ranges are chosen as [60 GeV, 1.5 TeV]. For the low states, the detected events only cover an energy range of [60 GeV, 800 GeV].

<sup>2</sup>Modified Julian Date, a common date format in astronomy. MJD counts days since 1858-11-17 00:00:00.0.



**Figure 9.3:** Number of events depending on their squared distance to the On and Off regions.  $N_{\text{ON}}$  and  $N_{\text{OFF}}$  denote the events below the  $\theta^2$ -cut at  $0.015 \text{ degree}^2$  (blue line).  $S$  denotes the Li&Ma-significance. For the flaring state, we detect a very strong signal. During the low state, the source is weak but detectable with a significance of more than  $5\sigma$ .



**Figure 9.4:** Number of excess events ( $N_{\text{ON}} - N_{\text{OFF}}$ ) binned in estimated energy. The grey dashed lines mark the energy boundaries for the following analysis. The lower limit is set at 60 GeV, the upper limit is set at 800 GeV for the low states and at 1.5 TeV for the flaring state.

### 9.2.2 Light Curve and SED

The SED for the three time intervals is modeled in the defined energy ranges. Since NGC 1275 is known as a weak source and only detected at comparably low energies, all datasets are modeled with a power law with exponential cutoff:

$$\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^\gamma e^{-\left(\frac{E}{E_c}\right)}, \quad (9.1)$$

The reference energy is fixed at 100 GeV for all states. The likelihood-fitting is performed with Gammapy and the resulting fit parameters are given in Table 9.1.

As expected and already reported [22], we see a higher flux in all energy bins during the flare and a lower flux before and after the flare. For the lower states, some of the flux points are calculated as upper limits ( $\sqrt{TS} < 3$ ), which must be kept in mind if deriving any statements about the SED and the behavior of the source. But we can see at least a strong hint that the mean flux was lower after the flare than before. This is in agreement with the fluxes reported by *Fermi*-LAT before and after the flare, which I will describe in detail in chapter 12.

Figure 9.6 shows the long-term light curve of NGC 1275 with the significant flare around New Year's Eve 2017. Apart from that flare, NGC 1275 turns out to be very weak and barely detectable, especially in the years after the flare. For these measurements, I obtain only upper limits (shown for bins with  $\sqrt{TS} < 3$ ).

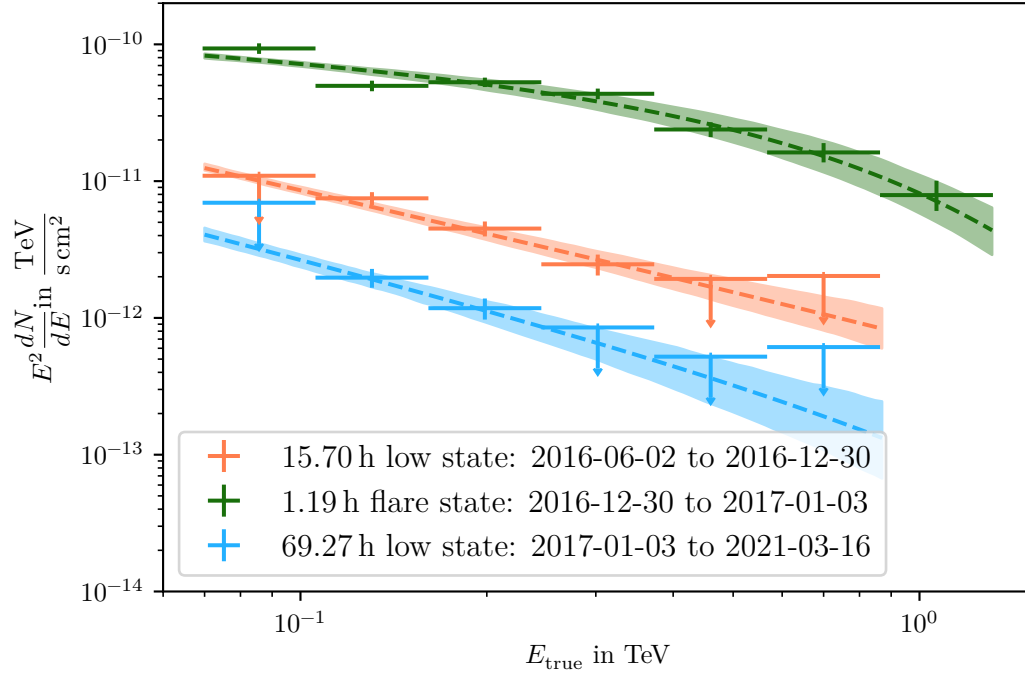
These results fit the reported flux values from *Fermi*-LAT in the MeV range, where the flux decreases rapidly after its maximum in the mid of 2018, see Figure 12.2 in chapter 12. We expect a similar behavior in the VHE range, since the energy ranges of *Fermi*-LAT and MAGIC are slightly overlapping.

Furthermore, the flux values I obtain with autoMAGIC and Gammapy are in good agreement with the values reported by Ansoldi et al. [22]; though I want to stress here that the flux values are not directly comparable since they are performed at different data sets and partly different software tools.

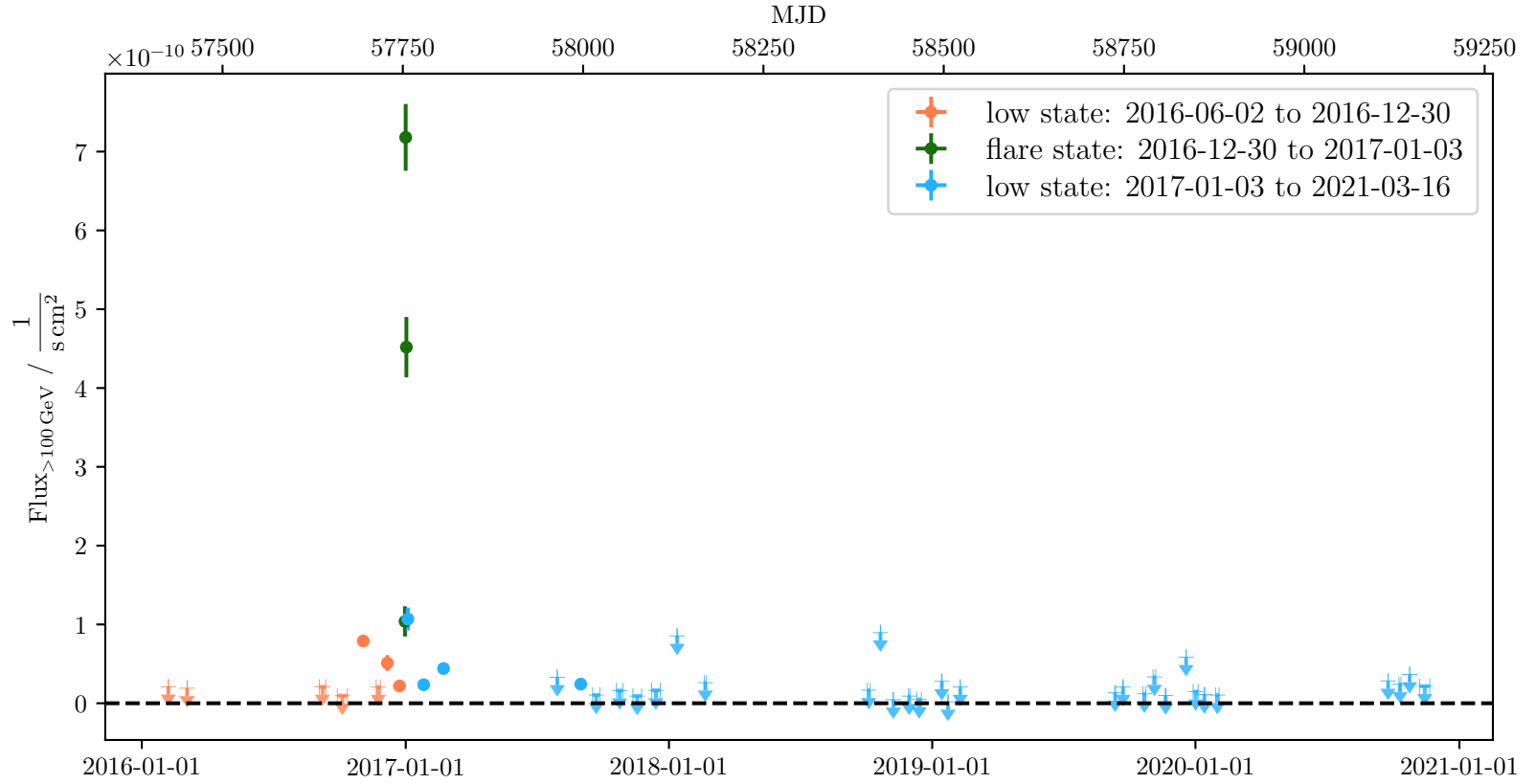
All in all, 303954 jobs were necessary to analyze this dataset and the corresponding MC simulations measured over 4.5 years and under 23 different combinations of observation conditions (moonlight, zenith range, MC period).

**Table 9.1:** Parameters of the fit for the power-law model with exponential cutoff fitted to the datasets of NGC 1275 and Perseus-MA for the three states of activity.

Parameter	Low before	Flare	Low after
$E_0$ in TeV	$0.10 \pm 0.00$	$0.10 \pm 0.00$	$0.10 \pm 0.00$
$N_0$ in $10^{-10} \frac{1}{\text{TeV cm}^2 \text{ s}}$	$8.66 \pm 0.84$	$85.98 \pm 5.47$	$2.84 \pm 0.49$
$\gamma$	$3.04 \pm 0.22$	$2.24 \pm 0.15$	$3.14 \pm 0.38$
$E_C$ in $\frac{1}{\text{TeV}}$	$0.10 \pm 0.62$	$1.80 \pm 0.48$	$0.68 \pm 1.27$



**Figure 9.5:** SED and binned flux points for NGC 1275 measured with MAGIC for three states of activity. The green points show the SED during the flaring state (4 runs), whereas the SED during the low state before and after the flare is colored in orange and blue. For the low states, we see that NGC 1275 is in a lower state after the flare than before.



**Figure 9.6:** Long-term light curve of NGC 1275 from 2016 to 2020 produced with Gammapy. Upper limits are shown for bins with  $\sqrt{TS} < 3$ . The flare (green) was reported in [22]. Besides the flaring state, NGC 1275 is very weak and barely detectable.

## **Part III**

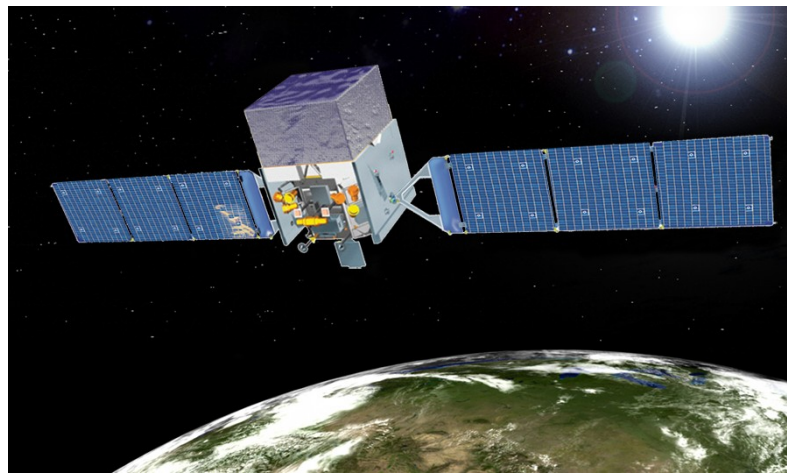
### ***Fermi-LAT***





## Space-Based Gamma-Ray Detector: *Fermi*-LAT

If we want to measure the gamma radiation without reconstructing a photon's properties from its airshower, the detector has to be above Earth's atmosphere. This requires a space mission and a lot more money, compared to IACTs. One of the few currently operating gamma-ray telescopes, where this effort was made, is the Large Area Telescope (LAT) onboard of the *Fermi* Gamma-Ray Space Telescope; a satellite, which started its mission in June 2008 [26]. The *Fermi*-LAT is capable of detecting gamma rays in the energy range from 20 MeV to 300 GeV, which perfectly covers the energy range below that of all operating IACTs. Compared to IACTs, the LAT has a very large field of view of 2.4 sr, which covers about 20 % of the sky [26].



**Figure 10.1:** Computer-generated image of the *Fermi* Gamma-Ray Space Telescope carrying the Large Area Telescope. The *Fermi* satellite was launched in 2008. Image credit: NASA.

The gamma-ray detector consists of 16 identical particle detectors composed of a tracker and a calorimeter module. The tracker modules, which are made of layers of silicon strip detectors and tungsten foils, are responsible for tracking the path of a charged particle. The calorimeter modules, made of CsI(Tl) crystals, completely absorb the incoming particles and measure their energy.

An incoming gamma ray will produce an electron and a positron by pair production when hitting a foil. The electrons and positrons ionize the silicon atoms on their path through the detector. By tracking the path and measuring the energy deposited in the detector, the gamma ray's energy and position can be reconstructed very precisely.

## 10 Space-Based Gamma-Ray Detector: *Fermi-LAT*

The tracker and calorimeter modules are enclosed by an Anti-Coincidence Detector (ACD), which detects events initiated by non-photon background particles. If a background event is spotted by the ACD, a veto signal is issued so that the signal within the tracker and calorimeter modules will not be recorded.

Apart from the LAT, the *Fermi* satellite carries another instrument, the Gamma-Ray Burst Monitor (GBM), which is designed to detect very short and powerful outbursts of gamma-ray emission, such as gamma-ray bursts and solar flares. The GBM is capable of monitoring almost the whole sky at the same time and is sensitive for gamma rays in the energy range from 8 keV to 40 MeV.

A detailed review of the *Fermi* satellite, its mission and construction can be found, for example, at [6, 26, 72]. The data analysis will be described in the next chapter.

## Fermi-LAT Data Analysis

---

Analogous to the data reduction for an IACT like MAGIC, the data analysis of the *Fermi*-LAT is event-based and therefore very similar. As already described in [chapter 5](#), the reconstructed features can not be treated as truth and are convoluted with the detector's IRFs. Furthermore, also *Fermi*-LAT has to deal with background events.

In contrast to an IACT, the background events from charged CR particles are suppressed by the ACD before they are even recorded. Nevertheless, a background remains that has to be considered for the later analysis.

Another difference is the relatively large FoV of the *Fermi*-LAT, which prevents the observation of just a single source. This is why *Fermi*-LAT data is not organized in observations of a specific source but as a constant flow of events categorized by their arrival time and direction. To analyze a specific source, all events within a certain radius around ( $10^\circ$  to  $20^\circ$ ) the source position are selected. This means that all additional sources around the target source have to be modeled as well.

### 11.1 Data Flow and Levels

An event that was not vetoed by the ACD and left a track in the detector will be analyzed by the Data Acquisition System (DAQ) in order to perform a first on-board raw data reduction and separation. Roughly 10000 events per second can be processed of which above 400 remain [\[72\]](#). Only two to five of them will be selected as photon signal later. These events are stored and sent to Earth in packages six to seven times a day with a typical latency of 12 hours.

The data downlinked from the satellite is received at the Mission Operations Center (MOC) which is located at the Goddard Space Flight Center (GSFC) in Maryland, USA. As a first step, the MOC takes care of cleaning up the data, e.g. removing broken or duplicated events, and arranges the events in the correct time order. The result is called *Level 0 data*. The Level 0 data is transferred to the respective instrument operations center for the LAT and the GBM data. At the LAT Instrument Science and Operations Center (LISOC), located in California, event lists are created, containing the reconstructed primary type, source direction, and energy (*Level 1 data*, comparable to MARS' *Melibe*a output). With these event lists, a scientist can now perform target-tailored analysis and produce *Level 2 data* containing high-level results like spectra and light curves.

As we learned from the MAGIC analysis, the high-level results we obtain from a data set depend on the cuts we specify and the quality of the IRFs. For different target types, different cuts are required, which affects for example the amount of background events in

the data and results consequently in different IRFs. LISOC provides three different classes, suitable for different use cases:

- TRANSIENT: loose background suppression, increased  $A_{\text{eff}}$  and PSF
- SOURCE: default for point and slightly extended sources
- ULTRACLEANVETO: strong background suppression, smaller  $A_{\text{eff}}$

For this work, the SOURCE class is used. The data is stored as FITS files and includes the event lists as well as a *spacecraft file* which contains the IRFs and information about the pointing and the actual on-time of the detector. Data of Level 1 can now be analysed with the *Fermitools* or *FTOOLS* [31] to obtain scientific results as SEDs and light curves. The analysis process is quite similar to the one described in section 5.5 but has some extensions.

Since the FoV of the LAT is very large, the spatial distribution of all gamma-like events can not be estimated as a single (Gaussian) distribution, but as a superposition of multiple sources distributed all over the observed sky. Furthermore, these sources can be extended which means that they can not be modeled by a simple Gaussian distribution, anymore.

We must, therefore, always model the spectral *and* the spatial distribution for the analysis of *Fermi*-LAT data. The modeling itself is performed with a likelihood maximization, as described in section 5.5. Additionally to the *binned* likelihood fit as used for the MAGIC data analysis, the software tools developed for *Fermi*-LAT can perform an *unbinned* likelihood fit. For the binned likelihood approach in Equation (5.14), the events are binned in estimated energy first and the likelihood is build as product over these bins. The unbinned likelihood, in turn, is the product over the probabilities of all single events. The latter approach might lead to better results, but is only suitable for small data sets since the computation time increases with the amount of events.

### 11.2 Fermi-LAT Data Analysis Tools

The data used for this work is processed with *fermipy* [176], which is a Python package that is based on the *Fermitools* [47]. All packages needed for the full data analysis can be downloaded as a Docker Image<sup>1</sup>. For this thesis, the Docker Image tagged with 2019-05-14 is used, which includes *fermipy* version 0.17.4 and *Fermitools* version 1.0.5. In contrast to MAGIC, the LAT collaboration provides ready-to-use IRFs, version P8R3\_SOURCE\_V2\_V1 is used in this case. To complete the setup, we need a galactic interstellar emission model, which is given in *gll\_iem\_v07.fits*<sup>2</sup> and an isotropic model, *iso\_P8R3\_SOURCE\_V2\_v1.txt*. Both are included in the Docker Image. To calculate the SED, light curve, and spatial distribution, a binned likelihood fit is used.

Further details about the software tools and analysis details like the formulation of the likelihood can be found in the *Fermitools* online documentation [47].

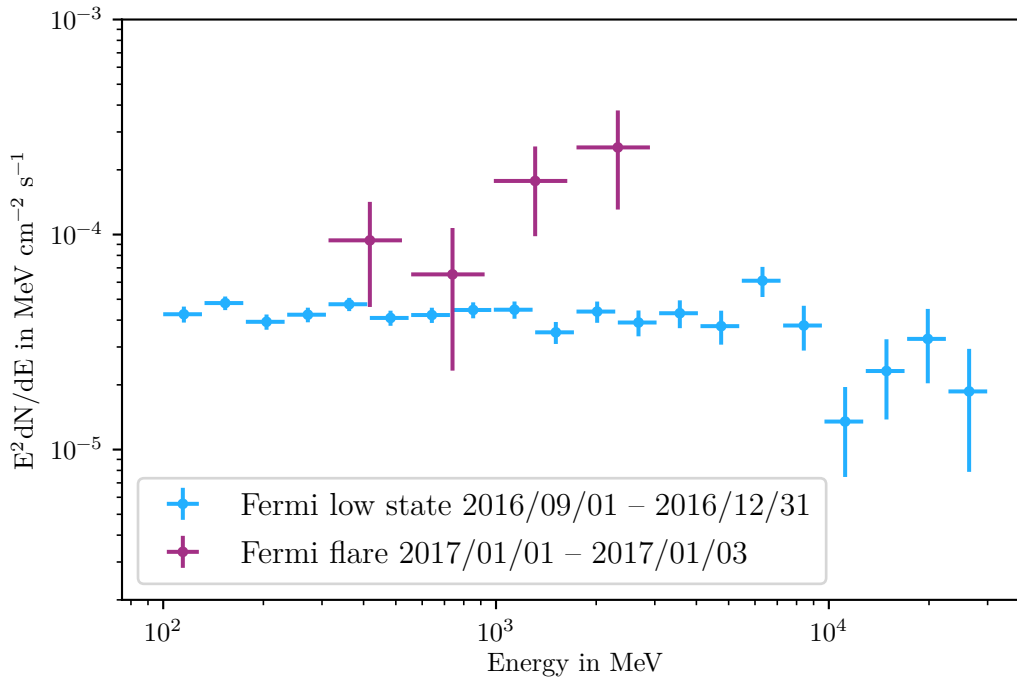
<sup>1</sup><https://hub.docker.com/r/fssc/fermibottle>

<sup>2</sup>[https://fermi.gsfc.nasa.gov/ssc/data/analysis/software/aux/4fgl/gll\\_iem\\_v07.fits](https://fermi.gsfc.nasa.gov/ssc/data/analysis/software/aux/4fgl/gll_iem_v07.fits)

## Fermi-LAT Results for NGC 1275

*Fermi*-LAT observed NGC 1275 since 2008 with no large interruptions. For this work, I use LAT data for two different studies: First, to investigate the optical depth of the BLR, described in section 16.2, and second, for the light curve cross-correlation performed in section 17.3. In the first case, the SED is needed to perform a joint fit with MAGIC data, in the latter case, I am interested in the light curve.

To perform the joint SED fit, two data sets are needed according to the time intervals defined by Ansoldi et al. in [22], who investigated two states of activity: a low state from 2016-09-01 to 2016-12-31 and a flaring state from 2017-01-01 to 2017-01-03. The flux points are shown in Figure 12.1.

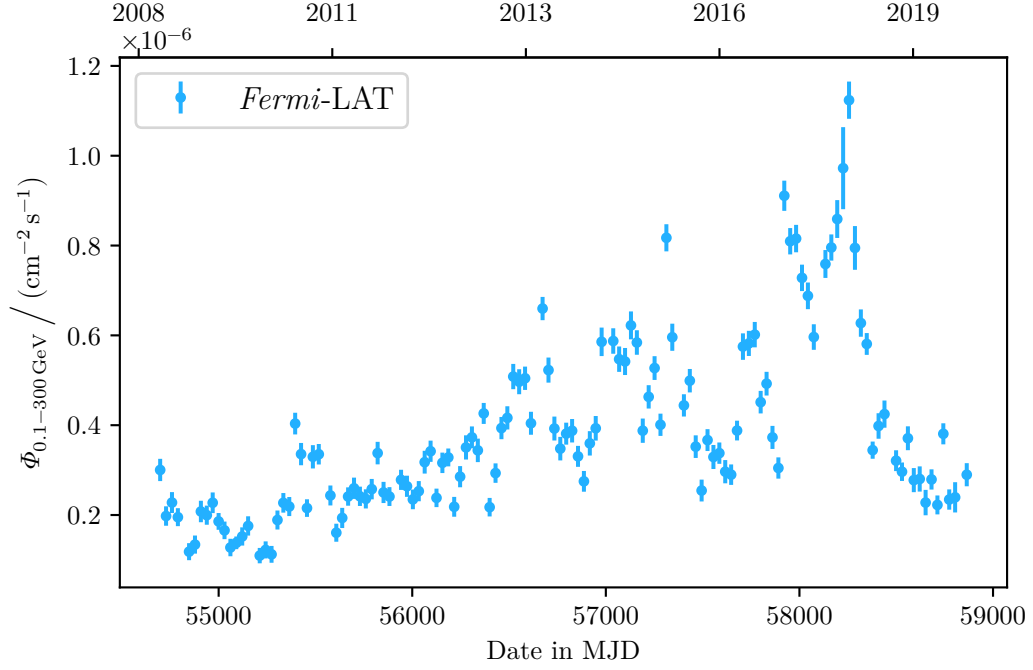


**Figure 12.1:** Fluxpoints of NGC 1275 measured with *Fermi*-LAT for two states of activity: The low state from September to December 2016 and the flaring state at the beginning of January 2017.

Furthermore, I calculated a long-term light curve using data from the beginning of *Fermi*'s mission on 2008-08-04 to 2019-06-26. The light curve is used for cross-correlation studies together with radio data from the VLBA, which is why the *Fermi*-LAT light curve is binned

monthly accordingly to the radio data. The result is shown in Figure 12.2 and reveals a long-rising trend from 2008 to 2018 accompanied by short flares. After 2018-05-18, we see a rapidly decreasing flux. Since then, the flux remains on a stable but low flux level.

A light curve with finer binning can be downloaded directly from the LAT collaboration's web page<sup>1</sup>.



**Figure 12.2:** Light curve of NGC 1275 measured by *Fermi*-LAT from 2008-08-04 to 2019-06-26 in monthly binning.

<sup>1</sup>[https://fermi.gsfc.nasa.gov/ssc/data/access/lat/msl\\_lc/source/NGC\\_1275](https://fermi.gsfc.nasa.gov/ssc/data/access/lat/msl_lc/source/NGC_1275)

## **Part IV**

### **VLBA**

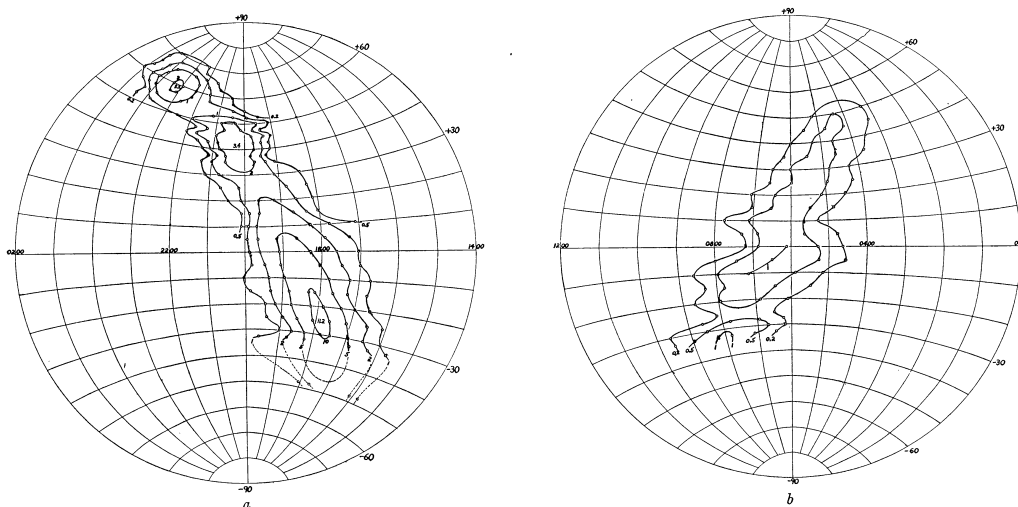




## Radio Astronomy

On the lower side of the electromagnetic spectrum, the wave nature of the emitted radiation requires different detector techniques than far gamma-ray emission. Dipole or dish antennas are used to detect radio waves at wavelengths of several millimeters to meters, which would correspond to energies of neV or even peV. Unlike in gamma-ray astronomy, it is very uncommon to use these units in radio astronomy. In radio astronomy, energies are given in units of frequency or wavelength.

Earth's atmosphere is transparent for radio waves at most wavelengths (see [Figure 2.1](#)), and ground-based antennas are much easier to build than a space-based gamma-ray detector or an IACT. This is why the history of radio astronomy started already in the 1930s. Karl Guthe Jansky discovered radio emission at 14.6 m (20.5 MHz) coming from the Milky Way [\[98\]](#). He was followed by Grote Reber, who built a single dish antenna in his backyard in Wheaton, Illinois, operating at 160 MHz. Reber published the first-ever radio map in 1944, which already revealed the strong radio sources Cygnus A and Cassiopeia A, shown in [Figure 13.1](#).



**Figure 13.1:** First radio contour maps of the Milky Way measured by Grote Reber in 1944, showing both hemispheres of the sky. Cygnus A is located on the left map at  $40^\circ$  and Cassiopeia A on the same map at  $60^\circ$  [\[136, 135\]](#).

His multi-frequency observations in the following years revealed the non-thermal nature of the radio emission, which was later explained by the synchrotron radiation of accelerated electrons [\[67\]](#). During World War II, Reber was the World's only radio astronomer, before

others joined him in the 1950s and further radio telescopes were built all over the world, e. g. in 1957 the Lovell Telescope in England [116] or the Stockert Radio Telescope in Germany in 1956 [23].

Another milestone was the detection of the 21 cm emission line from neutral hydrogen by Ewen and Purcell in 1951, which was predicted by van der Hulst in 1944 [69]. This discovery allowed for scanning the hydrogen density in the Milky Way and its spiral structure was unveiled for the first time.

Several years later, in 1965, Penzias and Wilson experimented with a horn antenna build for the detection of large microwaves. More or less accidentally, they discovered the CMB [175] as a faint isotropic electromagnetic radiation corresponding to a thermal black body spectrum at 2.72 K (see [section 2.6](#)).

## 13.1 Radio Interferometry

After the first proof of concept for radio telescopes, the astronomers gained more and more access to the hidden radio signals of the universe by pushing the technical possibilities to the limits. In the 1950s and 1960s, bigger and bigger single-dish telescopes were built to improve the angular resolution  $\theta$ , which can be understood as the minimal distance at which two points can be separated by the observer. The angular resolution is given by the Rayleigh criterion

$$\theta = 1.22 \frac{\lambda}{D} \quad (13.1)$$

with the observed wavelength  $\lambda$  and the diameter  $D$  of a telescope's dish. Improving a telescope's resolution or observing longer wavelengths with the same resolution therefore requires bigger dishes. At a certain point (and that point was definitely reached when the huge dish of the Green Bank's 300-foot telescope collapsed under its own weight), building bigger dishes was not an option anymore. The answer to that problem was interferometry: combining multiple antennas pairs several meters or even kilometers apart from each other to a virtual extremely large dish and correlate their received signals. The hardware for these antennas can be quite different, as for example, parabolic dishes, dipole antennas, or special-shaped sheets of metal are used. The first radio interferometers were built in the 1970s, for example, the Karl G. Jansky Very Large Array (JVLA) in New Mexico or the Westerbork Synthesis Radio Telescope (WSRT) in the Netherlands.

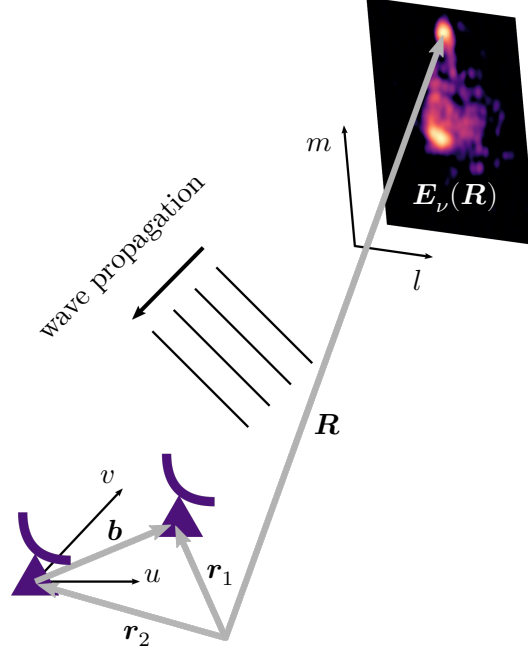
Radio interferometry is a topic of great complexity, which requires profound knowledge that can be achieved in dedicated study programs and a lifetime of work in this field. This is why I will concentrate on the very basic ideas and techniques needed to understand roughly the way from an electromagnetic wave emitted by a source to an image of the source's morphology and leave the rest to the real radio astronomers.

### 13.1.1 Van Cittert-Zernike Theorem

The Van Cittert-Zernike theorem, established by the Dutch physicists Pieter Hendrik van Cittert and Frits Zernike in the 1930s, describes how the electromagnetic radiation of an

incoherent source corresponds to the electric field measured by two correlated antennas. Originally, their formalism was derived for optical light, but is also valid for radio waves [44, 180]. Briefly said, the emitted radiation and the measured electric field are connected via a Fourier transform. The brightness distribution emitted by an astronomical source can be calculated by the Fourier transform of a correlated signal measured by two antennas on Earth. Following Clark [45], this idea will be described in detail in the following. Boldface symbols indicate vectors.

Consider a source of electromagnetic radiation far away from an observer at location  $\mathbf{R}$ . The electric field coming from that source can be described as  $\mathbf{E}(\mathbf{R}, t)$ .



**Figure 13.2:** Propagation of a radio wave component  $E_\nu(\mathbf{R})$  from a source at location  $\mathbf{R}$  to two antennas at  $\mathbf{r}_1$  and  $\mathbf{r}_2$ . The vector  $\mathbf{b} = (\mathbf{r}_2 - \mathbf{r}_1) = (u, v)$  between  $\mathbf{r}_1$  and  $\mathbf{r}_2$  is called *baseline*.  $(u, v)$  denote the array coordinates,  $(l, m)$  are the coordinates of the celestial sphere where the brightness distribution of the source is projected on.

This electric field can now be expressed as a sum of its complex Fourier components  $\mathbf{E}_\nu(\mathbf{R})$ , so-called *quasi-monochromatic components* which removes the time-dependency. Every component will now be observed at a location  $\mathbf{r}$  far away from  $\mathbf{R}$ . Summing up all components leads to the general form of

$$\mathbf{E}_\nu(\mathbf{r}) = \int \int \int P_\nu(\mathbf{R}, \mathbf{r}) \mathbf{E}_\nu(\mathbf{R}) \, dx \, dy \, dz , \quad (13.2)$$

with the propagator  $P_\nu(\mathbf{R}, \mathbf{r})$ , describing how the electric field at location  $\mathbf{r}$  is affected by changes in the electric field at  $\mathbf{R}$ . For this very general form, Clark makes several simplifications: First, the vector nature of electromagnetic radiation is neglected, and it

is considered as a scalar field  $E$  at every point. Here, this assumption is taken for the sake of simplicity and has to be reverted later. The second simplification lies in the nature of astronomical observations itself and is true for all radio interferometry measurements: for very distant sources, we will always measure a source in two dimensions, the so-called *brightness distribution*. The third dimension (depth) can not be resolved. Therefore, we can consider this surface of the source to be projected to a celestial sphere of a very large radius  $|\mathbf{R}|$ . The projected electric field is denoted with  $\epsilon_\nu(\mathbf{R})$ . As the third simplification, the space between the source at  $\mathbf{R}$  and the observer at  $\mathbf{r}$  is considered to be empty. This is, of course, not true for real-life observations, where amongst others, the ionosphere or human-made radiation will always influence the electromagnetic waves coming from an astronomical object.

After all these assumption, Equation (13.2) becomes

$$E_\nu(\mathbf{r}) = \int \epsilon_\nu(\mathbf{R}) \frac{e^{2\pi i \nu |\mathbf{R}-\mathbf{r}|/c}}{|\mathbf{R}-\mathbf{r}|} dS \quad (13.3)$$

with  $dS$  the surface element of the celestial sphere.

Equation (13.3) describes a general relation how an electric field at  $\mathbf{R}$  (in the sky) would be measured at a certain point  $\mathbf{r}$  (on Earth). We can now measure the electric field at two (telescope) locations  $\mathbf{r}_1$  and  $\mathbf{r}_2$  and correlate the signal, like

$$V_\nu(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{E}_\nu(\mathbf{r}_1) \mathbf{E}_\nu^*(\mathbf{r}_2) \rangle , \quad (13.4)$$

which will resume in

$$V_\nu(\mathbf{r}_1, \mathbf{r}_2) = \int \langle |\epsilon_\nu(\mathbf{R})|^2 \rangle |\mathbf{R}|^2 \frac{e^{2\pi i \nu |\mathbf{R}-\mathbf{r}_1|/c}}{|\mathbf{R}-\mathbf{r}_1|} \frac{e^{-2\pi i \nu |\mathbf{R}-\mathbf{r}_2|/c}}{|\mathbf{R}-\mathbf{r}_2|} dS . \quad (13.5)$$

We will denote the unit vector  $\frac{\mathbf{R}}{|\mathbf{R}|}$  as  $\mathbf{s}$  and the observed intensity of the electric field as  $\langle |\epsilon_\nu(\mathbf{R})|^2 \rangle |\mathbf{R}|^2 = I_\nu(\mathbf{s})$ . The surface element  $dS$  is replaced by  $|\mathbf{R}|^2 d\Omega$ , ignoring small terms of order  $\frac{|\mathbf{r}|}{|\mathbf{R}|}$ , assuming that the source is at a far greater distance to Earth than the telescopes to each other, which is true for all applications in radio astronomy. Equation (13.5) becomes

$$V_\nu(\mathbf{r}_1, \mathbf{r}_2) = \int I_\nu(\mathbf{s}) e^{-2\pi i \nu (\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{s} / c} d\Omega . \quad (13.6)$$

At this point, Equation (13.6) depends only on the distance of the two telescope positions  $\mathbf{r}_1$  and  $\mathbf{r}_2$ . The term  $\mathbf{r}_1 - \mathbf{r}_2$  is called *baseline*  $\mathbf{b}$  in radio-astronomy language. By correlating  $N$  telescopes,  $\frac{N(N-1)}{2}$  baselines are available. For successful radio interferometry, varying baselines are needed to measure different Fourier components since one baseline covers one Fourier component. This is why antennas of radio interferometers are not evenly positioned and look like they were randomly placed by the construction workers.

A baseline is always given in the reference frame of the source. If Earth rotates, the positions of the antennas change with respect to the position of the source. This way, more and more baselines are covered during the observation time just by the rotation of the Earth, see Figure 13.3. These baselines are called *projected* baselines. The more baselines exist,

the more Fourier components of the source's brightness distribution can be detected and the better the radio image will be.

A baseline can now be expressed in terms of the wavelength  $\lambda$  by  $(\mathbf{r}_1 - \mathbf{r}_2) = \lambda \cdot (u, v, w = 0)$ . The third dimension  $w$  is set to zero if we assume all telescopes to be located at the same plane. Compared to the distances of a source from Earth, this assumption even holds true for telescopes distributed on the spherical surface of Earth.

The components of the unit vector can be stated as  $\mathbf{s} = (l, m, \sqrt{1 - l^2 - m^2})$  which represent the coordinates on the sky. Equation (13.6) then becomes

$$V_\nu(u, v) = \int \int I_\nu(l, m) \frac{e^{-2\pi i(ul+vm)}}{\sqrt{1-l^2-m^2}} dl \, dm . \quad (13.7)$$

Equation (13.7) offers the secret, how single, spatially distributed antennas are able to image a brightness distribution at the sky: The measured electric field at a certain frequency  $\nu$   $V_\nu(u, v)$ , also called *visibility*, is the Fourier transform of the brightness distribution  $I_\nu(l, m)$  of a distant radio source. Since a Fourier transform can be inverted, the brightness distribution at the sky can be expressed by the measured visibilities on Earth:

$$I_\nu(l, m) \propto \int \int V_\nu(u, v) e^{2\pi i(ul+vm)} du \, dv . \quad (13.8)$$

The more telescopes and, therefore, baselines at  $(u, v)$  are available, the more components of the Fourier sum of the original electric field can be detected, and the more precise the resulting image will be. Ideally, there would be a baseline for all combinations of  $u$  and  $v$  available. In that case, the image of the source brightness could be restored without losses.

In reality, the  $uv$  space is sampled sparsely since the number of baselines is always limited. Additionally, influences from Earth's ionosphere, other noise, and detector issues affects the measurement. We can therefore not easily apply Equation (13.8) and treat the result directly as an image of the source. Similar to gamma-ray astronomy, we have an inverse problem, which sounds familiar from section 5.5. The true brightness distribution  $I_\nu(l, m)$  is convoluted with a detector response  $B(l, m)$ , which is called *synthesized beam* or PSF, analogous to IACTs. The resulting image is called *dirty image*  $I_\nu^D(l, m)$ :

$$I_\nu^D(l, m) = I_\nu(l, m) * B(l, m) \propto \int \int V_\nu(u, v) S(u, v) e^{2\pi i(ul+vm)} du \, dv \quad (13.9)$$

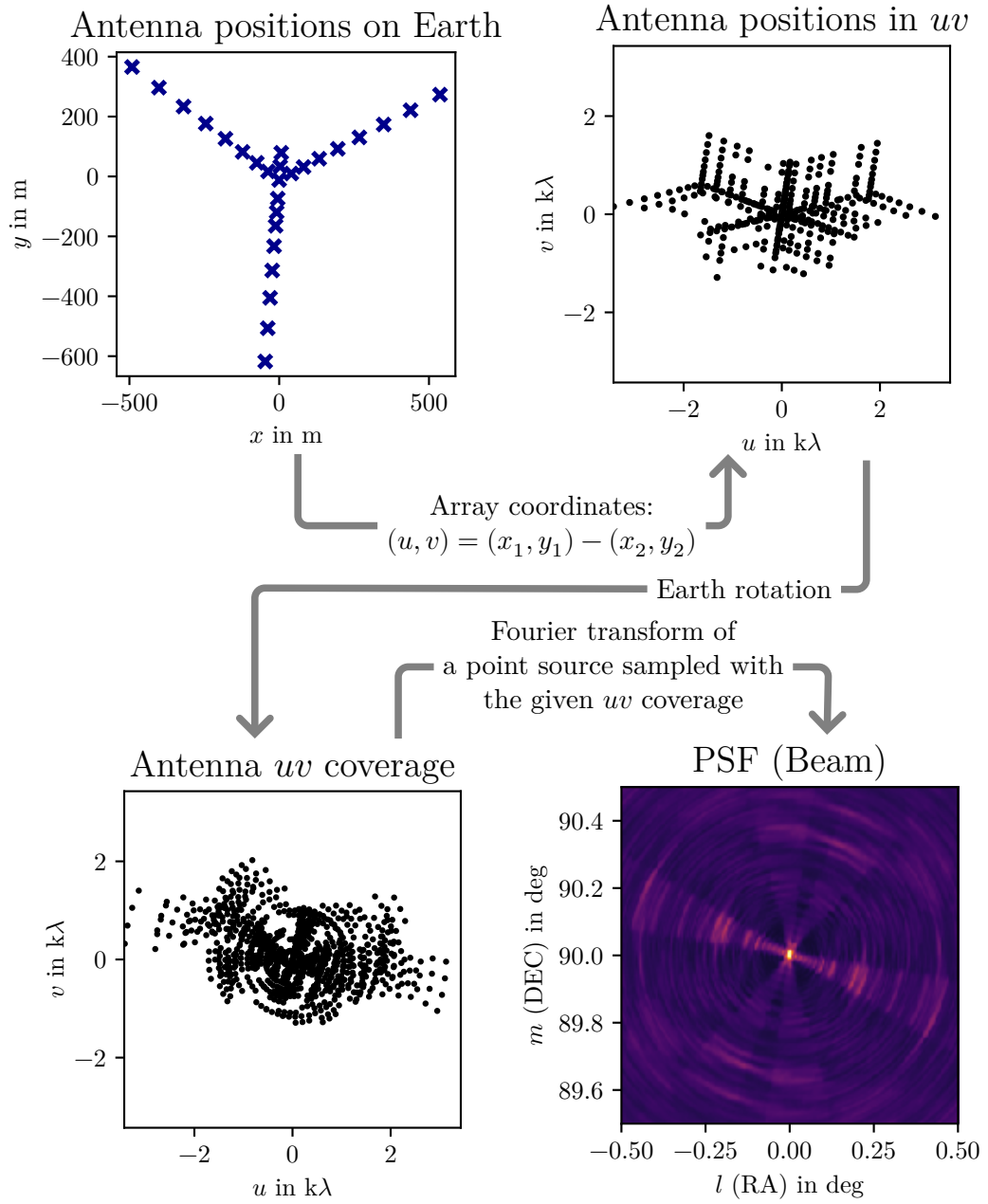
with

$$B(l, m) = \int \int S(u, v) e^{2\pi i(ul+vm)} du \, dv \quad (13.10)$$

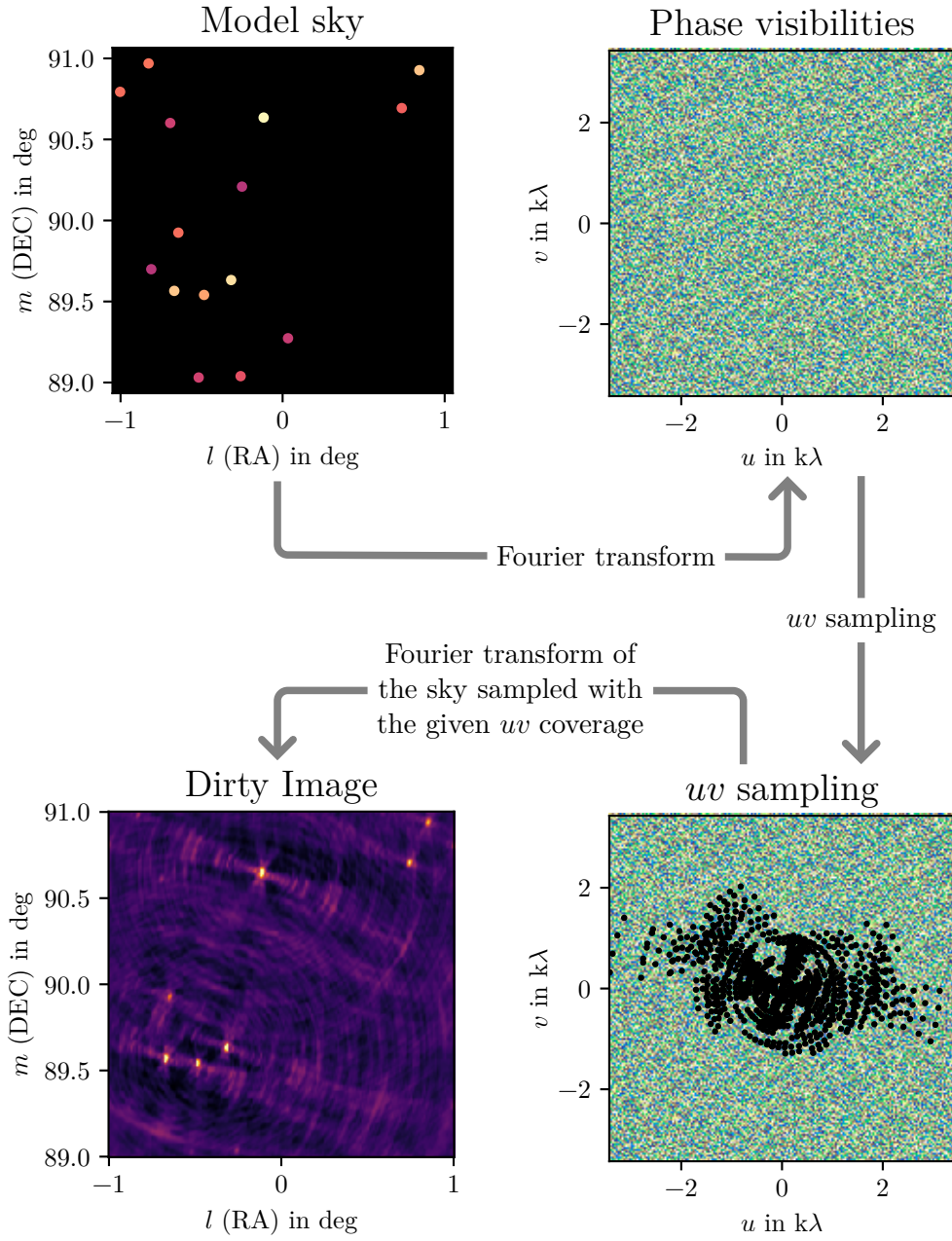
with the sampling function  $S(u, v)$  expressing the limited number of baselines.

After obtaining this dirty image from the measurements, further analysis steps are necessary to solve the inverse problem, remove all artifacts and restore the original brightness distribution as good as possible. This process is described in subsection 14.1.1.

Figure 13.4 and Figure 13.4 provide an overview of the described  $uv$  sampling in radio interferometry.



**Figure 13.3:** Radio interferometry in a nutshell, part I: The antenna positions in  $x, y$  (upper left, taken from the JVLA) are converted to the baseline vectors in  $u, v$  (upper right). If the Earth rotates, additional projected baselines are created (lower left). Sampling a point source at the covered baselines results in the *dirty beam* or PSF (lower right). Images here and on the next page were created based on [134].



**Figure 13.4:** Radio interferometry in a nutshell, part II: The Fourier transform (upper right) of the sky brightness distribution (upper left) is sampled at all projected baselines (lower right). The dirty map (lower left), obtained by sampling the visibilities, can now be considered as a convolution of the true sky brightness with the dirty beam (last page, lower right). The dirty image contains a lot of noise and artefacts that have to be removed later. For the sake of a simpler arrangement, the color bars are not shown. The flux in radio astronomy is always given in Jy/beam. Images were created based on [134].



### 13.1.2 Units in Radio Astronomy

If we integrate the brightness distribution of the source over a certain solid angle, we end up with the *flux density*. The spectral flux density in radio astronomy is given in the non-SI unit of Jy named after Karl Guthe Jansky:

$$1 \text{ Jy} = 10^{-26} \frac{\text{W}}{\text{Hz m}^2} . \quad (13.11)$$

A radio map is a grid of quadratic pixels that are smaller than the beam. The pixel value is usually given in Jy/beam, which is the flux density that we would measure if we point a certain beam exactly at this pixel. The sum over all pixels in a radio map does therefore not result in the integral over the surface brightness but must be corrected for the beam size, which is represented as an ellipse with axes of the size of the Full Width Half Maximum (FWHM) of the beam.

## 13.2 Very Long Baseline Array

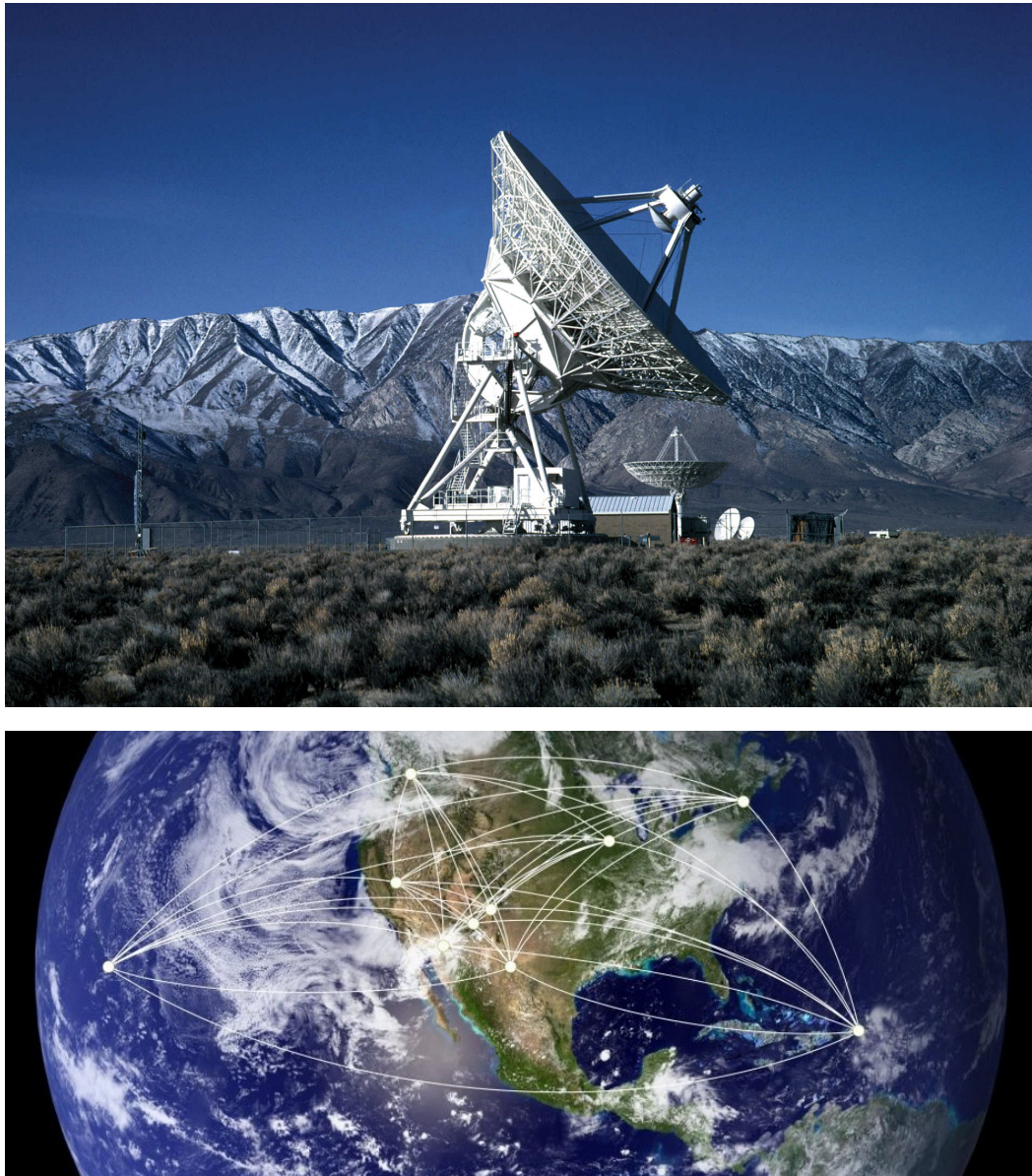
The Very Long Baseline Array (VLBA) of the National Radio Astronomy Observatory (NRAO) is one of the world's largest radio interferometers. It consists of ten parabolic antennas spread over the United States [121]. All ten dishes are of identical design with a diameter of 25 m and sensitive to radio waves of 90 cm to 3 mm which is equivalent to 312 MHz to 96 GHz. This uniform design reduces the amount of work necessary for the correlation and calibration of the data. The longest baseline with 8611 km spans between Hawaii and the Virgin Islands. The VLBA was built between 1986 and 1993 and is operating since then without large interruptions.

The radio interferometer is not only used for observations of astronomical sources. By observing pulsars, the Earth orientation parameter UT1-UTC is determined by three antennas of the VLBA once a day. With this parameter, the current rotation of the Earth around its own axis can be described and, therefore, the exact time. Furthermore, geophysical studies are made by precisely measuring the distances of the different antennas with pulsar observations. This way, movements of Earth's crust or the effects of earthquakes can be studied [34].

### 13.2.1 Boston University Blazar Monitoring Program

The data used for this thesis is obtained by the BU-BLAZAR program, a monitoring program managed by the University of Boston that started in 2007 [99]. Within this program, 34 blazars and three radio galaxy jets are observed once per month at 43 GHz. One of them is 3C 84, which is observed since 2010. The clean images are created by the Boston team and are publicly available, see appendix F.





**Figure 13.5: Top:** VLBA radio antenna in Owens Valley, California, with a diameter of 25 m. All antennas of the VLBA are of identical design. Image Credit: NRAO/AUI/NSF

**Bottom:** Locations of the ten VLBA antennas spread over the USA. The longest baseline runs from the Mauna Kea on Hawaii to St. Croix on the Virgin Islands. Further antennas are located at Hancock (New Hampshire), North Liberty (Iowa), Fort Davis (Texas), Los Alamos (New Mexico), Pie Town (New Mexico), Kitt Peak (Arizona), Owens Valley (California), and Brewster (Washington). Image credit: Jeff Hellerman, NRAO/AUI/NSF



## Radio Data Analysis

---

In contrast to astroparticle physics, the data obtained by radio interferometers requires an image-based analysis. The signal of the observed source is not mapped by individual events but by the features of the electromagnetic waves we detect with the radio interferometer. The low-level part of this data analysis, especially the calibration with all its details and challenges, will not be discussed in this thesis, but I will depict the general idea behind radio imaging briefly.

### 14.1 Imaging

In [subsection 13.1.1](#), I described the importance of the Fourier transform for radio interferometry and the need for a good  $uv$ -coverage. [Equation \(13.9\)](#) gives the mathematical formulation of the dirty image, which is the true brightness distribution convoluted with the sampling function, the beam. In principle, we have the same underlying problem here as for event-based gamma-ray astronomy: We have a ground truth folded with an unknown detector function, which we can not simply separate from each other. In contrast to gamma-ray astronomy, where MC simulations play an important role, image deconvolution in radio astronomy works differently. I will present here a simple and classical approach, which is used by the Boston team to create the images used for this thesis. More advanced techniques are slowly making their way into the VLBI community.

The classical imaging in radio interferometry can be done with `Difmap` [\[147\]](#), a well-established (published in 1994) but still used and simple program to manually deconvolve the dirty images and obtain the clean maps. In the next section, I will briefly discuss how this is done.

#### 14.1.1 Difmap and CLEAN

We start with the dirty image in [Equation \(13.9\)](#), which contains the true brightness distribution convoluted with the dirty beam:

$$I_\nu^D(l, m) = I_\nu(l, m) * B(l, m) \quad (14.1)$$

$$I_\nu^D(l, m) \propto \int \int V_\nu(u, v) S(u, v) e^{2\pi i(ul+vm)} du dv . \quad (14.2)$$

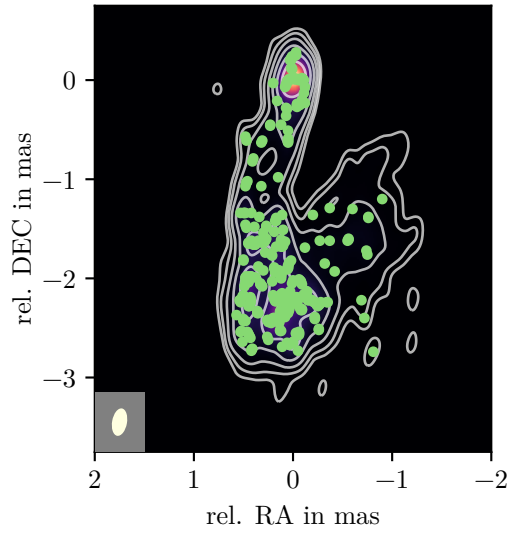
The basic idea is now to iteratively model the flux density with delta functions, beginning with the most prominent features and subtract the model from the dirty image. In the next step, new model components are added by modeling the remaining flux density with additional delta functions. For every step, the model is convolved with the beam of the

detector and subtracted from the dirty image. Step by step, a source model is created and a residual map remains. The iteration stops when the residual map has reached a certain flux level. To create the *clean map*, the model is convoluted with the beam and the residual noise is added. These clean maps are now our estimation of the true brightness distribution and can be used for any further scientific high-level analysis.

If a clean map was produced successfully, it is mostly presented in publications as a contour map with lines in logarithmic distance. The lowest contour is usually set to  $5\sigma$  where  $\sigma$  is the RMS of the noise in an area of the image where no signal from the observed source is expected. Furthermore, the size of the PSF, or *beam*, which was convoluted with the model components, is given. This way, a researcher can estimate the resolution of an image and compare the size of the PSF to features in the map.

Figure 14.1 shows a clean map together with the position of Difmap’s delta components.

**Figure 14.1:** Radio map of 3C 84 on 2013-06-16 observed by the VLBA-BU-BLAZAR program at 43 GHz. Contourlines are shown at 0.009, 0.02, 0.06, 0.15, 0.40, 1.05 and 2.76 Jy/beam. The beam size is depicted in the bottom left corner. The green dots mark the positions of the delta components computed by Difmap used to build the flux model. Only components with a flux value  $< 5\sigma$  are shown.



The method implemented in Difmap, is called *Difference Mapping* and was described as CLEAN algorithm by Jan Högbom in 1974 [96]. His work had a great influence on the data analysis in radio interferometry and is still the foundation for most of the state-of-the-art deconvolution frameworks. (See, for example, [50] for further reading.)

During this cleaning process, the so-called *self-calibration* can be performed to reduce corruptions in the measured visibilities from hardware issues or atmospheric changes. Self-calibration only works for datasets with a high signal-to-noise ratio. To self-calibrate an image, a model of the source is created with the process described above. Between the cleaning steps, the amplitudes and phases of the measured visibilities are re-calibrated to fit the model. After some iterations, when a noise threshold is reached, the imaging process starts from scratch but using the self-calibrated phases and amplitudes. Self-calibration has some caveats but can heavily increase the effective sensitivity. For further reading see [129, 164].

In addition to the mentioned issues, further effects originating from the antennas, the baselines, the atmosphere, and the software have to be considered.

Based on the final model resulting from the cleaning and calibration, the scientific high-level analysis is performed. A common goal is to estimate the flux density in a radio hot spot or another bright feature in an image and its movement over time. To derive this information from a radio image, Difmap is capable of fitting multiple 2D Gaussian distributions to the model and integrate the estimated flux. For this approach, a user has to set starting parameters for the position where this Gaussian should be fitted.

The attentive reader might have noticed that the described procedure has one huge disadvantage: the source model is built from delta components that are placed in a region of the image that was selected by a user. The user decides which parts of the image they label as the source region and what is considered to be noise and artifacts from the sparse  $uv$ -coverage. For the Gaussian modeling, also start values have to be given by a user. Difmap relies on a lot of human interaction, which makes the results non-reproducible and the analysis of large data sets very hard. Furthermore, for one image, roughly ten to hundred iterations are needed to obtain a good result.

Since its invention in the 1970s, a lot attempts were made to improve this algorithm and automatize the choice of model components, for example CASA [157, 114] and wsclean [126]. New approaches try to attack the deconvolution issue with ML methods, which would substantially speed up the data reconstruction and reduces any amount of human interaction to a minimum.

## 14.2 Image Segmentation: Random Walker

As mentioned in the last section, Difmap offers the possibility to model bright regions in an image with Gaussian distributions and calculate their flux density. For this work, I aim to study the evolution of the radio components C1 and C3 of 3C 84 over several years, c.f. Figure 2.8. This results in 91 images from which I have to select the regions of the radio components and calculate their flux density. To reduce the human interaction at this analysis step and enforce reproducibility, I used another approach than Difmap here to obtain a component's flux density from the model created with the CLEAN algorithm.

Selecting a certain region of interest in an image is a classical problem for image-processing algorithms. In this case, the subgroup of image-segmentation algorithms seems favorable to solve this task. A widely used image-segmentation algorithm is the *Random Walker Algorithm* [88], which uses a random walker to assign every pixel of an image with pre-defined labels. The random walk is a statistical model to build a chain of movements created from independent random variables  $Z_n$  in arbitrary dimensions. Every variable is drawn from the same distribution. For the application to image segmentation, a random walk is performed on the pixel grid of an image. The probability for a step from one pixel to another is calculated from the pixel values. Grady [88] shows that the solutions of the combinatorial Dirichlet problem can be used to define these probabilities. In this work, the detailed calculations are not elaborated further, but can be found in [88].

In the context of image segmentation, the random walker algorithm is used as follows, see also Figure 14.2:

1. For every segment (or: radio component), one pixel which surely belongs to this

segment has to be marked (e. g. with ascending integer numbers).

2. For every unmarked pixel, a probability is calculated for a random walker to reach a labeled pixel starting from this unlabeled pixel.
3. For every pre-defined segment, all pixels are labeled with an associated probability above a certain threshold.

For this work, I use the random walker algorithm that is implemented in the `scikit-image` package [171] which provides image processing tools in `Python`.

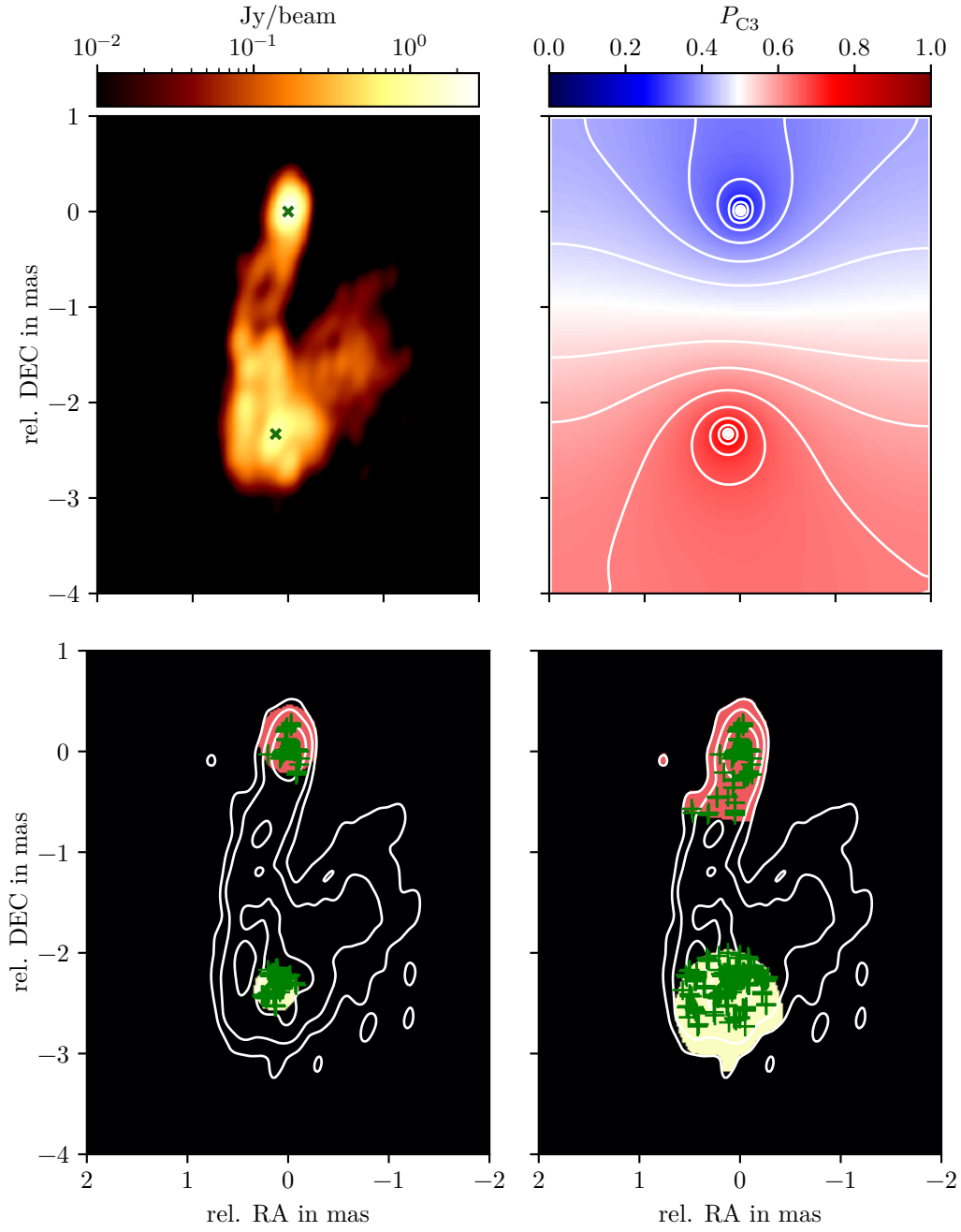
For selecting these regions, a pixel is marked in both components (C1 and C3) that certainly belongs to the respective component. After the random walker algorithm was applied to the image, every pixel is associated with a probability between 0 and 1 to belong to C1 or C3. By defining a threshold on these probabilities, pixels belonging to C1 and C3 are selected. To measure the flux density deposited in the selected regions, all model components (delta functions) of the `Difmap` model that lie within a segment are summed up.

In contrast to the traditional technique of fitting Gaussian distributions to the brightness distribution, my approach comes with several advantages. As mentioned, image segmentation can be executed automatically and does not require user interaction for every image. Furthermore, the selected regions do not have to be of elliptical shape but can have any possible form. This includes structures like the line-shaped outer regions of an edge-bright jet that can be selected from an image this way.

The flux density of a bright region calculated via image segmentation depends on the amount of delta components that are assigned to a region. Usually, the model components in the center of a bright region are orders of magnitude brighter than components at the edge of the segment. Therefore, the amount of these edge-positioned components do not heavily affect the total flux in a segment. To verify this statement and to provide uncertainties for the so obtained total flux values, the parameters of the random walker were varied to create a sample of different segmentations.

With the described approach, I am able to automatically monitor the flux density of the two radio components of 3C 84, C1 and C3, over a long period of time. The resulting light curves can then be used for further scientific studies combined with data from other wavelengths, as for example, cross-correlation studies which are presented in [chapter 15](#).





**Figure 14.2:** Random walker image segmentation to define the region of a radio component within a clean map. Upper left: Clean map plotted in logarithmic color scale with seeds for the cluster marked with the x. Upper right: Probability for every pixel to belong to component C3. Lower panels: Segmented areas selected by different probability thresholds for  $P_{C3}$ . The included model components, which contribute to the flux density of the selected area, are marked with green crosses.





## Flux Variations in 3C 84's Radio Components

---

### 15.1 Identification of the Radio Components in VLBA Radio Images

In [section 2.4](#) the prominent radio features of 3C 84 were already described. The core of the AGN, which is assumed to host the BH, is mapped as the radio component C1, a component quite steady in size and flux. This component is mostly resolved as of the size of the beam, which leads to the conclusion that the true radiating region can be much smaller than depicted in a radio map. The component C2 was detectable in the 2010s only as a diffuse remnant of a former component and has nearly vanished these days. This is why I do not undertake any flux measurement for C2 in this work. The component C3 was ejected from the core around 2005 and moved southwards until it broadens and fades out around 2018, see [Figure 2.8](#).

Especially this component is of great interest when we study the gamma-ray emission region since we observe an increase in gamma-ray flux around the same time with *Fermi*-LAT, see [Figure 12.2](#).

The obvious question, which arises after inspecting the light curves measured at the radio and gamma-ray range, is if these light curves and the mechanisms responsible for it are connected. This question can be addressed by a cross-correlation study, as described in [chapter 17](#).

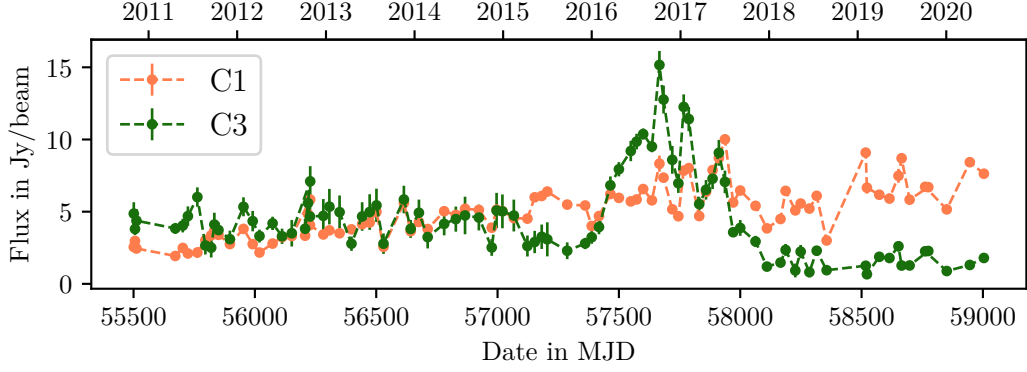
To be able to perform this cross-correlation study, the flux density of the individual components must be extracted from the radio maps first.

This task is performed as described in [section 14.2](#) and returns a possibility for every pixel in a radio map to belong to each component. Using the `scikit-image` implementation, the parameter `beta` can be used to modify the algorithm by scaling the gradients. The greater `beta`, the more difficult is the random walker motion and the lower the possibility for pixels at a great distance to a marker pixel to belong to that associated segment. For this work, `beta` is sampled uniformly as between 10 and 30 to vary the subset of flux components per segment. This is used to estimate the uncertainty of the resulting total flux density.

To select C1 and C3 in an image, we take all pixels with a probability to belong to a certain segment above a given threshold. Since this threshold also directly affects the amount of delta components accounted to the component, the threshold is sampled uniformly between 0.6 and 6.5 for C1 and between 0.65 and 0.7 for C3. These limits are hand-tuned so that the brightest parts of the radio knots are always included in the selected region, but the regions do not fill the whole  $5\sigma$  contour. In the future, an automated or rule-based approach to set these boundaries has to be developed.

The random walker will, by default, label every pixel in an image. In the case of radio maps, only the source region is of interest, which is typically defined as flux density above five

times the standard deviation of the image noise ( $5\sigma$ ). Therefore, a second filter, additional to the probability threshold, is applied by this flux density threshold. Again, to estimate the systematical errors, the noise factor is sampled uniformly between 5 and 8. Having all these steps done, a group of pixels is selected for C1 and C3 for every epoch and all model components from the `Difmap` flux model can be summed up to calculate the total flux in every component. The uncertainties are obtained by repeating the segmentation 100 times with different parameters. The results are shown in Figure 15.1.



**Figure 15.1:** Flux density in Jy/beam for radio components C1 and C3 for all epochs of the BU-BLAZAR program's 43 GHz data.

## 15.2 Total Flux Calibration

The radio light curves created with the random walker image segmentation show strong fluctuations, especially in the epochs of 2016/2017, see Figure 15.1. These fluctuations outrange the uncertainties and are very unusual in AGN radio light curves. It is therefore likely that these strong fluctuations are caused by calibration issues and not by true flux variabilities.

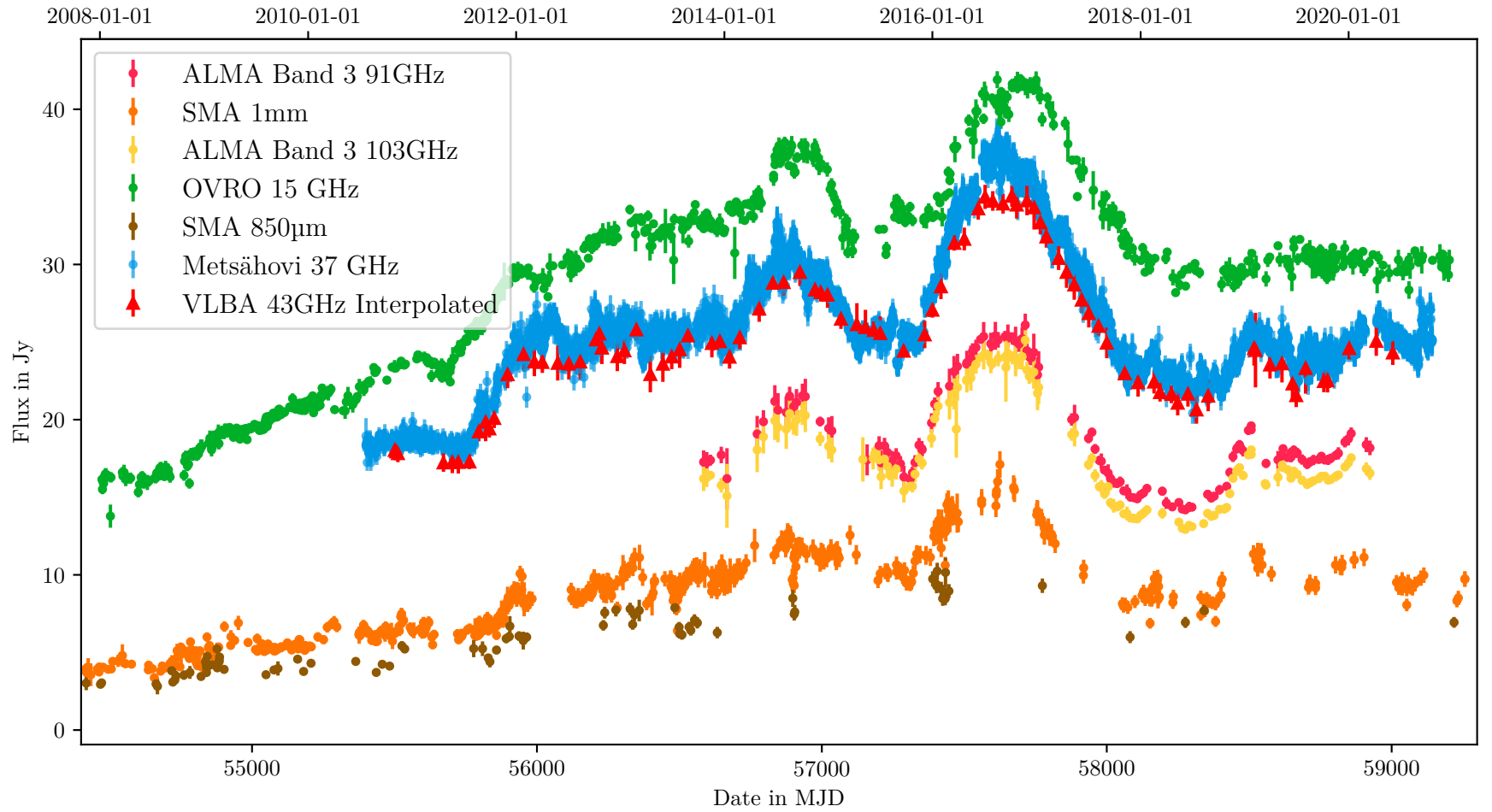
During a usual radio observation, so-called *calibrator sources* with a steady and well-known flux are observed to calibrate the detector for measuring the target source. Additionally, the self-calibration can be performed, as described in subsection 14.1.1. In the case of the present data, both approaches seem to fail. Informal communication with the BU-BLAZAR program team confirmed that there were some total flux calibration issues during the high flux state of 3C 84 in 2016/2017, which was probably caused by mis-calculating the diffuse emission of the extended source.

Fortunately, 3C 84 was observed by other radio telescopes at other wavelengths, which offers the opportunity to re-calibrate the data with an interpolation of flux densities from other measurements. For this work, data from the Atacama Large Millimeter/submillimeter Array (ALMA), the Metsähovi Radio Observatory, the Owens Valley Radio Observatory (OVRO) and the Submillimeter Array (SMA) is used. The references and data usage statements can be found in appendix F. From these datasets, quasi-simultaneous radio observations with a maximal time delay of 15 days are selected. Every subset is then fitted

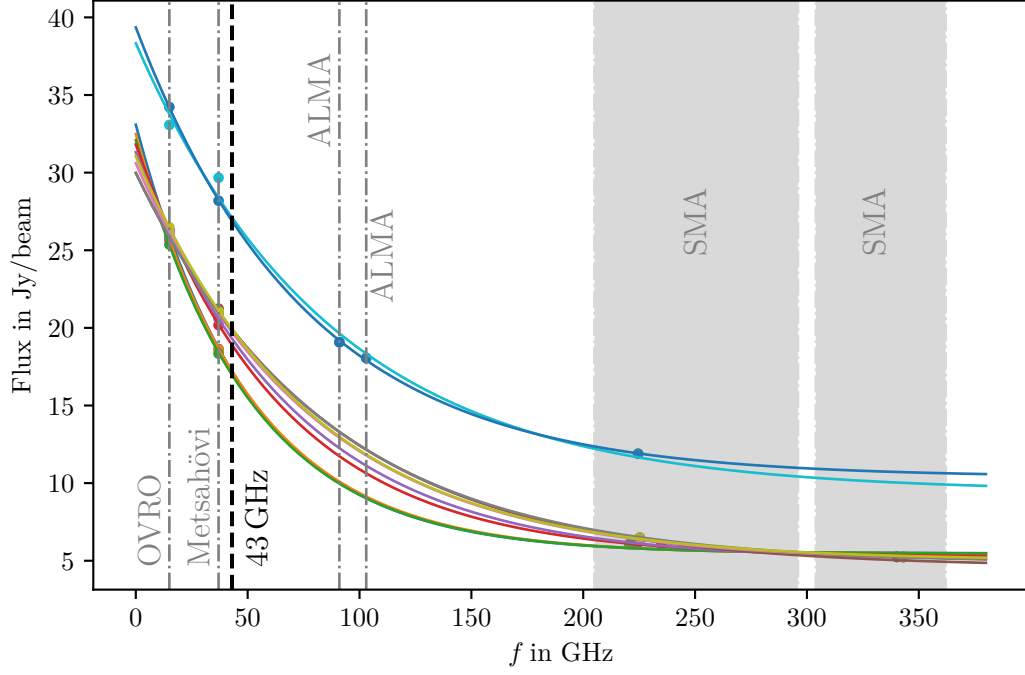
with a power-law function to obtain the flux at 43 GHz. The used measurements and their wavelengths are shown in Table 15.1, the exponential fit for some example epochs is shown in Figure 15.3.

**Table 15.1:** Data used for total flux calibration of the VLBA data at 43 GHz.

Observatory	Location	Wavelength in GHz
ALMA	Atacama Desert, Chile	91, 103
Metsähovi	Finland	22, 37
OVRO	California, USA	15
SMA	Mauna Kea, Hawaii	205-296 (1 mm), 304-362 (850 nm)



**Figure 15.2:** Radio light curves of 3C 84 from various instruments that are used to calibrate the flux density at 43 GHz. The interpolated flux obtained from this calibration is plotted with red triangles.

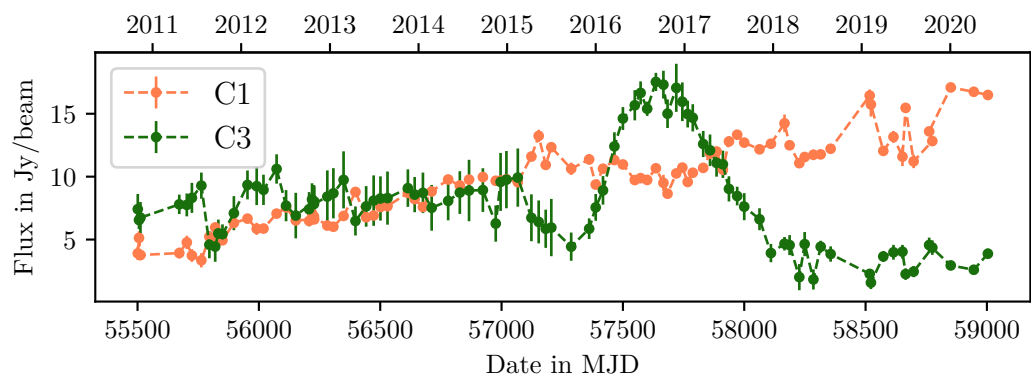


**Figure 15.3:** Exponential fit to calibrate the flux at 43 GHz (black dashed line). Dot-dashed lines mark the frequencies of the datasets used for calibration.

The total uncalibrated flux density is given by the sum over all model components above the threshold used for the random walker image segmentation. The calibration factor between the interpolated and the observed total flux is given as

$$g = \frac{F_{\text{interpolated}}}{F_{\text{observed}}} . \quad (15.1)$$

The calibrated flux values for C1 and C3 are then given as the product of the uncalibrated flux and the calibration factor. The calibrated radio light curves are shown in [Figure 15.4](#).



**Figure 15.4:** Calibrated flux in Jy/beam for radio components C1 and C3 for all epochs of the BU-BLAZAR program's 43 GHz data.

## **Part V**

### **Multiwavelength Analysis**





# Constraining the Gamma-Ray Emission Region in NGC 1275

# 16

## 16.1 Theoretical Groundwork

As stated in subsection 2.3.4, the BLR can be visualized as collection of gas clouds surrounding and rotating with the BH. The photons contained in this region serve as *seed photons* and can be scattered to very high energies by electrons moving at relativistic speeds. Such electrons are provided in the jet's plasma. Other seed-photon fields exist from the dust torus and the accretion disk. The same photon populations which are accelerated via IC scattering can, in turn, absorb gamma rays by pair production.

Imagine a blob of plasma moving at relativistic speed along the jet axis with a distance  $r$  to the BH and a distance  $x$  to a *reprocessing region*, where radiation from the accretion disk is absorbed and re-radiated as line emission. This reprocessing region has a distance  $R_{\text{re}}$  from the center of the BH in an angle  $\theta_{\text{re}}$  to the jet axis, see Figure 16.1. The angle between the blob-reprocessing region distance  $x$  and the jet axis is denoted as  $\theta_*$ . All angles are expressed as  $\mu \equiv \cos \theta$ .

Photons that would be accelerated by the plasma blob very close to the BH have to pass the gas clouds of the BLR. On their way, the high energetic photons might interact with BLR photons and get absorbed due to pair production, depending on their energy, the characteristics of the BLR, and, most importantly, on their point of emission along the jet. Gamma rays emitted very close to the BH have to travel a longer distance through the BLR and have, therefore, a higher probability to get absorbed compared to photons that are produced at the outer edge of the BLR.

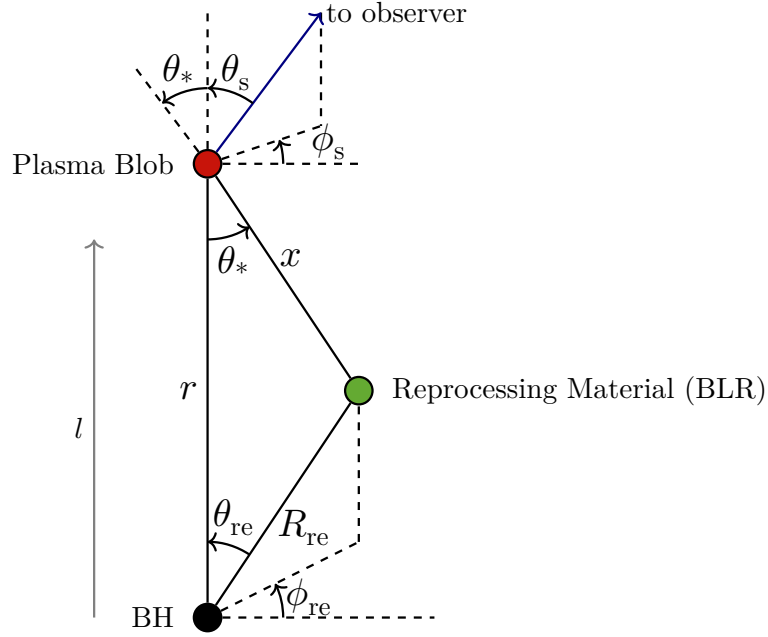
Taking this absorption into account, the intrinsic energy spectrum at high energies is modified by a factor of  $e^{-\tau_{\gamma\gamma}(\epsilon_1)}$ , with the optical depth  $\tau_{\gamma\gamma}$ . The analytical expression of  $\tau_{\gamma\gamma}$  was derived by Finke for the absorption of gamma-ray emission by the BLR, the dust torus, and the accretion disk [73]. In case of the BLR,  $\tau_{\gamma\gamma}$  depends on the energy of the particle, the distance of the emission region from the BH and the geometric shape of the BLR.

Following [73], the general optical depth for photoabsorption is given by

$$\tau_{\gamma\gamma}(\epsilon_1) = \int_r^\infty dl \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \cdot (1 - \cos\Psi) \int_0^\infty d\epsilon \frac{u(\epsilon, \Omega; l)}{\epsilon m_e c^2} \cdot \sigma_{\gamma\gamma} \left( \frac{\epsilon \epsilon_1 (1 + z)}{2} (1 - \cos\Psi) \right) \quad (16.1)$$

with

- $\epsilon_1$  the energy of the gamma ray
- $r$  distance of the plasma blob from the BH



**Figure 16.1:** Geometry of the reprocessed emission of the BLR as seed photons for Compton scattering in the plasma blob, as described in [73].  $l$  denotes the direction of the jet.

- $l$  distance from the BH along the jet
- $\phi$  azimuth angle of the gamma ray's direction
- $\mu \equiv \cos \theta$  polar angle of the gamma ray's direction, with  $\theta$  as angle between jet axis and line of sight
- $\Psi$  angle between the gamma ray direction and the direction of a BLR photon
- $u(\epsilon, \Omega; l)$  energy density of the seed photon field, depending on the energy  $\epsilon$  and solid angle, for a distance  $l$  from the BH
- $m_e$  electron mass
- $c$  speed of light
- $z$  redshift

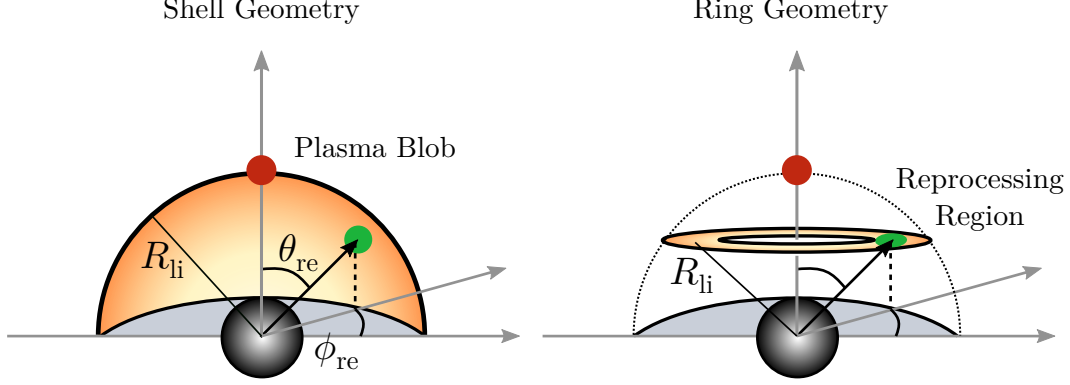
The Compton cross-section is given by

$$\sigma_{\gamma\gamma}(s) = \frac{3}{8} \sigma_T (1 - \beta_{\text{cm}}^2) \left[ (3 - \beta_{\text{cm}}^4) \ln \left( \frac{1 + \beta_{\text{cm}}}{1 - \beta_{\text{cm}}} \right) - 2\beta_{\text{cm}} (2 - \beta_{\text{cm}}^2) \right] \quad (16.2)$$

with the Thomson cross-section  $\sigma_T$  and

$$\beta_{\text{cm}} = \sqrt{1 - s^{-1}}. \quad (16.3)$$

Specifying Equation (16.1) for the BLR, the energy density  $u(\epsilon, \Omega; l)$  for a seed photon field of photons from line emission has to be inserted. This energy density is different for the two geometries introduced in subsection 2.3.4, the *flattened ring geometry* and the *spherical shell geometry*, see Figure 16.2.



**Figure 16.2:** Sketch of the *shell* and *ring* geometries of the BLR as used to describe  $u(\epsilon, \Omega; l)$  in [73]. The green area marks the reprocessing region, the region where the seed photon field is produced. The red dot marks the position of the plasma blob, where the seed photons are scattered to very high energies.

For a **shell geometry**, we have

$$u(\epsilon, \Omega; l) = \frac{\xi_{li} L_{\text{disk}}}{(4\pi)^2 c} \delta(\epsilon - \epsilon_{li}) \int_{-1}^1 \frac{d\mu_{re}}{x^2} \delta(\mu - \mu_*) \quad (16.4)$$

with

$$\mu_* = \left[ 1 - \left( \frac{R_{li}^2}{x} \right)^2 (1 - \mu_{re}^2) \right]^{\frac{1}{2}} \quad (16.5)$$

and

- $\xi_{li}$ : the fraction of radiation from the accretion disk that is re-radiated by the BLR
- $L_{\text{disk}}$ : the accretion disk luminosity
- $\epsilon_{li}$ : the energy of the re-radiated photons for a specific line at a specific distance from the BH
- $R_{li}$ : radius of the shell where s specific line emission is produced

The delta approximation  $\delta(\epsilon - \epsilon_{li})$  expresses the single line energy that all photons from the same shell are assumed to have. This is why the final  $\tau_{\gamma\gamma}(\epsilon_1)$  has to be calculated for a known strong line of a source's BLR.

Inserting Equation (16.4) in Equation (16.1) results in

$$\tau_{\gamma\gamma}(\epsilon_1) = 900 \frac{\xi_{\text{li}} l_{\text{Edd}}}{\epsilon_{\text{li}}} \int_{\tilde{r}}^{\infty} d\tilde{l} \int_{-1}^1 \frac{d\mu_{\text{re}}}{\tilde{x}^2} \left[ \frac{\sigma_{\gamma\gamma}(\tilde{s})}{\sigma_{\text{T}}} \right] (1 - \mu_{\star}) \quad (16.6)$$

with  $\tilde{r} = \frac{r}{R_g}$  and  $\tilde{l} = \frac{l}{R_g}$ , where  $R_g$  is the gravitational radius of the BH and  $l_{\text{Edd}}$  is the Eddington luminosity. In this notation,  $\mu_{\star}$  becomes

$$\mu_{\star} = \left[ 1 - \left( \frac{R_{\text{li}}}{R_g \tilde{x}} \right)^2 (1 - \mu_{\text{re}}^2) \right]^{\frac{1}{2}}. \quad (16.7)$$

The rescaled distance between the blob and the reprocessing region  $\tilde{x}$  is given by

$$\tilde{x}^2 = \frac{R_{\text{li}}^2 + l^2 - 2lR_{\text{li}}\mu_{\text{re}}}{R_g^2} \quad (16.8)$$

and the Compton cross-section depends on

$$\tilde{s} = \frac{\epsilon_{\text{li}} \epsilon_1 (1 + z)(1 - \mu_{\star})}{2}. \quad (16.9)$$

Approximating the BLR's shape with a **ring geometry**, the electron density in Equation (16.1) is given by

$$u(\epsilon, \Omega; l) = \frac{\xi_{\text{li}} L_{\text{disk}}}{(4\pi)^2 c x^2} \delta(\epsilon - \epsilon_{\text{li}}) \delta\left(\mu - \frac{r}{x}\right) \quad (16.10)$$

with

$$x^2 = R_{\text{li}}^2 + r^2. \quad (16.11)$$

The optical depth becomes

$$\tau_{\gamma\gamma}(\epsilon_1) = 900 \frac{\xi_{\text{li}} l_{\text{Edd}}}{\epsilon_{\text{li}}} \int_{\tilde{r}}^{\infty} \frac{d\tilde{l}}{\tilde{x}^2} \left( 1 - \frac{\tilde{l}}{\tilde{x}} \right) \left[ \frac{\sigma_{\gamma\gamma}(\tilde{s})}{\sigma_{\text{T}}} \right] \quad (16.12)$$

with

$$\tilde{x}^2 = \left( \frac{R_{\text{li}}}{R_g} \right)^2 + \tilde{l}^2 \quad (16.13)$$

and

$$\tilde{s} = \frac{\epsilon_{\text{li}} \epsilon_1 (1 + z)(1 - \frac{\tilde{l}}{\tilde{x}})}{2}. \quad (16.14)$$

For the purposes of this study, we consider the radius of the Ly $\alpha$  line as the effective border of the BLR, because it causes the dominant contribution to the opacity inside this radius.

Having this specific line, the parameters in Equation (16.6) to Equation (16.14) can be obtained.  $\xi_{\text{li}}$  and  $\epsilon_{\text{li}}$  in Equation (16.6) and Equation (16.12) can be approximated using

Table 5 in the appendix of [73], which lists BLR parameters for various lines in units of the  $H\beta$  line. For NGC 1275 the luminosity of the  $H\beta$  line is given in [133] as

$$L(H\beta) = 8.94 \times 10^{40} \frac{\text{erg}}{\text{s}}. \quad (16.15)$$

The radius of the  $H\beta$  line can be calculated using the approximation given in [73, Appendix]

$$R(H\beta) = 10^{16.94 \pm 0.03} \left( \frac{L(5100 \text{ \AA})}{1 \times 10^{44} \text{ erg/s}} \right)^{0.533 \pm 0.035} \text{ cm} \quad (16.16)$$

$$(16.17)$$

with the relation from Greene and Ho [89]

$$\left( \frac{L(5100 \text{ \AA})}{1 \times 10^{44} \text{ erg/s}} \right) = \left( \frac{L(H\beta)}{(1.425 \pm 0.007) \times 10^{42} \text{ erg/s}} \right)^{0.8826 \pm 0.0039} \quad (16.18)$$

which results in

$$R(H\beta) = 2.37 \times 10^{16} \text{ cm}. \quad (16.19)$$

With this radius, the radius of the  $\text{Ly}\alpha$  is [73, Appendix, Table 5]:

$$R(\text{Ly}\alpha) = 0.27 R(H\beta) = 6.39 \times 10^{15} \text{ cm} = R_{\text{li}}. \quad (16.20)$$

For the fraction of re-radiated luminosity from the disk  $\xi_{\text{li}}$  from [73, Appendix, Table 5] gives:

$$\xi_{\text{Ly}\alpha} = \frac{12 \cdot L(H\beta)}{L_{\text{disk}}} \quad (16.21)$$

with a disk luminosity of  $L_{\text{disk}} = 1.26 \times 10^{46} \text{ erg/s}$ , which is the Eddington luminosity of a BH with  $M = 10^8 M_{\odot}$ . The Schwarzschild radius is given as  $R_g = 1.5 \times 10^{13} \text{ cm}$ .

The  $\tau_{\gamma\gamma}(\epsilon_1)$  obtained using the flattened ring geometry is, in general, lower and pair production starts at higher energies. This is the result of a lower flux, which is produced by Compton-scattering of the BLR photons, assuming a ring geometry.

For this work, the implementation of Finke's calculations by Alexander Sandrock is used [143].

## 16.2 Modeling the Photo-Absorbed SED

Having the analytical form of  $e^{-\tau_{\gamma\gamma}(\epsilon_1)}$  enables us to modify an SED in the gamma-ray regime with a factor of  $e^{-\tau_{\gamma\gamma}(E, R_{\text{blob}})}$ , to model the absorption in the BLR and, finally, constrain the gamma-ray emission region. This is done via fitting a modified log-parabola spectrum (Equation (5.17)) to data measured by *Fermi*-LAT and MAGIC between 100 MeV to 1 TeV:

$$\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^{-(\alpha + \beta \log(\frac{E}{E_0}))} \cdot e^{-\tau_{\gamma\gamma}(E, R_{\text{blob}})} \quad (16.22)$$

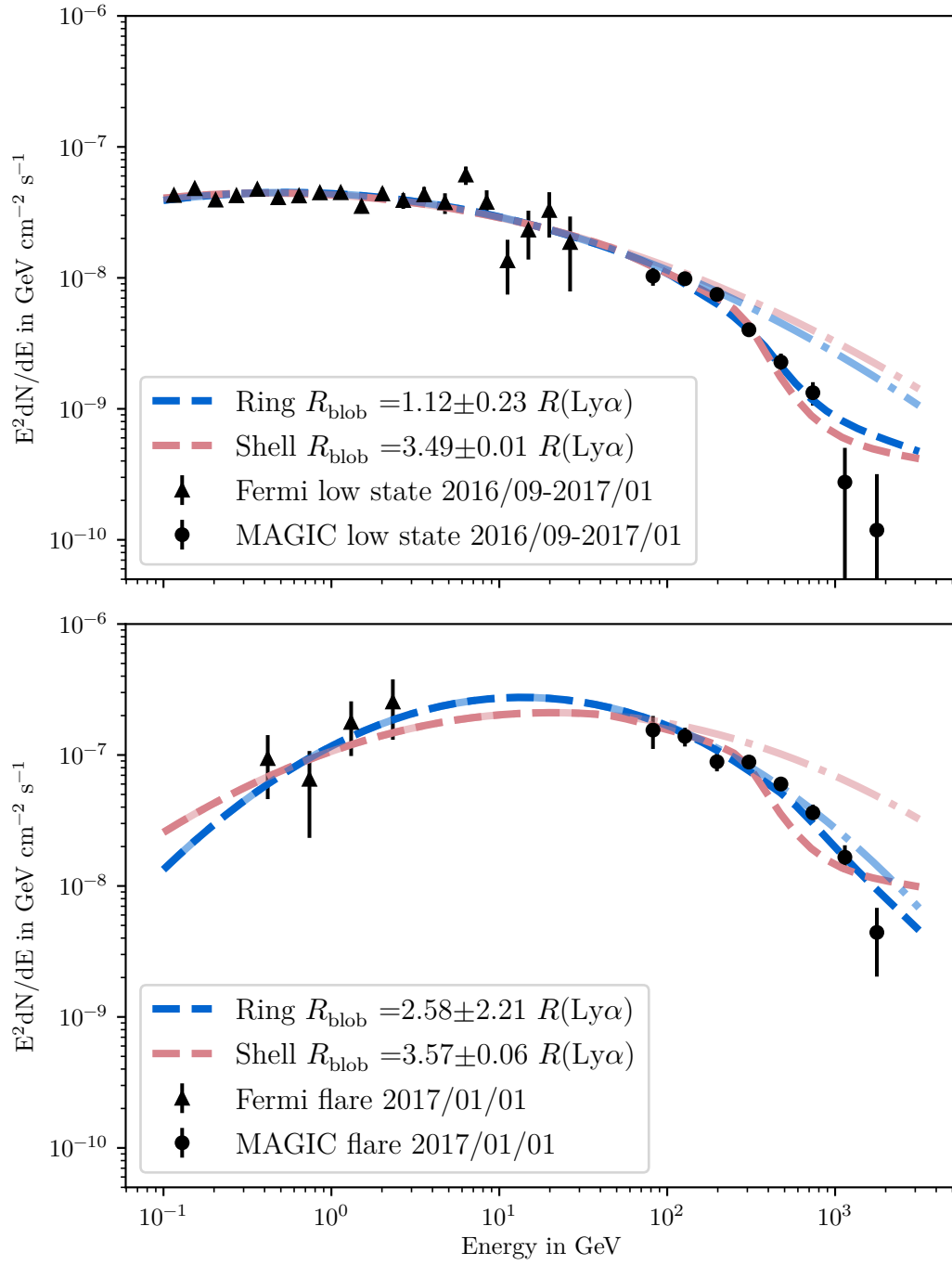
with the fit parameter  $R_{\text{blob}}$  as distance of the emission region (blob) from the BH, denoted as lower integral bound  $r$  in Equation (16.1). In section 16.1, I consider only the BLR photons as absorbing photon field. Following Finke, photon fields originating from the accretion disk and the dust torus can attenuate the VHE-spectrum additionally. In this case, the  $R_{\text{blob}}$  we obtain by the fit has to be considered as a lower limit for the distance between the BH and the emission region, since the gamma-ray emission could be produced further away from the BH and suffer a stronger absorption as considered by this calculation.

For the joint SED fit of *Fermi*-LAT and MAGIC data, I use data from two time intervals, where NGC 1275 was in a low flux state (September 2016 to December 2016) and in a flaring state (January, 1st, 2017). The analysis of the *Fermi*-LAT data in the same time intervals was performed by me, see chapter 12. The differential flux points for MAGIC were taken from Ansoldi et al. [22].

The joint fit is performed for both BLR geometries and the two flux states of the source. As mentioned in section 16.1, the photo-absorption is calculated for the strongest emission line in the BLR, the  $\text{Ly}\alpha$  line.

Fitting the data obtained by *Fermi*-LAT and MAGIC results in the fit parameters given in Table 16.1. The resulting SEDs are shown in Figure 16.3.

To get a rough estimation of the scales of  $R_{\text{blob}}$  compared to the radio images from the 43 GHz BU-BLAZAR program, I calculate  $R_{\text{blob}}$  in units of mas, using the distance of the source as  $d = 62.5 \text{ Mpc}$  [165]. For this task, the angle between the line of sight and the jet axis has to be known. As mentioned in section 2.4, a very wide range of values can be found in the literature for 3C 84's viewing angle. Here, I use the maximum and minimum reported values, with  $65^\circ$  [77] and  $17^\circ$  [22], respectively.



**Figure 16.3:** SED of NGC 1275 fitted with an absorbed power-law for the low state (above) and the flaring state (below). The absorption term is calculated using a shell (yellow) and a ring (red) geometry to describe the BLR's shape. The dot-dashed lines show the unabsorbed spectrum.

**Table 16.1:** Fit parameter for log-parabola function with absorption during the low state 2016/09 to 2016/12 (upper table) and the flaring state on 2017/01/01 (lower table). These results are also published in [109].

Parameter	Ring (low)	Shell (low)
$N_0$ in $10^{-8} \frac{1}{\text{GeV cm}^2 \text{s}}$	$4.39 \pm 0.18$	$4.31 \pm 0.17$
$E_0$ in GeV	1	1
$\alpha$	$2.06 \pm 0.03$	$2.07 \pm 0.02$
$\beta$	$0.05 \pm 0.01$	$0.04 \pm 0.01$
$R_{\text{blob}}$ in R(Ly $\alpha$ )	$1.12 \pm 0.23$	$3.49 \pm 0.01$
$R_{\text{blob}}$ in mas, $\theta = 17^\circ$	0.002	0.007
$R_{\text{blob}}$ in mas, $\theta = 65^\circ$	0.007	0.022
Parameter	Ring (flare)	Shell (flare)
$N_0$ in $10^{-8} \frac{1}{\text{GeV cm}^2 \text{s}}$	$11.69 \pm 4.09$	$10.67 \pm 6.63$
$E_0$ in GeV	1	1
$\alpha$	$1.35 \pm 0.20$	$1.55 \pm 0.33$
$\beta$	$0.13 \pm 0.03$	$0.07 \pm 0.05$
$R_{\text{blob}}$ in R(Ly $\alpha$ )	$2.58 \pm 2.21$	$3.57 \pm 0.06$
$R_{\text{blob}}$ in mas, $\theta = 17^\circ$	0.005	0.007
$R_{\text{blob}}$ in mas, $\theta = 65^\circ$	0.016	0.022

### 16.3 Results

From the fit values in Table 16.1 we can set the gamma-ray emission region in context to the BLR and the source morphology. In case of the shell geometry, I find  $R_{\text{blob}} > 1$  which places the emission region outside the BLR<sup>1</sup>. For the ring geometry, the values are not conclusive. As expected,  $R_{\text{blob}}$  is smaller for the ring geometry in both states of flux activity. In the case of the low flux activity state,  $R_{\text{blob}}$  is close to the Ly $\alpha$  radius, hence the emission region can not originate from near the black hole.

The calculation of  $R_{\text{blob}}$  in mas allows for comparing the obtained distances with the source morphology. The currently best resolved image of the central region of 3C 84 taken by the RadioAstron [87] and the Global Millimeter VLBI Array [101] achieved a beam size of  $\approx 0.05 \times 0.05 \text{ mas}^2$ . Scales at the obtained fit values in the order of  $10^{-3} \text{ mas}$  can not be resolved. Assuming the central BH to be at the center of the radio component C1<sup>2</sup>, the emission region could be located within this core component. As stated above, the values for  $R_{\text{blob}}$  have to be considered as lower limits on the distance between the BH and the emission region, since additional absorption could take place due to other photon fields.

This result excludes all acceleration mechanisms placing the gamma-ray emission region very close to the BH. Possible models for VHE gamma-ray production are discussed in the

<sup>1</sup>Assuming the Ly $\alpha$  emission originating from the outer edge of the emission line. More lines with greater radii exist, but with much fainter luminosities. The BLR has, of course, no sharp edge.

<sup>2</sup>Which is a topic of debate.



next section.

These results are also published in Linhoff et al. [109].

## 16.4 Acceleration Mechanisms

I will discuss briefly the findings of [section 16.3](#) in context of existing models for gamma-ray emission from AGN jets. Some of them place the gamma-ray emission region close to the black hole and can therefore be considered to be very unlikely.

### 16.4.1 Magnetospheric Model

The Magnetospheric Model [9] predicts ultra-high energy gamma-ray emission due to gaps in the magnetosphere of an AGN's central black hole. These gaps occur at low accretion rates when not enough charged particles are induced to provide a full screening of the electric field, which is caused by the rotation of the compact black hole. A particle accelerated to very high energies by the electric field can leave the magnetosphere. In this scenario, the gamma-ray emission has its origin in the innermost core of the source, which is excluded as an emission region by the findings from [section 16.3](#). Furthermore, the magnetospheric model was also considered to be unlikely in Ansoldi et al. [22] because the flux measured by MAGIC exceeds the upper limit of flux predicted with this model by a factor of  $\approx 3$ .

### 16.4.2 Spine-Layer Model

Another model, compatible with the findings of this work, was proposed by Tavecchio and Ghisellini [156]. This model was already mentioned in [section 2.3](#). It consists of a cylindrical fast spine with a Lorentz factor of  $\Gamma_{\text{spine}} = 10 - 20$  and slower layer ( $\Gamma_{\text{layer}} = 2 - 4$ ). The two components in this model are associated with the two bumps in a blazar's SED: The spine produces emission from radio up to the X-ray regime, while the layer is responsible for the high-energy emission. As mentioned in [subsection 2.3.2](#), this model is supported by VLBA and space-VLBI radio images, which reveal the cylindrical jet structure in the upper part of the 3C 84's jet, e.g. [87, 118]. What strikes here is the constraint on the viewing angle that Tavecchio and Ghisellini state. Due to the optical absorption, the viewing angle must be  $< 25^\circ$  to allow emission above a few tens of GeV.

Since the MAGIC collaboration reported the discovery of photons with energies  $> 1 \text{ TeV}$  [22], which I can confirm with my analysis, this upper limit is probably quite conservative. The viewing angle was estimated to be  $17^\circ$  based on the data taken by MAGIC, which is in agreement with the spine-layer model. In former publications, the viewing angle was calculated to be even larger, e.g.  $30^\circ$  to  $55^\circ$  in [170],  $65^\circ$  in [77] and  $25^\circ$  in [3]. Fujita and Nagai [77] even reported the discovery of a sub-parsec counter jet in 3C 84 based on a VLBA observations from December 2015 and January 2016. These findings apparently exclude small viewing angles that are needed to explain gamma-ray emission by this model.

### 16.4.3 Internal Shocks and Magnetic Reconnection

Shocks play a major role for the particle acceleration and the production of gamma rays. Acceleration mechanisms trying to explain especially gamma-ray flares, as happened for NGC 1275, have a long history and are developed and refined until today. Rees proposed the *Internal Shocks Model* in 1978 [138]. In this model, internal shocks are caused by instabilities at the jet's base, in the accretion flow, or through an external medium. Perturbations in the magnetic field lead to shock fronts with different speeds. If a faster shock front catches up on a slower shock front, gamma rays can be emitted in these collisions. This theory is driven by the observation of radio knots in M 87. These knots could be the result of such shock collisions.

In case of NGC 1275, no major radio component was detected or emerged from the core, even years later. There is no evidence that the flare MAGIC detected was emitted by a large plasma collision along the jet.

Over three decades, this idea was slightly modified and expanded, in most cases to explain the observations of particular sources, for example in [137, 150, 83, 82]. Giannios, Uzdensky, and Begelman proposed a *Jets-in-Jet Model*, declaring reconnection of the magnetic field lines as the drive of VHE emission. Due to magnetic reconnection plasma is heated and compressed to plasmoids, emerging the reconnection region at relativistic speed. Each of these blobs produces an SSC spectrum and moves with a higher Lorentz factor compared to the bulk shock. This way, the VHE emission can escape the radiation from the bulk and would not be absorbed. Furthermore, these mini jets point in random directions, which makes the emerging gamma-ray detectable at Earth, even if the jet is pointed elsewhere.

For NGC 1275, acceleration mechanisms based on shocks along the jet seem likely and are supported by the findings of this work.

## Multi-Wavelength Cross-Correlation Analysis

A common approach to combine multi-wavelength data in a joint analysis is to study the correlation of light curves at different wavebands. The time lag between two correlated light curves provides us with hints about the underlying mechanisms driving the radiation at all parts of the electromagnetic spectrum.

The data we have to deal with as astronomers and astroparticle physicists always comes with two main features: the time series is discrete and unevenly sampled. Furthermore, data products from different experiments do not cover the same time range and sampling. Interpolating these time series is never an option as we can not state anything about the time intervals where no measurements were taken.

### 17.1 Discrete Cross-Correlation Functions

A widely used method to account for the mentioned features is the Unbinned Discrete Correlation Function (UDCF) which was proposed by Edelson and Krolik in 1988 [64]. The UDCF for two discrete timeseries  $a$  and  $b$  is defined as:

$$\text{UDCF}_{ij} = \frac{(a_i - \bar{a})(b_j - \bar{b})}{\sqrt{(\sigma_a^2 - e_{a,i}^2)(\sigma_b^2 - e_{b,j}^2)}} , \quad (17.1)$$

where  $(a_i, b_j)$  are all combinations of measured flux points,  $\bar{a}$  and  $\bar{b}$  are the mean flux values and  $\sigma_a$  and  $\sigma_b$  the respective standard deviations.  $e_a$  and  $e_b$  denote the measurement errors.

The  $\text{UDCF}_{ij}$  are then binned in time lags  $\Delta t_{ij} = t_j - t_i$  such that  $\tau - \Delta\tau/2 \leq \Delta t_{ij} < \tau + \Delta\tau/2$  and averaged over  $M$  pairs per bin. The final Discrete Correlation Function (DCF) is then defined as

$$\text{DCF}(\tau) = \frac{1}{M} \sum_{ij} \text{UDCF}_{ij} . \quad (17.2)$$

This approach comes with a certain disadvantage for (non-stationary) AGN light curves: high variabilities and flaring activities might dominate the sample mean values  $\bar{a}$  and  $\bar{b}$ , which is why the mean over all flux points is not a good estimate for all time scales. Amongst others, Welsh [173] propose to calculate this mean value only for a subset of flux points belonging to an interval  $\Delta\tau$ . Originally defined for evenly sampled data, Max-Moerbeck et al. [112] used the Local Cross-Correlation Function (LCCF) as proposed by Welsh to calculate cross-correlation also for unevenly sampled light curves:

$$\text{LCCF}(\tau) = \frac{1}{M} \frac{\sum_i \sum_j (a_i - \bar{a}_\tau)(b_j - \bar{b}_\tau)}{\sigma_{a\tau} \sigma_{b\tau}}, \quad (17.3)$$

where  $\bar{a}_\tau$  and  $\bar{b}_\tau$  are the means over  $M$  samples that contribute with their  $\Delta t_{ij}$  to a certain time lag  $\tau = k\Delta t$ . In the same manner,  $\sigma_{a\tau}$  and  $\sigma_{b\tau}$  are the given standard deviations. Max-Moerbeck et al. and Welsh evaluated a better performance in favour of the LCCF over the UDCF, which is why this method is used in this work.

## 17.2 Light Curve Simulation

The LCCF obtained for two measured light curves with Equation (17.3) is meaningless without confidence levels, as we can never state anything from the absolute LCCF value. To evaluate if a correlation is significant, many ( $\sim 10^6$ ) correlations of uncorrelated light curves have to be calculated to obtain confidence levels for every time lag  $\tau$ . Only if a correlation exceeds that confidence level for a certain time lag, we can state a significant correlation for the two data sets with this time lag.

The two main characteristics of every time series are the Power Spectral Density (PSD) and the Probability Distribution Function (PDF). To be able to compare the correlations of the simulated light curves with the correlation of the originals, the simulated light curves must feature the same PSD and PDF as the original time series.

The PSD is given by the Fourier spectrum of the true and continuous underlying flux variation and can never be measured exactly as every discrete sampling modifies the Fourier spectrum. In this work, I use the term PSD for the unknown underlying true function and the term “periodogram” for the Fourier spectrum of a discrete and finite time series, which is sampled from the underlying true PSD.

The PDF represents the distribution of the absolute flux values and is of very specific shape for different types of sources. For AGN, usually, a superposition of two functions is observed representing the quiet and the flaring state.

To create artificial light curves representing the same PSD as the original light curve, the method proposed by Timmer and Koenig in 1995 is widely used. Since this method does not take into account the underlying PDF of a measured light curve, Emmanoulopoulos, McHardy, and Papadakis [68] expanded the Timmer&Koenig algorithm to create artificial light curves featuring both, PSD and PDF.

For this work, I implemented the algorithm invented by Emmanoulopoulos, McHardy, and Papadakis with some improvements as proposed by Max-Moerbeck et al. in [112], e. g. resampling the original light curve and applying a Hanning-window function before calculating the periodogram. For further details, see appendix E, [163, 68, 112] and my code, which is open source to use for the scientific community.

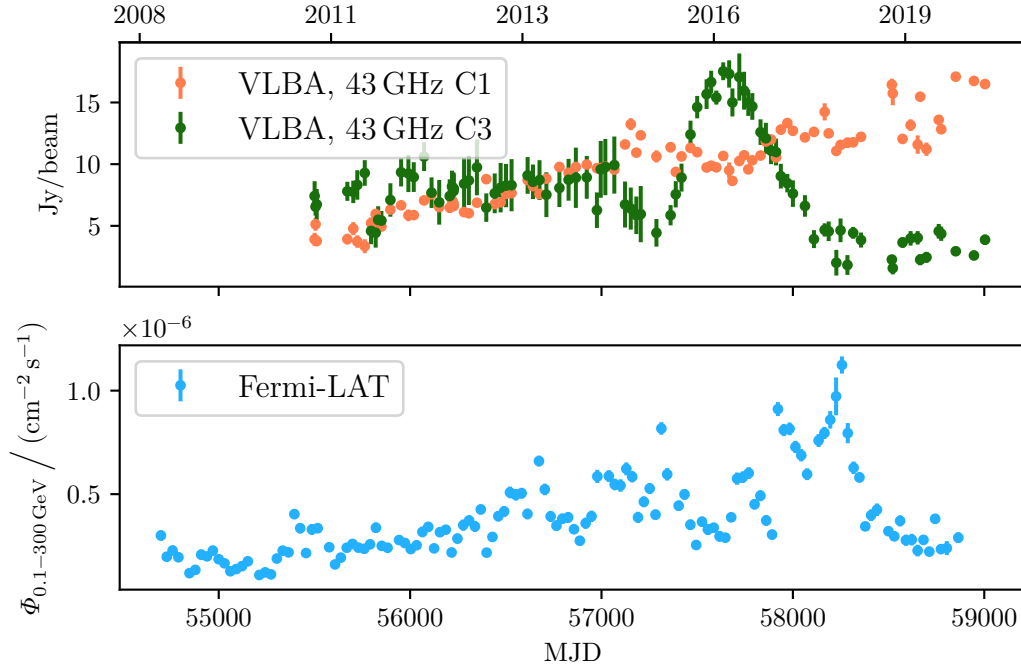
 <https://github.com/lena-lin/emmanoulopoulos>

The cross-correlation function, as well as the simulated light curves, are used in the next section to study the correlation of gamma-ray and radio data from 3C 84.

### 17.3 Cross-Correlation Analysis for *Fermi*-LAT and VLBA Light Curves

For cross-correlation study of 3C 84, I correlate the light curves computed for the LAT data in [chapter 12](#) with the light curves for the two radio components I derived in [section 15.2](#). Since the long-term light curves obtained with MAGIC does mostly contain upper limits, this data is not suitable for any cross-correlation studies.

The *Fermi*-LAT data is binned monthly to fit to the sampling coming with the VLBA data, see [chapter 12](#).



**Figure 17.1:** Light curves of 3C 83/NGC 1275 produced by *Fermi*-LAT and the VLBA at 43 GHz for two components C1 and C3. The VLBA data is calibrated as described in [section 15.2](#).

The cross-correlation for the *Fermi*-LAT light curve and the VLBA-C1 and VLBA-C3 data is shown in [Figure 17.2](#). To estimate the confidence intervals, 1000 artificial light curves per waveband are simulated featuring the same PSD and PDF as the original light curves (see [appendix E](#) for further details on the computation of the simulated light curves). In total,  $10^6$  cross-correlations are used to calculate the confidence intervals at  $3\sigma$ .

For the cross-correlations of *Fermi*-LAT and the VLBA C3 component, no correlation significant above  $3\sigma$  is found. The global maximum for this correlation is around  $-400 \pm 50$  days, which is probably dominated by the prominent peaks of the radio light curve in 2016 and the gamma-ray peak in 2018. (Positive values for  $\tau$  indicate the gamma-ray flux leading the radio light curve and vice versa for negative values.)

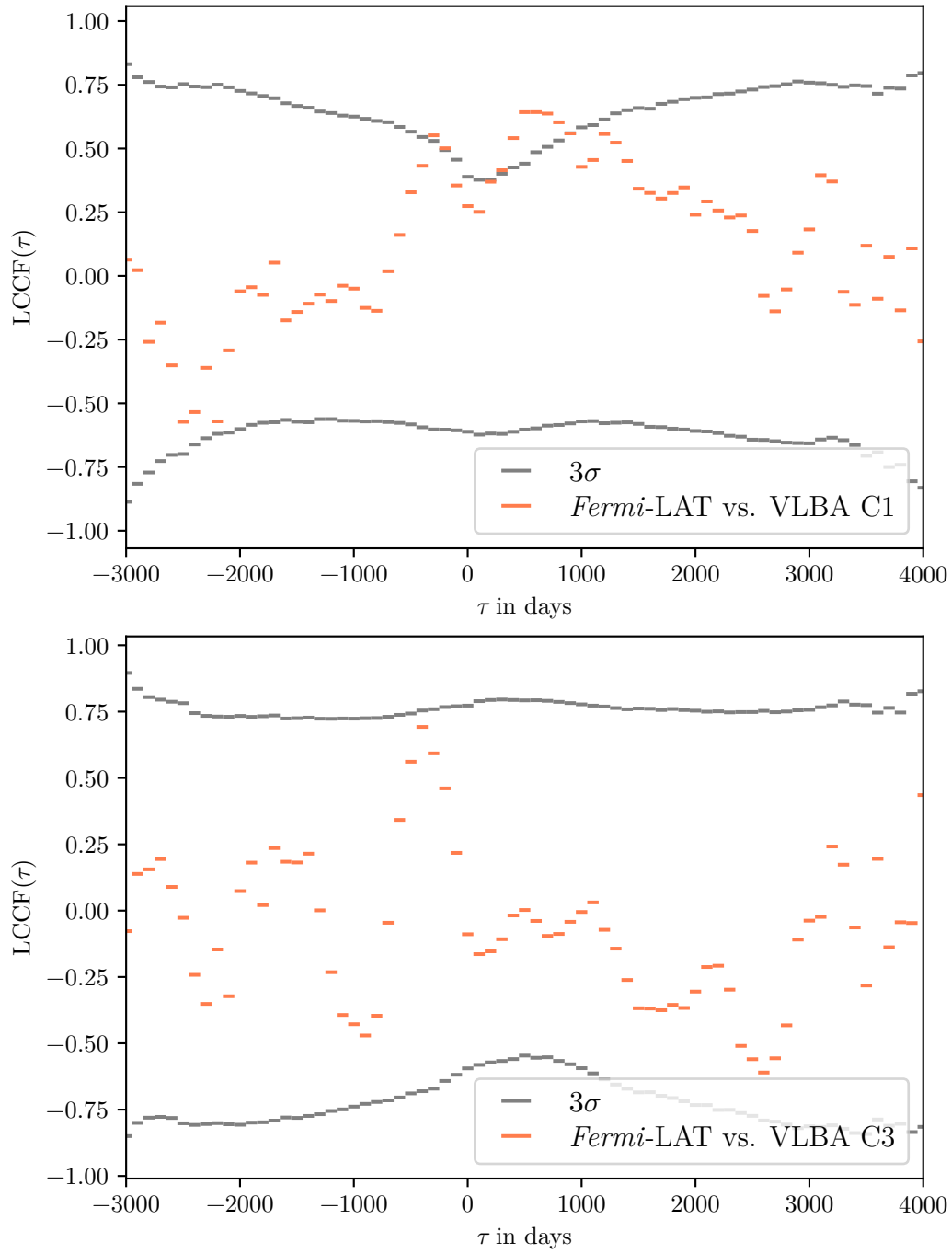
The correlation analysis of *Fermi*-LAT and VLBA C1 shows peak correlations around  $550 \pm 100$  and  $-300 \pm 50$  days, which exceed the  $3\sigma$  confidence level. These results are

in line with the correlations found by Hodgson et al., who correlate the *Fermi*-LAT light curve (2008 – 2016) with radio data measured at 225 GHz by the SMA (2002 – 2016) [95]. Hodgson et al. found significant correlations at 99.97 % ( $3.5\sigma$ ) for a time lag of  $-370 \pm 120$  days and  $290 \pm 155$  days.

Since the SMA data can not resolve the morphology of a source, the contribution of the components could not be separated by Hodgson et al. From the results of the *Fermi*-LAT/VLBA correlation, it seems that the correlations found by Hodgson et al. are dominated by the C1 component, whereas the C3 component might contribute to the negative time lag only.

Summing up, the cross-correlation analysis results in favor of the C1 component hosting the gamma-radio connection, which is not in conflict with the results from the BLR study in [chapter 16](#). Still, it is striking that the evolution of the C3 component and the gamma-ray flux increase and rapid fall happen in a similar time range, and that after the fade-out of C3 also the gamma-ray flux remain at a very low level.

If the C3 component would not be connected to the gamma-ray emission, another mechanism must be responsible for the rapidly falling flux detected by *Fermi*-LAT after May 2018. An alternative idea to explain the decrease is given in the next chapter.



**Figure 17.2:** LCCF for *Fermi*-LAT versus VLBA C1 (above) and C3 (below) light curves in bins of 100 days. Gray lines indicate the  $3\sigma$  confidence levels. Positive  $\tau$  indicate that the gamma-ray flux leads the radio flux. For C1, correlations at  $550 \pm 100$  and  $-300 \pm 50$  days exceed the  $3\sigma$  confidence level. C3 shows a peak correlation at  $-400 \pm 50$  days which is below the  $3\sigma$  threshold.

## 17.4 Possible Gamma-Ray Suppression and Neutrino Production?

As a last thought, I will combine all aspects about 3C 84 found during the course of this thesis. The following construct is highly speculative at the moment but worth to be considered for future studies.

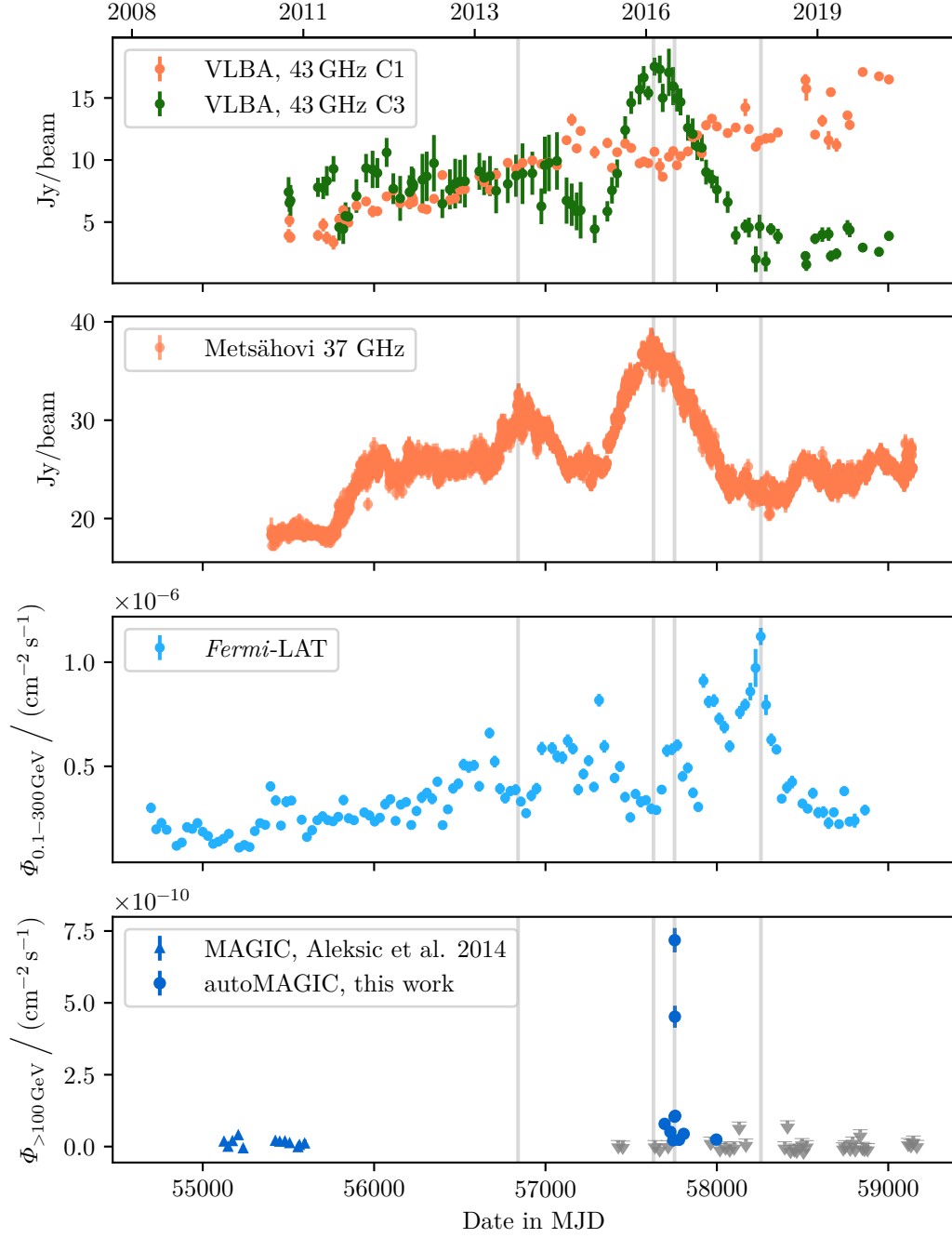
Kun et al. [105] and Halzen, Kheirandish, Weisgarber, and Wakely [90] proposed the idea of gamma-ray suppressed states of AGN due to photo-meson production. If high-energy photons accelerated in an AGN's jet hit a proton-field, pions are produced. The decays of charged pions further produce neutrinos, which can be detected on Earth. In these states, the observed source would be opaque to gamma-rays, while radio emission is able to leave the AGN unimpeded. Kun et al. found neutrino events detected by the neutrino detector IceCube simultaneously with decreases in gamma-ray flux measured by *Fermi*-LAT and MAGIC for several sources [105].

Projecting this model onto 3C 84, we would expect a rapid decrease in the gamma-ray light curve, while radio flux remains stable. And indeed, this is the case in 2018, as stated before and shown in Figure 17.3.

The additional photon field could be explained by a second object that builds a binary system with 3C 84. This binary system is also proposed as a reason for the jet-precession we found, described in Dominik, Linhoff, Elsässer, and Rhode [61]. The second object could cover 3C 84 and its jet during the rotation and provide the protons needed for increasing the pion production and, therefore, the opacity to gamma-rays.

So far, no neutrino discovery from the position of the Perseus cluster was published. For the future, it might be worth the effort to investigate the time interval around the rapid decrease in gamma-ray flux for neutrino events.





**Figure 17.3:** Multiwavelength light curves of 3C 84 between 2008 and 2021 measured by the VLBA at 43 GHz, the Metsähovi radio telescope at 37 GHz *Fermi*-LAT and MAGIC. Global flux maxima for all wave bands are marked with gray lines.



## **Part VI**

### **Conclusion & Outlook**



## Conclusions and Outlook

---

### 18.1 New Insights into 3C 84/NGC 1275

During this work, I analyzed new data of 3C 84/NGC 1275 recorded with MAGIC and combined them with data and scientific results from *Fermi*-LAT and the VLBA.

The main phenomena of 3C 84 discussed in this work are:

- **From radio observations:** 3C 84 features a bright radio component associated with the core region of the AGN where the BH is located. In 2005, the new component C3 emerged from the core, moved south along the jet with a brightness maximum around 2016, and faded out since then to not be detectable anymore.
- **In the MeV/GeV range:** Starting from 2008, when *Fermi*-LAT started data taking, NGC 1275 showed a slowly rising trend over the years. In 2018, the gamma-ray flux reached its maximum and rapidly decreased afterwards.
- **In the GeV/TeV range:** MAGIC detected a huge orphan flare in 2017 that lasted for two days. Apart from that, NGC 1275 is in a very low flux state.

The questions to address, as stated in the introduction, are:

- Where is the origin of the gamma-ray emission?
- Is there a connection between the emission at radio wavelengths and the VHE gamma rays?

#### 18.1.1 Origin of the Gamma-Ray Emission

The first question is addressed by combining the *Fermi*-LAT and MAGIC data to a joint spectrum with the aim to study the absorption taking place in the BLR, see [chapter 16](#). With this study, I can constrain the gamma-ray emission to be outside the BLR, which excludes gamma-ray emission mechanisms that take place very close to the BH. I place the minimal distance of the gamma-ray emission from the BH  $R_{\text{blob}}$  I derived in context with the source morphology that we see in radio images. Compared to the scales in the VLBA radio maps,  $R_{\text{blob}}$  has a size of  $10^{-2}$  to  $10^{-3}$  mas, depending on the viewing angle. This result stated the C3 region as a possible origin of the VHE emission. But also the C1 component can not be excluded. The scales of  $R_{\text{blob}}$  can not be resolved properly by the telescopes operating today, hence the edge of the BLR lies within C1.

### 18.1.2 Correlation Between Gamma Rays and Radio Components

To answer the second question, I performed a cross-correlation study with the light curves of both radio components extracted from the VLBA images and the *Fermi*-LAT long-term light curve, see [chapter 17](#). The light curve measured by MAGIC unfortunately contains only upper limits for most of the time, which makes it insufficient for correlation studies.

The cross-correlation of the light curves of *Fermi*-LAT and the VLBA C1 component results in two significant correlations. First, for the radio emission leading the gamma-rays with a delay of  $300 \pm 50$  days, second, vice-versa, the gamma-ray emission leading the radio emission with a delay of  $550 \pm 100$  days. These correlations do not directly explain the connection between gamma rays and radio emission, but are a hint that the radio emission from C1 is connected to the emission detected by the *Fermi*-LAT.

The most obvious approach to connect the most prominent features of 3C 84, namely the peak in the *Fermi*-LAT light curve and the evolution of the C3 component, is not strengthened by the correlation study.

### 18.1.3 Long-Term Analysis of the MAGIC Data

Apart from the flare in 2017, the analysis of NGC 1275 I performed for this work results in a series of upper limits after the flare, see [chapter 9](#). In the context of multi-wavelength analysis, we can learn something from this, anyhow. In 2017 and 2018, *Fermi*-LAT reported a constantly enhanced flux, while my analysis of the MAGIC data still results in upper limits far below the flux values during the flare. In contrast to that, the TeV flare happened more than a year before the flux in the MeV/GeV regime reached its maximum. This is a hint that the TeV flaring activity is not directly correlated with the flux at lower energies and we rather see a combination of multiple mechanisms.

### 18.1.4 Quo Vadis, 3C 84?

What we detect from 3C 84 is a superposition of multiple processes at various wavelengths and scales. On the bigger scales, we have the precessing motion of the jet, on intermediate scales we see the radio components emerging from the core and on the smaller scales, which we can not resolve anymore, gamma rays are produced. All these processes have to be taken into account for further studies concerning 3C 84.

Although I was able to gain new insights into 3C 84, we are far away from a complete understanding of the complex processes happening in this galaxy. An AGN is a living thing which will provide us with new incidents, certainly. We can therefore expect more and more data and new discoveries from 3C 84 in the next years helping us to understand the mechanisms of AGN in general.

The methods I used in this thesis are not restricted to the data I applied them to. It will be worth the effort to repeat and extend those analyses in a few years with additional data or maybe for new radio components that might emerge from the core of 3C 84 in the next decades. Additionally, they can be applied to other TeV radio galaxies.

The new detectors that are currently under construction, for example, CTA and the Square Kilometer Array (SKA), will deliver new insights and keep a whole generation of

scientists working.

## 18.2 Towards an Automated and Reproducible MAGIC Data Analysis

The main technical focus of this work is on the development of `autoMAGIC` to orchestrate an automated and reproducible data analysis with little human interaction. This approach enables the reproducible analysis of long-term data observed by MAGIC for the first time.

To demonstrate the capabilities of `autoMAGIC` and `Gammapy`, I performed a long-term analysis of the Crab Nebula, which I compared to the standard software tools of MARS, see [section 8.4](#). I found both analyses to be in good agreement, which is the first step to establish `autoMAGIC` as a sufficient tool for the production of the “MAGIC legacy”.

As mentioned in [section 8.5](#), `autoMAGIC` is an active software project and there are always aspects that can be improved or additional ideas that have to be implemented in the software. But for the future, `autoMAGIC` will enable us to handle the huge amount of data and computational work that comes with long-term analyses, which has been a great obstacle for such studies in the past. From now on, MAGIC is equipped for further long-term analyses and multi-wavelength studies and will contribute to the decryption of the complex processes going on in NGC 1275 and many other gamma-ray sources.





# **Part VII**

## **Appendix**



## Appendix A

### The autoMAGIC Database Structure

Besides the information mentioned in [section 8.3](#), additional tables are needed to map the relationships of all jobs, runs, observation conditions, and software versions. All in all, our database contains about 30 tables with far more than a million entries in total.

To illustrate some basic dependencies and to clarify the wording, I will describe the terms I use for the measured and the simulated data.

#### Observation, Run, SubRun

The overview over the related tables are shown in [Figure A.1](#). An **Observation** is a set of runs recorded on a specific date (night) with a specific pointing. A **Run** is the equivalent to a data run, which is about 20 minutes long. Accordingly, a **SubRun** is a roughly 2 minutes long subset of a run. For every subrun, one calibrated file is produced by the data acquisition system, which is the fixed starting point of our analysis. This is why the calibrated file names are listed in the database. For a run, multiple **Superstar** files can be processed, for example with different **MARS** versions. File names are therefore always connected to a job and not to a run. All relations are shown in [Figure A.1](#).

#### Monte Carlo Production, Set, Run and Analysis Period

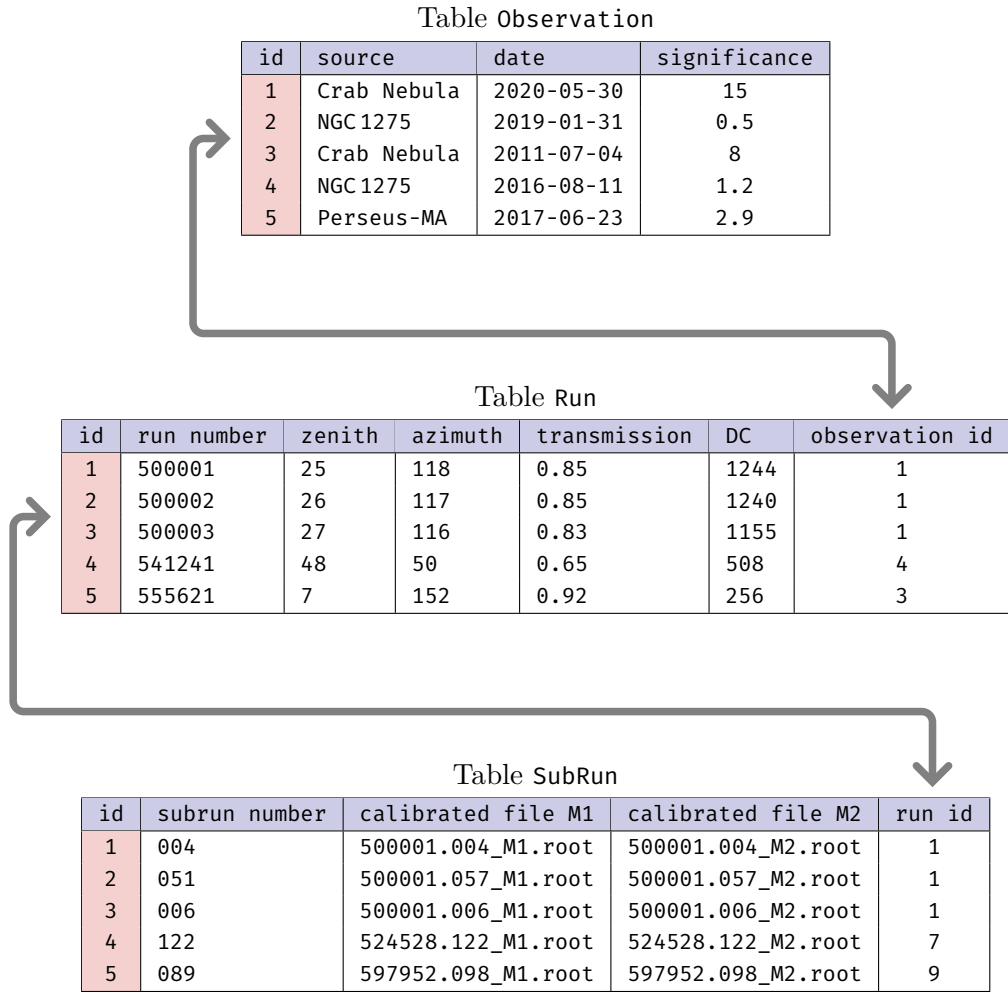
Every time the hardware conditions change due to updates, wear, or weather phenomena like Calima, a new **Monte Carlo Production** is produced. This means simulating events with **Corsika** and adjust the **MARS** camera and reflector simulators for adequate hardware parameters. A **Monte Carlo Production** contains multiple **Monte Carlo Sets** for different observation conditions, like zenith angle or trigger type. Every **Monte Carlo Set** has a lot of **Monte Carlo Runs** containing the actual events. Typically, about 5000 **Monte Carlo Runs** belong to one **Monte Carlo Set**. A **Monte Carlo Production** is usually connected to a single time range defined by a start date and a stop date. This is called **Analysis Period** in **MAGIC** terms. It can happen that some hardware characteristics are valid for more than one time range, for example, after the Calima dust is washed off the mirrors by the rain and the reflectivity of the mirrors is the same as before. All relations are shown in [Figure A.2](#).

#### State and Exit Code

Every job has a column to monitor the exit code of the executed program, a timestamp, and a state. If the job is created, which means inserted as a row into the database, the state would be **created**. Jobs waiting to be submitted to the cluster are marked as **queued**.

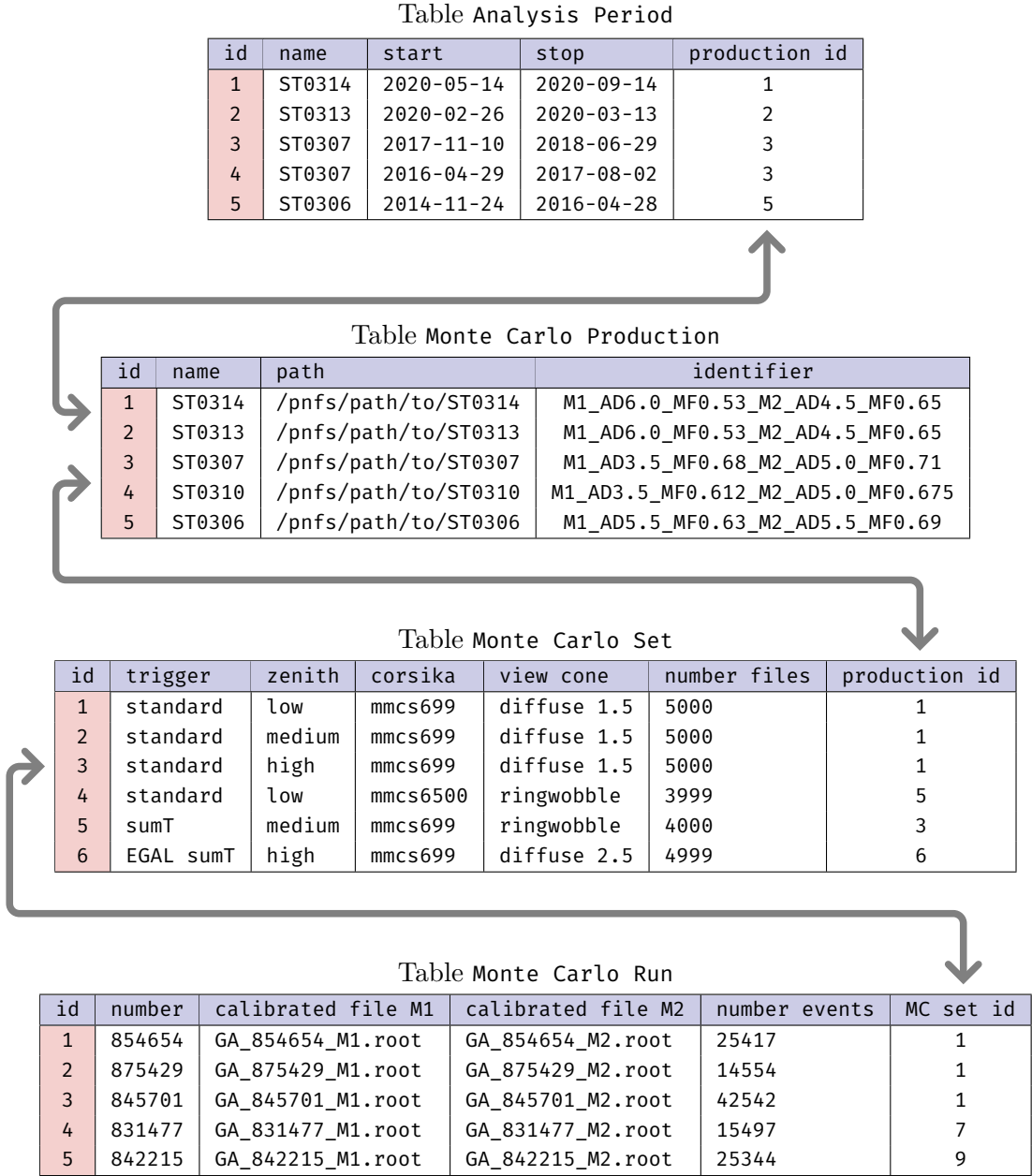
## *Appendix A The autoMAGIC Database Structure*

Jobs at the cluster, but waiting to be started are `idle` and `running` afterwards. If a job ends with exit code `0`, the job is set to state `success`, and to `error` otherwise. In case some problems with the cluster occur and the job could not be finished properly, it is marked as `stopped` and will be submitted again. A timestamp is set during the job is running to monitor when the job was submitted. The relations for a job table are shown in [A.3](#).

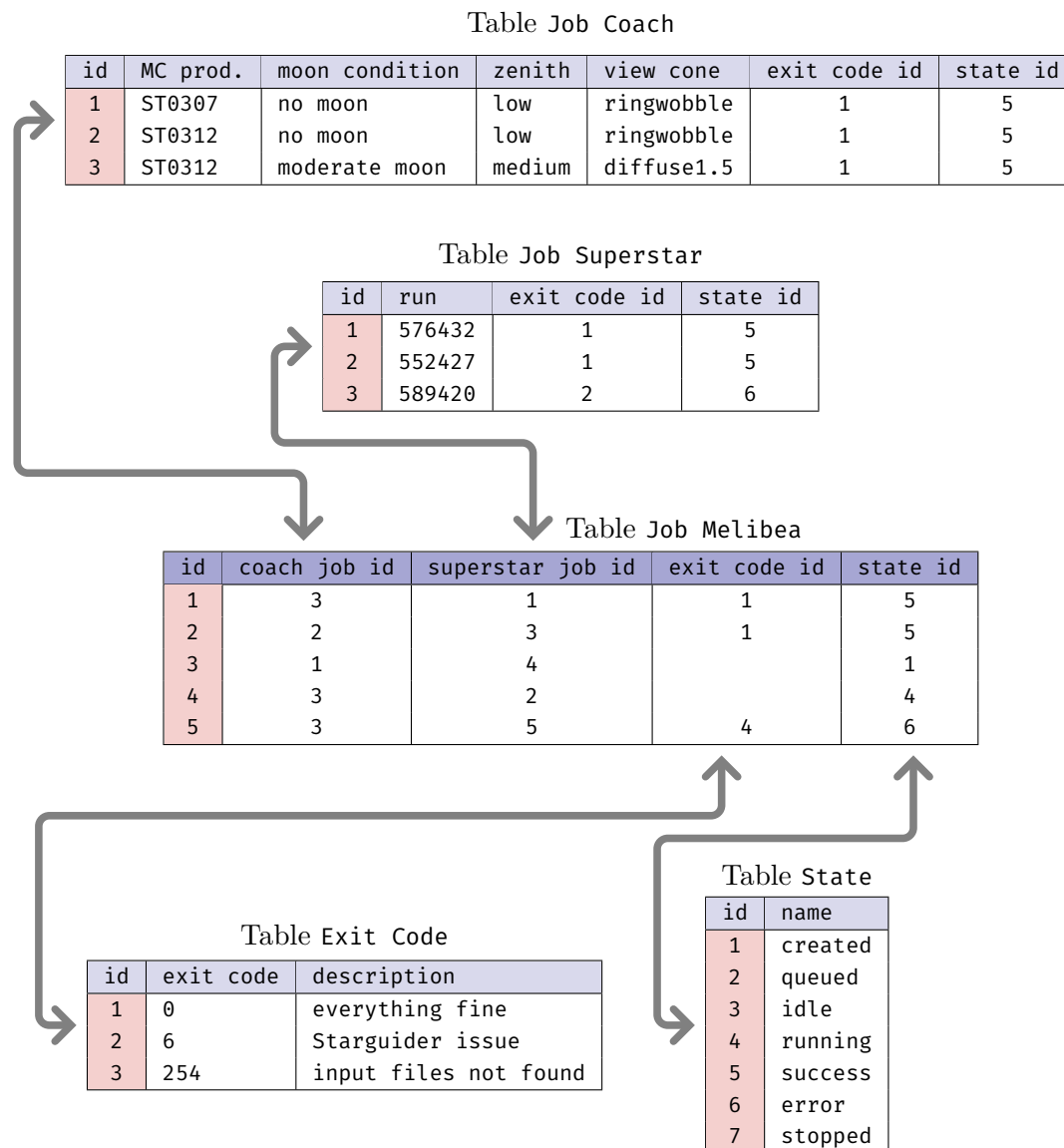


**Figure A.1:** Database relations for Observation, Run and SubRun. An Observation is a set of Runs recorded on a specific date (night) with a specific pointing. A Run possesses all parameters which are used for quality checks and grouping the data by their observation conditions. The transmission gives information about the amount of clouds above the telescopes and the DC (direct current) indicates the level of NSB, for example by moonlight. A SubRun is a roughly 2 minutes long subset of a run, and is taken by both telescopes separately. The tables are connected via the `id` column, the so-called foreign key. This scheme is a very simplified version of the database structure to illustrate the terms and their connections. The actual tables are more complex.

## Appendix A The autoMAGIC Database Structure



**Figure A.2:** Database relations for Monte Carlo Monte Carlo Production, Monte Carlo Set and Monte Carlo Run. A Monte Carlo Production is defined by a set of hardware parameters and contains multiple Monte Carlo Sets for different observation conditions. A Monte Carlo Set contains of roughly 5000 Monte Carlo Runs which are represented by two files each (for M1 and M2). The View Cone reflects the maximum angle between the origin of the simulated particle to the center of the FoV. *ringwobble* denotes a fixed distance of 0.4 degree (suitable for most observations) and *diffuse 1.5/2.5* are events simulated with a distance between 0 and 1.5 degree or 2.5 degree, respectively. Every Monte Carlo Production is connected to at least one time range defined by a start date and stop date, the Analysis Period. It happens that a Monte Carlo Production is valid for more than one Analysis Period, as in case of ST0307. Again, this is a very simplified scheme of the real database structure to clarify the terms I use in the following. To fully map all MC related features, many more tables are necessary.



**Figure A.3:** Database relations for a job entry, in this example a **Melibea** job. **Melibea** classifies all events of a **Superstar** output file using a RF produced by a **Coach** job. Therefore, every **Melibea** job has a relation to a **Coach** job and a **Superstar** job. To monitor the state of a job, every job table has a column **state**. If a job is finished, the **exit code id** column is filled with the **id** of the program's exit code. Usually, 0 means that everything is ok. The exit codes are very useful to understand what went wrong if a job fails. If a job has not been finished yet, the **exit code id** is empty. Of course, the shown tables are simplified examples and much more columns are needed to map the complexity of the analysis workflow.

## A.1 PostgreSQL and SQLite

Our database is a set of tables containing all relevant information about the measured and simulated data and analysis-related parameters. To communicate with this database a dedicated language is needed. For autoMAGIC, we use Structured Query Language (SQL) a widely used language for managing relational databases. SQL allows us to access the database, get very specific or aggregated information and to insert or delete objects. For everything we want to know from the database, a *query* has to be generated and submitted to the database. Since autoMAGIC is written in Python, we use the Python package `SQLAlchemy` [28] to integrate the SQL commands into the Python code.

For implementing the database itself, there are many different Relational Database Management Systems (RDBMSs), which are used for different purposes. We use a `PostgreSQL` [152, 153] database for the production system (the actual autoMAGIC database), which is hosted at the PIC. `PostgreSQL` is free and open sources and allows for many requests at the same time, which is great for big data processing, but is a bit more complex to handle. This is why we additionally use `SQLite` for developing purposes. `SQLite` is another RDBMSs, but in contrast to `PostgreSQL` it can be hosted on a local computer and the database is stored in a simple file. In turn, `SQLite` has some restrictions, for example, it is not capable of answering to many requests at the same time, which is problematic for the real production, but not an issue for developing and small test cases. In the usual workflow, new autoMAGIC features are developed and tested with a local `SQLite` database and then implemented into the `PostgreSQL` production database at the PIC. For further reading, see

- ❶ `PostgreSQL` Documentation: <https://www.postgresql.org>
- ❷ `SQLite` Documentation: <https://www.sqlite.org>
- ❸ `SQLAlchemy` Documentation: <https://www.sqlalchemy.org>



## Appendix B

### Job Submission with HTCondor

The job submission at the PIC is handled with the workload management system HTCondor [160], which is integrated into autoMAGIC via Python bindings. HTCondor takes care of submitting jobs to a cluster and starting them as soon as a free node is available. Furthermore, it is responsible for managing the logfiles, which contain the executed program's standard output and error messages.

HTCondor is developed at the University of Wisconsin-Madison since 1988. For further reading, see

📖 HTCondor Documentation: <https://research.cs.wisc.edu/htcondor>

Despite the large capacities of the PostgreSQL database and the HTCondor system, we have to limit the maximum number of jobs that is submitted to the cluster at once. Otherwise, we would overburden the cluster and the database in some cases, for example, when ten thousands of MC runs have to be processed. Working with the PostgreSQL database, this limit is roughly 200 jobs. For a local test database, only 20-50 jobs should be submitted at once to ensure smooth data processing.



## Appendix C

### MARS vs. Gammapy: Study on $\theta^2$ Cuts

As described in [section 8.4](#) Flute offers the possibility to apply  $\theta^2$  and hadronness cuts energy-dependent. In Gammapy, there are only global cuts feasible. To study the effect the different cut methods have on the high-level results, I choose two small data sets of the Crab Nebula with different characteristics and produce a light curve and a SED with Flute/Fold and magicDL3/Gammapy. These data sets consist of a couple of runs taken in two different analysis periods. To avoid biases in the high-level results induced by the observing conditions, only good-quality data is selected, meaning low zenith and good weather ( $\text{transmission}^1 > 0.9$ ).

The second data set consists of runs measured at moderate moonlight to test the ability of autoMAGIC for analyzing data recorded under moonlight conditions. Furthermore, MC sets produced with different view cones were used for the two analyses. The first set uses MC events originating in a distance between 0 and  $2.5^\circ$  from the FoV's center, whereas the second analysis is performed with ringwobble MCs, which means that all simulated events have the same distance of  $0.4^\circ$  from the FoV's center as the observed object.

The light curves and spectra are compared to a reference spectrum from Aleksić et al. [13]. The two test sets are described in [Table C.1](#).

**Table C.1:** Characteristics of the data sets used for a first cross check between MARS and Gammapy.

Test Set	Period	MC View Cone	Moon (DC)	Zenith
1	ST0307	diffuse $2.5^\circ$	$0\ \mu\text{A}$ to $2200\ \mu\text{A}$	$5^\circ$ to $35^\circ$
2	ST0311	ringwobble	$3300\ \mu\text{A}$ to $5500\ \mu\text{A}$	$5^\circ$ to $35^\circ$

There are three ways to define a hadronness or  $\theta^2$  cut:

- Choose a fixed cut value globally for all energy bins.
- Choose a cut globally “from efficiency”, meaning that the value is chosen where a user-defined percentage of MC events survives the cut.
- Choose a cut from efficiency individually for every energy bin.

Obviously, an energy-dependent cut is the best choice in most cases since it accounts for the fact that the PSF varies for different energies.

<sup>1</sup>transmission is a measure of the amount of clouds covering the sky in the FoV of the telescopes. The transmission is determined with the Light Detection and Ranging (LIDAR), a laser pointing to the sky and measuring the light reflected by the clouds.

Flute offers all of these options to specify a hadronness and  $\theta^2$  cut. `magicDL3` (used to calculate the IRFs and convert the data into DL3 files) currently applies only global cuts (user-defined or from efficiency), because `Gammapy` can not handle energy-dependent cuts, yet.

To investigate the impact of these different cut-choice methods, I performed the same analysis for both software frameworks with a global user-defined cut and a cut from efficiency. For `Flute`, the cut from efficiency is energy-dependent, but not for `magicDL3`.

The global cuts are chosen as  $\theta^2 < 0.02 \text{ deg}^2$  and hadronness  $< 0.28$ . The efficiencies are set to 75% for  $\theta^2$  and 90% for hadronness in `Flute` and for `Gammapy` an efficiency of 90% is used for both cuts. Since `Flute` applies the cuts energy-dependent in the latter case, the resulting subsets will not be identical.

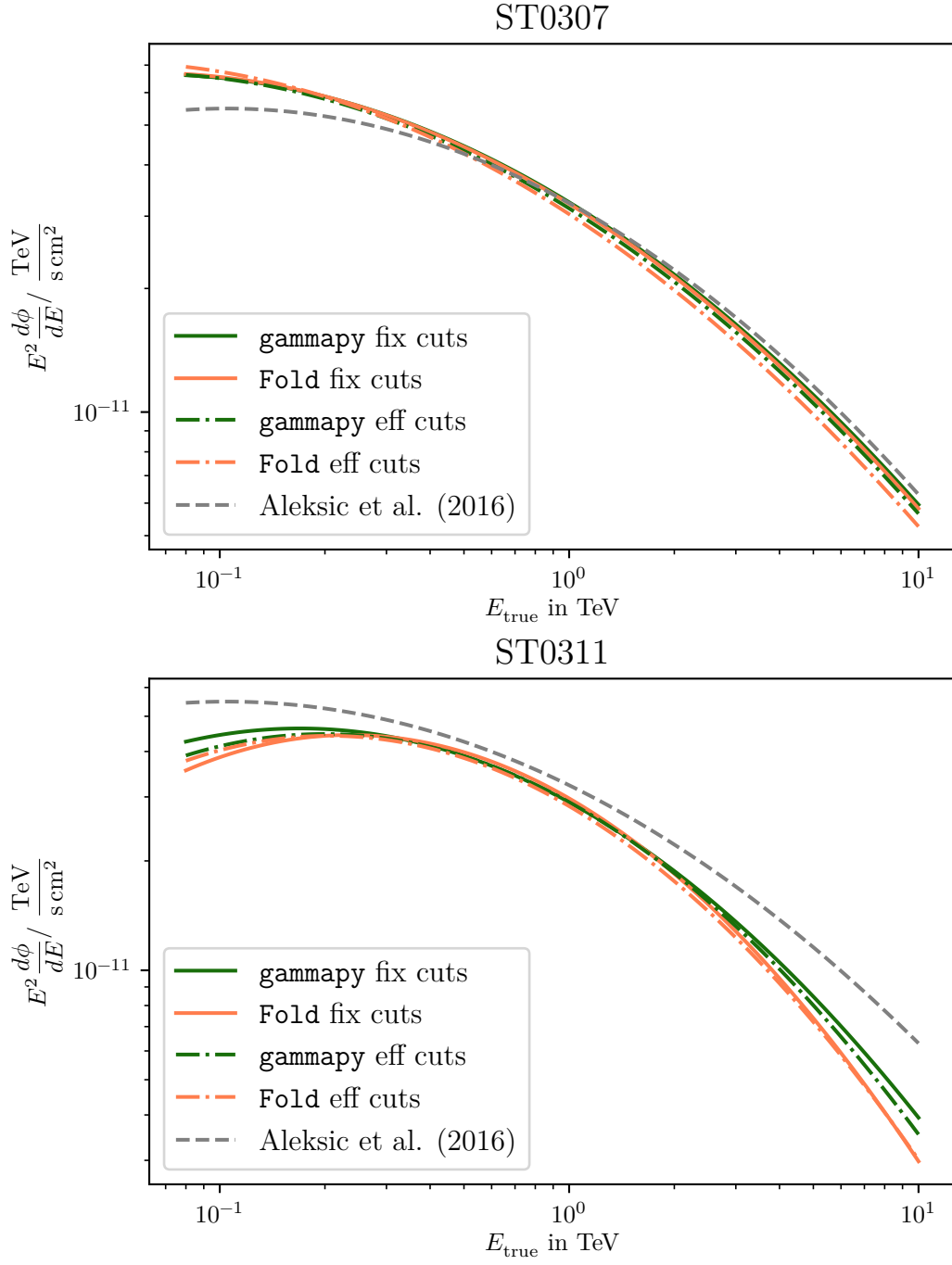
## C.1 Spectral Energy Distribution

To compare the results for the SED, I use `Fold` which computes a forward folding which is similar (but not identical) to the method that is implemented in `Gammapy`. The result is show in [Figure C.1](#) and the fit parameter for the log-parabola spectrum ([Equation \(5.17\)](#)) are listed in [Table C.2](#).

The SEDs produced with `Gammapy` and `Fold` are very similar and match the reference spectrum quite well. We do not expect to exactly reproduce the reference spectrum, since the data set is very small. The small discrepancies between `Fold` and `Gammapy` can be explained by slightly different minimization approaches and a different treatment of the background term in the likelihood.

**Table C.2:** Fit parameter for the log-parabola spectrum, calculated by `Gammapy` and `Fold`. The reference energy  $E_0$  is fixed at 300 GeV.

Period	Software	Cut	$N_0$ in $10^{-10} \frac{1}{\text{TeV cm}^2 \text{ s}}$	$\alpha$	$\beta$
ST0307	<code>Fold</code>	global	$5.87 \pm 0.17$	$2.30 \pm 0.03$	$0.09 \pm 0.00$
ST0307	<code>Fold</code>	efficiency	$5.77 \pm 0.17$	$2.34 \pm 0.03$	$0.09 \pm 0.00$
ST0307	<code>Gammapy</code>	global	$5.88 \pm 0.17$	$2.29 \pm 0.04$	$0.09 \pm 0.03$
ST0307	<code>Gammapy</code>	efficiency	$5.77 \pm 0.17$	$2.31 \pm 0.04$	$0.09 \pm 0.03$
ST0311	<code>Fold</code>	global	$4.88 \pm 0.16$	$2.09 \pm 0.03$	$0.19 \pm 0.00$
ST0311	<code>Fold</code>	efficiency	$4.79 \pm 0.16$	$2.13 \pm 0.03$	$0.18 \pm 0.00$
ST0311	<code>Gammapy</code>	global	$4.90 \pm 0.16$	$2.17 \pm 0.05$	$0.15 \pm 0.03$
ST0311	<code>Gammapy</code>	efficiency	$4.83 \pm 0.17$	$2.14 \pm 0.05$	$0.17 \pm 0.04$



**Figure C.1:** Log-parabola SED computed by `Gammapy` and `Fold` for two test sets from analysis periods ST0307 (2016-10-12) and ST0311 (2018-12-18). The SED is fitted for the datasets produced with  $\theta^2$  and hadronness cuts from efficiency or user-defined global cuts.

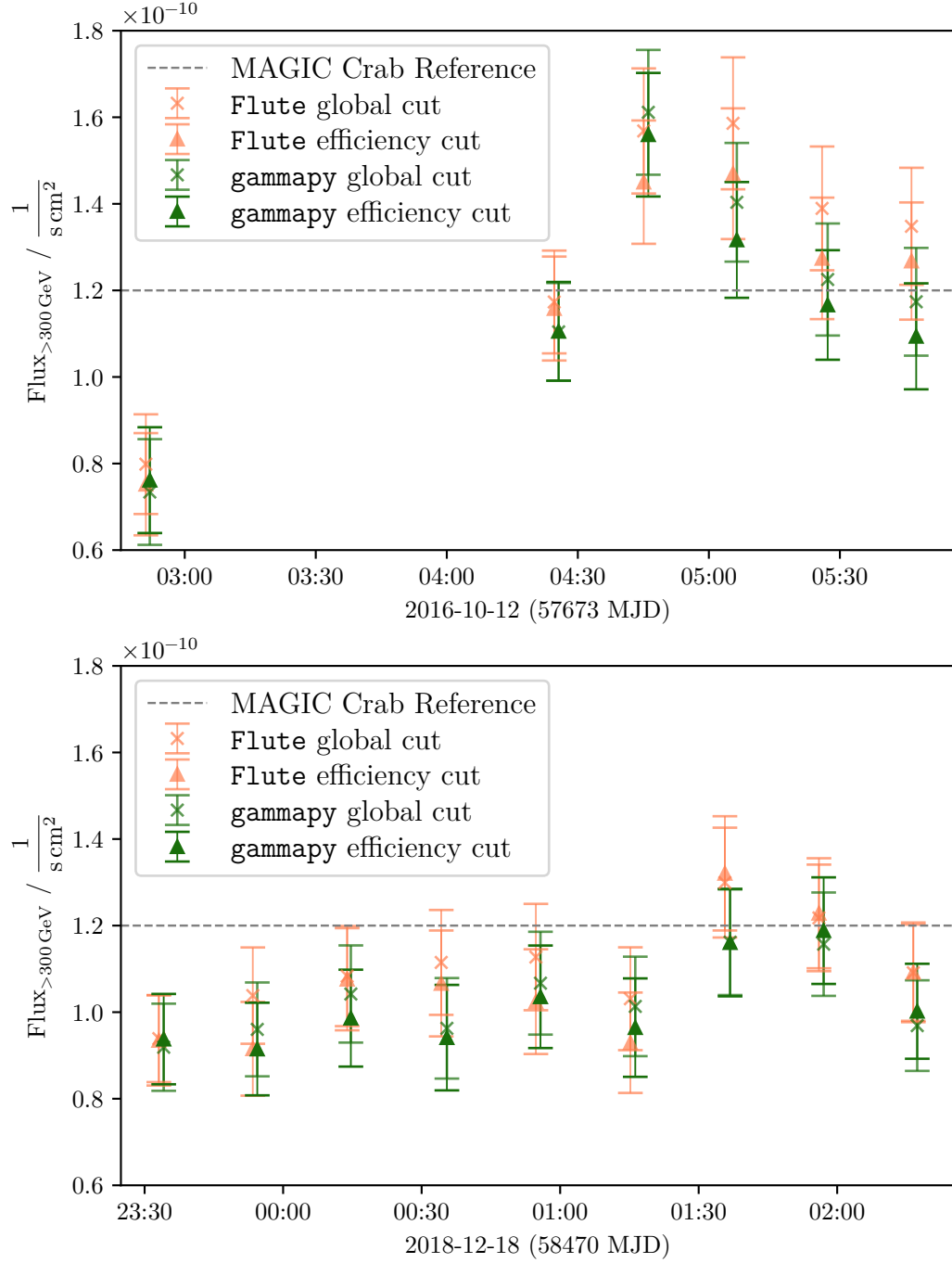
## C.2 Light Curves

The light curves obtained from the two test sets described in Table C.1 are shown in Figure C.2. In all cases, the flux points are calculated using a spectrum with fixed parameters  $\alpha$  and  $\beta$  (from [13]) to keep the differences between the analyses as small as possible:

$$\frac{dN}{dE} = N_0 \left( \frac{E}{300 \text{ GeV}} \right)^{-(2.31+0.26 \log(\frac{E}{300 \text{ GeV}}))}. \quad (\text{C.1})$$

The light curves in Figure C.2 show slight fluctuations around the reference value, which is as expected for the same reasons as listed above. We see furthermore a slightly different computation of the mean time value for a run, where the marker is set, which is probably caused by different handling of the start and stop times by both programs. But this small shift is very well within the duration of a single run, which is about 20 minutes.

All in all, we see a very good agreement between `Flute` and `DL3+Gammapy` for computing the light curve of the two test data sets. No discrepancies are visible even for data measured under moonlight conditions.



**Figure C.2:** Light curves computed with *Flute* and *Gammamap* from data of the Crab Nebula observed on 2016-10-12 at low zenith and no moon (above) and on 2018-12-18 at low zenith and moderate moon. Each software is executed with two different methods to adjust the  $\theta^2$  and hadronness cut. The *global cut* refers to a user-defined fixed cut which is used for all events. The *triangles* represent data selected with a *cut from efficiency*, which is evaluated on the MC set. *Flute* calculates the cut from efficiency for every energy bin separately, which is why these efficiency cuts are not comparable for *Gammamap* and *Flute*. Nevertheless, the analysis results are in very good agreement and, furthermore, in agreement with the reference value for the Crab Nebula's flux reported in [13].

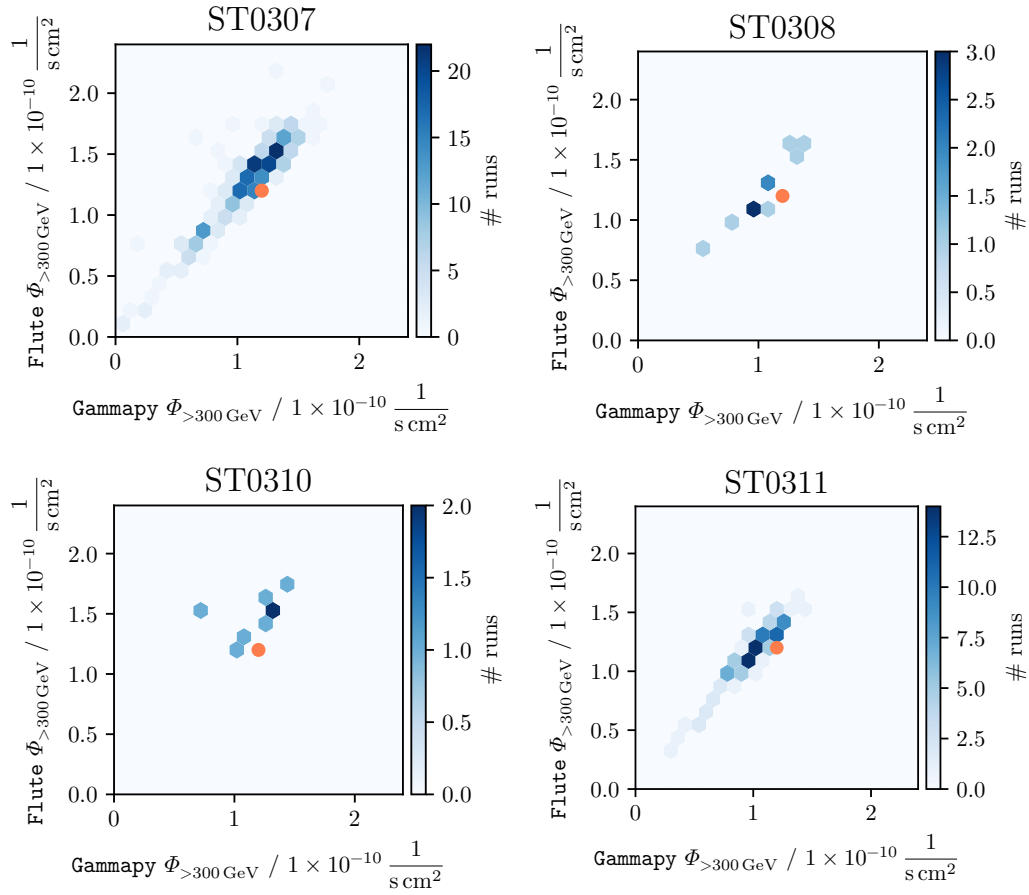




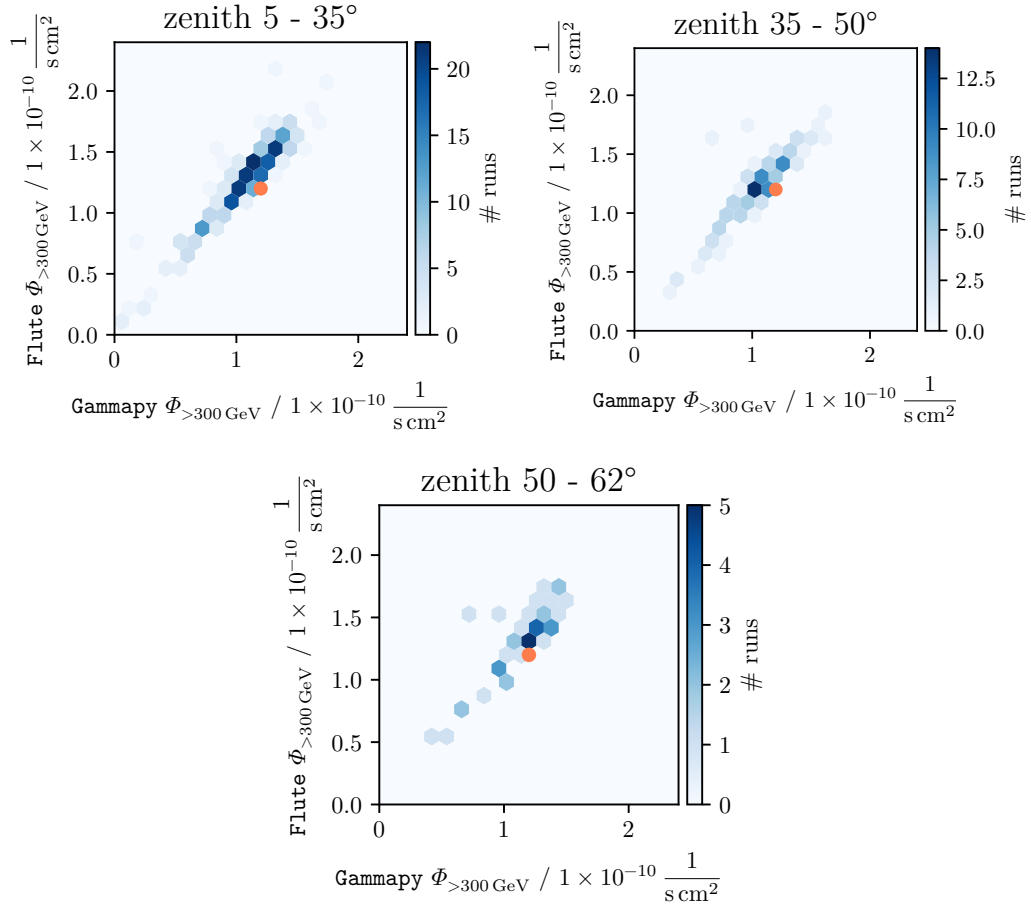
## Appendix D

### MARS vs. Gammapy: Comparison by Observation Conditions

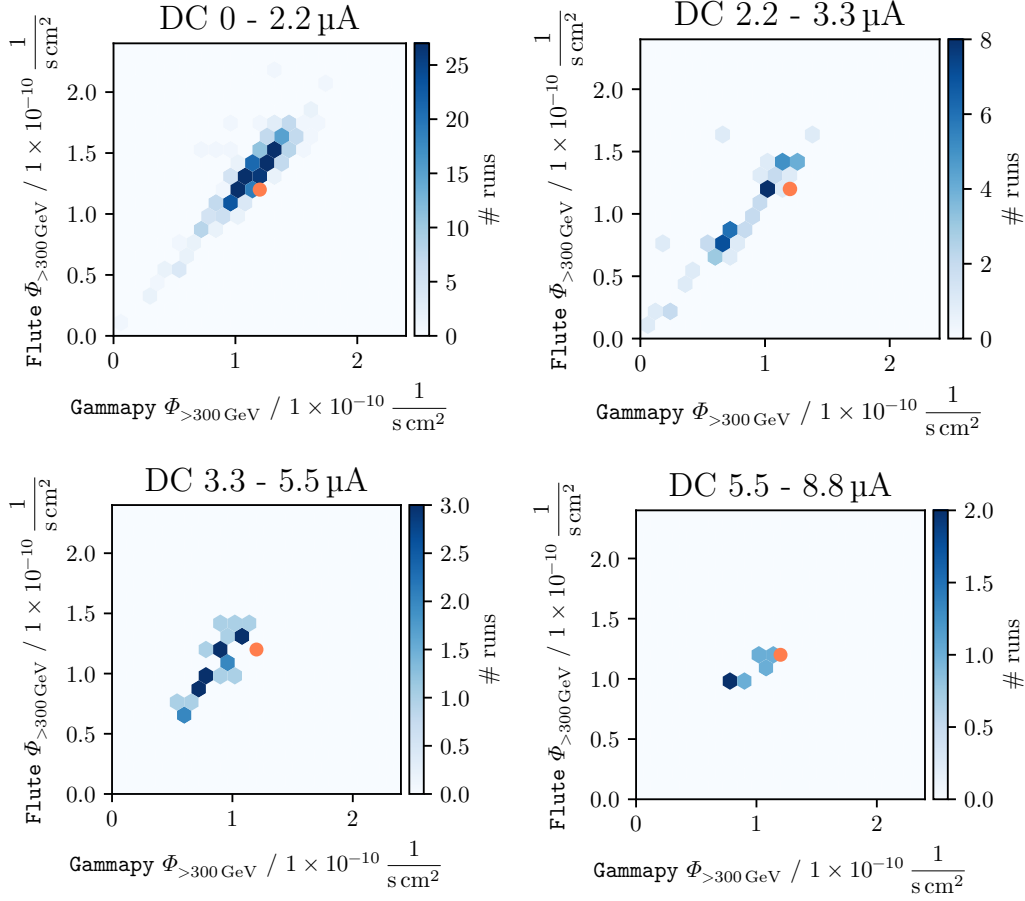
To evaluate the performance of the high-level analysis, the flux points computed by *Flute* and *Gammapy* for the run-wise binned Crab Nebula light curve are plotted in 2D-histograms. Each histogram represents a subset of all runs for a special observation condition or analysis period.



**Figure D.1:** 2d-binned flux points for the Crab Nebula computed by *Flute* and *Gammapy* for different MC productions. The orange dot marks the reference flux reported by Aleksić et al., [14].



**Figure D.2:** 2d-binned flux points for the Crab Nebula computed by Flute and Gammapy for different zenith ranges. The orange dot marks the reference flux reported by Aleksić et al., [14].



**Figure D.3:** 2d-binned flux points for the Crab Nebula computed by Flute and Gammapy for different moon ranges. The orange dot marks the reference flux reported by Aleksić et al., [14].



## Appendix E

### Light Curve Simulation

In this work, simulated light curves are used to estimate the confidence interval of the cross-correlation of two light curves obtained from the same source at different wavebands. To obtain the correct confidence intervals, the simulated light curves have to be uncorrelated but must feature the same Power Spectral Density (PSD) and Probability Distribution Function (PDF) as the original light curve. The approach used here follows the methods proposed by Emmanoulopoulos, McHardy, and Papadakis [68] and Timmer and Koenig [163] with some additional features proposed by Max-Moerbeck et al. [112].

The method proposed by Timmer and Koenig results in artificial light curves following the original PSD only. Emmanoulopoulos, McHardy, and Papadakis extend this approach to an algorithm for generating light curves with the correct PSD *and* PDF. Furthermore, Max-Moerbeck et al. suggest some additional steps for estimating the underlying PSD from the periodogram of the original light curve.

The terms Probability Distribution Function, Power Spectral Density and periodogram are defined as follows:

- PSD: Continuous Fourier spectrum of the true function (in our case: flux over time). In case of AGN light curves, the PSD follows a  $1/f^\alpha$  spectrum. The true distribution can never be measured since we are limited by the discrete sampling of the time series. For astronomy data, this is always the case because we usually have a very sparse time sampling with one data point per day, week, or even months, often interrupted by large gaps.
- Periodogram: The discrete Fourier spectrum of the time series sampled from the true function. The better the time resolution, the better the PSD can be estimated by the periodogram.
- PDF: Distribution of the absolute (flux) values of the time series.

#### E.1 Discrete Fourier Transform and Periodogram

Following [68], the periodogram is calculated based on the Discrete Fourier Transform (DFT) of a evenly sampled<sup>1</sup> time series  $x(t_k)$  of length  $N$  and bin width  $t_{\text{bin}}$ :

---

<sup>1</sup>Light curves are never evenly sampled, which I will address in a minute.

## Appendix E Light Curve Simulation

$$\text{DFT}(j) = \sum_{k=1}^N x(t_k) \exp\left(\frac{2\pi i(k-1)j}{N}\right) \quad (\text{E.1})$$

$$\text{for } j = 0 \dots N-1. \quad (\text{E.2})$$

The DFT results in a series of complex numbers. The corresponding Fourier frequencies for even and odd  $N$  are defined as:

**$N$  even:**

$$f_j = \begin{cases} f_0 = 0 & \text{for } j = 0 \\ f_j^+ = j/Nt_{\text{bin}} & \text{for } j = 1 \dots \frac{N}{2} - 1 \\ f_{N/2} = f_{\text{Ny}} = 1/2t_{\text{bin}} & \text{for } j = \frac{N}{2} \\ f_j^- = -(N-j)/Nt_{\text{bin}} & \text{for } j = \frac{N}{2} + 1 \dots N-1 \end{cases} \quad (\text{E.3})$$

**$N$  odd:**

$$f_j = \begin{cases} f_0 = 0 & \text{for } j = 0 \\ f_j^+ = j/Nt_{\text{bin}} & \text{for } j = 1 \dots \frac{N-1}{2} \\ f_j^- = -(N-j)/Nt_{\text{bin}} & \text{for } j = \frac{N+1}{2} \dots N-1 \end{cases} \quad (\text{E.4})$$

$$\text{The Nyquist frequency } f_{N/2} \text{ does not exist for odd } N. \quad (\text{E.5})$$

$f^+$  and  $f^-$  denote the positive and the negative Fourier components, respectively.

Amplitude  $\mathcal{A}$  and phase  $\Phi$  of the complex numbers are defined as:

$$\mathcal{A}_j = \frac{1}{N} \sqrt{\text{Re}[\text{DFT}(j)]^2 + \text{Im}[\text{DFT}(j)]^2} \quad (\text{E.6})$$

and

$$\Phi_j = \arg[\text{DFT}(j)] = \arctan\{\text{Im}[\text{DFT}(j)], \text{Re}[\text{DFT}(j)]\}. \quad (\text{E.7})$$

The periodogram  $P(f_j)$  can now be calculated as the squared amplitude of the  $j$ -th component:

$$P(f_j) = \mathcal{A}_j^2 = \frac{1}{N^2} \{\text{Re}[\text{DFT}(j)]^2 + \text{Im}[\text{DFT}(j)]^2\} \quad (\text{E.8})$$

$$\text{for } j = 0 \dots N-1. \quad (\text{E.9})$$

In our case, the Fourier transformed time series (the light curve) contains only real valued numbers,  $x(t_k) \in \mathbb{R}$ . Therefore, the positive DFT components are the complex conjugated negative DFT components and their amplitudes are equal:

$$\text{DFT}(j^-) = [\text{DFT}(j^+)]^* \quad (\text{E.10})$$

$$\mathcal{A}_{j+} = \mathcal{A}_{j-}. \quad (\text{E.11})$$

The periodogram of a real valued time series can be written as

$$P(f_j) = \mathcal{A}_j^2 = \frac{2}{N^2} \{ \text{Re}[\text{DFT}(j)]^2 + \text{Im}[\text{DFT}(j)]^2 \} \quad (\text{E.12})$$

$$\text{with } f_j = \frac{j}{N t_{\text{bin}}} \quad (\text{E.13})$$

$$j = 0 \dots \frac{N}{2} \quad \text{for even } N \quad (\text{E.14})$$

$$j = 0 \dots \frac{N-1}{2} \quad \text{for odd } N. \quad (\text{E.15})$$

Following [68], the periodogram is normalized with the fractional root mean square factor  $N t_{\text{bin}} / \mu^2$  as proposed by Vaughan, Edelson, Warwick, and Uttley [166].  $\mu$  denotes the mean of the original time series values. The periodogram is finally defined as:

$$P(f_j) = \frac{2 t_{\text{bin}}}{\mu^2 N^2} \{ \text{Re}[\text{DFT}(j)]^2 + \text{Im}[\text{DFT}(j)]^2 \}. \quad (\text{E.16})$$

### E.1.1 Sampling Effects

As stated at the beginning, the time series that is Fourier transformed has to be evenly sampled, which is almost never the case for astronomical light curves. The light curves are therefore interpolated and re-sampled to a user defined bin width  $t_{\text{bin}}$ , which should be in the same range as the original sampling. Max-Moerbeck et al. [112] show that the periodogram of a re-binned evenly sampled light curve represents the underlying PSD of the light curve better as the periodogram of the unevenly sampled data. Note that we can never use the interpolated light curves to derive any statement of the absolute flux values during the times between the measurements! The interpolation just reduces the impact of the uneven time sampling on the periodogram. This effect produces a bigger difference between the PSF and the periodogram than the missing values in the light curve, which is why interpolating a light curve is fine for the PSD estimation.

Furthermore, there are two effects we have to be aware of if we work with a sampled time series of finite length: *Red Noise Leakage* and *Aliasing*.

Red noise leakage is caused by the finite length of the light curve, when power that is originally deposited at wavelengths longer than the observation time is transferred to higher frequencies. This effect has an impact on the slope of the periodogram. Based on the original shape and slope of the underlying PSD, this impact can vary. For a PSD with a spectral slope of  $\alpha < 1.5$  the red noise leakage is usually negligible [166].

Aliasing refers to the transfer of power from higher frequencies to lower frequencies and is a result of the discrete sampling, which restricts the time resolution (see [102] for further details).

To reduce this effects for the PSD estimation, Max-Moerbeck et al. propose to convolute

## Appendix E Light Curve Simulation

the original light curve with a Hanning window<sup>2</sup>. The Hanning window is defined as

$$w_{\text{Hanning}}(t) = \begin{cases} \cos^2\left(\pi \frac{t-T/2}{T}\right), & 0 \leq t \leq T \\ 0, & t > T \end{cases} \quad (\text{E.17})$$

with the length  $T$  of the light curve.

Summing up the prequel for the PSD estimation we have to:

1. Apply a Hanning window  $w_{\text{Hanning}}(t)$  to the original data  $x(t_k)$ .
2. Interpolate and re-bin the data for an evenly sampled light curve.
3. Calculate the DFT( $j$ ) from the re-binned light curve using Equation (E.1).
4. Calculate the periodogram  $P(f_j)$  from the Fourier components using Equation (E.16).

For the PDF estimation, the interpolated flux (without applying the window function!) must be scaled to values between 0 and 10. From these values the histogram for the PDF estimation can be calculated.

## E.2 PSD and PDF Estimation

Before artificial light curves can be simulated that follow a certain PSD and PDF, these functions must be estimated from the original light curve. In the last section, the periodogram  $P(f_j)$  was obtained from the original data. We can now fit a PSD model  $\mathcal{P}(f_j)$  to this periodogram. Furthermore, the PDF of the original light curve has to be fitted. Both fits are executed using a maximum likelihood estimation (similar to section 5.5).

### E.2.1 PSD Likelihood

The PSD model used in [68] is a bending power-law modified with an additional constant Poisson noise  $c$ :

$$\mathcal{P}(f; \gamma, c) = \frac{A f^{-\alpha_{\text{low}}}}{1 + (f/f_{\text{bend}})^{\alpha_{\text{high}} - \alpha_{\text{low}}}} + c \quad (\text{E.18})$$

with  $\gamma = (A, f_{\text{bend}}, \alpha_{\text{low}}, \alpha_{\text{high}})$  representing the model parameters. The Poisson noise  $c$  is added later to the model and is not a fit parameter.

To obtain the likelihood function, we assume the components of the periodogram  $P(f_j)$  to be asymptotically  $\chi^2$  distributed around the true  $\mathcal{P}(f_j)$ :

$$P(f_j) = \begin{cases} \frac{1}{2} \chi_2^2 \mathcal{P}(f_j), & j = 1 \dots N/2 - 1 \text{ for even } N \\ \frac{1}{2} \chi_1^2 \mathcal{P}(f_{N/2}), & j = N/2 \text{ for even } N \\ \frac{1}{2} \chi_2^2 \mathcal{P}(f_j), & j = 1 \dots (N-1)/2 \text{ for odd } N. \end{cases} \quad (\text{E.19})$$

---

<sup>2</sup>Named after Julius von Hann, also called Hann function. Not to be confused with the Hamming function, another window function named after Richard Hamming.



$\chi_\nu^2$  denotes the  $\chi^2$  distribution with  $\nu$  degrees of freedom.

From these distributions we can express probability to obtain a single periodogram component  $P(f_j)$  if the underlying PSD is given as  $\mathcal{P}(f_j; \gamma)$ :

$$\lambda [P(f_j)|\mathcal{P}(f_j; \gamma)] = \begin{cases} \Gamma[\frac{\nu-2}{2}, \mathcal{P}(f_j; \gamma)] & j = 1 \dots N/2 - 1 \text{ for even } N \\ \Gamma[\frac{\nu-1}{2}, \mathcal{P}(f_j; \gamma)] & j = N/2 \text{ for even } N \\ \Gamma[\frac{\nu-2}{2}, \mathcal{P}(f_j; \gamma)] & j = 1 \dots (N-1)/2 \text{ for odd } N. \end{cases} \quad (\text{E.20})$$

The unbinned likelihood for obtaining a full periodogram from a certain PSD is then given as

$$\mathcal{L} = \prod_{i=1}^{\begin{matrix} N/2 \text{ (even } N) \\ (N-1)/2 \text{ (odd } N) \end{matrix}} \lambda_j [P(f_j)|\mathcal{P}(f_j; \gamma)] . \quad (\text{E.21})$$

The best model estimation for  $\gamma$  is now obtained by minimizing the negative logarithmic likelihood function.

## E.2.2 PDF Likelihood

For the PDF estimation, the flux values are re-scaled to values between 0 and 10 which is a suitable range for the standard gamma and log-norm distributions. The normed flux is then binned in a histogram to obtain the flux distribution, which has to be fitted.

The PDF of an AGN's light curve can often be modeled as a superposition of two distributions representing two states of activity: The low state and the flaring activity. For this work, the PDF is modeled with a gamma and a log-normal distribution:

$$\text{PDF}(x; \boldsymbol{\eta}) = w_1 \gamma(x, a) + w_2 \mathcal{N}_{\log}(x, s, \mu, \sigma) \quad (\text{E.22})$$

$$= w_1 \cdot \frac{x^{a-1} e^{-x}}{\Gamma(a)} + w_2 \cdot \frac{1}{s(x - \mu) \sqrt{2\pi}} \exp \left( -\frac{\log^2(\frac{x-\mu}{\sigma})}{2s} \right) \quad (\text{E.23})$$

with the weights  $w_{1,2}$  so that  $w_1 + w_2 = 1$  and the fit parameters  $\boldsymbol{\eta} = (a, s, \mu, \sigma)$ . The unbinned likelihood is given as

$$\mathcal{L} = \prod_{k=1}^N \text{PDF}(x_k | \boldsymbol{\eta}) . \quad (\text{E.24})$$

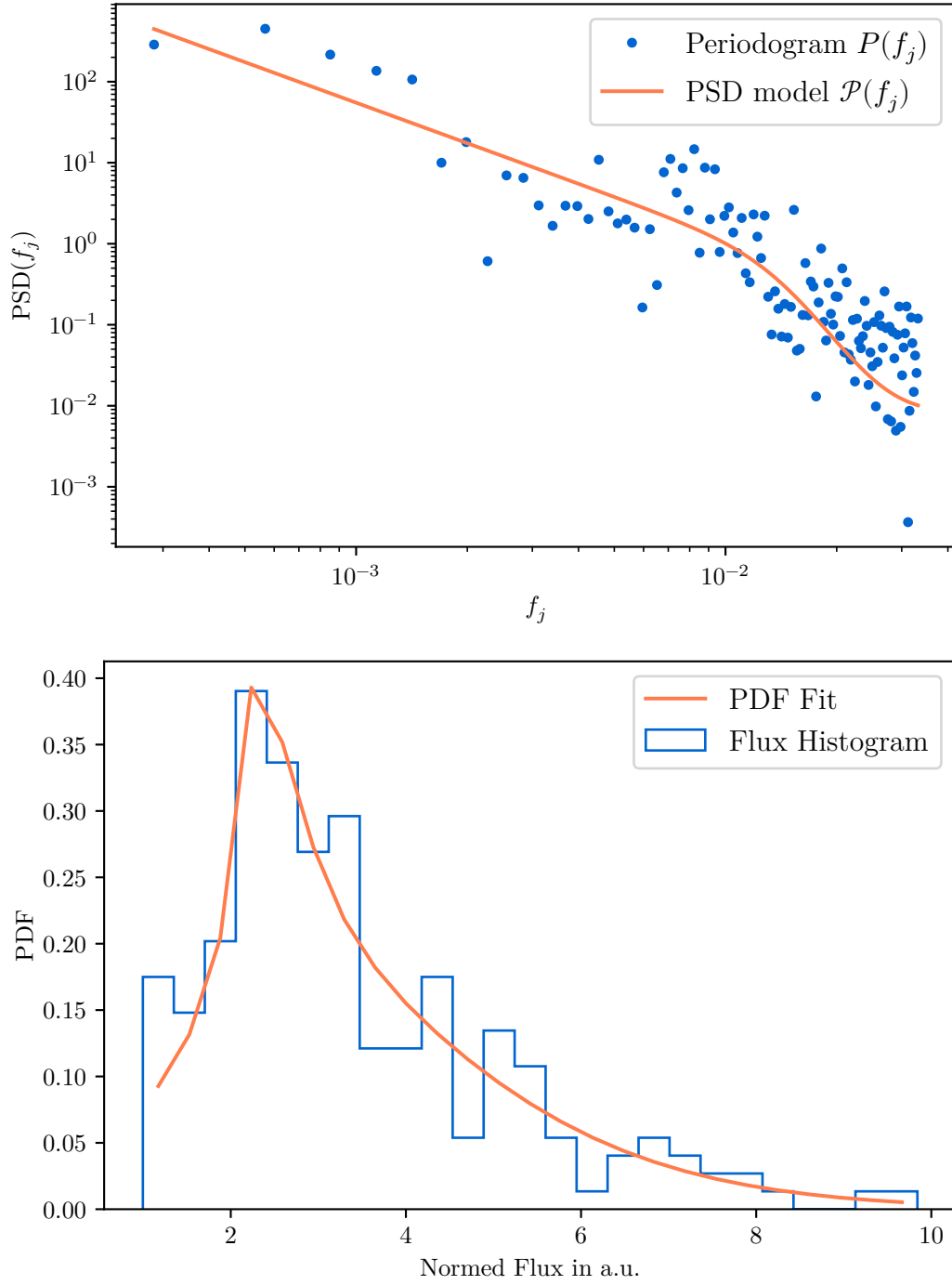
For a *binned* likelihood approach, the flux values of the light curve are binned in  $n_{\text{bin}}$  bins of the size  $\Delta x$ . The bin value can now be estimated as the integral of the PDF in the corresponding bin edges. The integral of the PDF is given by the Cumulative Distribution Function (CDF), in this case the weighted sum of the CDFs of the gamma and the log-norm distributions as used in Equation (E.22). To estimate the value in a bin, the CDF must be evaluated at the bin edges. The binned likelihood is then given as

$$\mathcal{L} = \prod_{i=1}^{n_{\text{bins}}} \text{CDF}((i+1)\Delta x | \boldsymbol{\eta}) - \text{CDF}(i\Delta x | \boldsymbol{\eta}) . \quad (\text{E.25})$$

## *Appendix E Light Curve Simulation*

If the binned or the unbinned approach is used depends on the amount of data. For an unbinned fit, the computation time scales with the number of flux points, which is why it makes sense to use a binned fit for long light curves. For light curves with few flux points, a binned likelihood fit might fail and an unbinned approach should be chosen.

Having the likelihood defined, we have to take care of the maximization, respectively, the minimization of the logarithmic likelihood. For this work, the minimization is performed with the `minuit` [97] minimizer, which is implemented as Python package `iminuit` [56].



**Figure E.1: Above:** Periodogram of a light curve (blue dots) and the likelihood fitted PSD for the model described in Equation (E.18). **Below:** Flux histogram with the likelihood estimation for the PDF. The model used here is described in Equation (E.22). The flux is scaled to values between 0 and 10.

### E.3 Timmer & König Algorithm

As already mentioned, Timmer and Koenig developed a method to generate artificial light curves featuring the same PSD as the original light curve. This approach is also used by Emmanoulopoulos et al., which is why I will describe it briefly. To obtain a time series from a given PSD  $\mathcal{P}(f_j)$ , given Fourier frequencies  $f_j$  and length  $N_{\text{TK}}$  of the light curve, the following steps have to be performed:

1. Draw two normal distributed random numbers from  $\mathcal{N}(\mu = 0, \sigma = 1)$  for  $j = 0 \dots N_{\text{TK}}/2$  (even) or  $j = 0 \dots (N_{\text{TK}} - 1)/2$  (odd) and multiply them by  $\sqrt{\frac{1}{2}\mathcal{P}(f_j)}$ . The resulting numbers are used as real and imaginary part of the complex numbers that form the DFT of the simulated light curve. If  $N_{\text{TK}}$  is even,  $\text{DFT}(f_{N_y} = f_{N_{\text{TK}}/2})$  is real valued and only one random number has to be sampled.
2. To create a real valued light curve, calculate the complex conjugates of the sampled complex numbers for  $j = (N_{\text{TK}}/2) + 1 \dots N_{\text{TK}}$  (even) or  $j = (N_{\text{TK}} + 1)/2 \dots N_{\text{TK}}$  (odd).
3. The simulated light curve can now be calculated as the inverse Fourier transform of the sampled components.

Analog to the PSD estimation, we have to take care of red noise leakage and aliasing when we sample a light curve from a given PSD. The PSD of the simulated light curve can be different from the given one because of these effects.

To avoid red noise leakage in the simulated data, we create a light curve that is much longer than the original light curve ( $N_{\text{TK}} = 100 \dots 1000 \cdot N$ ) and take a random sequence of length  $N$  from the simulated light curve. The aliasing can be taken into account by simulating a light curve with smaller time bins  $t_{\text{bin}}$  and re-bin the result afterwards. How strong the effect of red noise leakage and aliasing is in the individual case, depends on the features of the original light curve and its sampling. Anyhow, a user should always keep this effects in mind. Both approaches to correct for red noise leakage and aliasing are implemented in my code.

### E.4 Emmanoulopoulos Loop

After the PSD and PDF of a light curve are successfully estimated, the actual light curve sampling can finally start. For this purpose, I follow the algorithm proposed by Emmanoulopoulos, McHardy, and Papadakis.

1. Create an artificial light curve  $x_{\text{TK}}$  from a given PSD  $\mathcal{P}(f_j)$  with the method by Timmer & König. The input PSD should not contain the Poisson noise (denoted as  $c$  in Equation (E.18)), which is added later. This light curve features the correct PSD, but not the PDF. Calculate the DFT of  $x_{\text{TK}}$  and the amplitudes  $\mathcal{A}_{\text{TK}}$  and phases  $\Phi_{\text{TK}}$  for the Fourier components.
2. Sample  $N$  values from the PDF, denoted as  $x_{\text{sim}, 1}$  and also calculate the DFT of  $x_{\text{sim}, 1}$ , and  $\mathcal{A}_{\text{sim}, 1}$  and  $\Phi_{\text{sim}, 1}$  from the Fourier components.

3. **Spectral adjustment:** Create new Fourier components by combining  $\mathcal{A}_{\text{TK}}$  and  $\Phi_{\text{sim},1}$  and calculate a new light curve  $x_{\text{adjust}}$  from the inverse Fourier transform of the combined components. This light curve still features the correct PSD, but a different PDF.
4. **Amplitude adjustment:** Order the  $x_{\text{sim},1}$  based on the ranking of  $x_{\text{adjust}}$ : The highest value of  $x_{\text{adjust}}$  is replaced with the highest value of  $x_{\text{sim},1}$  and so on. The resulting light curve now features the desired PDF, but the PSD has changed.
5. Repeat the process from step 2, sampling  $x_{\text{sim},2}$  and replacing the values from  $x_{\text{sim},1}$  with  $x_{\text{sim},2}$ . This is done several times until  $x_{\text{sim},k}$  does not change anymore and the  $P_k(f_j)$  follows the given  $\mathcal{P}(f_j)$ .

After a time series  $x$  is sampled this way, the Poisson noise has to be added to the light curve to take the count statistics of the detector into account. The light curve becomes

$$x_{\text{sim, Poisson}}(t_i) \propto \frac{\text{Poisson}[\mu = x_{\text{sim}}(t_i)\Delta t]}{\Delta t} \quad \text{for } i = 1 \dots N \quad (\text{E.26})$$

with the time bin width  $\Delta t$  and the Poisson distribution with mean  $x_{\text{sim}}(t_i)$ .

After this step is done, the final simulated light curve is obtained by interpolating and re-sampling the evenly sampled light curve  $x_{\text{sim, Poisson}}(t_i)$  in the same way as the original time series and re-scaling the flux values. For the purpose of cross-correlation studies, (several) thousand light curves have to be simulated to calculate the confidence levels.

The complete implementation of the described method is implemented in Python and can be found at GitHub.

 <https://github.com/lena-lin/emmanoulopoulos>



## Appendix F

### Funding Acknowledgement and Data Usage Statement

#### DFG Funding

Part of the work on this thesis has been supported by Deutsche Forschungsgemeinschaft (DFG) – project number 124020371 – within the Collaborative Research Center SFB 876 “Providing Information by Resource-Constrained Analysis”, DFG project number 124020371, SFB project C3.

#### MAGIC

The data published by Ansoldi et al. [22] used in section 16.2 is available at <http://vobs.magic.pic.es/fits/>.

#### Fermi-LAT

The *Fermi*-LAT data is available at the LAT Data Server at <https://fermi.gsfc.nasa.gov/cgi-bin/ssc/LAT/LATDataQuery.cgi>. For this work, three data sets are were used for the source NGC 1275:

- 2016-09-01 to 2016-12-31: MAGIC low state
- 2017-01-01 to 2017-01-03: MAGIC flaring state
- 2008-08-04 to 2019-06-26: long-term light curve

All data sets cover an energy range from 100 MeV to 300 GeV.

#### Alma

The ALMA Calibrator Source Catalogue is available at <https://almascience.eso.org/sc/>.

This study makes use of the following ALMA data: ADS/JAO.ALMA#2011.0.00001.CAL. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

#### Metsähovi

The 37 GHz observations were made with the 13.7 m diameter Metsähovi radio telescope. The observations are ON–ON observations, alternating the source and the sky in each feed horn. A typical integration time to obtain one flux density data point is between 1200 and

## *Appendix F Funding Acknowledgement and Data Usage Statement*

1400 s. The detection limit of our telescope at 37 GHz is on the order of 0.2 Jy under optimal conditions. Data points with a signal-to-noise ratio  $< 4$  are handled as non-detections. The flux density scale is set by observations of DR 21. A detailed description of the data reduction and analysis is given in Teräsranta et al. [159].

### **OVRO**

This research has made use of data from the OVRO 40-m monitoring program (Richards et al. [140]), supported by private funding from the California Institute of Technology and the Max Planck Institute for Radio Astronomy, and by NASA grants NNX08AW31G, NNX11A043G, and NNX14AQ89G and NSF grants AST-0808050 and AST-1109911.

### **MOJAVE: VLBA 15 GHz**

This research has made use of data from the MOJAVE database that is maintained by the MOJAVE team (Lister et al. [110]).

### **BU-BLAZAR Program: VLBA 43 GHz**

This study makes use of VLBA data from the VLBA-BU Blazar Monitoring Program (BEAM-ME and VLBA-BU-BLAZAR; <http://www.bu.edu/blazars/BEAM-ME.html>), funded by NASA through the Fermi Guest Investigator Program. The VLBA is an instrument of the National Radio Astronomy Observatory. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated by Associated Universities, Inc.

The data of 3C 84 is available at [https://www.bu.edu/blazars/VLBA\\_GLAST/0316.html](https://www.bu.edu/blazars/VLBA_GLAST/0316.html)



## Bibliography

For closed access publications, also an open-access version, usually from the [arXiv](#), is provided. Where an open access version could not be found, this is noted.

1. H. Abdalla et al. “The HESS Galactic Plane Survey”. *Astronomy & Astrophysics* 612, 2018, A1.  
DOI: [10.1051/0004-6361/201732098](#)
2. A. Abdo et al. “Gamma-Ray Flares from the Crab Nebula”. *Science* 331:6018, 2011, pages 739–742.  
DOI: [10.1126/science.1199705](#)
3. A. A. Abdo et al. “Fermi Discovery of Gamma-Ray Emission from NGC 1275”. *ApJ* 699:1, 2009, page 31.  
DOI: [10.1088/0004-637x/699/1/31](#)
4. J. Abraham et al. “Properties and Performance of the Prototype Instrument for the Pierre Auger Observatory”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 523:1-2, 2004, pages 50–95.  
DOI: [10.1016/j.nima.2003.12.012](#). <https://web.physik.rwth-aachen.de/~wiebusch/Publications/Auger/AUGER-nim-2003.pdf>.  
Note: Last visited: 2021-08-07
5. B. Acharya et al. (The Cherenkov Telescope Array Consortium Collaboration). “Science with the Cherenkov Telescope Array”, 2018.  
DOI: [10.1142/10986](#). [arXiv: 1709.07997](#)
6. M. Ackermann et al. “The Fermi Large Area Telescope on Orbit: Event Classification, Instrument Response Functions, and Calibration”. *The Astrophysical Journal Supplement Series* 203:1, 2012, page 4.  
DOI: [10.1088/0067-0049/203/1/4](#)
7. P. A. Ade et al. “Planck Early Results. I. The Planck Mission”. *Astronomy & Astrophysics* 536, 2011, A1.  
DOI: [10.1051/0004-6361/201116464](#)
8. N. Aghanim et al. “Planck 2018 Results-I. Overview and the Cosmological Legacy of Planck”. *Astronomy & Astrophysics* 641, 2020, A1.  
DOI: [10.1051/0004-6361/201833880](#)
9. F. Aharonian, M. Barkov, and D. Khangulyan. “Scenarios for Ultrafast Gamma-Ray Variability in AGN”. *ApJ* 841:1, 2017, page 61.  
DOI: [10.3847/1538-4357/aa7049](#)

10. A. Albert et al. “A Survey of Active Galaxies at TeV Photon Energies with the HAWC Gamma-Ray Observatory”. *The Astrophysical Journal* 907:2, 2021, page 67.  
DOI: [10.3847/1538-4357/abca9a](https://doi.org/10.3847/1538-4357/abca9a). arXiv: [2009.09039v1](https://arxiv.org/abs/2009.09039v1)
11. J. Albert et al. “Implementation of the Random Forest Method for the Imaging Atmospheric Cherenkov Telescope MAGIC”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 588:3, 2008, pages 424–432.  
DOI: [10.1016/j.nima.2007.11.068](https://doi.org/10.1016/j.nima.2007.11.068). arXiv: [0709.3719](https://arxiv.org/abs/0709.3719)
12. J. Albert et al. “VHE  $\gamma$ -Ray Observation of the Crab Nebula and its Pulsar with the MAGIC Telescope”. *The Astrophysical Journal* 674:2, 2008, pages 1037–1055.  
DOI: [10.1086/525270](https://doi.org/10.1086/525270)
13. J. Aleksić et al. “Measurement of the Crab Nebula spectrum over three decades in energy with the MAGIC telescopes”. *Journal of High Energy Astrophysics* 5-6, 2015, pages 30–38. ISSN: 2214-4048.  
DOI: [10.1016/j.jheap.2015.01.002](https://doi.org/10.1016/j.jheap.2015.01.002). <https://www.sciencedirect.com/science/article/pii/S2214404815000038>
14. J. Aleksić et al. “Performance of the MAGIC stereo system obtained with Crab Nebula data”. *Astroparticle Physics* 35:7, 2012, pages 435–448. ISSN: 0927-6505.  
DOI: [10.1016/j.astropartphys.2011.11.007](https://doi.org/10.1016/j.astropartphys.2011.11.007). arXiv: [1108.1477v2](https://arxiv.org/abs/1108.1477v2)
15. J. Aleksić et al. “The major upgrade of the MAGIC telescopes, Part I: The hardware improvements and the commissioning of the system”. *Astroparticle Physics* 72, 2016, pages 61–75. ISSN: 0927-6505.  
DOI: [10.1016/j.astropartphys.2015.04.004](https://doi.org/10.1016/j.astropartphys.2015.04.004). arXiv: [1409.6073](https://arxiv.org/abs/1409.6073)
16. J. Aleksić et al. “The major upgrade of the MAGIC telescopes, Part II: A performance study using observations of the Crab Nebula”. *Astroparticle Physics* 72, 2016, pages 76–94. ISSN: 0927-6505.  
DOI: [10.1016/j.astropartphys.2015.02.005](https://doi.org/10.1016/j.astropartphys.2015.02.005). arXiv: [1409.5594v3](https://arxiv.org/abs/1409.5594v3)
17. J. Aleksić et al. “Contemporaneous Observations of the Radio Galaxy NGC 1275 from Radio to Very High Energy  $\gamma$ -Rays”. *A&A* 564, 2014, A5.  
DOI: [10.1051/0004-6361/201322951](https://doi.org/10.1051/0004-6361/201322951)
18. J. Aleksić et al. “MAGIC gamma-ray Telescope Observation of the Perseus Cluster of Galaxies: Implications for Cosmic Rays, Dark Matter, and NGC 1275”. *The Astrophysical Journal* 710:1, 2010, pages 634–647.  
DOI: [10.1088/0004-637x/710/1/634](https://doi.org/10.1088/0004-637x/710/1/634)
19. R. Alfaro et al. “All-Particle Cosmic Ray Energy Spectrum Measured by the HAWC Experiment from 10 to 500 TeV”. *Physical Review D* 96:12, 2017. ISSN: 2470-0029.  
DOI: [10.1103/physrevd.96.122001](https://doi.org/10.1103/physrevd.96.122001). arXiv: [1710.00890](https://arxiv.org/abs/1710.00890)

20. R. A. Alpher and R. C. Herman. “On the Relative Abundance of the Elements”. *Physical Review* 74:12, 1948, page 1737.  
DOI: [10.1103/PhysRev.74.1737](https://doi.org/10.1103/PhysRev.74.1737).  
Note: No open-access version available.
21. H. Anderhub et al. “Design and Operation of FACT—the First G-APD Cherenkov Telescope”. *Journal of Instrumentation* 8:06, 2013, P06008.  
DOI: [10.1088/1748-0221/8/06/P06008](https://doi.org/10.1088/1748-0221/8/06/P06008)
22. S. Ansoldi et al. “Gamma-Ray Flaring Activity of NGC 1275 in 2016-2017 Measured by MAGIC”. *A&A* 617, 2018, A91.  
DOI: [10.1051/0004-6361/201832895](https://doi.org/10.1051/0004-6361/201832895)
23. Astropheiler Stockert e.V.  
<https://astropheiler.de>
24. Astropy Collaboration et al. “Astropy: A Community Python Package for Astronomy”. *A&A* 558, A33, 2013, A33.  
DOI: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068). arXiv: [1307.6212](https://arxiv.org/abs/1307.6212) [[astro-ph.IM](https://arxiv.org/archive/astro)]
25. Astropy Collaboration et al. “The Astropy Project: Building an Open-Science Project and Status of the v2.0 Core Package”. *AJ* 156:3, 123, 2018, page 123.  
DOI: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f). arXiv: [1801.02634](https://arxiv.org/abs/1801.02634) [[astro-ph.IM](https://arxiv.org/archive/astro)]
26. W. B. Atwood et al. “The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission”. *The Astrophysical Journal* 697:2, 2009, pages 1071–1102. ISSN: 1538-4357.  
DOI: [10.1088/0004-637x/697/2/1071](https://doi.org/10.1088/0004-637x/697/2/1071)
27. D. Barber, W. Donaldson, G. K. Miley, and H. Smith. “Angular Diameters of Some Radio Sources with Anomalous High-Frequency Spectra”. *Nature* 209:5025, 1966, pages 753–755
28. M. Bayer. “SQLAlchemy”. In: *The Architecture of Open Source Applications Volume II: Structure, Scale, and a Few More Fearless Hacks*. Ed. by A. Brown and G. Wilson. aosabook.org, 2012.  
<http://aosabook.org/en/sqlalchemy.html>
29. A. Bell. “The Acceleration of Cosmic Rays in Shock Fronts—I”. *Monthly Notices of the Royal Astronomical Society* 182:2, 1978, pages 147–156.  
DOI: [10.1093/mnras/182.2.147](https://doi.org/10.1093/mnras/182.2.147)
30. A. Bennett and F. Simth. “The Preparation of the Revised 3C Catalogue of Radio Sources”. *Monthly Notices of the Royal Astronomical Society* 125:1, 1962, pages 75–86.  
DOI: [10.1093/mnras/125.1.75](https://doi.org/10.1093/mnras/125.1.75)
31. J. Blackburn, R. Shaw, H. Payne, J. Hayes, et al. “FTOOLS: A general Package of Software to Manipulate FITS Files”. *Astrophysics Source Code Library*, 1999, ascl-9912

## Bibliography

32. R. Blandford and D. Payne. “Hydromagnetic Flows from Accretion Discs and the Production of Radio Jets”. *Monthly Notices of the Royal Astronomical Society* 199:4, 1982, pages 883–903.  
DOI: [10.1093/mnras/199.4.883](https://doi.org/10.1093/mnras/199.4.883)
33. R. D. Blandford and R. L. Znajek. “Electromagnetic Extraction of Energy from Kerr Black Holes”. *MNRAS* 179:3, 1977, pages 433–456.  
DOI: [10.1093/mnras/179.3.433](https://doi.org/10.1093/mnras/179.3.433)
34. D. A. Boboltz et al. *US Naval Observatory VLBI Analysis Center*. Technical report. NAVAL OBSERVATORY WASHINGTON DC, 2009
35. L. Breiman, J. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. Chapman & Hall/CRC, 1984
36. L. Breiman. “Random forests”. *Machine learning* 45:1, 2001, pages 5–32.  
DOI: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
37. K. Brügge. “Unmasking the Gamma-Ray Sky, Comprehensive and Reproducible Analysis for Cherenkov Telescopes”. PhD thesis. Dortmund: TU Dortmund University, 2019
38. R. Brun and F. Rademakers. “ROOT—an Object Oriented Data Analysis Framework”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389:1-2, 1997, pages 81–86
39. R. Brun et al. *root-project/root: v6.18/02*. Version v6-18-02. 2019.  
DOI: [10.5281/zenodo.3895860](https://doi.org/10.5281/zenodo.3895860). [10.5281/zenodo.3895860](https://doi.org/10.5281/zenodo.3895860)
40. J. B. Buß. “Bad Moon Rising? Studies on the Performance of the First G-APD Cherenkov Telescope under Bright Light Conditions Using SiPMs for Gamma-Ray Observations”. PhD thesis. Dortmund: TU Dortmund University, 2020
41. A. Celotti, G. Ghisellini, and A. C. Fabian. “Bulk Comptonization Spectra in Blazars”. *Monthly Notices of the Royal Astronomical Society* 375:2, 2007, pages 417–424.  
DOI: [10.1111/j.1365-2966.2006.11289.x](https://doi.org/10.1111/j.1365-2966.2006.11289.x)
42. M. Cerruti, A. Zech, C. Boisson, and S. Inoue. “A Hadronic Origin for Ultra-High-Frequency-Peaked BL Lac Objects”. *Monthly Notices of the Royal Astronomical Society* 448:1, 2015, pages 910–927.  
DOI: [10.1093/mnras/stu2691](https://doi.org/10.1093/mnras/stu2691)
43. P. A. Cherenkov. “Visible Light from Clear Liquids Under the Action of Gamma Radiation”. *Comptes Rendus (Doklady) de l’Académie des Sciences de l’URSS* 2:8, 1934, pages 451–454.  
Note: No open-access version available.

44. P. H. van Cittert. “Die wahrscheinliche Schwingungsverteilung in einer von einer Lichtquelle direkt oder mittels einer Linse beleuchteten Ebene”. *Physica* 1:1-6, 1934, pages 201–210.  
DOI: [10.1016/S0031-8914\(34\)90026-4](https://doi.org/10.1016/S0031-8914(34)90026-4)
45. B. Clark. “Coherence in Radio Astronomy”. In: *Synthesis Imaging in Radio Astronomy II*. Vol. 180. 1999, page 1
46. E. H. T. Collaboration et al. “First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole”. *Astrophys. J. Lett* 875:1, 2019, page L1.  
DOI: [10.3847/2041-8213/ab0ec7](https://doi.org/10.3847/2041-8213/ab0ec7)
47. F. Collaboration. *Fermitools Documentation*. 2021.  
<https://fermi.gsfc.nasa.gov/ssc/Data/Analysis/documentation/>
48. P. A. collaboration et al. “The Pierre Auger Observatory and its Upgrade”. *Science Reviews-from the end of the world* 1:4, 2020, pages 8–33.  
DOI: [10.52712/sciencereviews.v1i4.31](https://doi.org/10.52712/sciencereviews.v1i4.31)
49. J. Comella et al. “Crab Nebula Pulsar NP 0532”. *Nature* 221:5179, 1969, pages 453–454.  
DOI: [10.1038/221453a0](https://doi.org/10.1038/221453a0).  
Note: No open-access version available.
50. T. Cornwell. “Hogbom’s CLEAN algorithm. Impact on Astronomy and Beyond-Commentary on: Högbom JA, 1974, A&AS, 15, 417”. *Astronomy & Astrophysics* 500:1, 2009, pages 65–66.  
DOI: [10.1051/0004-6361/200912148](https://doi.org/10.1051/0004-6361/200912148)
51. R. D. Cousins, J. T. Linnemann, and J. Tucker. “Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 595:2, 2008, pages 480–501. ISSN: 0168-9002.  
DOI: [10.1016/j.nima.2008.07.086](https://doi.org/10.1016/j.nima.2008.07.086). arXiv: [physics/0702156v4](https://arxiv.org/abs/physics/0702156v4). <http://dx.doi.org/10.1016/j.nima.2008.07.086>
52. R. A. Daly and A. P. Marscher. “The Gasdynamics of Compact Relativistic Jets”. *Astrophys. J.; (United States)* 334, 1988.  
DOI: [10.1086/166858](https://doi.org/10.1086/166858). <https://www.osti.gov/biblio/6392250>
53. A. J. Deason et al. “The Edge of the Galaxy”. *Monthly Notices of the Royal Astronomical Society* 496:3, 2020, pages 3929–3942.  
DOI: [10.1093/mnras/staa1711](https://doi.org/10.1093/mnras/staa1711). <https://dro.dur.ac.uk/31569/1/31569.pdf?DDD25+>
54. C. Deil et al. “Gammapy - A Prototype for the CTA Science Tools”. In: *35th International Cosmic Ray Conference (ICRC2017)*. Vol. 301. International Cosmic Ray Conference. 2017, 766, page 766.  
arXiv: [1709.01751](https://arxiv.org/abs/1709.01751) [astro-ph.IM]

## Bibliography

55. C. Deil et al. “Open High-level Data Formats and Software for Gamma-Ray Astronomy”. In: *AIP Conference Proceedings*. Vol. 1792. 1. AIP Publishing LLC. 2017, page 070006
56. H. Dembinski and P. O. et al. “scikit-hep/iminuit”, 2020.  
DOI: [10.5281/zenodo.3949207](https://doi.org/10.5281/zenodo.3949207)
57. K. Denney et al. “Diverse Kinematic Signatures from Reverberation Mapping of the Broad-Line Region in AGNs”. *The Astrophysical Journal Letters* 704:2, 2009, page L80.  
DOI: [10.1088/0004-637X/704/2/L80](https://doi.org/10.1088/0004-637X/704/2/L80)
58. W. Dent and F. Haddock. “A New Class of Radio Source Spectra”. *Nature* 205:4970, 1965, pages 487–488.  
DOI: [10.1038/205487a0](https://doi.org/10.1038/205487a0)
59. P. Doll et al. *The Karlsruhe Cosmic Ray Project KASCADE*. Technical report. 1990
60. A. Dominguez et al. “ExtraGalactic Background Light Inferred from AEGIS Galaxy-SED-type Fractions”. *Monthly Notices of the Royal Astronomical Society* 410:4, 2011, pages 2556–2578.  
DOI: [10.1111/j.1365-2966.2010.17631.x](https://doi.org/10.1111/j.1365-2966.2010.17631.x)
61. R. M. Dominik, L. Linhoff, D. Elsässer, and W. Rhode. “3C 84: A Possibly Precessing Jet in 43-GHz Observations”. *Monthly Notices of the Royal Astronomical Society* 503:4, 2021, pages 5448–5454.  
DOI: [10.1093/mnras/stab799](https://doi.org/10.1093/mnras/stab799)
62. R. M. Dominik. “3C84/NGC1275: Jet Precession in 43 GHz Data on Sub-Parsec Scales”. Technische Universität Dortmund, 2020
63. J. L. E. Dreyer. *A New General Catalogue of Nebulae and Clusters of Stars: Being the Catalogue of the Late Sir John FW Herschel*. Royal Astronomical Society, 1888
64. R. Edelson and J. Krolik. “The Discrete Correlation Function-A New Method for Analyzing Unevenly Sampled Variability Data”. *ApJ* 333, 1988, pages 646–659.  
<http://adsabs.harvard.edu/pdf/1988ApJ...333..646E>
65. D. O. Edge et al. “A survey of radio sources at a frequency of 159 Mc/s.” *Memoirs of the Royal Astronomical Society* 68, 1959, page 37.  
<http://astro.vaporia.com/start/threec.html>
66. F. Eisenhauer et al. “A Geometric Determination of the Distance to the Galactic Center”. *The Astrophysical Journal Letters* 597:2, 2003, page L121.  
DOI: [10.1086/380188](https://doi.org/10.1086/380188)
67. R. Ekers. “Non-Thermal Radio Astronomy”. *Astroparticle Physics* 53, 2014, pages 152–159.  
DOI: [10.1016/j.astropartphys.2013.05.012](https://doi.org/10.1016/j.astropartphys.2013.05.012)

68. D. Emmanoulopoulos, I. McHardy, and I. Papadakis. “Generating Artificial Light curves: Revisited and Updated”. *Monthly Notices of the Royal Astronomical Society* 433:2, 2013, pages 907–927.  
DOI: [10.1093/mnras/stt764](https://doi.org/10.1093/mnras/stt764)
69. H. Ewen and E. Purcell. “Sep. 1951. Observation of a Line in the Galactic Radio Spectrum: Radiation from Galactic Hydrogen at 1,420 Mc./sec”. *Nature* 168, page 356.  
<https://phys.cst.temple.edu/~tuf43817/PHYS3701/S19/Purcell121cmDiscovery.pdf>.  
Note: last visited: 2021-08-07
70. A. Fabian et al. “A Very Deep Chandra Observation of the Perseus Cluster: Shocks, Ripples and Conduction”. *Monthly Notices of the Royal Astronomical Society* 366:2, 2006, pages 417–428.  
DOI: [10.1111/j.1365-2966.2005.09896.x](https://doi.org/10.1111/j.1365-2966.2005.09896.x)
71. E. Fermi. “Galactic Magnetic Fields and the Origin of Cosmic Radiation.” *The Astrophysical Journal* 119, 1954, page 1.  
DOI: [10.4159/harvard.9780674366688.c105](https://doi.org/10.4159/harvard.9780674366688.c105)
72. *Fermitools: Cicerone*. 2021.  
<https://fermi.gsfc.nasa.gov/ssc/Data/Analysis/documentation/Cicerone/>.  
Note: Online documentation of the Fermitools and overview of the LAT and GRB detector, last visited: 2021-08-09
73. J. D. Finke. “External Compton Scattering in Blazar Jets and the Lotion of the Gamma-Ray Emitting Region”. *ApJ* 830:2, 2016, page 94.  
DOI: [10.3847/0004-637X/830/2/94](https://doi.org/10.3847/0004-637X/830/2/94)
74. J. D. Finke, C. D. Dermer, and M. Böttcher. “Synchrotron Self-Compton Analysis of TeV X-Ray-Selected BL Lacertae Objects”. *The Astrophysical Journal* 686:1, 2008, pages 181–194.  
DOI: [10.1086/590900](https://doi.org/10.1086/590900)
75. D. J. Fixsen. “THE TEMPERATURE OF THE COSMIC MICROWAVE BACKGROUND”. *The Astrophysical Journal* 707:2, 2009, pages 916–920.  
DOI: [10.1088/0004-637x/707/2/916](https://doi.org/10.1088/0004-637x/707/2/916)
76. A. Franceschini, G. Rodighiero, and M. Vaccari. “Extragalactic Optical-Infrared Background Radiation, its Time Evolution and the Cosmic Photon-Photon Opacity”. *Astronomy & Astrophysics* 487:3, 2008, pages 837–852.  
DOI: [10.1051/0004-6361:200809691](https://doi.org/10.1051/0004-6361:200809691)
77. Y. Fujita and H. Nagai. “Discovery of a New Subparsec Counterjet in NGC 1275: The Inclination Angle and the Environment”. *MNRAS* 465:1, 2016, pages L94–L98.  
DOI: [10.1093/mnrasl/slw217](https://doi.org/10.1093/mnrasl/slw217)



## Bibliography

78. C. M. Gaskell. “Direct Evidence for Gravitational Domination of the Motion of Gas within One Light-week of the central Object in NGC 4151 and the determination of the Mass of the probable Black Hole”. *The Astrophysical Journal* 325, 1988, pages 114–118.  
<http://adsabs.harvard.edu/pdf/1988ApJ...325..114G>.  
Note: last visited: 2021-08-07
79. C. M. Gaskell. “What Broad Emission Lines Tell Us About How Active Galactic Nuclei Work”. *New Astronomy Reviews* 53:7-10, 2009, pages 140–148.  
DOI: [10.1016/j.newar.2009.09.006](https://doi.org/10.1016/j.newar.2009.09.006)
80. M. Georganopoulos and D. Kazanas. “Decelerating Flows in TeV Blazars: A Resolution to the BL Lacertae-FR I Unification Problem”. *The Astrophysical Journal* 594:1, 2003, pages L27–L30.  
DOI: [10.1086/378557](https://doi.org/10.1086/378557)
81. A. M. Ghez et al. “The First Measurement of Spectral Lines in a Short-Period Star Bound to the Galaxy’s Central Black Hole: A Paradox of Youth”. *The Astrophysical Journal* 586:2, 2003, pages L127–L131.  
DOI: [10.1086/374804](https://doi.org/10.1086/374804)
82. G. Ghisellini, F. Tavecchio, G. Bodo, and A. Celotti. “TeV Variability in Blazars: How Fast can it Be?”. *Monthly Notices of the Royal Astronomical Society: Letters* 393:1, 2009, pages L16–L20.  
DOI: [10.1111/j.1745-3933.2008.00589.x](https://doi.org/10.1111/j.1745-3933.2008.00589.x)
83. G. Ghisellini, F. Haardt, and G. Matt. “Aborted Jets and the X-Ray Emission of Radio-Quiet AGNs”. *Astronomy & Astrophysics* 413:2, 2004, pages 535–545.  
DOI: [10.1051/0004-6361:20031562](https://doi.org/10.1051/0004-6361:20031562)
84. G. Ghisellini and P. Madau. “On the Origin of the  $\gamma$ -Ray Emission in Blazars”. *Monthly Notices of the Royal Astronomical Society* 280:1, 1996, pages 67–76.  
DOI: [10.1093/mnras/280.1.67](https://doi.org/10.1093/mnras/280.1.67)
85. D. Giannios, D. A. Uzdensky, and M. C. Begelman. “Fast TeV Variability in Blazars: Jets in a Jet”. *MNRAS* 395:1, 2009, pages L29–L33.  
DOI: [10.1111/j.1745-3933.2009.00635.x](https://doi.org/10.1111/j.1745-3933.2009.00635.x)
86. R. C. Gilmore, F. Prada, and J. Primack. “Modelling Gamma-Ray Burst Observations by Fermi and MAGIC Including Attenuation Due to Diffuse Background Light”. *Monthly Notices of the Royal Astronomical Society* 402:1, 2010, pages 565–574.  
DOI: [10.1111/j.1365-2966.2009.15909.x](https://doi.org/10.1111/j.1365-2966.2009.15909.x)
87. G. Giovannini et al. “A Wide and Collimated Radio Jet in 3C84 on the Scale of a Few Hundred Gravitational Radii”. *Nature Astronomy* 2:6, 2018, page 472.  
DOI: [10.1038/s41550-018-0431-2](https://doi.org/10.1038/s41550-018-0431-2). arXiv: [1804.02198v1](https://arxiv.org/abs/1804.02198v1)



88. L. Grady. “Random Walks for Image Segmentation”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28:11, 2006, pages 1768–1783.  
DOI: [10.1109/TPAMI.2006.233](https://doi.org/10.1109/TPAMI.2006.233). <http://leogrady.net/wp-content/uploads/2017/01/grady2006random.pdf>
89. J. E. Greene and L. C. Ho. “Estimating Black Hole Masses in Active Galaxies Using the H $\alpha$  Emission Line”. *ApJ* 630:1, 2005, page 122.  
DOI: [10.1086/431897](https://doi.org/10.1086/431897)
90. F. Halzen, A. Kheirandish, T. Weisgarber, and S. P. Wakely. “On the Neutrino Flares from the Direction of TXS 0506+ 056”. *The Astrophysical Journal Letters* 874:1, 2019, page L9.  
DOI: [10.3847/2041-8213/ab0d27](https://doi.org/10.3847/2041-8213/ab0d27)
91. C. R. Harris et al. “Array Programming with NumPy”. *Nature* 585:7825, 2020, pages 357–362.  
DOI: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)
92. D. Heck et al. *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*. Technical report FZKA 6019. Forschungszentrum Karlsruhe, 1998.  
<https://digbib.ubka.uni-karlsruhe.de/volltexte/fzk/6019/6019.pdf>
93. O. Hervet et al. “Shocks in Relativistic Transverse Stratified Jets-A New Paradigm for Radio-Loud AGN”. *Astronomy & Astrophysics* 606, 2017, A103.  
[10.1051/0004-6361/201730745](https://doi.org/10.1051/0004-6361/201730745)
94. A. M. Hillas. “Cerenkov Light Images of EAS Produced by Primary Gamma”. In: *19th International Cosmic Ray Conference (ICRC19), Volume 3*. Vol. 3. 1985.  
<http://adsabs.harvard.edu/pdf/1985ICRC....3...445H>
95. J. A. Hodgson et al. “KVN Observations Reveal Multiple Gamma-Ray Emission Regions in 3C 84?” *MNRAS* 475:1, 2018, pages 368–378.  
DOI: [10.1093/mnras/stx3041](https://doi.org/10.1093/mnras/stx3041)
96. J. Högbom. “Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines”. *Astronomy and Astrophysics Supplement Series* 15, 1974, page 417.  
<http://adsabs.harvard.edu/pdf/1974A&AS...15..417H>
97. F. James and M. Roos. “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations”. *Comput. Phys. Commun.* 10, 1975, pages 343–367.  
DOI: [10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9). <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.158.9157&rep=rep1&type=pdf>
98. C. Jansky. “The Discovery and Identification by Karl Guthe Jansky of Electromagnetic Radiation of Extraterrestrial Origin in the Radio Spectrum”. *Proceedings of the IRE* 46:1, 1958, pages 13–15.  
DOI: [10.1109/JRPROC.1958.286704](https://doi.org/10.1109/JRPROC.1958.286704).  
Note: No open-access version available

## Bibliography

99. S. Jorstad and A. Marscher. “The VLBA-BU-BLAZAR Multi-Wavelength Monitoring Program”. *Galaxies* 4:4, 2016, page 47.  
DOI: [10.3390/galaxies4040047](https://doi.org/10.3390/galaxies4040047)
100. R. P. Kerr. “Gravitational Field of a Spinning Mass as an Example of Algebraically Special Metrics”. *Phys. Rev. Lett.* 11, 5 1963, pages 237–238.  
DOI: [10.1103/PhysRevLett.11.237](https://doi.org/10.1103/PhysRevLett.11.237).  
Note: No open-access version available
101. J.-Y. Kim et al. “Spatially Resolved Origin of Millimeter-Wave Linear Polarization in the Nuclear Region of 3C 84”. *Astronomy & Astrophysics* 622, 2019, A196
102. J. W. Kirchner. “Aliasing in  $1/f \propto$  Noise Spectra: Origins, Consequences, and Remedies”. *Physical Review E* 71:6, 2005, page 066110.  
DOI: [10.1103/PhysRevE.71.066110](https://doi.org/10.1103/PhysRevE.71.066110). [http://seismo.berkeley.edu/~kirchner/reprints/2005\\_74\\_spectral\\_aliasing.pdf](http://seismo.berkeley.edu/~kirchner/reprints/2005_74_spectral_aliasing.pdf)
103. K. Kosack et al. *ctapipe*.  
DOI: [10.5281/zenodo.4746317](https://doi.org/10.5281/zenodo.4746317). <https://github.com/cta-Observatory/ctapipe/tree/v0.11.0>
104. P. Kumar, R. Hernández, Ž. Bošnjak, and R. B. Duran. “Maximum Synchrotron Frequency for Shock-Accelerated Particles”. *Monthly Notices of the Royal Astronomical Society: Letters* 427:1, 2012, pages L40–L44.  
DOI: [10.1111/j.1745-3933.2012.01341.x](https://doi.org/10.1111/j.1745-3933.2012.01341.x)
105. E. Kun et al. “Cosmic Neutrinos from Temporarily Gamma-Suppressed Blazars”. *The Astrophysical Journal Letters* 911:2, 2021, page L18.  
DOI: [10.3847/2041-8213/abf1ec](https://doi.org/10.3847/2041-8213/abf1ec). arXiv: 2009.09792v2
106. V. Kurilchik. “The Knots in the M87 Jet”. *Soviet Astronomy Letters* 5, 1979, pages 301–303.  
<http://adsabs.harvard.edu/pdf/1979SvAL....5..301K>
107. M. F. L’Annunziata. *Nuclear Radiation, its Interaction with Matter and Radioisotope Decay*. AcadEmic Press San Diego, 2003, pages 1–122.  
DOI: [10.1016/B978-0-12-436603-9.X5000-5](https://doi.org/10.1016/B978-0-12-436603-9.X5000-5)
108. T.-P. Li and Y.-Q. Ma. “Analysis Methods for Results in Gamma-Ray Astronomy”. *The Astrophysical Journal* 272, 1983, pages 317–324.  
<http://adsabs.harvard.edu/pdf/1983ApJ...272..317L7>
109. L. Linhoff et al. “Excluding possible sites of high-energy emission in 3C 84”. *Monthly Notices of the Royal Astronomical Society* 500:4, 2021, pages 4671–4677.  
DOI: [10.1093/mnras/staa3521](https://doi.org/10.1093/mnras/staa3521)
110. M. Lister et al. “MOJAVE. XV. VLBA 15 GHz Total Intensity and Polarization Maps of 437 Parces-Scale AGN Jets from 1996 to 2017”. *The Astrophysical Journal Supplement Series* 234:1, 2018, page 12

111. F. Mannucci, M. Salvati, and R. Stanga. “Line Profile and Variability Data to Probe the Broad-Line Region Geometry: Of Disks and Nests”. *The Astrophysical Journal* 394, 1992, pages 98–103.  
<http://adsabs.harvard.edu/pdf/1992ApJ...394...98M>
112. W. Max-Moerbeck et al. “A Method for the Estimation of the Significance of Cross-Correlations in Unevenly Sampled Red-Noise Time Series”. *Monthly Notices of the Royal Astronomical Society* 445:1, 2014, pages 437–459.  
DOI: [10.1093/mnras/stu1707](https://doi.org/10.1093/mnras/stu1707)
113. J. C. McKinney. “General Relativistic Magnetohydrodynamic Simulations of the Jet Formation and Large-Scale Propagation from Black Hole Accretion Systems”. *Monthly Notices of the Royal Astronomical Society* 368:4, 2006, pages 1561–1582.  
DOI: [10.1111/j.1365-2966.2006.10256.x](https://doi.org/10.1111/j.1365-2966.2006.10256.x)
114. J. McMullin et al. “CASA Architecture and Applications”. In: *Astronomical Data Analysis Software and Systems XVI*. Vol. 376. ASP Conf. Ser. Astron. Soc. Pac., San Francisco. 2007, page 127.  
<http://adsabs.harvard.edu/pdf/2007ASPC..376..127M>.
115. R. Mirzoyan and E. Lorenz. “On the Calibration Accuracy of Light Sensors in Atmospheric Cherenkov Fluorescence and Neutrino Experiment”. In: *International Cosmic Ray Conference*. Vol. 7. 1997, pages 265–268.  
<http://adsabs.harvard.edu/pdf/1997ICRC....7..265M>
116. I. Morison. “50 Years of the Lovell Telescope”. *Astronomy & Geophysics* 48:5, 2007, pages 5–23.  
DOI: [10.1111/j.1468-4004.2007.48523.x](https://doi.org/10.1111/j.1468-4004.2007.48523.x)
117. N. Murray, J. Chiang, S. Grossman, and G. Voit. “Accretion Disk Winds from Active Galactic Nuclei”. *The Astrophysical Journal* 451, 1995, page 498.  
<http://adsabs.harvard.edu/pdf/1995ApJ...451..498M>
118. H. Nagai et al. “Nature of Radio Feature Formed by Re-Started Jet Activity in 3C 84 and its Relation with  $\gamma$ -Ray Emissions”. *Astronomische Nachrichten* 337:1-2, 2016, pages 69–72.  
DOI: [10.1002/asna.201512267](https://doi.org/10.1002/asna.201512267)
119. H. Nagai et al. “VLBI Monitoring of 3C 84 (NGC 1275) in Early Phase of the 2005 Outburst”. *PASJ* 62:2, 2010, pages L11–L15.  
DOI: [10.1093/pasj/62.2.L11](https://doi.org/10.1093/pasj/62.2.L11)
120. M. Nakamura and K. Asada. “The Parabolic Jet Structure in M87 as a Magnetohydrodynamic Nozzle”. *The Astrophysical Journal* 775:2, 2013, page 118.  
DOI: [10.1088/0004-637x/775/2/118](https://doi.org/10.1088/0004-637x/775/2/118)
121. P. J. Napier et al. “The Very Long Baseline Array”. *Proceedings of the IEEE* 82:5, 1994, pages 658–672.  
DOI: [10.1109/5.284733](https://doi.org/10.1109/5.284733). <http://adsabs.harvard.edu/pdf/1994IAUS..158..117N>

## Bibliography

122. R. Narayan and E. Quataert. “Black Hole Accretion”. *Science* 307:5706, 2005, pages 77–80.  
DOI: [10.1126/science.1105746](https://doi.org/10.1126/science.1105746)
123. C. Nigro et al. “Towards Open and Reproducible Multi-Instrument Analysis in Gamma-Ray Astronomy”. *Astronomy & Astrophysics* 625, A10, 2019, A10.  
DOI: [10.1051/0004-6361/201834938](https://doi.org/10.1051/0004-6361/201834938). arXiv: [1903.06621](https://arxiv.org/abs/1903.06621) [astro-ph.HE]
124. C. Nigro. “Study of Persistent and Flaring Gamma-Ray Emission from Active Galactic Nuclei with the MAGIC Telescopes and Prospects for Future Open Data Formats in Gamma-Ray Astronomy”. PhD thesis. Berlin: Humboldt-Universität zu Berlin, 2019
125. M. Nöthe. “Monitoring the High Energy Universe, Open, Reproducible, Machine Learning Based Analysis for the First G-APD Cherenkov Telescope”. PhD thesis. Dortmund: TU Dortmund, 2020
126. A. Offringa et al. “WSCLEAN: an Implementation of a Fast, Generic Wide-Field Imager for Radio Astronomy”. *Monthly Notices of the Royal Astronomical Society* 444:1, 2014, pages 606–619.  
DOI: [10.1093/mnras/stu1368](https://doi.org/10.1093/mnras/stu1368)
127. D. E. Osterbrock. “Observational Model of the Ionized Gas in Seyfert and Radio-Galaxy Nuclei”. *Proceedings of the National Academy of Sciences* 75:2, 1978, pages 540–544.  
DOI: [10.1073/pnas.75.2.540](https://doi.org/10.1073/pnas.75.2.540)
128. F. N. Owen, P. E. Hardee, and T. Cornwell. “High-Resolution, High dynamic Range VLA Images of the M87 Jet at 2 Centimeters”. *The Astrophysical Journal* 340, 1989, pages 698–707.  
DOI: <http://adsabs.harvard.edu/pdf/1989ApJ...340..698O>
129. T. Pearson and A. Readhead. “Image Formation by Self-Calibration in Radio Astronomy”. *Annual Review of Astronomy and Astrophysics* 22:1, 1984, pages 97–130.  
DOI: [10.1146/annurev.aa.22.090184.000525](https://doi.org/10.1146/annurev.aa.22.090184.000525)
130. R. Penrose. “Gravitational Collapse: The Role of General Relativity”. *Nuovo Cimento Rivista Serie* 1, 1969, page 252.  
Note: No open-access version available.
131. R. Penrose and R. Floyd. “Extraction of Rotational Energy from a Black Hole”. *Nature Physical Science* 229:6, 1971, pages 177–179.  
DOI: [10.1038/physci229177a0](https://doi.org/10.1038/physci229177a0).  
Note: No open-access version available.
132. B. Peterson. “The Broad-Line Region in Active Galactic Nuclei”. In: *Physics of Active Galactic Nuclei at all Scales*. Springer, 2006, pages 77–100.  
DOI: [10.1007/3-540-34621-X\\_3](https://doi.org/10.1007/3-540-34621-X_3). [http://www.astro.osu.edu/~peterson/Docs/chapter\\_3.pdf](http://www.astro.osu.edu/~peterson/Docs/chapter_3.pdf)

133. B. Punsly et al. “Revealing the Broad Line Region of NGC 1275: The Relationship to Jet Power”. *ApJ* 869:2, 2018, page 143.  
DOI: [10.3847/1538-4357/aaec75](https://doi.org/10.3847/1538-4357/aaec75)
134. R. U. C. for Radio Astronomy Techniques & Technologies (RATT). *Fundamentals of Radio Astronomy*.  
[https://github.com/ratt-ru/fundamentals\\_of\\_interferometry](https://github.com/ratt-ru/fundamentals_of_interferometry)
135. G. Reber. “Cosmic Static”. *Astrophysical Journal* 100, 1944, pages 279–287.  
<http://adsabs.harvard.edu/pdf/1944ApJ...100..279R>
136. G. Reber. “Cosmic Static at 144 Meters Wavelength”. *Journal of the Franklin Institute* 285:1, 1968, pages 1–12.  
DOI: [10.1016/0016-0032\(68\)90463-8](https://doi.org/10.1016/0016-0032(68)90463-8).  
Note: No open-access version available.
137. M. J. Rees and P. Mészáros. “Unsteady Outflow Models for Cosmological Gamma-Ray Bursts”. *The Astroparticle Journal*, 1994.  
DOI: [10.1086/187446](https://doi.org/10.1086/187446). <http://adsabs.harvard.edu/pdf/1994ApJ...430L..93R>
138. M. Rees. “The M87 Jet: Internal Shocks in a Plasma Beam?” *Monthly Notices of the Royal Astronomical Society* 184:1, 1978, 61P–65P.  
DOI: [10.1093/mnras/184.1.61P](https://doi.org/10.1093/mnras/184.1.61P)
139. O. Reimer, M. Pohl, P. Sreekumar, and J. Mattox. “EGRET Upper Limits on the High-Energy Gamma-Ray Emission of Galaxy Clusters”. *The Astrophysical Journal* 588:1, 2003, page 155.  
DOI: [10.1086/374046](https://doi.org/10.1086/374046)
140. J. L. Richards et al. “Blazars in the FERMI Era: The OVRO 40 m Telescope Monitoring Program”. *The Astrophysical Journal Supplement Series* 194:2, 2011, page 29.  
DOI: [10.1088/0067-0049/194/2/29](https://doi.org/10.1088/0067-0049/194/2/29)
141. W. A. Rolke, A. M. Lopez, and J. Conrad. “Limits and Confidence Intervals in the Presence of Nuisance Parameters”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 551:2-3, 2005, pages 493–503.  
DOI: [10.1016/j.nima.2005.05.068](https://doi.org/10.1016/j.nima.2005.05.068). eprint: [physics/0403059v5](https://arxiv.org/abs/physics/0403059v5)
142. R. Ruffini and J. A. Wheeler. *Relativistic Cosmology and Space Platforms*. Technical report. 1971.  
[https://inis.iaea.org/collection/NCLCollectionStore/\\_Public/03/024/3024932.pdf](https://inis.iaea.org/collection/NCLCollectionStore/_Public/03/024/3024932.pdf)
143. A. Sandrock. *Blazar Delta*. Version 1.0.0. 2021.  
[https://github.com/asandrock/blazar\\_delta](https://github.com/asandrock/blazar_delta)
144. R. Schödel et al. “A Star in a 15.2-Year Orbit around the Supermassive Black Hole at the Centre of the Milky Way”. *Nature* 419:6908, 2002, pages 694–696.  
DOI: [10.1038/nature01121](https://doi.org/10.1038/nature01121). [https://homepages.spa.umn.edu/~twj/ast8011\\_seminar/nature01121.pdf](https://homepages.spa.umn.edu/~twj/ast8011_seminar/nature01121.pdf).  
Note: Last visited: 2021-08-07

## Bibliography

145. F. Seward, W. Tucker, and R. Fesen. “Faint X-Ray Structure in the Crab Pulsar Wind Nebula”. *The Astrophysical Journal* 652:2, 2006, page 1277.  
DOI: [10.1086/508532](https://doi.org/10.1086/508532)
146. C. K. Seyfert. “Nuclear Emission in Spiral Nebulae”. *The Astrophysical Journal* 97, 1943, page 28.  
<http://adsabs.harvard.edu/pdf/1943ApJ....97...28S>
147. M. Shepherd, T. Pearson, and G. Taylor. “DIFMAP: An Interactive Program for Synthesis Imaging”. In: *Bulletin of the American Astronomical Society*. Vol. 26. 1994, pages 987–989.  
<http://adsabs.harvard.edu/pdf/1997ASPC..125...77S>
148. G. F. Smoot. “COBE Observations and Results”. In: *AIP Conference Proceedings CONF-981098*. Vol. 476. 1. American Institute of Physics. 1999, pages 1–10.  
DOI: [10.1063/1.59326](https://doi.org/10.1063/1.59326). eprint: [astro-ph/9902027](https://arxiv.org/abs/astro-ph/9902027)
149. D. Sobczyńska. *MMCS from CORSIKA 6.014*. Technical report TDAS 02-10. University of Lodz, 2002.  
Note: Internal MAGIC document
150. M. Spada, G. Ghisellini, D. Lazzati, and A. Celotti. “Internal Shocks in the Jets of Radio-Loud Quasars”. *Monthly Notices of the Royal Astronomical Society* 325:4, 2001, pages 1559–1570.  
DOI: [10.1046/j.1365-8711.2001.04557.x](https://doi.org/10.1046/j.1365-8711.2001.04557.x)
151. F. R. Stephenson and D. A. Green. “Was the supernova of AD 1054 reported in European history?” *Journal of Astronomical History and Heritage* 6, 2003, pages 46–52.  
<http://articles.adsabs.harvard.edu/pdf/2003JAHH....6...46S>
152. M. Stonebraker. *The Design of the Postgres Storage System*. Technical report. California University Berkeley Electronics Research Lab, 1987.  
<https://apps.dtic.mil/sti/pdfs/ADA187244.pdf>.  
Note: Last visited: 2021-08-07
153. M. Stonebraker, L. A. Rowe, and M. Hirohama. “The Implementation of POSTGRES”. *IEEE transactions on knowledge and Data engineering* 2:1, 1990, pages 125–142.  
DOI: [10.1109/69.50912](https://doi.org/10.1109/69.50912). <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.4391&rep=rep1&type=pdf>.  
Note: Last visited: 2021-08-07
154. A. Strong and G. Bignami. “Gamma-Ray Observations Toward NGC 1275 and the Origin of the Emission in the Infrared, X-Rays, and Gamma-Rays”. *The Astrophysical Journal* 274, 1983, pages 549–557.  
<http://adsabs.harvard.edu/pdf/1983ApJ...274..549S>

155. F. Tavecchio and G. Ghisellini. “Flat Broad Line Region and Gamma-Ray Absorption in Blazars”, 2012.  
arXiv: [1209.2291](#).  
Note: arXiv preprint
156. F. Tavecchio and G. Ghisellini. “On the Spine-Layer Scenario for the Very High-Energy Emission of NGC 1275”. *MNRAS* 443:2, 2014, pages 1224–1230.  
DOI: [10.1093/mnras/stu1196](#)
157. M. Taylor, P. Shopbell, M. Britton, and R. Ebert. *ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV*. 2005.  
Note: No open-access version available.
158. A. Tchekhovskoy and O. Bromberg. “Three-Dimensional Relativistic MHD Simulations of Active Galactic Nuclei Jets: Magnetic Kink Instability and Fanaroff-Riley Dichotomy”. *Monthly Notices of the Royal Astronomical Society: Letters* 461:1, 2016, pages L46–L50.  
DOI: [10.1093/mnrasl/slw064](#)
159. H. Teräsranta et al. “Fifteen Years Monitoring of Extragalactic Radio Sources at 22, 37 and 87 GHz”. *Astronomy and Astrophysics Supplement Series* 132:3, 1998, pages 305–331.  
DOI: [10.1051/aas:1998297](#)
160. D. Thain, T. Tannenbaum, and M. Livny. “Distributed Computing in Practice: The Condor Experience.” *Concurrency - Practice and Experience* 17:2-4, 2005, pages 323–356.  
DOI: [10.1002/cpe.938](#)
161. *The Nobel Prize in Physics 2020*. 2020.  
<https://www.nobelprize.org/prizes/physics/2020/penrose/facts>
162. A. N. Tikhonov and V. Y. Arsenin. “Solutions of Ill-Posed Problems”. *New York* 1:30, 1977, page 487.  
DOI: [10.2307/2006360](#)
163. J. Timmer and M. Koenig. “On Generating Power Law Noise.” *Astronomy and Astrophysics* 300, 1995, page 707.  
<http://adsabs.harvard.edu/pdf/1995A%26A...300..707T>
164. S. van der Tol, B. D. Jeffs, and A.-J. van der Veen. “Self-Calibration for the LOFAR Radio Astronomical Array”. *IEEE Transactions on Signal Processing* 55:9, 2007, pages 4497–4510.  
DOI: [10.1109/TSP.2007.896243](#). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.76.8733&rep=rep1&type=pdf>.  
Note: Last visited: 2021-08-07
165. R. B. Tully et al. “Cosmicflows-2: The Data”. *The Astronomical Journal* 146:4, 2013, page 86.  
DOI: [10.1088/0004-6256/146/4/86](#)



## Bibliography

166. S. Vaughan, R. Edelson, R. Warwick, and P. Uttley. “On Characterizing the Variability Properties of X-Ray Light Curves from Active Galaxies”. *Monthly Notices of the Royal Astronomical Society* 345:4, 2003, pages 1271–1284.  
DOI: [10.1046/j.1365-2966.2003.07042.x](https://doi.org/10.1046/j.1365-2966.2003.07042.x)
167. R. Vermeulen, A. Readhead, and D. Backer. “Discovery of a Nuclear Counterjet in NGC 1275: A New Way to Probe the Parsec-Scale Environment”. *The Astrophysical Journal* 430, 1994, pages L41–L44.  
<http://adsabs.harvard.edu/pdf/1994ApJ...430L..41V>
168. P. Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. *Nature Methods* 17, 2020, pages 261–272.  
DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
169. R. C. Walker et al. “The Structure and Dynamics of the Subparsec Jet in M87 Based on 50 VLBA Observations over 17 Years at 43 GHz”. *The Astrophysical Journal* 855:2, 2018, page 128.  
DOI: [10.3847/1538-4357/aaafcc](https://doi.org/10.3847/1538-4357/aaafcc)
170. R. Walker, J. Romney, and J. Benson. “Detection of a VLBI Counterjet in NGC 1275: A possible Probe of the Parsec-Scale Accretion Region”. *ApJ* 430, 1994, pages L45–L48.  
<http://adsabs.harvard.edu/pdf/1994ApJ...430L..45W>
171. S. van der Walt et al. “scikit-image: Image processing in Python”. *PeerJ* 2, 2014, e453. ISSN: 2167-8359.  
DOI: [10.7717/peerj.453](https://doi.org/10.7717/peerj.453). [10.7717/peerj.453](https://doi.org/10.7717/peerj.453)
172. T. C. Weekes et al. “Observation of TeV Gamma-Rays from the Crab Nebula Using the Atmospheric Cherenkov Imaging Technique”. *Astrophysical Journal* 342, 1989, pages 379–395.  
DOI: [10.1086/167599](https://doi.org/10.1086/167599)
173. W. Welsh. “On the Reliability of Cross-Correlation Function Lag Determinations in Active Galactic Nuclei”. *Publications of the Astronomical Society of the Pacific* 111:765, 1999, page 1347.  
DOI: [10.1086/316457](https://doi.org/10.1086/316457). arXiv: [astro-ph/9911112](https://arxiv.org/abs/astro-ph/9911112)
174. R. Weymann, J. Scott, A. Schiano, and W. Christiansen. “A Thermal Wind Model for the Broad Emission Line Region of Quasars”. *The Astrophysical Journal* 262, 1982, pages 497–510.  
<http://adsabs.harvard.edu/pdf/1982ApJ...262..497W>
175. R. W. Wilson. “The Cosmic Microwave Background Radiation”. *Reviews of Modern Physics* 51:3, 1979, page 433.  
DOI: [10.1016/S0370-1573\(00\)00025-9](https://doi.org/10.1016/S0370-1573(00)00025-9)
176. M. Wood et al. “Fermipy: An Open-Source Python Package for Analysis of Fermi-LAT Data”. In: *Proceedings of Science*. ICRC 2017. 2017, page 824.  
arXiv: [1707.09551v1](https://arxiv.org/abs/1707.09551v1)



177. K. Yamamoto et al. “Development of Multi-Pixel Photon Counter (MPPC)”. In: *2007 IEEE Nuclear Science Symposium Conference Record*. Vol. 2. IEEE. 2007, pages 1511–1515.  
DOI: [10.1109/NSSMIC.2006.356038](https://doi.org/10.1109/NSSMIC.2006.356038). <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1076.1986&rep=rep1&type=pdf>.  
Note: Last visited: 2021-08-07
178. S. Y. Yong et al. “The Kinematics of Quasar Broad Emission Line Regions Using a Disk-Wind Model”. *Publications of the Astronomical Society of Australia* 34, 2017.  
DOI: [10.1017/pasa.2017.37](https://doi.org/10.1017/pasa.2017.37)
179. R. Zanin et al. “MARS, the MAGIC Analysis and Reconstruction Software”. In: *Proc. of the 33rd International Cosmic Ray Conference, Rio de Janeiro, Brasil*. Citeseer. 2013.  
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.717.5452&rep=rep1&type=pdf>.  
Note: Last visited: 2021-08-07
180. F. Zernike. “The Concept of Degree of Coherence and its Application to Optical Problems”. *Physica* 5:8, 1938, pages 785–795.  
DOI: [10.1016/S0031-8914\(38\)80203-2](https://doi.org/10.1016/S0031-8914(38)80203-2).  
Note: No open-access version available.



# Glossary

- 3C** Third Cambridge Catalogue of Radio Sources. 7
- ACD** Anti-Coincidence Detector. 90, 91
- AGN** Active Galactic Nucleus. 6–14, 17, 21, 24, 113, 114, 129, 131, 132, 136, 141, 142, 165, 169
- ALMA** Atacama Large Millimeter/submillimeter Array. 114, 115, 175
- BH** Black Hole. 9–15, 18, 19, 113, 121–126, 128, 141
- BLR** Broad-Line Region. 12, 13, 22, 93, 121–128, 134, 141
- BU-BLAZAR program** Boston University Blazar Monitoring Program. 16, 104, 114, 118, 126
- CDF** Cumulative Distribution Function. 169
- CMB** Cosmic Microwave Background. 17, 18, 98
- COBE** Cosmic Background Explorer. 17
- CoG** Center of Gravity. 36, 37, 40
- CORSIKA** Cosmic Ray Simulations for Kascade. 42
- CR** Cosmic Rays. 27, 47, 49, 91
- CTA** Cherenkov Telescope Array. 31, 57, 59, 142
- DAQ** Data Acquisition System. 91
- DC** Direct Current. 66, 68, 69, 74, 75, 155
- DCF** Discrete Correlation Function. 131
- DFT** Discrete Fourier Transform. 165, 166, 172
- DL3** Data Level 3. 54, 55, 57–59, 61–63, 65, 69, 74, 76, 77, 79, 80, 156, 158
- DRS4** Domino Ring Sampler version 4. 33, 52
- EBL** Extragalactic Background Light. 17, 48
- EC** External Compton. 22
- EGRET** Energetic Gamma Ray Experiment Telescope. 21, 23, 79
- EHT** Event Horizon Telescope. 10

## *Glossary*

- ESA** European Space Agency. 17
- FACT** First G-APD Cherenkov Telescope. 58
- FITS** Flexible Image Transport System. 58, 61, 62, 92
- FoV** Field of View. 31, 33, 38, 43–45, 49–51, 69, 79, 80, 91, 92, 150, 155
- FR** Fanaroff-Riley. 7
- FSRQ** Flat Spectrum Radio Quasar. 7
- FWHM** Full Width Half Maximum. 104
- GADF** Gamma Astro Data Format. 57–59, 62–64
- GBM** Gamma-Ray Burst Monitor. 90, 91
- GRB** Gamma-Ray Burst. 77
- GSFC** Goddard Space Flight Center. 91
- GTI** Good Time Intervals. 57
- HAWK** High-Altitude Water Cherenkov Gamma-Ray Observatroy. 31
- H.E.S.S.** High Energy Stereoscopic System. 58, 59
- IACT** Imaging Air Cherenkov Telescope. 30, 31, 36, 38, 40, 43, 46, 89, 91, 97, 101
- IC** Inverse Compton. 21, 121
- IR** infrared. 17
- IRF** Instrument Response Function. 44–46, 53, 54, 57, 59, 61–63, 91, 92, 156
- JVLA** Karl G. Jansky Very Large Array. 98, 102
- KASCADE** Karlsruhe Shower Core and Array Detector. 42
- LAT** Large Area Telescope. 1, 2, 13, 21, 23, 27, 57–59, 79, 84, 89–94, 113, 125, 126, 133–137, 141, 142, 175
- LCCF** Local Cross-Correlation Function. 131, 132, 135
- LIDAR** Light Detection and Ranging. 155
- LISOC** LAT Instrument Science and Operations Center. 91, 92
- LST** Large-Sized Telescope. 31
- LUT** Look-Up Table. 53, 62, 66

- MAGIC** Major Atmospheric Gamma Imaging Cherenkov. 2, 19, 24, 30, 31, 33, 35, 36, 38, 40–44, 52, 53, 57–59, 61, 62, 65, 66, 69, 71, 76, 77, 79–81, 84, 85, 91–93, 125, 126, 129, 130, 133, 136, 137, 141–143, 147, 175
- MARS** MAGIC Analysis and Reconstruction Software. 42, 52, 54, 55, 57–59, 61–66, 69, 70, 76, 91, 143, 147, 155
- MC** Monte Carlo. 38, 40, 42, 44–47, 53, 55, 62, 63, 65, 66, 69, 74, 76, 77, 80, 84, 107, 150, 153, 155, 159, 161
- MHD** magnetohydrodynamic. 11
- ML** Machine Learning. 36, 38, 109
- MMCS** MAGIC Monte Carlo Software. 42
- MOC** Mission Operations Center. 91
- MOJAVE** Monitoring of Jets in Active Galactic Nuclei with VLBA Experiments. 14, 176
- MOLA** MAGIC Online Analysis. 61
- MST** Mid-Sized Telescope. 31
- NASA** National Aeronautics and Space Administration. 58
- NGC** New General Catalogue of Nebulae and Clusters of Stars. 79
- NRAO** National Radio Astronomy Observatory. 104
- NSB** Night Sky Background. 34, 62, 149
- OVRO** Owens Valley Radio Observatory. 114, 115
- PCA** Principal Components Analysis. 36
- PDF** Probability Distribution Function. 132, 133, 165, 168, 169, 171–173
- PIC** Port d’Informació Científica. 65, 69, 80, 152, 153
- PMT** Photomultiplier Tube. 30, 33
- PSD** Power Spectral Density. 132, 133, 165, 167–169, 171–173
- PSF** Point Spread Function. 43, 44, 53, 58, 92, 101, 102, 108, 155, 167
- PWN** Pulsar Wind Nebula. 18–20
- RDBMS** Relational Database Management System. 152
- RF** Random Forest. 37–42, 45, 53, 61–63, 66, 67, 69, 74, 76, 77, 151
- RMS** Root Mean Square. 36, 108
- SED** Spectral Energy Distribution. 17, 21–23, 48, 63, 69, 70, 73, 74, 84, 85, 92, 93, 125–127, 129, 155–157

## *Glossary*

- SiPM** Silicon Photomultiplier. 30
- SKA** Square Kilometer Array. 142
- SMA** Submillimeter Array. 114, 115, 134
- SQL** Structured Query Language. 152
- SSC** Synchrotron Self-Compton. 22, 24, 130
- SST** Small-Sized Telescope. 31
- TS** Test Statistics. 49, 53
- UDCF** Unbinned Discrete Correlation Function. 131, 132
- UV** ultraviolet. 13, 17, 28
- VERITAS** Very Energetic Radiation Imaging Telescope Array System. 58, 59
- VHE** Very-High Energy. 17, 22, 31, 42, 84, 126, 128, 130, 141
- VLBA** Very Long Baseline Array. 1, 2, 16, 93, 104, 105, 115, 129, 133–135, 137, 141, 142
- VLBI** Very Long Baseline Interferometry. 12, 107, 129
- WSRT** Westerbork Synthesis Radio Telescope. 98

## Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel „Multiwavelength Analysis of the TeV-Radio Galaxy 3C 84/NGC 1275“ selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

---

Ort, Datum

---

Unterschrift

### Belehrung

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50 000 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§ 63 Abs. 5 Hochschulgesetz –HG–).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z. B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen.

---

Ort, Datum

---

Unterschrift

# Acknowledgements

Zu allererst gilt mein Dank unserer wundervollen Lehrstuhl-Sekretärin Andrea Teichmann. Andrea, Du hältst den Laden zusammen und hast immer ein offenes Ohr. Ohne Dich wäre niemand von uns arbeitsfähig. Danke!

Ich bedanke mich bei Prof. Dr. Dr. Wolfgang Rhode und Dr. habil. Dominik Elsässer, für Eure Unterstützung, die vielen Erfahrungen, die Ihr mir ermöglicht habt, und Eure positive Lesart der Dinge, wenn ich zu skeptisch bin.

Vielen Dank an Prof. Dr. Matthias Kadler und die Würzburger Gruppe für Eure Gastfreundschaft und die Bereitschaft uns als Radio-Novizen anzulernen.

Ein großes Dankeschön an Prof. Dr. Kröninger, für die Begutachtung dieser Arbeit.

Danke an alle aktuellen und ehemaligen Menschen am Lehrstuhl E5b, die mich in den letzten Jahren über kürzere oder längere Distanzen begleitet haben. Ich habe viel von Euch allen gelernt!

Danke Max, wir wissen beide, dass weder diese Arbeit, noch autoMAGIC ohne deine Hilfe je zustande gekommen wären. Dein immenses Wissen, deine kognitiven Fähigkeiten und deine Hartnäckigkeit sind ebenso förderlich wie frustrierend.

Alex, Danke für die Idee zu unserem Paper, deine Geduld mit mir und deine vielen vielen kleinen und großen Weisheiten im Alltag.

Kai, ich weiß nicht mehr genau warum, aber irgendwie bist du schuld an allem. Danke für alles.

Danke Kevin, für deine Unterstützung in allen Lebenslagen, deinen kompromisslosen Humor und deine unaufgeregte Gradlinikeit.

Simone, danke, dass wir alle Freuden und Leiden teilen können.

Danke Henning, dafür dass du zu groß denkst, wenn ich zu klein denke. Im Mittel passt es dann schon irgendwie.

Danke Marie, für deine Freundschaft, deine positive Energie und deinen Optimismus. Du bist immer da, wenn's brennt oder ein Kuchen gebacken werden muss.

Ali, du bist die beste Schreibtischnachbarin der Welt! Danke für deine Hilfe beim Suchen nach Antworten auf die großen Fragen bei MAGIC und des Lebens.

Cosimo, thank you for your patience with me, for answering all my questions (sometimes twice), and for always supporting me and autoMAGIC!

Danke an alle, die sich die Mühe gemacht haben, diese Arbeit Korrektur zu lesen: Alex, Alicia, Kevin, Simone und natürlich Max.

Danke an Jan und Alex, für eure Geduld mit mir am Samstagmorgen und dass ihr mir all die unwichtigen Dinge für die Prüfung beigebracht habt.

Laura, ohne dich hätte ich das Studium niemals geschafft! Danke, dass du diese Jahre mit mir in allen Höhen und Tiefen durchgestanden hast!

Ich bedanke mich bei meinem Physiklehrer Herrn Päßler, der vor vielen Jahren die Initialzündung für diesen Irrsinn gestartet hat.



Ein tief empfundenes Dankeschön gilt meinen wunderbaren Eltern. Ihr habt mir ein witterungsbeständiges Fundament für mein sorgenfreies Leben gegeben. Das muss euch erst mal einer nachmachen.

Danke an die tollste Frau in meinem Leben: meine großartige Schwester Maike.

Ich danke meinen Mädels Carina, Helena, Johanna, Sarah für den großen gemeinsamen Schatz, den wir uns über die Jahre zusammen ge...ammelt haben.

And finally, danke Max. Es ist meistens ein großes Glück mein Leben mit dir zu teilen.