

Thesis for obtaining the academic degree
Bachelor of Science

Finding optimal hyperparameters for cleaning
algorithms for the Cherenkov Telescope
Array

Anno Knierim
born in Hagen

2022

Chair of Experimental Physics V

Department of Physics

TECHNISCHE UNIVERSITÄT DORTMUND

Reviewer: Prof. Dr. Dr. Wolfgang Rhode
Co-reviewer: Dr. Dirk Wiedner
Submission date: 7 September 2022

This thesis is set in Libertinus (Serif, Sans and Math) and Fira Code typeset using L^AT_EX with LuaT_EX from T_EXLive 2021.
Title graphic by **TITLE GRAPHIC AUTHOR**.

Abstract

The abstract is a short summary of the thesis in English, and together with the German summary, it has to fit on this page.

Kurzfassung

Hier steht eine Kurzfassung der Arbeit in deutscher Sprache inklusive der Zusammenfassung der Ergebnisse. Zusammen mit der englischen Zusammenfassung muss sie auf diese Seite passen.

Contents

1	Gamma-Ray Astronomy	1
2	IACTs and the Cherenkov Telescope Array	3
2.1	Imaging Air Cherenkov Telescopes	3
2.2	The Cherenkov Telescope Array	4
3	Data processing	7
4	Finding Optimal Hyperparameters for the Cleaning Algorithms	9
4.1	Cleaning Algorithms	9
4.2	Hyperparameters	9
5	Results	10
6	Conclusions and Outlook	11
Bibliography		12
Glossary		13
Appendix		14
Acknowledgements		15

1

Gamma-Ray Astronomy

Astronomy, being one of the oldest sciences, is a vast field of study dating back to the Babylonians. From the earliest days of civilization, astronomers have been studying the stars and the planets to understand the universe. It, therefore, is no surprise that astronomy spawned a great number of discoveries throughout the centuries. Whereas first observations were made by eye only, we now have access to a multitude of experiments and telescopes that deepen our understanding of the universe. With the discovery of cosmic rays (CR) by Victor Hess in the early 20th century, the new field of astroparticle physics was born [8].

From then on we found many different types of cosmic messengers, the most recent being the discovery of gravitational waves in 2015 [1].

rewrite this
(maybe)

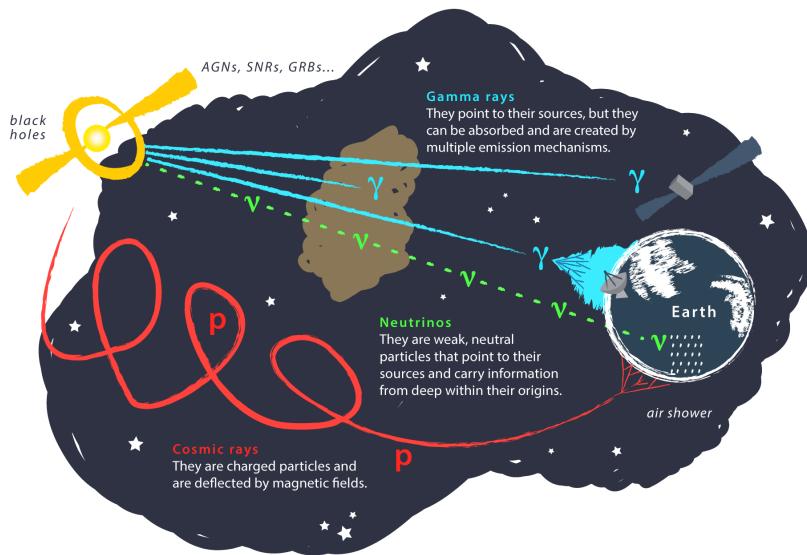


Figure 1.1: Different types of cosmic rays on their way to Earth. Charged particles like protons and electrons are deflected by magnetic fields and therefore making it hard to pinpoint the source. Only the origin of photons and neutrinos can be reconstructed directly since they are uncharged particles and therefore travel in straight lines. However, photons can be absorbed or created in multiple mechanisms. Since neutrinos only rarely interact with matter via the weak force, their detection is significantly harder than for photons [2].

CR come in different types, that are either charged or uncharged, as shown in Figure 1.1. Charged particles like electrons, protons or atomic nuclei are difficult to trace back to their origin as they are

deflected by the cosmic electromagnetic fields. Uncharged particles like photons or neutrinos, however, travel in straight lines, making it easier to reconstruct their origins, although photons can be absorbed by dust clouds in their way.

Since neutrinos are harder to detect due to their weak interaction with matter, photons are easier to study with space- and ground-based experiments.

Therefore, in recent years, gamma-ray astronomy has become an important research field in astroparticle physics. The term gamma-rays is generally denoted as photons with energies above 100 keV [7]. Due to this high-energy nature, gamma rays pose some of the most powerful CR in the universe and since photons at such energies cannot be produced by thermal processes, their origin can be described by higher order processes involving charged particles.

For the past two decades, ground-based Imaging Air Cherenkov Telescope (IACT) experiments like the Major Atmospheric Gamma-Ray Imaging Cherenkov (MAGIC) telescopes, the Very Energetic Radiation Imaging Telescope Array System (VERITAS) and the High Energy Stereoscopic System (H. E. S. S.) have been monitoring these very-high-energy gamma rays (VHE gamma rays) to gain an understanding of their production. This allowed us to determine different source classes inside and outside our galaxy, with the most important source class inside our galaxy being supernova remnants (SNRs) such as the Crab Nebula.

IACTs and the Cherenkov Telescope Array

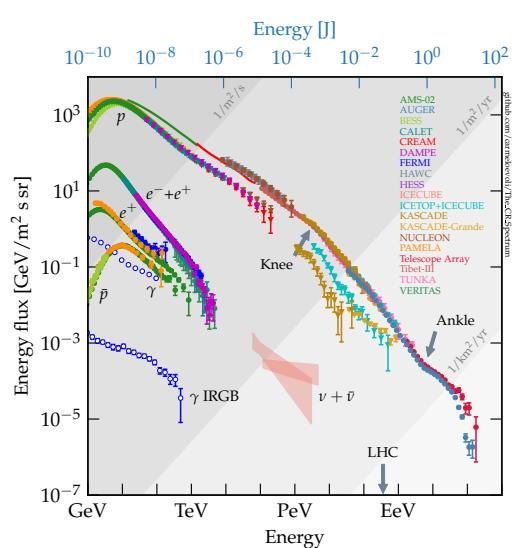
2

Most modern gamma-ray observations are performed with either space-based experiments or with Imaging Air Cherenkov Telescopes (IACTs), which are ground-based telescopes or arrays of telescopes that use the Cherenkov light emitted by extensive air showers (EASs) in the atmosphere. In the following sections I will introduce IACTs and the Cherenkov Telescope Array (CTA) and explain the mechanisms that make it possible to observe gamma rays with these types of experiments.

2.1 Imaging Air Cherenkov Telescopes

Because of their ground-based setup, IACTs are taking advantage of the Earth's atmosphere to get a larger effective area than any space-based instrument. This is especially helpful for energies above 100 GeV, where the gamma-ray flux is low compared to lower energies that have higher fluxes. This means, that space-based experiments with their small effective area will see fewer high-energy events compared to ground-based experiments. The cosmic ray flux in Figure 2.1 shows this well: The flux decreases rapidly with higher energies, with only one high-energy photon per day and square meter reaching Earth from the Crab Nebula [9] and even fewer photons for energies in the EeV domain, where only 1 particle per square kilometer per year reaches Earth.

For high-energy gamma rays, Earth's atmosphere is opaque, so ground-based experiments rely on cascades of sub-particles in so-called EASs. For gamma-ray astronomy, the electromagnetic component of these EASs, shown in Figure 2.2a, is more relevant than the hadronic component. When a gamma ray interacts with Earth's atmosphere, it decays into an electron and a positron via pair production. These charged particles then emit more photons via bremsstrahlung. These photons, again, produce more charged particles, which in turn emit more photons. This process continues until any of these processes reaches energies below a certain threshold, i. e. 1022 keV for the pair production process. For these processes to happen, the particles have to be in the fields of the atoms of the Earth's atmosphere. As the particles travel at speeds faster than



confirm values
and add source

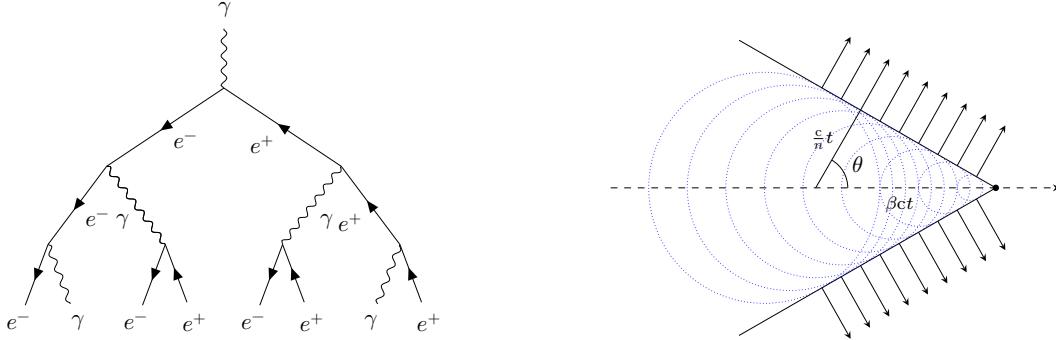
wording

Figure 2.1: The cosmic ray flux as a function of energy. For very high energies the flux becomes very small with only 1 particle per square meter per year reaching Earth. For even higher energies in the domain of EeV this flux becomes 1 particle per square kilometer per year [6].

light in the medium of the atmosphere, they emit Cherenkov light at a fixed angle θ with respect to the refraction index n of the atmosphere and the factor $\beta = v/c$. The angle can be determined trigonometrically as

$$\cos \theta = \frac{1}{\beta n}, \quad (2.1)$$

as can be seen in Figure 2.2b.



(a) Simplified Heitler model of the electromagnetic component of an extensive air shower. A gamma ray induces an electron and a positron via pair production that then emit more photons via bremsstrahlung.

(b) Cherenkov radiation (outward-pointing arrows) from a charged particle traveling at uniform velocity, faster than the speed of light in the medium. The dashed line shows the path of the particle over time.

Figure 2.2: The electromagnetic component of an extensive air shower (see Figure 2.2a) is induced by gamma rays that then decay into electrons and positrons via pair production. These charged particles will emit more gamma rays via bremsstrahlung, which again will produce new electrons and positrons. This will continue until e.g. the energy of the gamma rays is below the threshold of 1022 keV for pair production. Another process is the production of Cherenkov light from the electrons and positrons, which is shown in Figure 2.2b. The Cherenkov light is the result of the charged particles traveling faster than the speed of light in the atmosphere, thus emitting photons in a fixed angle θ w.r.t. the refraction index n of the medium and the factor β .

Cherenkov light emitted by EASs can then be collected by an IACT's mirrors and detected by its camera within a timeframe of an order of nanoseconds. The resulting image will show the shape of the impact of the air shower and can be used to determine the shower's primary particles' properties, and the type and reconstruct its origin. As the hadronic component of EASs produce electromagnetic subshowers, IACTs have a dominant hadronic background, which has to be separated from the gamma ray-induced showers.

Adjust figure size of Figure 2.2

maybe write something about ML

2.2 The Cherenkov Telescope Array

The CTA is a new generation of IACTs that will consist of two sites, one of which will be built at the Observatorio del Roque de los Muchachos (ORM) on the Canarian island of La Palma while the other will be built in the southern hemisphere at the European Southern Observatorys (ESO) Paranal Observatory in the Atacama desert of northern Chile. In this work, I will focus on the northern array, also called CTA north, on La Palma.

The CTA consists of three types of telescopes, shown in Figure 2.3, each being sensitive to a different energy range [3]:

Large-Sized Telescope (LST): The largest telescopes in the CTA have primary deflector diameter of 23 m with a field of view of 4.3 deg and are sensitive to energy ranges from 20 GeV to 150 GeV.

Medium-Sized Telescope (MST): The MSTs have a primary deflector diameter of 11.5 m with a field of view of 7.7 deg for their NectarCam and are sensitive to energy ranges from 150 GeV to 5 TeV.

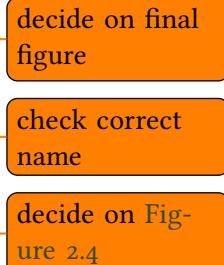
Small-Sized Telescope (SST): Being the smallest telescope type in the CTA, the SSTs have a primary deflector diameter of 4.3 m, a secondary deflector diameter of 1.8 m with a field of view of 10.5 deg and are sensitive to energy ranges from 5 TeV to 300 TeV.



Figure 2.3: A rendered image of the 3 telescope prototypes in CTA. From left: The SST, two types of MST (SCT and MST) and the LST. This work will not feature the SCT, as the used simulation data contains only MST data.

Since the southern hemisphere has more galactic high-energy sources than the northern hemisphere, only CTA south will feature the SSTs along the other two telescope types. CTA north will feature the MSTs and LSTs. Figure 2.4 shows the planned full layout of both sites of the CTA where CTA south is displayed with its 70 SSTs, 25 MSTs and the 4 LSTs in the center. CTA north will feature 15 MSTs and 4 LSTs. Due to the simulation data used, this work will focus on the reduced layout, also called Prod5b alpha configuration, for the northern site, consisting of 9 MSTs and 4 LSTs.

The MST's and LST's cameras are based on photo multiplier tube (PMT) photodetectors, while the SST's camera is based on a silicon photomultiplier (SiPM) photodetector.



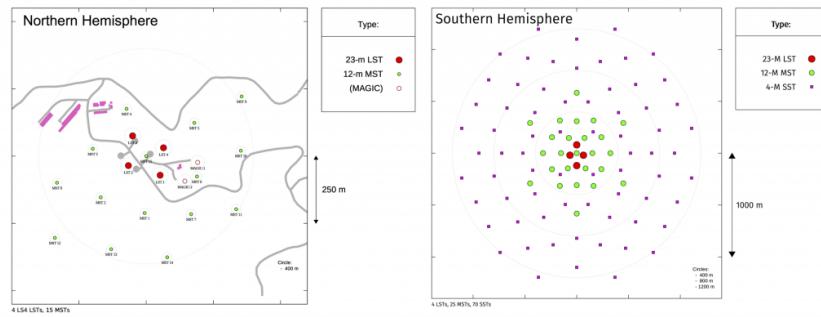


Figure 2.4: Alpha Configuration layout of the CTA north and south sites. CTA south would feature 70 SSTs, 25 MSTs and 4 LSTs [5], while CTA north would feature 9 MSTs and 4 LSTs [4].

Data processing

3

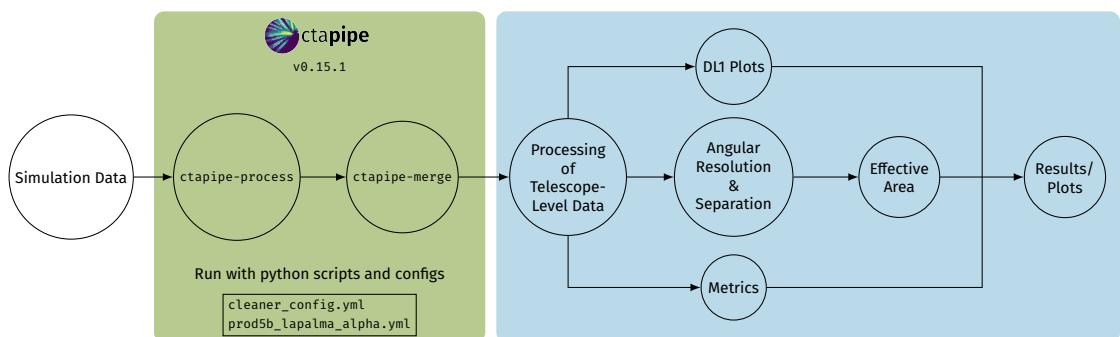


Figure 3.1: Schematic overview of the data pipeline used for this work. Single runs of simulation data are processed with `ctapipe-process` and then merged via `ctapipe-merge`. The merged data is then processed on array- or telescope data level resulting in scores for metrics as well as `dl1` images and plots for the angular resolution and the effective area.

```
def metrics(events, output_file_metrics, unique_file_id):
    """Calculates the metrics for the given telescope type.

    Parameters:
    -----
    events: astropy.table.Table
        The events table. Must contain the following columns:
        - true_image: The true image of the event.
        - image: The event image containing background noise.
    output_file_metrics: pathlib.Path
        The path to the output file for the metrics.
    unique_file_id: int
        The unique file id for the input file.
    """

    # initialize the metrics calculator
    metrics_calc = TprFprCalculator(
        true_image=events["true_image"],
        image=events["image"],
        clean_mask=events["image_mask"]
    )

    # calculate the metrics
    metrics_data = metrics_calc.tpr_fpr()

    # write metrics data to DataFrame and then to a file
    metrics = pd.DataFrame(data=metrics_data)
    metrics.insert(loc=0, column='unique_file_id', value=unique_file_id)
    metrics.to_csv(
        output_file_metrics,
        index=False,
        mode='a',
        header=not output_file_metrics.exists()
    )
```

Finding Optimal Hyperparameters for the Cleaning Algorithms

4

4.1 Cleaning Algorithms

Write about cleaning algorithms

4.2 Hyperparameters

Write about the hyperparameters

Results

5

Conclusions and Outlook

6

Bibliography

1. B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration Collaboration). “Observation of Gravitational Waves from a Binary Black Hole Merger.” *Phys. Rev. Lett.* 116, 6 2016, page 061102.
DOI: [10.1103/PhysRevLett.116.061102](https://doi.org/10.1103/PhysRevLett.116.061102). <https://link.aps.org/doi/10.1103/PhysRevLett.116.061102>
2. J. A. Aguilar, J. Yang, and S. Bravo. *Neutrinos and gamma rays, a partnership to explore the extreme universe*. IceCube/WIPAC. 2016.
<https://icecube.wisc.edu/news/view/455> visited on 2022-08-05
3. *CTA Technology*. Cherenkov Telescope Array Consortium.
https://www.cta-observatory.org/wp-content/uploads/2019/12/CTA-Specifications_v08_formatted.pdf visited on 2022-08-10
4. *CTAO’s site in the northern hemisphere*. Cherenkov Telescope Array Consortium.
<https://www.cta-observatory.org/about/array-locations/la-palma/> visited on 2022-08-10
5. *CTAO’s site in the southern hemisphere*. Cherenkov Telescope Array Consortium.
<https://www.cta-observatory.org/about/array-locations/chile/> visited on 2022-08-10
6. C. Evoli. *The Cosmic-Ray Energy Spectrum*. 2020.
DOI: [10.5281/zenodo.4396125](https://doi.org/10.5281/zenodo.4396125). <https://doi.org/10.5281/zenodo.4396125>
7. S. Funk. “Ground- and Space-Based Gamma-Ray Astronomy.” *Annual Review of Nuclear and Particle Science* 65:1, 2015, pages 245–277.
DOI: [10.1146/annurev-nucl-102014-022036](https://doi.org/10.1146/annurev-nucl-102014-022036). <https://doi.org/10.1146/annurev-nucl-102014-022036>
8. M. S. Longair. *High energy astrophysics: an informal introduction for students of physics and astronomy*. Cambridge University Press, Cambridge [England] ; New York, 1981. ISBN: 9780521235136 9780521280136
9. M. Nöthe. “Monitoring the High Energy Universe. Open, Reproducible, Machine Learning Based Analysis for the first G-APD Cherenkov Telescope.” PhD thesis. Technische Universität Dortmund, 2020

Glossary

CR cosmic rays. 1, 2

CTA Cherenkov Telescope Array. 3–6

EAS extensive air shower. 3, 4

ESO European Southern Observatory. 4

H.E.S.S. High Energy Stereoscopic System. 2

IACT Imaging Air Cherenkov Telescope. 2–4

LST Large-Sized Telescope. 5, 6

MAGIC Major Atmospheric Gamma-Ray Imaging Cherenkov. 2

MST Medium-Sized Telescope. 5, 6

ORM Observatorio del Roque de los Muchachos. 4

PMT photo multiplier tube. 5

SiPM silicon photomultiplier. 5

SNR supernova remnant. 2

SST Small-Sized Telescope. 5, 6

VERITAS Very Energetic Radiation Imaging Telescope Array System. 2

VHE gamma rays very-high-energy gamma rays. 2

Appendix

Hier könnte ein Anhang stehen, falls Sie z. B. Code, Konstruktionszeichnungen oder Ähnliches mit in die Arbeit bringen wollen. Im Normalfall stehen jedoch alle Ihre Resultate im Hauptteil der Bachelorarbeit und ein Anhang ist überflüssig.

Acknowledgements

Eidesstattliche Versicherung

(Affidavit)

Name, Vorname
(surname, first name)

Bachelorarbeit
(Bachelor's thesis)

Titel
(Title)

Matrikelnummer
(student ID number)

Masterarbeit
(Master's thesis)

Ich versichere hiermit an Eides statt, dass ich die vorliegende Abschlussarbeit mit dem oben genannten Titel 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.

I declare in lieu of oath that I have completed the present thesis with the above-mentioned title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution before.

Ort, Datum
(place, date)

Unterschrift
(signature)

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 50.000,00 € 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:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offense can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offenses of this type is the Chancellor of TU Dortmund University. In the case of multiple or other serious attempts at deception, the examinee can also be unenrolled, Section 63 (5) North Rhine-Westphalia Higher Education Act (*Hochschulgesetz, HG*).

The submission of a false affidavit will be punished with a prison sentence of up to three years or a fine.

As may be necessary, TU Dortmund University will make use of electronic plagiarism-prevention tools (e.g. the "turnitin" service) in order to monitor violations during the examination procedures.

I have taken note of the above official notification.*

Ort, Datum
(place, date)

Unterschrift
(signature)

***Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the Bachelor's/ Master's thesis is the official and legally binding version.**