Silicon Photomultipliers for UHECR Observation from Space

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Tobias Johannes Jammer
aus Stuttgart

Tübingen
2018

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen
Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	18.12.2018
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Josef Jochum
2. Berichterstatter:	Prof. Dr. Tobias Lachenmaier

Zusammenfassung

Der Ursprung hochenergetischer kosmischer Strahlung und die zugrunde liegenden Beschleunigungsmechanismen sind Bestandteil aktueller Forschung in der Astroteilchen-Physik. Um diesen Fragen auf den Grund zu gehen, muss eine große Anzahl von Teilchenschauern gemessen werden, die beim Eindringen kosmischer Strahlung in die Atmosphäre enstehen. Die Idee des Extreme Universe Space Observatory (EUSO) ist deshalb, die Atmosphäre aus dem Weltall zu beobachten, um auf einer großen Fläche Teilchenschauer und deren Fluoreszenzlicht aufzuzeichnen. Dies erfordert empfindliche Lichtdetektoren. Neben den herkömmlich benutzten Photomultiplier-Tubes sind Silizum-Photomultiplier (SiPM) eine vielversprechende Alternative, da sie keine Hochspannung benötigen und zudem leichter sind. Derzeit gibt es noch keine experimentellen Erfahrungen mit SiPM im Weltall, insbesondere im Hinblick auf die hohe thermische Dunkelrate und starke Temperaturempfindlichkeit der Detektoren.

Die vorliegende Arbeit besteht aus zwei Teilen: Im ersten Teil wird mithilfe eines experimentellen Aufbaus der Einfluss der Temperatur auf die Verstärkung von SiPM untersucht. Dazu wird ein Algorithmus zur Echtzeitmessung der Verstärkung entwickelt, der es erlaubt die Verstärkung ohne externe Temperaturmessung zu stabilisieren. Im zweiten Teil werden Daten einer SiPM-Prototyp-Kamera analysiert, die während eines EUSO-Ballonflugs im Frühling 2017 aufgenommen wurden. Eine statistische Auswertung der Daten ergibt eine Abschätzung des Untergrundes für die Messung kosmischer Strahlung, sowohl durch thermisches Rauschen im Detektor als auch durch den physikalischen UV-Hintergrund der Atmosphäre. Außerdem werden zeitlich veränderliche Signale identifiziert, die zum Beispiel durch Wolken verursacht werden können. Zum Schluss wird ein Vergleich der SiPM-Prototyp-Kamera mit der Hauptkamera des Ballons gezeigt, bei der herkömmliche Photomultiplier-Tubes benutzt wurden.

Silizium-Photomultiplier, SiPM, Kosmische Strahlung, Astroteilchen-Physik

Abstract

The origin and the underlying acceleration mechanisms of ultra high energy cosmic rays (UHECR) are one of the topics of research in astroparticle physics. To answer these questions, measurements with high statistics are needed of the extensive air showers the UHECR produce in the earth's atmosphere. By going to space, the proposed Extreme Universe Space Observatory (EUSO) aims to detect atmospheric showers over a large area using the fluorescence method, which requires sensitive light detectors. Silicon photomultipliers (SiPM) have emerged as promising alternatives to conventional photomultiplier tubes, with various advantages such as lower weight and lower operating voltage. However, the ability of SiPM to detect light in a near-space environment has yet to be demonstrated, and the influence of the high rate of thermal noise and strong dependence on ambient temperature has to be assessed.

This work is divided into two parts. First, the temperature dependence of the SiPM gain is studied with an experimental setup. A novel algorithm is presented to measure the SiPM gain in real-time, which allows the stabilization of the gain without the need for external temperature measurements. In the second part, data of a prototype SiPM camera, which was flown on a EUSO super pressure balloon pathfinder in the spring of 2017, were analysed. This includes a statistical analysis of the background for UHECR search, both from thermal detector noise and from physical UV background in the atmosphere. Furthermore, transient events were identified, which hint at moving clouds. In the end, a brief comparison between the SiPM prototype and the main camera using photomultiplier tubes is shown.

silicon photomultiplier, SiPM, cosmic rays, astroparticle physics

Contents

Chapter 1

Introduction

Astroparticle physics is a young field at the interface between particle physics, astronomy and cosmology. It studies physics at the most extreme scales, as the evolution and structure of the universe at the cosmological scale is driven by the properties of elementary particles and interactions at the smallest scale [1]. As a result, it is possible (and even necessary) to study cosmological events to learn about fundamental particle physics [2, 3]. At the same time, new advances in particle physics can have huge implications on our understanding of the universe as a whole [4].

One messenger from the extreme universe are cosmic rays (CR), high energy particles from outside the solar system or even the Milky Way. They span several orders of magnitude in energy and are the highest-energy particles we can measure [5]. Although there has been a huge effort to study CR for the past 100 years [6, 7, 8, 9], there are still a lot of unanswered questions regarding their origin and the mechanisms, which accelerate particles to such energies. In order to detect CR, we exploit the fact that they create showers of secondary particles in the earth's atmosphere, such as electrons and muons [10]. As the secondary particles excite atmospheric nitrogen, the emitted fluorescence light can be used to learn about the energy, direction and composition of the CR [11]. Today, CR experiments like the Pierre Auger Observatory [12] and the Telescope Array [13] measure CR at the highest energies with fluorescence telescopes, which look at the atmosphere, and surface detectors, which measure the shower particles directly, at the same time. This is necessary, since high energy CR are very rare events. In order to collect enough statistics, which allows for meaningful measurements, large areas of the atmosphere need to be observed. As this is increasingly difficult from the ground, the proposed Extreme Universe Space Observatory (EUSO) [14] aims to go to space to observe larger areas of the atmosphere and thus find answers for the ques-

tions regarding the highest-energy CR. In order to reliably detect CR from above using the fluorescence method, a sensitive camera is needed to track the air shower and to distinguish the fluorescence light from the background light in the atmosphere. Conventionally, photomultiplier tubes (PMT) have been used as detectors in such cameras [15, 12]. Nearly two decades ago, silicon photomultipliers (SiPM) have emerged as promising solid-state alternatives to PMT with several advantages, especially to space experiments, such as lower weight, lower operating voltage and general robustness [16]. In recent years, SiPM have become mature devices, which have replaced PMT for many applications. However, SiPM have a few drawbacks, like a high rate of thermal noise, which mimics photon signals, called dark counts. Furthermore, some sensor properties depend heavily on the ambient temperature.
To demonstrate the EUSO idea of detecting CR by looking down to earth, there are several pathfinders. The latest pathfinder was EUSO-SPB, which consisted of a scaled down EUSO-like camera onboard a super pressure balloon, which flew at an altitude of $\sim 33\,\mathrm{km}$ [17]. It was launched in April 2017 from New Zealand and flew for 12 days before it was prematurely let down. In addition to a conventional PMT camera with 2304 pixels, it included the silicon photomultiplier elementary cell addon (SiECA), a prototype SiPM camera with 256 pixels [18]. This allows a very first look at the SiPM performance in near-space and can be used to estimate if the SiPM dark counts are a hindrance to CR detection or if SiPM can indeed be used for future experiments.

This work is structured as follows: In chapter 2, an introduction is given to CR physics and their detection by air showers in the atmosphere. Furthermore, the concepts of CR detection from above and the design of EUSO are presented. In chapter 3, the fundamentals of light detection with SiPM are outlined, and their properties are explained. This chapter also includes an overview of recent developments, which helped the detectors to reach maturity. In chapter 4, the discussion of the experimental work of this thesis begins. The temperature dependence of SiPM is studied, with an emphasis on the SiPM gain. Different methods to stabilize the gain are discussed and a novel approach is presented, which exploits the thermal dark counts as a measure of the detector temperature. In chapter 5, data of the SiECA prototype camera are analysed. This includes the definition of cuts to discard unphysical data, the separation of transient events from static background and a statistical analysis of the data content to evaluate the contribution of dark counts to atmospheric background and transient events. In the end, a brief comparison between SiECA and the PMT camera of EUSO-SPB is shown. Chapter 6 concludes this work and summarizes the results of this thesis.

Chapter 2

Cosmic Rays

Even in modern physics, the origin of ultra high energy cosmic rays (UHECR), the acceleration mechanisms and distribution of their sources are not fully understood. This chapter will introduce the physics of cosmic rays and present many of the unanswered questions, especially in the ultra high energy region. The detection of high energy cosmic rays will be outlined with a focus on space-bound or near-space experiments in the framework of the Extreme Universe Space Observatory (EUSO).

2.1 The Cosmic Ray Spectrum

Charged particles which hit the earth from space are called cosmic rays (CR). Most CR (99 %) are atomic nuclei, most notably protons (90 %), alpha particles (9 %) and heavier nuclei (1 %) up to iron. Electrons only make up a small fraction (1 %) of CR. Their energies range over many order of magnitude from 10^9 eV to the highest observed energies 10^{20} eV, and most likely higher [5, 19]. CR come from outside the solar system and are isotropic at most energies due to propagation in the galactic magnetic field [20]. Solar wind particles are typically not considered to be CR, since their energies are far lower ($<10^8$ eV). The flux of CR $\frac{dN}{dE}$ depends heavily on the CR energy E and can be approximately described by a power law

$$\frac{dN}{dE} \propto E^{-\alpha}, \tag{2.1}$$

with a spectral index of $\alpha \approx 2.7$. The all-particle spectrum of the CR flux for energies above 10^{13} eV is shown in Figure 2.1. The spectrum is multiplied by $E^{2.6}$ to reveal features in the changing of the spectral index α. The three most distinctive features are marked in Figure 2.1: the *knee* between 10^{15} eV and 10^{16} eV, which marks a steepening of the spectrum towards higher spectral

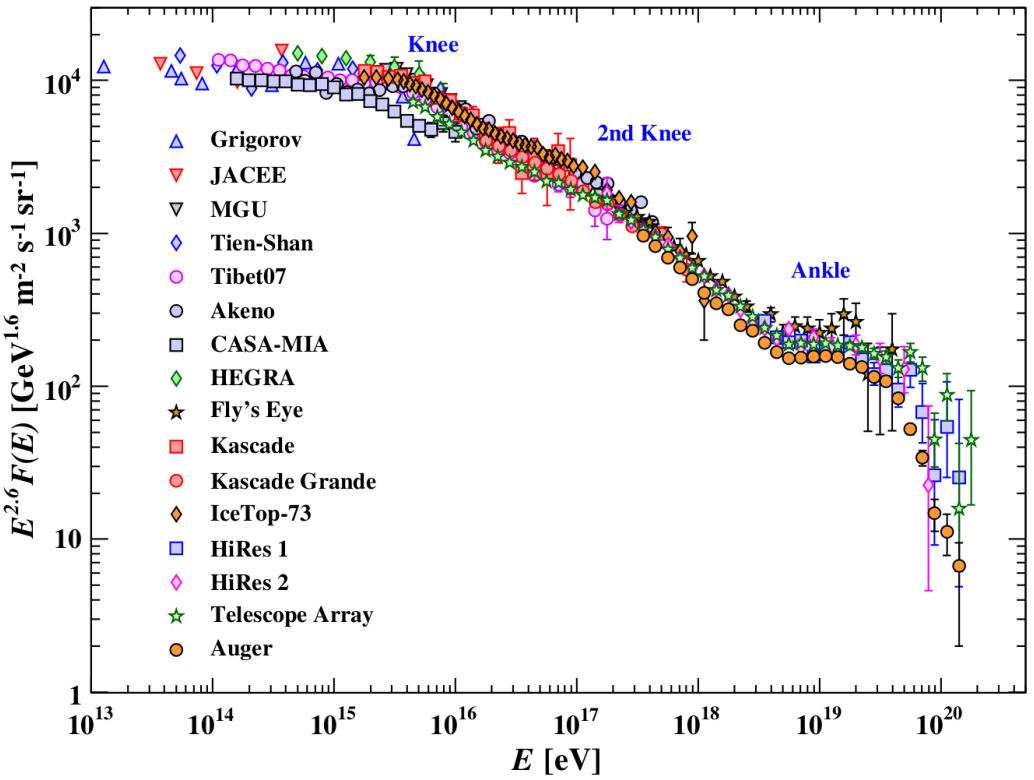


Figure 2.1: All particle spectrum of CR flux, scaled by $E^{2.6}$, against the primary particle energy E. The three features mark changes in the spectral index of the flux, from [22].

indices α, the *second knee*, observed by the KASCADE collaboration [21], where the spectrum steepens further, and the *ankle* around 2×10^{18} eV, where the spectrum flattens again. An additional feature is the strong suppression above 5×10^{19} eV. Due to the small statistics of the measurements and large systematic uncertainties (see section 2.2), the reasons for the spectral features are still not known with absolute certainty. In the following, some of the more popular models, which offer explanations for certain features, are presented.

Studying the abundances of primary CR nuclei below 10^{15} eV reveals large similarities with solar system abundances [10] and almost no dependence of the abundances on primary energy (with a few exceptions [23]). Thus, CR below the *knee* are assumed to be of galactic origin. The appearance of the *knee* could mark the limits of galactic accelerators, e.g. supernova remnants [22]. Since the shock front of a supernova only has a finite lifetime, CR can only be accelerated up to certain energies, which are near the *knee* region $\sim 10^{15}$ eV [24, 25]. Other acceleration mechanisms (e.g. γ-ray bursts) are also able to describe *knee*-like spectral features [26]. Apart from source spectra, the *knee* could also

be explained by propagation effects, such as leaky box models [27, 28], where the galactic magnetic field B fails to confine the CR with increasing rigidity

$$R = \frac{p}{Z} = r_L B, \tag{2.2}$$

where p and Z are the CR momentum and charge, and r_L is the Larmor radius. As CR escape the galaxy, the measured flux decreases, and the *knee* feature appears. In such a model, the *second knee* can be explained by a heavier primary component, which escape the galaxy at higher energies, whereas the *knee* is caused by the escape of lighter primaries, such as protons.

It is believed that the *ankle* marks the transition from galactic primaries to extragalactic primaries, i.e. the extragalactic flux starts to dominate the galactic flux. The CR in this energy regime are called ultra high energy cosmic rays (UHECR). However, the detailed composition (ratio of primary abundances) is still unclear, as is the contribution of galactic and extragalactic fluxes. Composition studies should be able to answer these questions, but are difficult due to the limited statistics and high systematic uncertainties. At the highest energies, a strong suppression is expected, as the primaries start to interact with the cosmic microwave background (CMB) via *photo-pion production*

$$\begin{aligned} p + \gamma_{\mathrm{CMB}} &\rightarrow \Delta^+ \rightarrow p + \pi^0 \\ p + \gamma_{\mathrm{CMB}} &\rightarrow \Delta^+ \rightarrow n + \pi^+. \end{aligned} \tag{2.3}$$

This is called the Greisen-Zatsepin-Kuzmin (GZK) cutoff [29, 30]. All experiments, which measure these UHECR, see a cutoff at similar energies [31, 32, 33, 34, 35]. However, to prove that this cutoff is due to the GZK effect, again composition studies need to be carried out, as the GZK models favour a very light composition over intermediate mass nuclei. In mixed composition models a similar cutoff is expected due to *photo-dissociation* of heavy nuclei [36]. So far, the data is compatible with both models [37]. Although the flux is very low at the highest energy, data [38, 39] agrees on a small iron component of primaries. This is puzzling, as iron has a high acceleration efficiency and is expected to be comparably stable in the CMB [11]. With the current experiments, any statement on the composition beyond the absence of heavy nuclei is very challenging. As extragalactic intermediate nuclei eventually fragment to protons on their way to earth, an intermediate nucleus component requires medium scale anisotropies of CR due to nearby sources. So far, there is only one hint at such a "hotspot" reported by the Telescope Array collaboration with $3.4\,\sigma$ [40].

At energies below 10^{15} eV, CR can be detected directly by balloon or satellite experiments. Since the flux drops rapidly with increasing energy, larger detectors would be needed to gather sufficient statistics. As this is not feasible, CR above 10^{15} eV are detected from ground exploiting the fact that high energy CR form extensive air showers (EAS) by creating secondary particles in the atmosphere. In the following section, the underlying mechanisms of EAS are explained.

2.2 Detection of Ultra High Energy Cosmic Rays

2.2.1 Extensive Air Showers

Above 10^{15} eV, the hadronic particles can interact with nuclei in the air like nitrogen, oxygen and argon at an altitude of 15 km — 35 km. These interactions produce a shower of secondary particles, called extensive air showers (EAS), which can be detected from the ground. The most abundant secondary particles are charged and neutral pions, $\pi^{\pm}$ and π^0. As the subsequent shower particles are created from the decay of these mesons, three shower components can be distinguished, a hadronic component, which contains long-lived mesons, an electro-magnetic component consisting of $e^{\pm}$ and γ, and a muonic component of $\mu^{\pm}$ and $\nu_\mu/\bar{\nu_\mu}$. A schematic overview of EAS is shown in Figure 2.2.

Electro-Magnetic Component

The electro-magnetic (em.) component consists of particles, which only interact electro-magnetically, namely $e^{\pm}$ and γ. The em. component is created by a decay of neutral mesons into two γ

$$\pi^0 \rightarrow \gamma + \gamma. \tag{2.4}$$

It is sustained by the creation of further em. particles via pair production of a e^+/e^--pairs and the production of γ by bremsstrahlung of $e^{\pm}$. The e^- are also subject to ionization energy loss, which is dominated by bremsstrahlung losses until the e^- reach a critical energy E_c, which is ~ 86 MeV in air. Some properties of the em. component can be described by the Heitler model [42]. In this model, after a certain distance λ_{em}, a particle is destroyed and its energy is distributed to two new particles. After n interactions, 2^n particles have been created. Thus, the number of particles N at a certain atmospheric depth X can be described as

$$N(X) = 2^{\frac{X}{\lambda_{\mathrm{em}}}}. \tag{2.5}$$

Obviously, this process cannot continue infinitely. Once particles reach their critical energy E_c, bremsstrahlung is no longer the dominant mechanism for

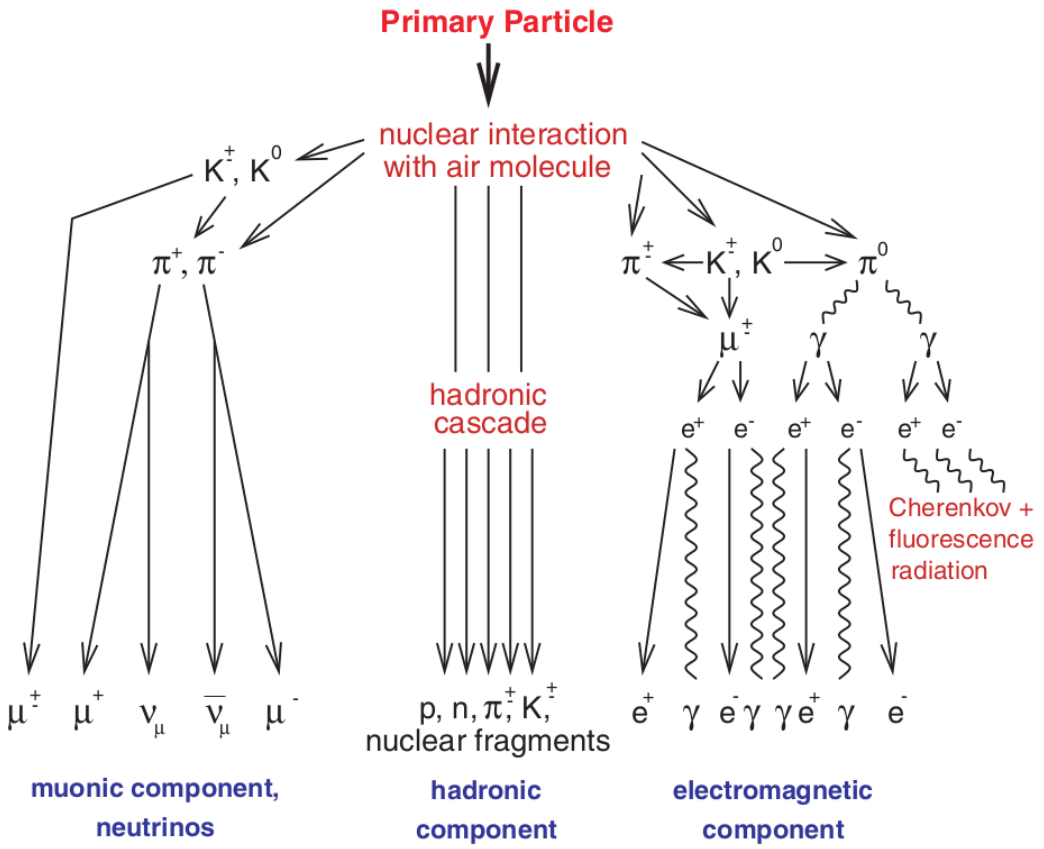


Figure 2.2: Schematic overview of EAS development. The three branches distinguish muonic, hadronic and electro-magnetic components and highlight the most important interactions, from [41].

energy loss and the shower runs out. Since the energy is evenly distributed, the total number of particles at this energy is

$$N_{\max} = \frac{E_0}{E_c}, \tag{2.6}$$

where E_0 is the total available energy for the em. shower. Therefore the atmospheric depth at the shower maximum $X_{\max}$ is given by

$$X_{\max}(E_0) \sim \lambda_{\mathrm{em}} \ln\left(\frac{E_0}{E_c}\right). \tag{2.7}$$

These predictions by the simple Heitler model are confirmed in more exhaustive models [43, 44] and detailed shower simulations [45].

Hadronic and Muonic Component

For the description of hadronic showers, an extended Heitler model can be used [46]. Here, the energy of the hadronic particles is split between charged

particles ($\frac{2}{3}$, $\pi^{\pm}$) and neutral particles ($\frac{1}{3}$, π^0). The neutral particles will rapidly decay into an em. shower (see Equation 2.4). After a certain distance λ_{had}, the charged particles either interact with nuclei in the air, or they will decay, if their energy is below a typical decay energy E_{dec}, and produce muons (see below). Since the em. particles outnumber other contributions (see Figure 2.3), the depth of the shower is mainly given by its em. component

$$X_{\max}(E_0) \sim \lambda_{\mathrm{had}} + X_0 \ln\Big(\frac{E_0}{2n_{\mathrm{tot}}E_c}\Big), \tag{2.8}$$

where n_{tot} is the number of particles produced in the shower. Muons are created in the decay of the charged particles in hadronic showers

$$\begin{aligned} \pi^+ &\to \mu^+ + \nu_\mu \\ \pi^- &\to \mu^- + \bar{\nu_\mu}. \end{aligned} \tag{2.9}$$

As one muon is produced per charged particle, the total number of muons is given by

$$N_\mu = \Big(\frac{E_0}{E_{\mathrm{dec}}}\Big)^{\alpha}, \tag{2.10}$$

where $\alpha = \frac{\ln n_{\mathrm{ch}}}{\ln n_{\mathrm{tot}}}$ is the fraction of charged particles produced in the hadronic shower, which depends on energy, air density and hadronic interactions.

So far in the description for hadronic showers, the primary particle was assumed to be a proton. For heavier primaries, the superposition model can be used [47]. Due to the high shower energies, a nucleus with mass A and energy E_0 can be treated as A independent primaries with energy $E_h = \frac{E_0}{A}$. This yields

$$\begin{aligned} N^A_{\max} &\approx A \cdot \frac{E_h}{E_c} = \frac{E_0}{E_c} = N_{\max} \\ X^A_{\max} &\approx X_{\max}(\frac{E_0}{A}) \\ N^A_\mu &\approx A^{1-\alpha} \cdot N_\mu. \end{aligned} \tag{2.11}$$

As can be seen, there is almost no difference in the number of particles in the shower for heavier nuclei (compare Equation 2.6). However, the position of the shower maximum and the number of muons depend on the mass of the primary. In this way, the mass of the primary and thus the composition of CR is accessible experimentally.

The lateral and longitudinal distributions of EAS are summarized in Figure 2.3. The lateral spread of the em. shower is mainly determined by Coulomb scattering of electrons. Muons have larger lateral spread as they are mostly produced high up in the atmosphere and small lateral momenta are translated

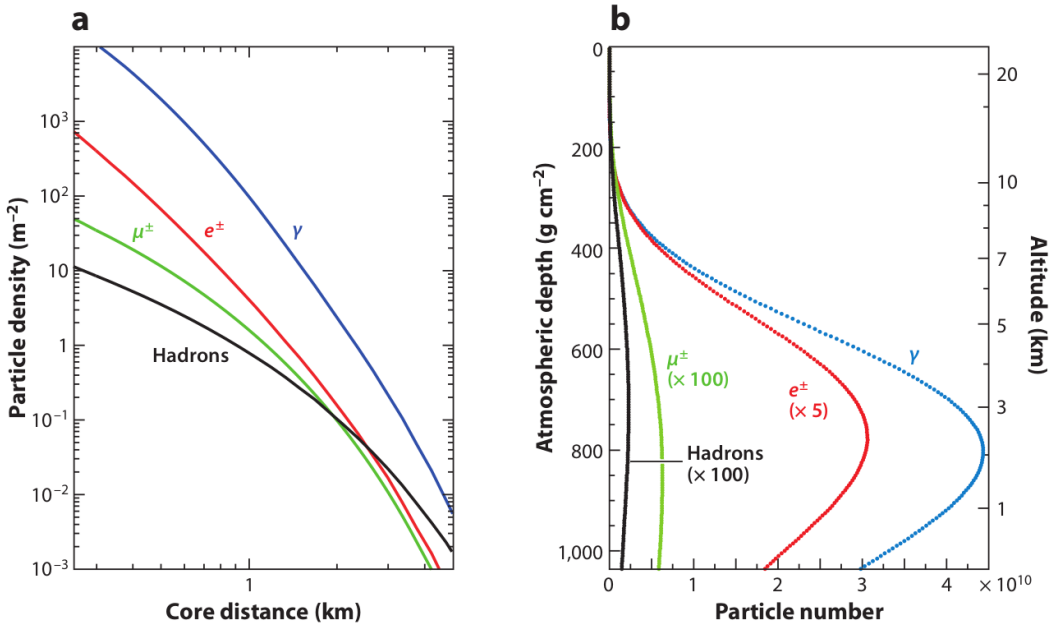


Figure 2.3: Simulated lateral (a) and longitudinal (b) shower profile for 10^{19} eV proton. The electro-magnetic particles are the most frequent, from [48].

to big footprints. In general, there are large fluctuations between showers, even for the same energy and primary. Showers of heavy primaries tend to show smaller shower-to-shower fluctuations [10], which can be understood from the superposition model.

2.2.2 Detection of Extensive Air Showers

To detect EAS from ground, the properties of the shower and its secondary particles are used. Over the years multiple detection strategies have emerged. The most obvious strategy — which was also how CR were initially discovered in 1912 by Hess [49] — is to measure the particles of the EAS directly. For this, modern UHECR experiments use so-called surface detectors, which measure electrons or muons created in the EAS. Secondly, since many particles created in the EAS are relativistic, Cherenkov radiation is created in the atmosphere, which allows the measurement of UV light from the ground. A third approach is to measure the fluorescence light from the excitation of nitrogen atoms in the atmosphere, which also falls in the UV spectrum. These methods will be discussed in more detail in the following section. There are also attempts to measure the radio emission from air showers. Recent developments of this technique are reviewed in [50].

Surface Detectors

Surface detectors are able to measure the charged secondary particles of EAS. Typically, this is achieved with scintillation detectors, where light is recorded from charged particles, which pass through the scintillating material. As an alternative to scintillation detectors, water Cherenkov detectors are used, which make use of the Cherenkov effect [51] in water of the secondary particles. Water Cherenkov detectors have a larger thickness ($\sim$ m instead of $\sim$ cm for scintillation detectors), which increases their sky coverage for large zenith angles. This is particularly useful for experiments which are interested in the highest energies (see the water Cherenkov array of the Pierre Auger Observatory [52]). Surface detectors measure particles from the em. shower component as well as muons, but the em. component can be blocked with appropriate shielding if only muons are of interest.
In order to register the lateral distribution of the shower, surface detector arrays track the arrival times of the incoming particles and thus are able to reconstruct the orientation of the shower and ultimately the orientation of the primary CR. The position of the shower core can be determined by measuring the incoming particle densities, which are fitted to lateral distribution functions, such as the Nishimura-Kamata-Greisen function [53]. For the reconstruction of the primary CR energy, several different methods can be used. As discussed before (see subsection 2.2.1), the number of muons can work as a measure for the primary energy, e.g. for KASCADE, the number of muons in a ring around the shower axis can be used as a measure for the primary CR energy [54]. Alternatively, one can use a correlation between the number of electrons and number of muons for energy reconstruction [55]. While all of the above methods depend on shower simulations, there are also methods independent of simulations, using the fact that primaries arrive isotropically, as used by the Pierre Auger collaboration [32]. To detect the mass of the primary, the ratio of muons and electrons can be used (see Equation 2.11).

Cherenkov Light Detectors

Cherenkov radiation is emitted when a particle moves in a dielectric medium faster than the medium's speed of light [51, 56]. The light is emitted with a continuous spectrum in forward direction. Roughly one third of the charged particles in EAS emit Cherenkov light, most abundantly electrons [57]. The Cherenkov light can be detected on ground with photomultiplier tubes (PMT) (or silicon photomultipliers, see chapter 3) looking towards the sky. Cherenkov light detectors can be distinguished into two categories: Light integrating detectors and imaging detectors. Light integrating detectors consist of an array

of PMT distributed over a large area, often equipped with Winston cones to maximize the collected light. They measure the lateral density distribution of Cherenkov light and thus the electrons of EAS. Similarly to surface detector arrays, this allows the measurement of primary energy, shower direction and mass of the primary. Examples of integrating Cherenkov detectors are BLANCA [58] and TUNKA [59].

An alternative to integrating detectors are imaging Cherenkov detectors, called imaging Cherenkov telescopes, where the light is recorded on a focal plane image of a pixelated detector. This allows for a geometric reconstruction of the shower profile, as the detected light follows the air shower. Thus, the method is much more independent from simulations than surface detectors or integrating Cherenkov detectors. Since only the projection of the shower on the focal plane is recorded, imaging Cherenkov telescopes generally have poor resolution for the direction of the primary CR. Modern experiments such as H.E.S.S. [60], MAGIC [61] and VERITAS [62] overcome this caveat by employing several telescopes at once and using stereo observation. Today, the imaging Cherenkov method is primarily used for γ-ray astronomy, which is mainly interested in γ-induced EAS.

Fluorescence Detectors

At energies above $\sim 10^{17}\,\mathrm{eV}$, particles in EAS can excite nitrogen atoms in the atmosphere. In their de-excitation, they isotropically emit UV photons, which can be used to measure the EAS from any direction. Due to the band structure of the nitrogen atom, which involves vibrational and rotational states, the UV photons are emitted with a characteristic spectrum of several lines, as shown in Figure 2.4. The number of emitted photons per deposited energy is called fluorescence yield and depends on atmospheric conditions like temperature and gas mixture. It is known only to a precision of $\sim 14\,\%$ [63]. However, measurements show that the fluorescence yield does not depend on the energy of the shower particles, thus allowing a calorimetric measurement of the deposited energy, from which the primary energy of the CR can be reconstructed.

Fluorescence light in the atmosphere is usually observed with telescopes, similar to the imaging Cherenkov telescopes described above. As such, they share the suboptimal angular resolution of single telescopes (4° to 5°), which is improved by using stereo observation to $\sim 0.6°$ [65]. Since fluorescence telescopes observe the development of the shower directly, measurements of the shower maximum X_{max} are much more precise than for surface detectors,

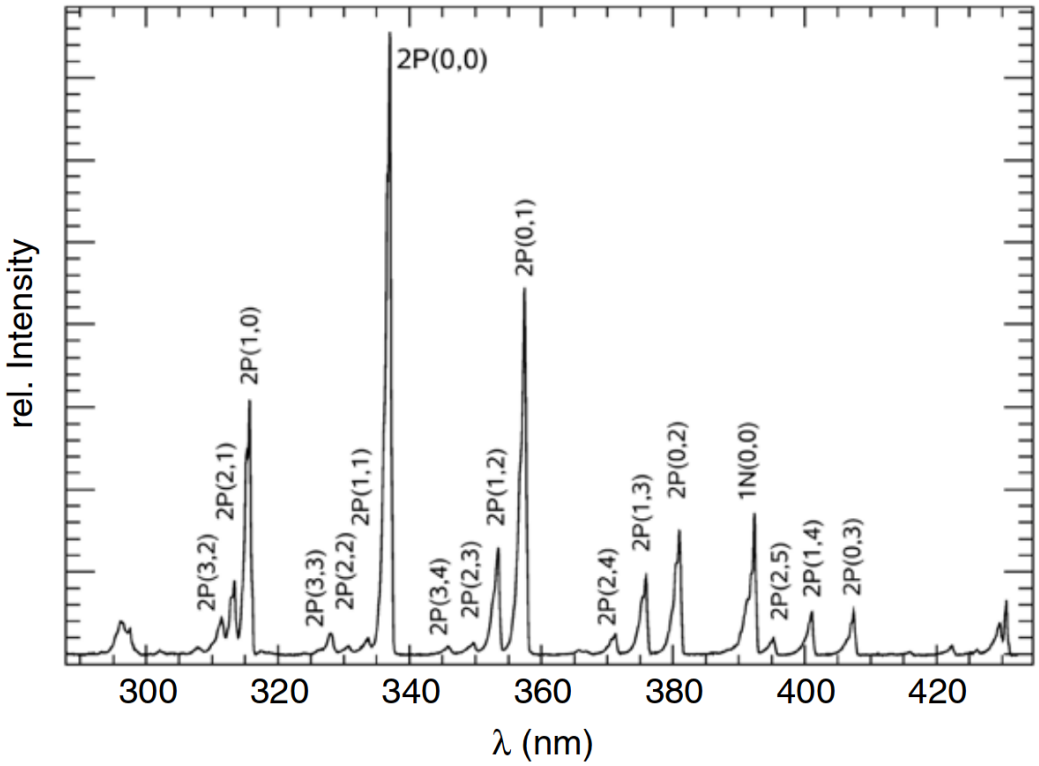


Figure 2.4: Spectrum of nitrogen in the atmosphere as measured by [64]. The majority of the intensity lies in the UV band.

which observe only the developed shower. The most prominent examples of fluorescence telescopes are the aforementioned Pierre Auger Observatory [12] and Telescope Array [13]. Due to the different strengths of surface detectors and fluorescence telescopes both experiments measure in "hybrid mode", i.e. they use both surface detectors and fluorescence telescopes. Since the surface detectors have a 100 % duty cycle, they contribute the most data to the energy spectrum. In contrast to this, fluorescence telescopes can only record data on clear, moon-less nights with a duty cycle of 15 % to 20 %. In order to minimize systematic uncertainties, e.g on the energy scale, fluorescence measurements are used to calibrate the surface detectors. For mass composition studies, fluorescence data is preferred, as the shower maximum X_{max} can be observed directly.

In order to improve statistics, especially at the highest energies, large areas of the atmosphere need to be observed, which requires giant experiments with many fluorescence telescopes. To overcome this obstacle, the JEM-EUSO experiment was proposed to observe the atmosphere from space, which allows for observation of a much larger area. The concepts behind UHECR observation

from above are outlined in the next section.

2.3 The EUSO Idea

The idea to detect CR in the atmosphere from above by looking down goes back to the 1970s [66]. Since then, several experiments have been proposed, but such an experiment has yet to be carried out. A brief historical overview can be found in [67]. The Extreme Universe Space Observatory (EUSO) is the latest attempt at measuring CR from above. To resolve the many open questions about CR (see above), especially at the highest energies, measurements of UHECR are needed with large exposure. Since they are detected using properties of EAS in the atmosphere, the exposure can be maximized by observing a larger area in the atmosphere. Thus, an experiment, which looks down onto the atmosphere, e.g. from the International Space Station (ISS) at an altitude of 400 km, can help in improving the understanding of CR. This principle is illustrated in Figure 2.5. Additionally, such an experiment offers uniform exposure for both northern and southern hemisphere, which eliminates systematic uncertainties, that arise in the comparison of different detectors (see [68]). In the following, the discussion will focus on the proposed EUSO experiment onboard the Japanese Experimental Module of the ISS, JEM-EUSO, which is described in detail in [14]. The projected exposure of JEM-EUSO is shown in Figure 2.6.

As discussed in section 2.2, the radiation of fluorescence light in EAS allows for detection of CR in the atmosphere from any direction. Thus, JEM-EUSO consists of two main parts: A UV camera, which records the fluorescence light of the EAS, and an atmospheric monitoring system to enable precise track reconstruction, since the fluorescence yield depends heavily on the atmosphere. The UV camera is designed in a modular way, where the detector is divided into 137 Photo Detection Modules (PDM), which consist of 36 multi-anode photomultipliers with 64 pixels each, giving each PDM ~ 2300 pixels. As each pixel has a size of $3\,\text{mm} \times 3\,\text{mm}$, they provide a spatial resolution of $\sim 0.5\,\text{km}$ on ground (at 400 km altitude). In order to properly track CR in the atmosphere, the time resolution of the camera is $2.5\,\mu\text{s} \equiv 1$ Gate Time Unit (GTU). Apart from direct fluorescence light, the camera can also detect reflected fluorescence and Cherenkov light from EAS. The standard mode of operation for JEM-EUSO is nadir mode, where the camera looks vertically onto an area of $1.4 \times 10^5\,\text{km}^2$. To further increase the observed area, the camera can be tilted [72].

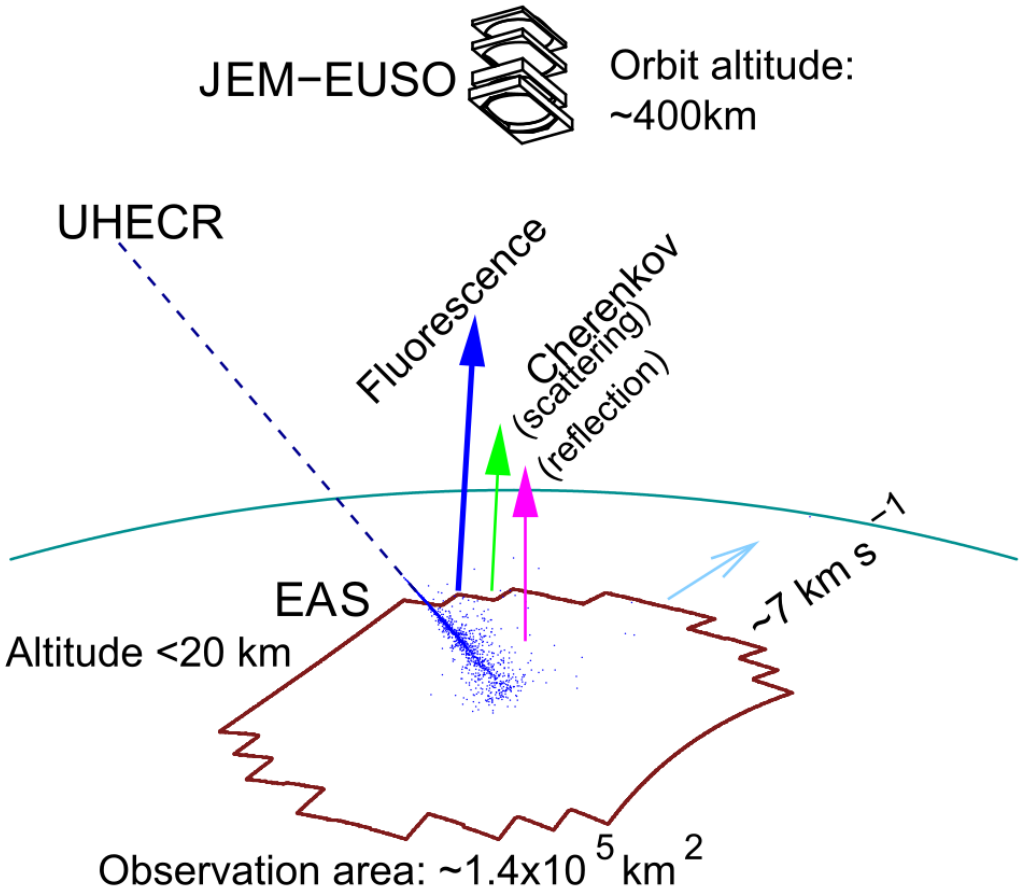


Figure 2.5: Schematic overview of the JEM-EUSO observation principle. Shown here is the nadir mode, where the telescope looks vertically down from 400 km in ISS orbit, from [69].

In comparison to ground-based observation, JEM-EUSO faces the additional challenge of a larger UV background, since the upward radiance is larger than the downward one [73]. As the whole experiment hinges on the ability to distinguish EAS light from background, the UV background is discussed in the following. In general, the UV background can be distinguished into a diffuse part and a transient part. The diffuse part is dominated by airglow in the atmosphere, e.g. de-excitation of nitrogen and oxygen atoms. Additionally, there is a component of reflected or scattered light from the downward radiance, such as moonlight, starlight and diffuse extragalactic light [73]. To minimize these contributions, observation is in general only possible during moonless nights. For the transient part, man-made light contributes heavily and is dominant above big cities [72]. Furthermore, there are a plethora of transient light events associated with lightnings in the atmosphere (see [74] for

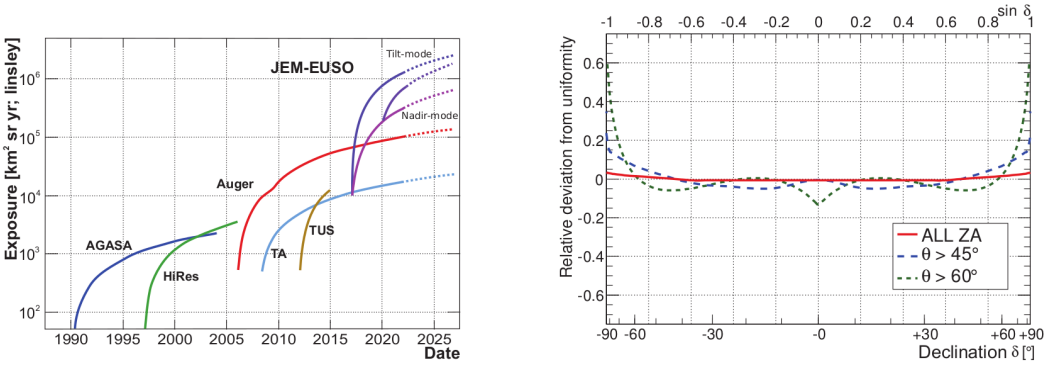


Figure 2.6: (a) Exposure of JEM-EUSO in nadir and tilt mode compared to other experiments. (b) Relative deviation of JEM-EUSO exposure for CR with different zenith angles θ, from [70, 71].

an introduction). While clouds also contribute to the background, JEM-EUSO is not impacted by them in a major way [75]. Low-altitude clouds can even help by attenuating man-made background.

So far, there is no launch date for the JEM-EUSO mission. Meanwhile, there are several pathfinders and prototypes to research the techniques and methods for a JEM-EUSO-like space mission. These pathfinders include a balloon flight in Timmins, Canada in 2014 named EUSO-BALLOON [76] and EUSO-TA [77], which is located at the Telescope Array site and and measures CR from below using a EUSO PDM. Furthermore, MINI-EUSO is a scaled down prototype of JEM-EUSO consisting of only one PDM, which will observe the atmosphere from the Russian module of the ISS [78]. A major part of this work is dedicated to EUSO-SPB, which is the second EUSO balloon mission, onboard a NASA superpressure balloon. It included a prototype camera using silicon photomultipliers, which is discussed in detail in chapter 5. In the following chapter, silicon photomultipliers will be introduced.

Chapter 3

The Silicon Photomultiplier

As this work is about silicon photomultipliers (SiPM), this section aims to give an overview of the working principle of the silicon-based light detector. Furthermore, advantages and drawbacks compared to conventional photomultipliers are outlined and recent developments in the SiPM sector are highlighted. For an exhaustive review of silicon light detectors with an emphasis on SiPM, the reader is referred to [16].

3.1 Photodiodes

The simplest form of a light detector made out of semiconducting material is a p-n junction diode. At the p-n junction there exists an electric field, which depletes the junction region of mobile charge carriers. The region is thus called depletion region. A photon with sufficient energy $\hbar\omega \geq E_g$, where E_g is the band gap of the material, can be absorbed and produce an electron-hole pair. Through the electric field in the depletion region, the charge carriers, i.e. electrons and holes, are separated and a current can be measured at the terminals of the diode. In principle, such a photodiode can be made from any semiconducting material, whose band gap E_g is low enough for the generation of electron-hole pairs with optical photons. A popular choice for photodiodes (and solid state electronics) is silicon, as the higher band gap (1.12 eV compared to 0.67 eV for germanium) generates less noise from thermal excitation. Additionally, it has other convenient properties, such as the easy formation of a passivization layer by silicon oxide [16]. This chapter will focus on silicon-based light detectors.

3.1.1 PIN diodes

The efficiency of a p-n junction as a light detector can be greatly improved by applying a reverse bias voltage, i.e. positive voltage applied to the n side. The applied voltage enhances the potential difference across the p-n junction, thus accelerating created charge carriers to their readout terminals. This reduces the loss of charge carriers due to trapping or recombination. Secondly, the size of the depletion region increases with the increasing applied voltage, resulting in a larger active volume. Another way to increase the size of the depletion region is the introduction of an undoped intrinsic semiconductor layer between the n and p layer, resulting in a PIN diode. As a light detector, the PIN diode has multiple advantages to the p-n diode, first of all the larger sensitive area. Secondly, the undoped intrinsic layer reduces the capacitance of the device, and thus electrical noise [16]. Additionally, the intrinsic layer has fewer impurities and therefore fewer energy states in the band structure. As a result, the dark current from the thermal generation of electron-hole pairs is reduced. Since the noise of a PIN diode is proportional to the area, its size typically does not exceed a few cm^2. PIN diodes need a minimum of several hundreds of photons per sensitive area to detect light signals.

3.1.2 Avalanche Photodiodes

As described above, the electric field across the p-n junction is enhanced by a reverse bias voltage. Therefore, for large applied voltages, the motion of the electrons and holes is dominated by the acceleration of the electric field. At strong electric fields, the charge carriers can collide with either the lattice or other carriers, which leads to the production of additional electron-hole pairs through impact ionization [79, 80]. As a result, the amount of charge carriers no longer follows the amount of absorbed photons, instead, the charge carriers are multiplied internally. Given high enough electric fields, the electron-hole pairs generated through impact ionization can in turn produce additional electron-hole pairs. This results in avalanche multiplication, which is used by avalanche photodiodes (APD). As the mobility of electrons is higher than the mobility of holes, the impact ionization for electrons can take place at lower electric fields. Thus, for lower voltages, only electrons take place in the impact ionization process and the avalanche develops only in the n direction of the diode and runs out if it reaches the readout terminal. For higher applied voltages both charge carriers contribute to the avalanche and a runaway avalanche happens, which develops in both directions. The threshold voltage is called the breakdown voltage V_b. To make use of the multiplication of charge carriers, APD are

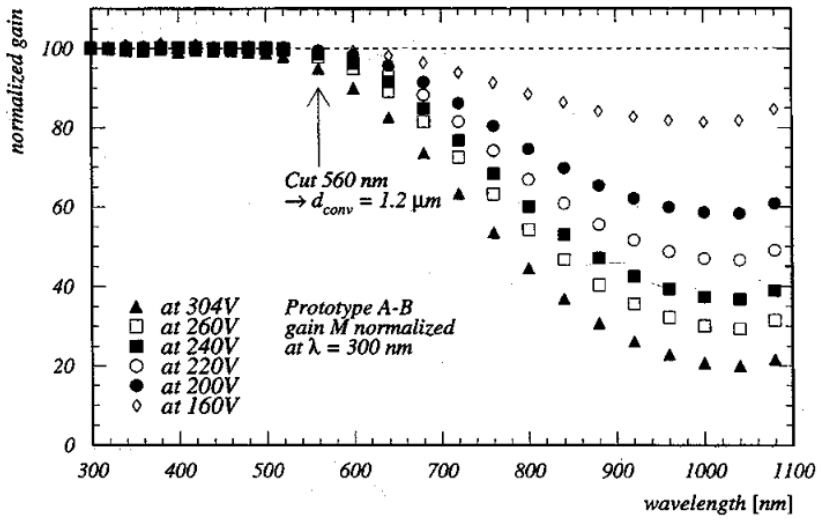


Figure 3.1: Gain of an APD (Hamamatsu S8148) against the incident photon wavelength, from [81].

operated below V_b. The multiplication factor is generally called the gain G of the device. As the avalanche in an APD develops linearly, electron-hole pairs, which are generated in the vicinity of the n side near the readout terminal, generate a smaller gain, since the avalanche runs out before it can undergo full amplification. Thus the gain of the APD depends on the wavelength of the incident photon, as shown in Figure 3.1. This also limits the sensitive area of APD to a few mm^2, since for good energy resolution, a uniform field distribution is needed. The gain of the APD changes exponentially with the applied bias voltage, therefore the relative change in gain is much higher for large applied voltages. A similar behaviour can be seen in the dependence of the gain on the temperature. Because of this, APD are usually operated at gains between 50 and 200 far below the breakdown voltage. They can detect as little as $10 - 20$ photons per sensitive area.

3.1.3 Single Photon Avalanche Diodes

To further increase the sensitivity to the single-photon level, the applied voltage can be increased above the breakdown voltage. In the resulting high electric fields across the p-n junction, holes will be accelerated enough to generate new electron-hole pairs through impact ionization. As the avalanche of holes develops in the opposite direction of the electron avalanche, and both new electrons and holes are generated as the avalanche continues, it is self-sustaining. Accordingly, the devices are called Geiger-mode photodiodes. Since the device

cannot detect further photons while in breakdown, the avalanche has to be quenched, e.g. by a voltage drop across a serial resistor or an active quenching circuit. Today, the devices are known as single photon avalanche diodes (SPAD). Since each generated electron-hole pair has a chance to start the breakdown, Geiger-mode photodiodes are binary devices (they detect *if* at least one photon was absorbed, not how many) with high gain. In contrast to APD, the gain is independent of the wavelength of the incident photon. They are the basic building block of the silicon photomultiplier, which will be discussed in the following. Due to the thermal generation of electron-hole pairs, the size of SPAD is limited to a few hundred µm^2.

3.2 Silicon Photomultipliers

Silicon photomultipliers (SiPM) were invented in 1998 by V. Golovin and Z. Sadygov in Russia [82, 83] and overcome the drawback of the SPAD's inability to count photons. In principle, a SiPM consists of many small SPAD. They are connected in parallel and thus share bias voltage and readout, but each diode is passively quenched and operated independently. Alternative names for SiPM are multi-pixel photon counter (MPPC) or Geiger-mode avalanche photodiodes (G-APD). A schematic view of a SiPM is shown in Figure 3.2. In the context of SiPM the SPAD are typically called microcells. Each microcell still works as a binary device, but due to parallel readout, photon counting is possible, as different photons can trigger avalanches in different microcells. Since each microcell avalanche produces the same signal a, the signal A of the SiPM is proportional to the number of fired cells N_{fired}

$$A = \sum_{N_{\text{fired}}} a = N_{\text{fired}} \cdot a. \tag{3.1}$$

During an ongoing avalanche in a microcell, no further photons can be detected. For a linear response of the SiPM to photons, it is crucial that the rate of incident photons is small compared to the inverse recovery time of the microcell, to allow for complete recovery before detection of the next photon. Similarly, there is a geometric nonlinearity if multiple photons hit the same microcell, which can be described as

$$N_{\text{fired}} \approx N_{\text{cells}}\left(1 - e^{-\frac{N_{\text{p.e.}}}{N_{\text{cells}}}}\right), \tag{3.2}$$

where N_{cells} is the number of microcells of the SiPM and $N_{\text{p.e.}}$ is the number of produced photoelectrons (avalanches). Modern commercial SiPM usually are $1 \times 1\,\text{mm}^2$ or $3 \times 3\,\text{mm}^2$ in size with a microcell pitch between 10 µm and 75 µm [85, 86], which guarantees sufficient dynamic range for typical applications. In the following, properties of SiPM are examined more closely.

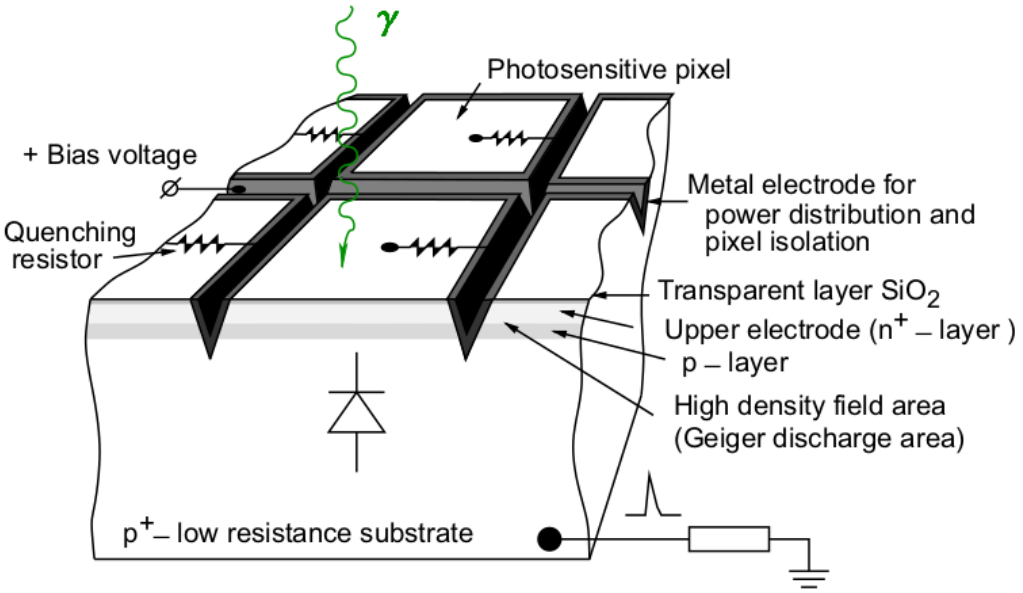


Figure 3.2: Schematic overview of the topology of a SiPM, from [84]. This SiPM has a n-on-p structure, which is suitable for detecting red and infrared photons. For the detection of UV photons, a p-on-n structure would be used (see subsection 3.3.2).

3.3 Properties of SiPM

3.3.1 High Gain

One of the most distinctive properties of SiPM compared to other silicon light detectors is its high gain G, which enables the detection of single photons. The gain G is the amplification factor of the original generated photoelectron, i.e. the amount of charge carriers, which are detected from one photon. It is proportional to the microcell capacitance C_{cell}

$$G = \frac{C_{\text{cell}}}{e} \cdot (V - V_b), \tag{3.3}$$

where e is the elementary charge, V is the applied voltage and V_b is the breakdown voltage. The quantity $V - V_b$ is called the overvoltage V_{over} and is one of the most important parameters of the SiPM, since many properties depend on the overvoltage, in addition to the gain. In order to achieve a signal with good resolution (compare Equation 3.1), fluctuations in microcell capacitance have to be kept to a minimum. Since the microcells are produced on the same wafer with the same process, this is not an issue for modern SiPM. The SiPM gain is comparable to the typical photomultiplier gain $G \approx 10^5 - 10^7$. Typical SiPM

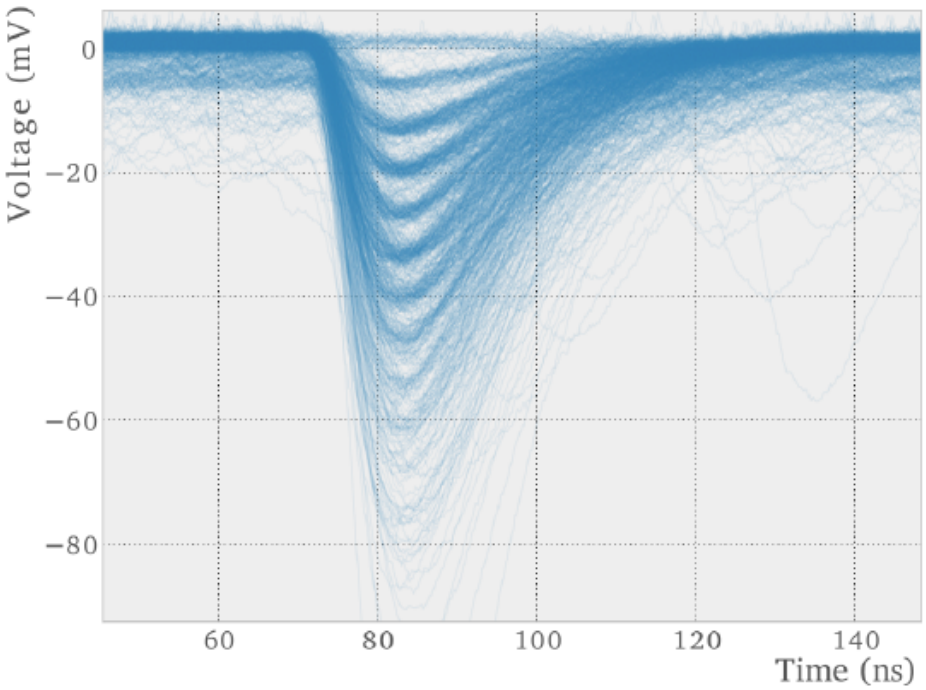


Figure 3.3: Superposition of 1000 triggered LED pulses from a SensL B-Series 30035 SiPM, taken from [87].

signals are shown in Figure 3.3. Due to the runaway avalanche (and passive quenching), the SiPM gain does not depend on the wavelength of the incident photons. Since each microcell signal produces the same signal, the gain can be accessed experimentally by comparing two SiPM signals of different height, i.e. the baseline and a 1 p.e. signal. This is easily measured as the difference of two peaks in a pulse integral spectrum, as seen in subsection 4.1.3.

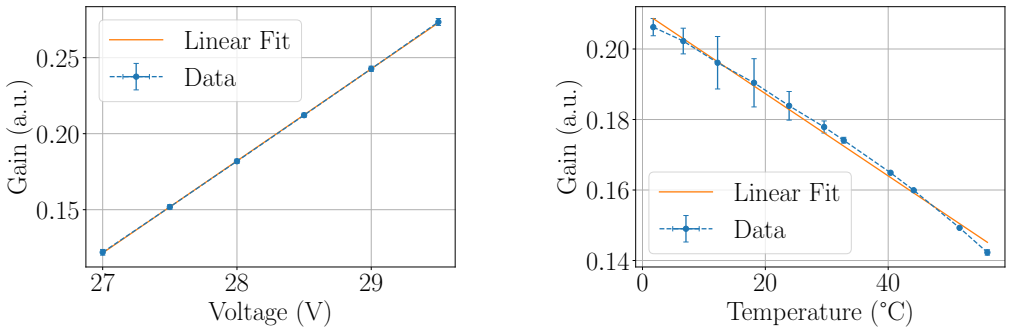
As shown in Equation 3.3, the SiPM gain depends proportionally on the applied voltage. This sensitivity of the gain G to the voltage can be parametrized as a voltage coefficient k_V

$$k_V = \frac{dG}{dV}. \tag{3.4}$$

An example measurement is shown in Figure 3.4a. The breakdown voltage itself depends on the temperature, which influences the mobility of the charge carriers. Therefore, the gain also depends on the temperature. A temperature coefficient k_T can be defined in a similar way to k_V

$$k_T = \frac{dG}{dT}. \tag{3.5}$$

An example measurement of this effect is shown in Figure 3.4b. As the temperature dependence of the gain can cause problems for experiments, techniques



(a) Dependence of SiPM gain on applied voltage.

(b) Dependence of SiPM gain on ambient temperature.

Figure 3.4: Dependence of SiPM gain on applied voltage (a) and surrounding temperature (b). The measurement of these data is detailed in subsection 4.1.3.

to stabilize the gain need to be used. A novel approach, which uses dark counts instead of external temperature readings to stabilize the gain is presented in chapter 4.

3.3.2 Photo Detection Efficiency

The photo detection efficiency (PDE) describes the probability of an incident photon to be detected by the SiPM. It is the product of the quantum efficiency QE of silicon, the fill factor ϵ_{geo} of the sensitive microcell area to the total SiPM area and the probability to trigger an avalanche ϵ_{Geiger}

$$\text{PDE} = \text{QE} \cdot \epsilon_{\text{geo}} \cdot \epsilon_{\text{Geiger}}. \tag{3.6}$$

The quantum efficiency of silicon is $\sim 95\,\%$ and cannot be increased significantly. Since the microcells are separated geometrically and the quenching resistors take up some space, a good fill factor is achieved by using large microcells. Typical values range from 0.4 to 0.8. As larger cell sizes (for the same area) limit the dynamic range, and introduce a higher recovery time due to higher capacity ($\tau = RC$), there is a trade-off to be made about the microcell size.

Due to their higher mobility, electrons are more likely to trigger an avalanche than holes. Therefore, photons, which are absorbed in the p layer, have a higher probability to detected [88], which introduces a wavelength dependence to the PDE. Depending on the topology of the SiPM (p-on-n vs n-on-p), the PDE can either peak near the UV or in the red waveband. PDE values for modern SiPM range from 30 % up to nearly 50 % at peak wavelength [89]. An

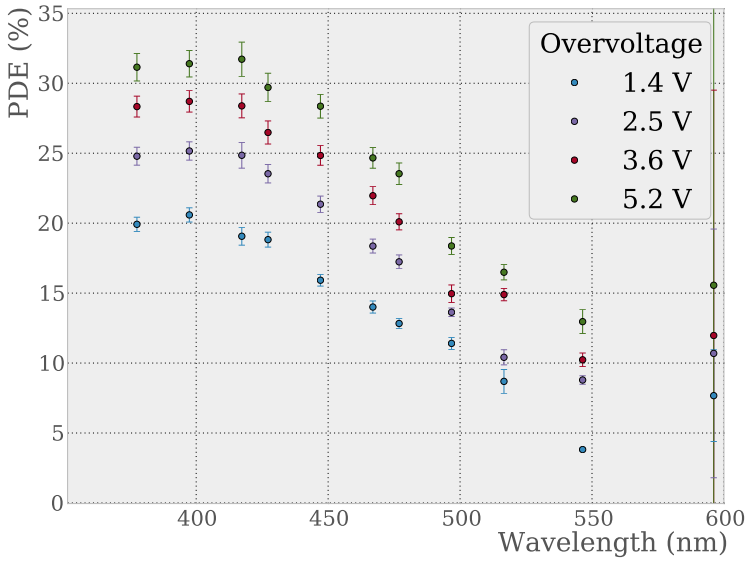


Figure 3.5: Measurement of the PDE of a SensL B-Series 30035 SiPM, taken from [87].

example measurement of SiPM PDE versus wavelength is shown in Figure 3.5. Since the trigger probability ϵ_{Geiger} increases with increasing acceleration of charge carriers, larger overvoltages increase the SiPM PDE.

3.3.3 Dark Counts

Photon signals in a SiPM are created from the generation of an electron-hole pair, which leads to an avalanche and thus a 1 p.e. signal. Due to this, each source of electron-hole pairs mimics photon signals. The resulting pulses, which do not come from incident photons are thus called dark counts. The greatest contributor to the dark counts are thermally generated electron-hole pairs. At room temperature, they lead to a dark count rate of $\sim 1\,\text{MHz}$ per $3 \times 3\,\text{mm}^2$ for recent devices. As they are thermally generated, cooling the device significantly reduces the dark count rate. Beside thermally generated electron-hole pairs, electron-hole pairs are generated by field-assisted tunneling [90]. However, the contribution to the dark count rate is small and only plays a significant role for devices, in which the thermal generation of electron-hole pairs is suppressed, e.g. by cooling.

As electrons are more likely to trigger an avalanche than holes, a thermal generation of an electron-hole pair on the p side has a better chance to produce a signal. As a result, p-on-n devices (which peak in the UV) show

lower dark count rates than n-on-p devices, since the p layer is much thinner (compare Figure 3.2).

3.3.4 Optical Crosstalk

During the avalanche, ~ 3 photons are emitted from the device per 10^5 charge carriers on average [91, 92]. These photons are either absorbed within the microcell, or they can leave the microcell to illuminate the environment or penetrate into other microcells. There, they can be absorbed like any external photon and trigger another avalanche, resulting in a signal >1 p.e., even when only one incident photon hit the detector. This effect is called optical crosstalk. Typical values range from 10 % up to 50 %, depending on the overvoltage. Since higher numbers of charge carriers lead to more emitted light, this effect is particularly pronounced at high gains, i.e. high overvoltages.

3.3.5 Interdependence of Properties

As seen above, all important SiPM parameters depend on the overvoltage, which makes finding an optimum working point a trade-off between desired and undesired properties. For example, high gain and PDE are desireable features, since they result in a good single photon resolution. Both features can be achieved by applying a high overvoltage, but this also increases the optical crosstalk and dark count rate, which both lead to an overestimation of the photon rate. Additionally, in contrast to conventional photomultipliers, gain and PDE are interconnected, as they cannot be changed independently. On a photomultiplier, the PDE is mostly given by the quantum efficiency of the photo-cathode and the gain can be tweaked by changing the supply voltage, with relatively little effect on other parameters. To satisfy the requirements for many experiments, there has been an ongoing effort to improve the SiPM properties in recent years, which will be discussed in the following.

3.4 Recent Developments

For many experiments, especially in the cosmic ray sector, high PDE is crucial to reach high sensitivity. Major improvements to the SiPM PDE have been made in the last years. As described in Equation 3.6, the PDE is made up of three factors, all of which have seen improvements. While the quantum efficiency of silicon cannot be changed, new materials for the passive layer on top of the SiPM are used. The use of silicone resin instead of epoxy to coat the SiPM provides a higher transparency in the UV waveband. Improvement

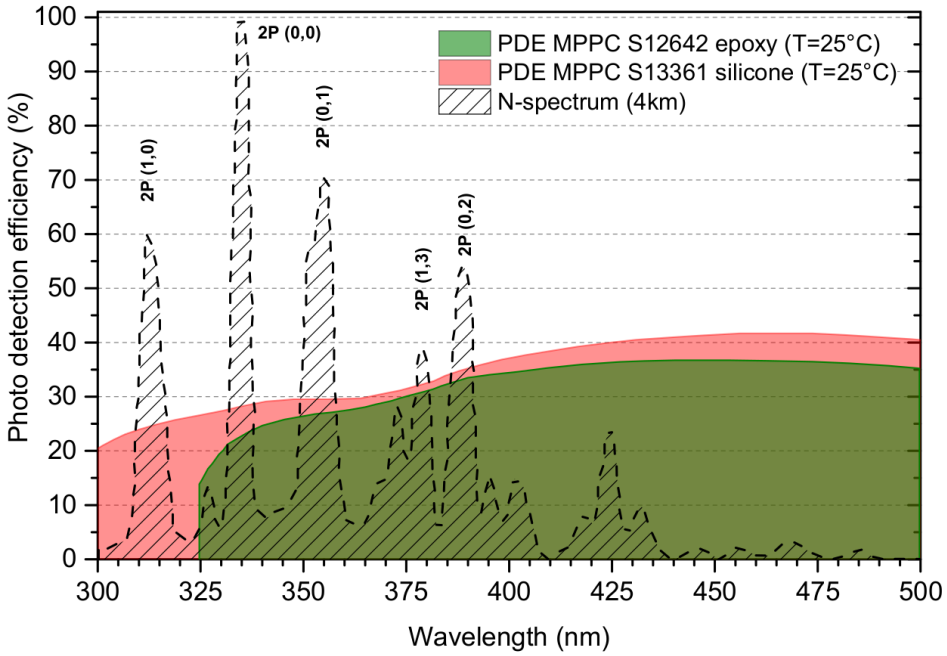


Figure 3.6: Expected PDE of two Hamamatsu SiPM with epoxy and with silicon window. The silicon window extends the PDE towards lower wavelengths. Taken from [93].

of the geometric fill factor was possible, as the use of "through-silicon via" technology (TSV) reduces the space needed for wiring on the photosensitive area, see Figure 3.7. Additionally, improvements to dark count rate and optical crosstalk allow operating the SiPM at higher overvoltage and thus higher avalanche trigger probability. Today, there are devices available with PDE of 50 %.

For the reduction of optical crosstalk, trenches were introduced between microcells, which limit propagation of photons from one microcell to the next. For further crosstalk reduction, a potential barrier can be introduced between the active layer of the SiPM and the bulk. This also reduces the dark count rate, as thermally generated electron-hole pairs can no longer diffuse into the active volume. Modern SiPM have a crosstalk probability of $\sim 10\,\%$ at their working point, and dark count rates of 100 kHz per mm^2 are common. As a bonus, the improved fabrication processes, which reduce dark counts, also improve other effect like the temperature dependence of the breakdown voltage.

Since UHECR experiments often need to track the propagation of light through the atmosphere, multi-pixel detectors are needed. In recent years, SiPM arrays have emerged to meet this requirement. Due to their fabrication

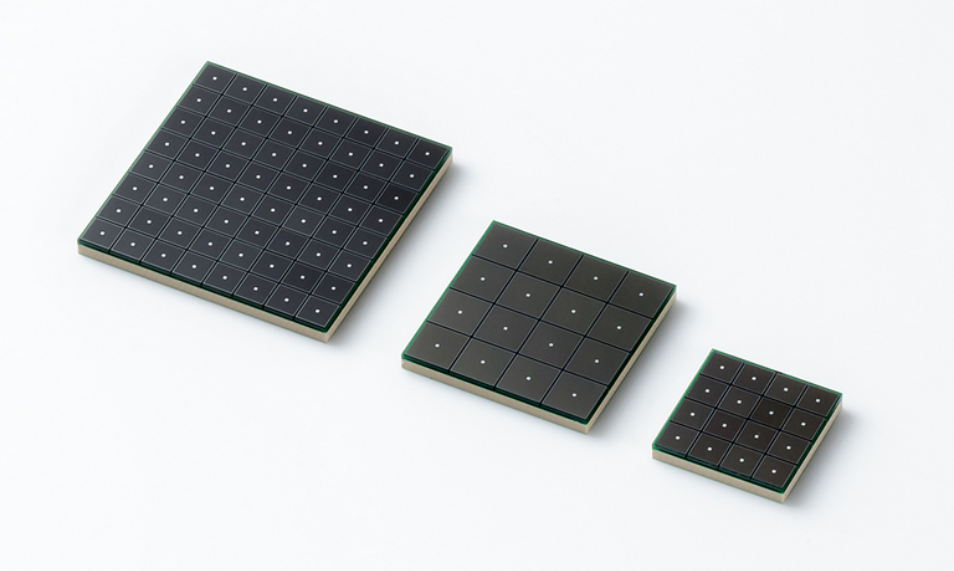


Figure 3.7: SiPM arrays by Hamamatsu. In this picture two arrays with $3 \times 3\,\mathrm{mm}^2$ pixel size are shown (left and right) and one array with $6 \times 6\,\mathrm{mm}^2$ pixel size. The dots in the middle of the pixels come from the use of TSV technology. Picture taken from [94].

process, they show excellent uniformity in SiPM properties such as breakdown voltage and gain while limiting dead space between the single pixels. The SiECA camera used for the EUSO-SPB balloon flight discussed in chapter 5 uses 8×8 SiPM arrays, other examples are shown in Figure 3.7.

To conclude, nearly all properties of SiPM have seen improvements. The maturity of SiPM makes them viable alternatives to conventional photomultiplier tubes for many experiments and applications. Especially for space or near-space experiments their reduced weight and robustness are attractive features, although the high dark count rates could hinder the search for rare UHECR events. In order to test SiPM in such an environment, a prototype SiPM camera was flown along the EUSO-SPB experiment. A detailed analysis of the SiPM data is presented in chapter 5. As the temperature in such an environment can change heavily, the next chapter discusses ways to stabilize the SiPM gain.

Chapter 4

Gain Stabilization of SiPM

In observational astroparticle physics experiments, light detectors can be subject to temperature changes in the order of 50 K (see e.g. Figure 5.17). Due to the SiPM detection principle, many SiPM parameters depend strongly on changes in temperature. In particular, a stable SiPM gain is crucial to good photon counting performance. In this chapter, the effect of temperature change on the SiPM gain is discussed as well as means to correct for this effect. A novel approach is presented, which measures the gain continuously from dark counts and eliminates the need for external temperature sensors for gain stabilization.
Section 4.1 discusses how to extract the gain of SiPM from dark count data by comparing the position of the 0 p.e. peak and the 1 p.e. peak in dark count spectra. Measurements at different temperatures are carried out to determine the temperature dependence of the SiPM gain and a correction factor is calculated to stabilize the gain with respect to temperature changes. After that, a method of simplifying these spectra is introduced, which allows for automated gain measurements with limited statistics. In section 4.2, the extracted correction factor is used to stabilize the gain of a SiPM by measuring the changing temperature. At last, a PID controller is utilized to stabilize the gain without a direct temperature measurement.

4.1 Gain Extraction

As described in section 3.3, the SiPM gain is defined as the average amount of detected charge carriers from the generation of one electron-hole pair. This charge can be measured by integrating the signal of a 1 p.e. pulse. An easy way to determine the gain is the creation of pulse height spectra or pulse integral spectra, where the gain can then be read off the difference between the baseline peak and the 1 p.e. peak. Since higher order p.e. peaks are just sums of single

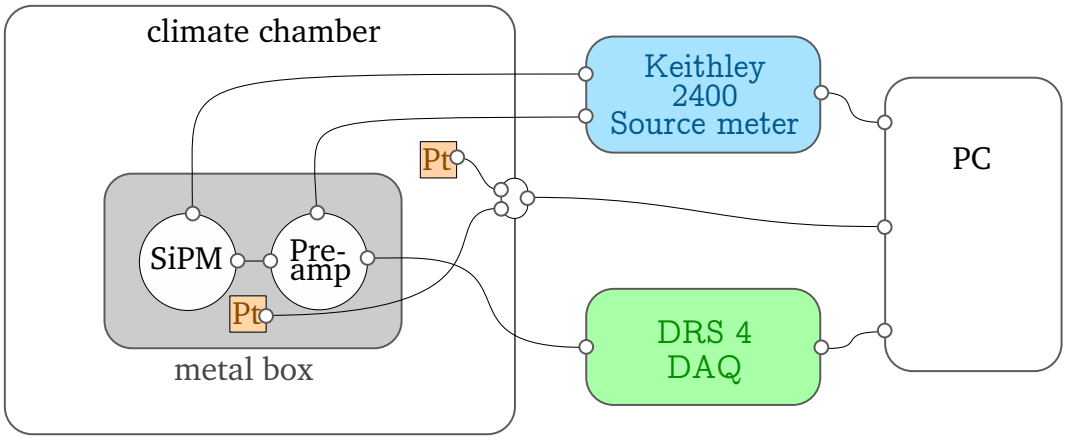


Figure 4.1: Experimental setup used in this chapter. Both SiPM and preamplifier are placed inside a light-tight metal box, which is depicted as grey. The box is placed inside a climate chamber, which also holds two Pt100 resistors to read out the temperature. The Pt100 are depicted in orange and read out over a serial interface by a PC. The preamplified SiPM signal is digitized by a DRS4 Evaluation board, shown in green, which is read out via USB. Finally, a Keithley 2400 source meter both provides the supply voltage for both SiPM and preamplifier.

p.e. pulses, it is possible to determine the gain from the different peak positions of any consecutive peaks in the spectrum, but the distinction of 0 p.e. peak and 1 p.e. peak is usually the easiest way, as these peaks tend to have the highest statistics.

4.1.1 Experimental Setup

While SiPM are essentially digital devices, which count photons, they still produce analog signals and are thus read out in an analog way. In order to reliably count photons, the charge released by a microcell avalanche, i.e. the gain has to be known, even if not in absolute units. For the measurements in this chapter no external light source was used. Instead, the SiPM were operated in darkness using the ever present dark counts from thermally excited electron-hole pairs. Since the generation of electron-hole pairs and thus dark counts is a stochastic process, data was taken using a random trigger scheme. The SiPM data was acquired with a *DRS4 Evaluation Board* [95] and processing of traces was done digitally. A single pixel C-Series SiPM from SensL [85] was used for all measurements in this chapter. It was preamplified with a custom made preamplifier described in [87], designed with this SiPM in mind. The C-Series SiPM from SensL come with the unique feature of two

output terminals. A fast terminal, which is capacitively coupled to the cells for time-critical measurements and a slow terminal, which is read out through the quench resistor in the standard way for SiPM. For this work, the slow output terminal was used. Both SiPM and preamplifier were placed in a light-tight metal box to prevent electronic pickup and to shield the sensor from incident light. The bias voltage was supplied by a Keithley 2400 source meter [96], which was controlled remotely from the same PC aquiring the data.
In order to measure temperature dependencies of the SiPM, the metal box with SiPM and preamplifier was put into a commercial climate chamber. The climate chamber used was the "Kälte-Wärme-Prüfschrank 150 d/70 DU" of Weiss Umwelttechnik, which was located in Institut für Astronomie und Astrophysik Tübingen. To further shield the SiPM from incident light, the entrance window of the climate chamber was closed with black foamed plastic. The temperature inside the climate chamber was read out with two Pt100 resistors, one inside the climate chamber, but outside the metal box and one resistor inside the metal box, right next to the SiPM. A schematic overview of the experimental setup is given in Figure 4.1.

4.1.2 Event Extraction

The traces were recorded by a *DRS4 Evaluation Board* with a sampling speed of 2 Gigasamples per second. Each trace consists of 1024 data points, thus $\sim$ 500 ns of data are recorded per trace. Since only dark counts were of interest, the *DRS4* chip was triggered randomly $\sim$ 450 times per second. An example of a typical trace[1] at 40 °C is shown in Figure 4.2. Due to their stochastic nature, dark counts are distributed randomly throughout the traces. In principle a dark count spectrum can be constructed by choosing an arbitrary interval either for integration or for determining the pulse height in each trace and building a histogram from this data. However, this has a few drawbacks, which are addressed in the following.

Since the gain is measured by examining the position of two different p.e. peaks in the pulse height spectrum (see subsection 4.1.3), it is vital to resolve different p.e. peaks with good certainty. This is easy enough by eye, but can get complicated quickly if the spectrum is noisy, since the baseline can shift and the gain is not always known a priori. While SiPM have a excellent single p.e. resolution, they are still subject to electrical noise, which broadens

[1] The preamplifier produces signals with negative polarity, which were used as is for the analyses. For the sake of clarity, the polarity of the pulses was changed in the description given here.

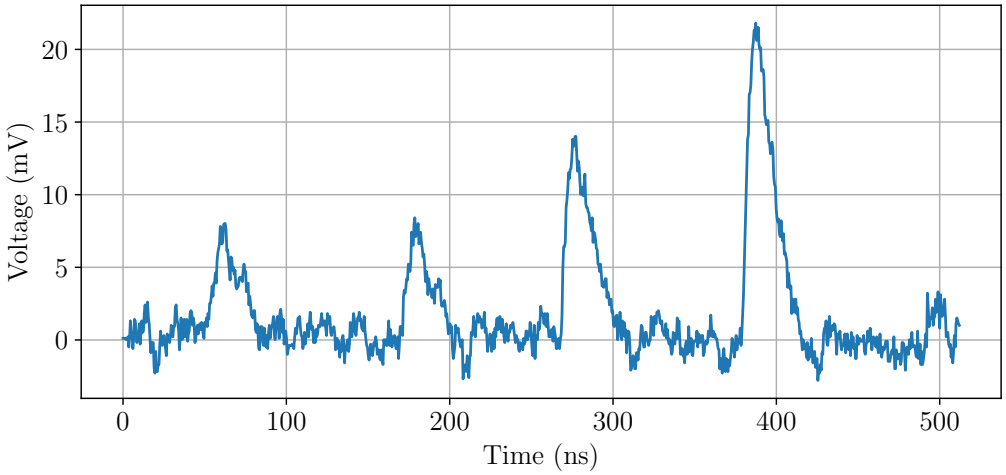


Figure 4.2: Typical trace recorded from a SensL C-Series SiPM @ 40 °C with 2 Gigasamples/s. The 1 p.e., 2 p.e. and 3 p.e. pulses can clearly be seen at 7 mV, 14 mV and 21 mV, respectively.

the peaks in the pulse height spectrum and makes distinction of peaks more difficult. A common way to combat such noise is digital signal processing by using filters, i.e. a convolution of a filter with data. Depending on the kind of noise, there exists an optimum filter, which maximizes the signal-to-noise ratio [97]. However, since the noise is in practice also time-dependant and therefore no single optimum filter exists, a compromise has to be found. For this chapter, a *matched filter* was used, which is the optimum filter for the case of white noise. Other filters were tested, but did not yield significantly better results. For a waveform $x(t)$, which consists of a noise component $n(t)$ and a signal component $s(t)$

$$x(t) = n(t) + s(t), \tag{4.1}$$

the matched filter $h(t) = s(-t)$ can be constructed directly from the expected signal $s(t)$. To extract a noise-free signal from the traces, a digital threshold trigger is run over sample traces to extract many single noisy signals. From these signals an average "template" pulse can be built, in which most noise is averaged out. For the digital trigger to fire, the waveform had to be above a certain threshold for at least 10 samples, i.e. 5 ns. The 1 p.e. trigger threshold was chosen by eye to ~ 0.5 p.e. and a limit was placed on the absolute pulse height to filter pulses > 1 p.e. The template pulse constructed from 90000 1 p.e. pulses is shown in Figure 4.3. Since the shape of the pulse stays the same for pulses > 1 p.e., the filter also matches these pulses. The filtered trace from Figure 4.2 is shown in Figure 4.4. It can be seen that pulses are much more pronounced in the filtered trace with regards to the baseline noise.

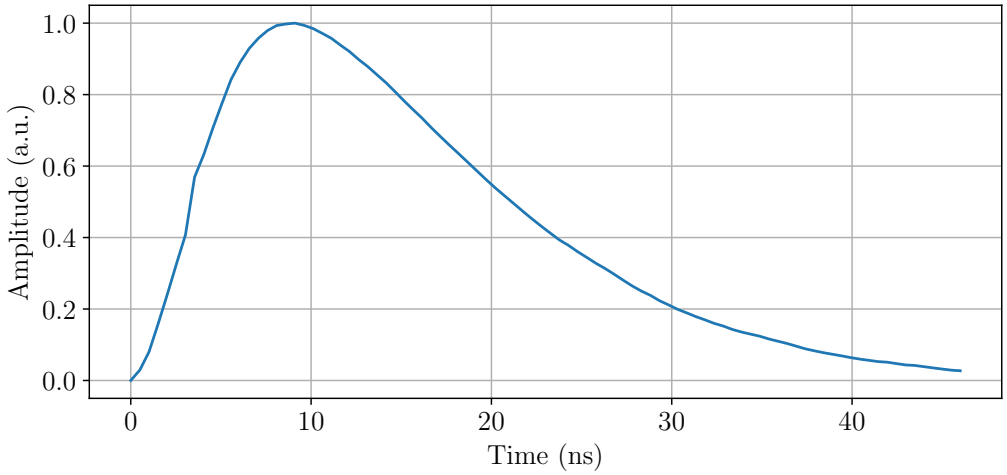


Figure 4.3: Template pulse constructed from 90000 1 p.e. pulses. This pulse is used as a matched filter in the event extraction from the traces.

For the creation of meaningful pulse height spectra, many pulses are needed. Due to the high dark count rate ($\sim \frac{\mathrm{MHz}}{\mathrm{mm}^2}$), there are traces with multiple pulses, which are lost if the spectrum is created using only one interval as naively suggested above. In order to make the most of every trace, care was taken to find every single pulse in the data. To do this, local maxima were found in the filtered traces in a window of a certain length l_w, as indicated by the orange lines in Figure 4.4. The values of the local maxima could then be used to build a pulse height spectrum. It should be noted that most of the local maxima are baseline events, which count into the 0 p.e. peak of the spectrum. This distorts the ratio of 0 p.e. counts and pulse counts and thus the Poisson nature of the spectrum. Since the position of the peaks stays the same, this is irrelevant for the gain measurement and more pulses are counted this way, which reduces the measurement time for gain measurements. About the length of the window l_w, there is a trade-off to be made. With a larger window, fewer 0 p.e. counts are picked up, which is desirable, since the spectrum is dominated by the 0 p.e. peak anyway and time is not spent processing redundant events. On the other hand, a larger window times make it more expensive computationally to find the local maxima, which becomes important in section 4.2, where performance is critical. In the end, a window length of $l_\mathrm{w} = 30\,\mathrm{samples} \mathrel{\hat{=}} 15\,\mathrm{ns}$ was chosen, which is about half the length of a pulse, as this provided a good compromise.

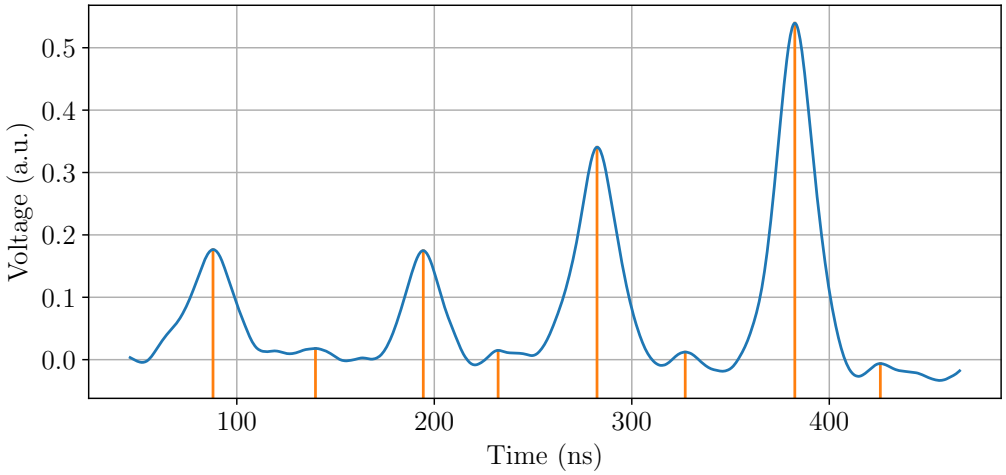


Figure 4.4: Trace from Figure 4.2 filtered with the matched filter in Figure 4.3. Only points in which filter and data overlap completely are given, thus the x-axis is a bit shorter compared to the unfiltered trace. The orange lines indicate the local maxima, which are extracted from the trace and used for measurements.

4.1.3 Dark Count Spectra

The values of the local maxima in the filtered traces (see orange lines in Figure 4.4) are used to construct a pulse height spectrum from dark counts. An example of such a spectrum can be seen in Figure 4.5. In the spectrum, single p.e. peaks are clearly separated, which makes it easy to measure the gain. Since the SiPM gain is defined as the amount of charge carriers detected from the generation of one electron-hole pair, it is equivalent to the integral of a 1 p.e. pulse. As higher order p.e. pulses are just sums of single p.e. pulses, the gain is also represented in the difference between two arbitrary consecutive peaks in the spectrum. The 0 p.e. peak and the 1 p.e. peak have the highest statistics, which makes is convenient to use these two peaks for gain measurement.

To precisely measure the positions of the peaks, normal distributions are fitted to the first two peaks

$$f(x) = A_0 \cdot e^{-\frac{1}{2}\left(\frac{x-\mu_0}{\sigma_0}\right)^2} + A_1 \cdot e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}. \tag{4.2}$$

The starting parameters A_i, μ_i and σ_i for the fit are estimated from the spectra by finding the location of the local maxima which correspond to the peaks. The gain G can then be obtained by the difference of the peak positions

$$G = \mu_1 - \mu_0. \tag{4.3}$$

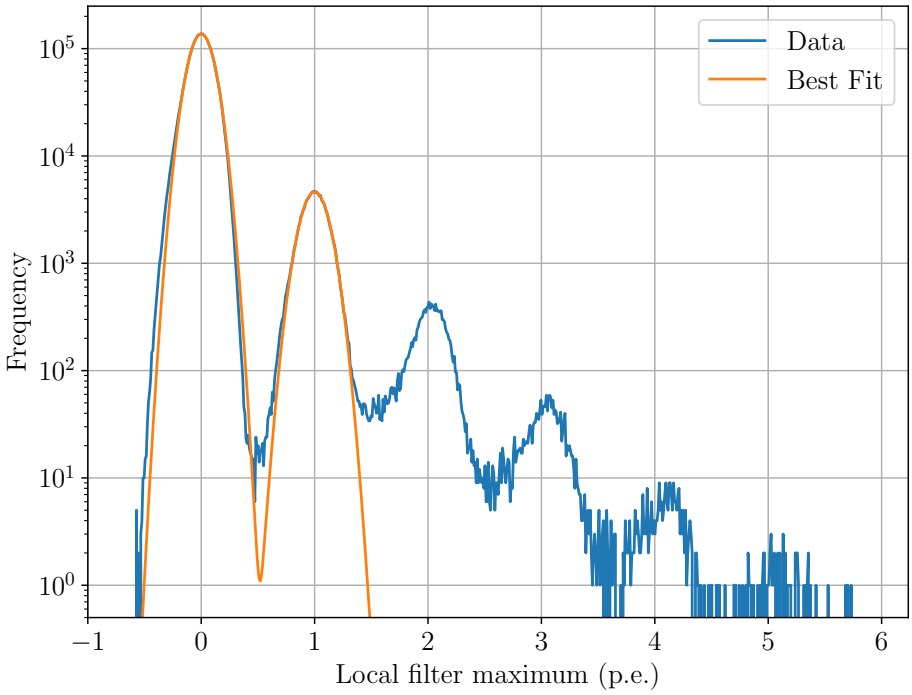


Figure 4.5: Dark count spectrum created from 350000 traces with a SensL C-Series SiPM at 28 V and 25.5 °C. The individual p.e. peaks are clearly separated. In orange, the best fit of a model with two Gaussians is shown, from which the gain can be measured as the difference between the Gaussian mean values μ_i.

An illustration of this is shown in Figure 4.5. Since higher order p.e. peaks are heavily suppressed, it suffices to fit two normal distributions. Ideally, one would fit as many Gaussians as can be seen in the spectrum, but in practice, this makes the fit more unstable and the benefit is minimal. It should be noted that this measures the gain in arbitrary units and not the physical charge collected in every pulse. However, this is sufficient to measure the temperature dependence of the gain and to stabilize the gain at different temperatures, which is what this chapter is mainly concerned with, therefore no further action was taken to translate this measurement into physical units.
This method to extract the gain was used to measure the temperature dependence on the gain k_T of a SensL C-Series SiPM as well as the dependence of the voltage on the gain k_V (see section 3.3). Since the SiPM gain depends linearly on both temperature and voltage, it is easy to determine a correction coefficient k_c, which can be used to correct the gain for temperature fluctuations by

adjusting the applied voltage.[2] The correction coefficient can be calculated as $k_c = \frac{k_T}{k_V}$.

For the measurement of the temperature dependence, the SiPM was held at a constant voltage of 28 V and the temperature was changed in 11 steps from 0 °C to 50 °C. Since it took some time for the temperature in the climate chamber to stabilize, a waiting period of 30 min was taken between setting the temperature to the next value and starting the measurement. Due to a miscalibration, there was a offset of ~ 1.7 K between the set temperature and the measured temperature at the inner Pt100. However, the correct temperature was read out and logged every 30 s, and the offset was the same at every temperature, therefore this did not cause any problems. Once the temperature was set, it was reasonably stable within $\sigma = \pm 0.13$ K. At every set temperature, 350000 traces were recorded. This value comes from a compromise between statistics, file size and measurement time. Due to its design, the *DRS4 Evaluation board* can record ~ 500 traces/s. Since each trace consists of 1024 16-bit numbers, it is at least 2 kB in size. Therefore, around 1 MB are recorded per second. From these traces, the pulses were processed as described in subsection 4.1.2 and the gain was determined as described above. The values of the fitted gains at different temperatures are shown in Figure 4.6. The influence of the temperature on the gain is often described as linear in literature (e.g. [98, 99]) and by the manufacturer [85]. Therefore a fit to a linear model was used to extract the temperature coefficient k_T, although the data in Figure 4.6 deviate slightly from this

$$k_T = (-0.117 \pm 0.003)\,\frac{\%}{\mathrm{K}}. \tag{4.4}$$

The uncertainties on the gain measurement were calculated by the fit of the p.e. peaks to the Gaussians, thus they are only statistical. The uncertainties on the temperature are very small and thus negligible, as can be seen in Figure 4.6, where the errorbars can barely be made out.

A similar measurement can be done for the voltage dependence, where the temperature is constant and the applied voltage is changed. This measurement was done at a temperature of 25 °C, and the voltage was changed from 27.0 V to 30.0 V in 6 steps. Apart from that, the settings used for the measurement

[2] Often in literature, the correction coefficient k_c is given as the temperature dependence of the breakdown voltage. However, to measure the breakdown voltage temperature dependently, measurements for multiple voltages have to be carried out for each temperature, which makes the method equivalent to the one presented here, where the gain measurement is taken for both dependencies.

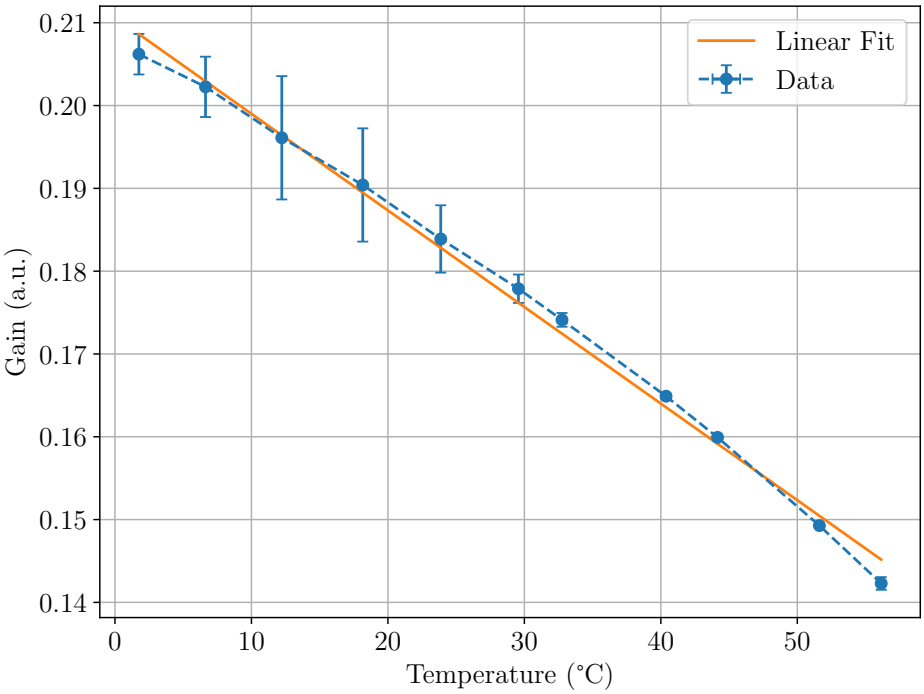


Figure 4.6: Measurement of gain with changing temperature. The errorbars represent the statistical uncertainty from the Gauss fit (see Figure 4.5) and from the temperature measurement.

were the same as describes above. The fitted gain at different applied voltages is shown in Figure 4.7. As in the measurement above, a linear model can be fitted to the data and the voltage coefficient k_V can be determined by the slope of this fit

$$k_V = (6.055 \pm 0.018)\,\frac{\%}{\mathrm{V}}. \tag{4.5}$$

From the two coefficients k_T and k_V, the correction coefficient k_c can be calculated as

$$k_c = \frac{k_T}{k_V} = (19.25 \pm 0.05)\,\frac{\mathrm{mV}}{\mathrm{K}}. \tag{4.6}$$

There is a small tension between with the value given by the manufacturer, who quotes a value of 21.5 $\frac{\mathrm{mV}}{\mathrm{K}}$ [85] and the one measured here. However, other groups [99] report a similar value for SensL SiPM of 18.9 $\frac{\mathrm{mV}}{\mathrm{K}}$. In section 4.2, this correction coefficient k_c will be used to stabilize the gain with changing temperatures.

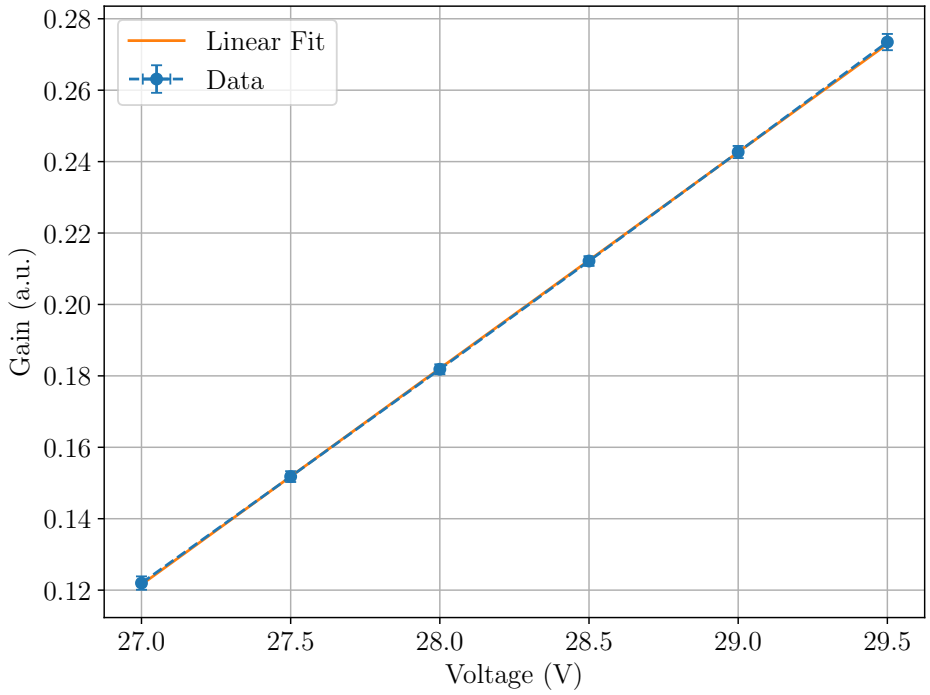


Figure 4.7: Measurement of gain with changing voltage.

4.1.4 Dynamic Histogram

While measuring the gain from the pulse height spectrum is a tried and reliable technique, it relies on static conditions for the duration of a measurement. Changes on SiPM parameters from external factors, e.g. a change in temperature or a slight shift in the baseline contribute to a broadening of the peaks in the pulse height spectrum and therefore to a decreased accuracy in the gain measurement. This can be mitigated by using a shorter duration for each measurement, but the loss in statistics leads to increased uncertainties on the gain measurements.
Here, an alternative approach to measuring gain is presented, which simplifies the creation of pulse height spectra in a continuous way and allows to take "snapshots" of the gain at various times during the measurement, while being less expensive computationally than a least square fit to the peaks of dark count spectra. This is important for live gain measurements, as will be seen below.

Since only the 0 p.e. peak and the 1 p.e. peak in a pulse height spectrum contribute to the gain measurement, a local maximum in a filtered trace (see sub-

section 4.1.2) can either add to the 0 p.e. peak, the 1 p.e. peak or can be discarded for the gain measurement. Therefore, instead of contributing to a histogram, in which the peaks are fitted later, the local maxima can be categorized immediately as either 0 p.e. or 1 p.e. (or higher) and the current gain can be given as the difference between the mean values of the already categorized 0 p.e. and 1 p.e. events

$$G_{\text{current}} = \overline{\mu_1} - \overline{\mu_0}, \tag{4.7}$$

where μ_1 and μ_0 are determined as described below. To measure the gain, the mean values of 0 p.e. pulses and 1 p.e. are known and they can be used to categorize the next set of values, and after a few traces a new current value for the gain is found etc. While this is similar to the creation of a histogram with little statistics and using the gain from the last measurement as a starting value for the fit, the method of "Dynamic Histograms" has the added benefit that the data for the mean values does not have to be thrown away by starting a new histogram. Instead, a new gain value can be calculated from the difference between running means of 0 p.e. and 1 p.e. values and gain can be tracked throughout a measurement. Additionally, changes in gain through temperature, or shifts in the baseline are automatically absorbed in the running means with no additional care needed. In the following, the implementation details used for this work are discussed.

For stability, not only two categories for 0 p.e. and 1 p.e. pulses are used, but also a category for 2 +p.e. pulses, which is needed to separate 1 p.e. pulses from higher ones and a category for outliers below the baseline. This results in four categories for classification of incoming pulses. Upon arrival of a new value from a local maximum of a filtered trace μ_{new}, the new value is put into the "nearest" category $\mathbf{C}$, i.e. the one with the lowest difference between the current running average value and the new value

$$\mathbf{C}(\mu_{\text{new}}) = \mathbf{C}\Big(\arg\min \|\overline{\mu_i} - \mu_{\text{new}}\| \Big). \tag{4.8}$$

This procedure is illustrated in Figure 4.8. Using μ_{new}, a new $\overline{\mu_i}$ is calculated for the respective category. While this could be done with a running mean by saving the last n values and calculating the mean value, in practice it was sufficient to calculate the new μ_i by an interpolation between the old value $\mu_{i,\text{old}}$ and the new value μ_{new} with a certain weight $0 < k < 1$, which is also simpler computationally

$$\mu_i = \mu_{i,\text{old}} \cdot (1 - k) + \mu_{\text{new}} \cdot k. \tag{4.9}$$

Numerically, this gives different values than a running mean with a width of n values, but in both cases, the μ_i converges to a certain new value, albeit

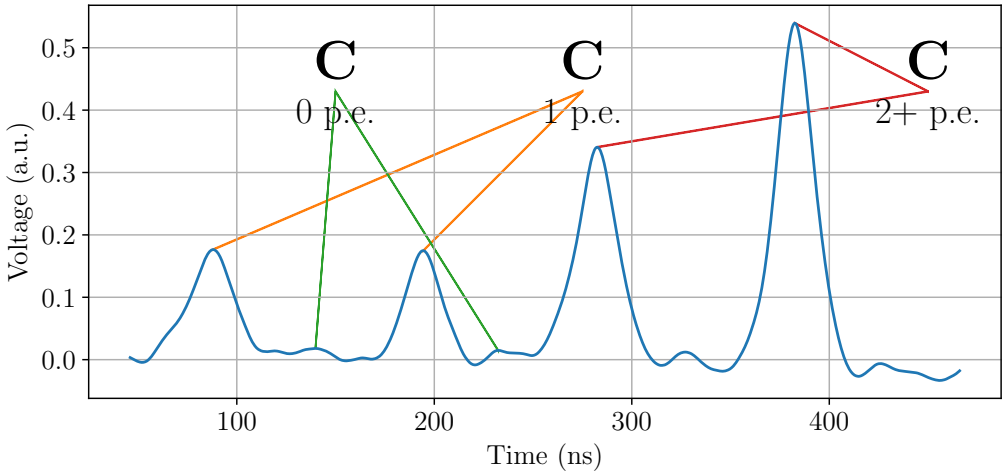


Figure 4.8: Sorting of local maxima into categories. The values are sorted depending on their distance to the current mean value of the categories, as described in Equation 4.8. As an example, the first local maximum at 88 ns is categorized as a 1 p.e. value and is sorted into the 1 p.e. category. There is one additional category for negative outliers, which was not needed in this example trace. Not illustrated are the two local maxima indicated in Figure 4.4, which would be sorted into the 0 p.e. category.

with different rates. The value of k can be tweaked, but in principle $\frac{1}{k}$ behaves like n, where for small values of k, changes in gain are picked up slower, but outliers do not have much of an effect and the values show less fluctuations. For the Dynamic Histogram to work best, a starting value μ_{S0} for the 0 p.e. value and for the gain G_S have to be given. The starting μ_i can then be initialized to the following values:

$$\begin{aligned} \mu_{0\,\text{p.e.}} &= \mu_{S0} \\ \mu_{1\,\text{p.e.}} &= \mu_{S0} + G_S \\ \mu_{2+\text{p.e.}} &= \mu_{S0} + 2 \cdot G_S \\ \mu_{\text{outlier}} &= \mu_{S0} - G_S \end{aligned} \tag{4.10}$$

Since the success of a gain measurement using the Dynamic Histogram hinges on the ability to characterize new pulses, care has to be taken for the "running mean" values μ_i not to drift apart. Therefore, changes in baseline not only change the $\mu_{0\,\text{p.e.}}$ value, but the change is propagated to all other categories as well. This ensures that in particular μ_{outlier} and $\mu_{2\,\text{p.e.}}$ are updated, since they receive the least updates from physical pulses, as high p.e. pulses only happen through random coincidence or optical crosstalk and outliers beyond the baseline ideally do not happen at all. In practice, only the values for 0 p.e.

and 1 p.e. are updated with the interpolation described in Equation 4.9. For the remaining two $\mu_{2\,\mathrm{p.e.}}$ and μ_{outlier}, new values are created at a certain rate according to the currently measured base line $\mu_{0\,\mathrm{p.e.}}$ and gain $\mu_{1\,\mathrm{p.e.}} - \mu_{0\,\mathrm{p.e.}}$ in the same way they were initialized (see Equation 4.10). This solves two problems, as the specific values are updated constantly, even if no events occur, i.e. no outliers happen, in a certain time. Secondly, the spacing between the categories is guaranteed to be optimal for the current gain, and the μ_i do not converge to different values, e.g. from high p.e. incident light.

To find suitable values for the weights k, the Dynamic Histogram was tested on the data used in subsection 4.1.3, in particular the data set used to determine k_T. This is convenient, since the gain values are already known and fluctuations in gain should be minimal throughout a set temperature, as the temperature and voltage were held constant. Here, the ability of the Dynamic Histogram method to accurately measure a known gain can be tested. Furthermore, this data set acts as an extreme test on how well the method handles rapid changes in gain. Since the temperature was changed, but care was taken not to measure until the next temperature was reached, gain in the data changes in an instant. For the Dynamic Histogram method, this is a real challenge, as the 0 p.e. and 1 p.e. values have to be tracked in order not to drift apart and in nature, the temperature changes continuously.

The gain as measured by the Dynamic Histogram method for various weights k can be seen in Figure 4.10. For this, a new measurement of the gain $G = \mu_{1\,\mathrm{p.e.}} - \mu_{1\,\mathrm{p.e.}}$ was taken every 2000 traces for the whole data set. The weights k were varied from $k = 0.5$ to $k = 1 \times 10^{-5}$. As a further test, the starting gain for the Dynamic Histogram was set to $G_S = 0.15$ a.u. to test how well such a misconfiguration would be handled. From Figure 4.10, it can be seen that for high values of k, the gain cannot be measured reliably and the fluctuations are very high, although the general trend of the gain can still be observed. The first value of k, where the steps in gain can be seen, is $k = 0.005$, but the gain measurement is still quite noisy. With decreasing k, clear time constants emerge, which come from the increasing "integration time", since past gain measurements are given more weight. Each temperature plateau lasted ~ 11 min, which means the Dynamic Histogram takes multiple minutes to measure the correct gain for smaller values of k. On the last value $k = 1 \times 10^{-5}$, the steps are barely visible. Since the dark count rate of SiPM increases with temperature, the frequency of 1 p.e. pulses increases as well, which can be seen in the amplitude of the fluctuations or the decrease in time constant. This also explains why for $k = 1 \times 10^{-5}$, the steps emerge again at

higher temperatures. As a compromise between an accurate gain measurement and a reasonably small time constant, a value of $k = 0.0005$ was chosen for the temperature control in section 4.2. The method seems stable enough to handle not only the instantly changing gain steps in most cases, but also the misconfiguration at the start without any problems.

A plot, in which both the gain measurements from the Dynamic Histogram with $k = 0.0005$ as well as the gain measurements from the spectra can be seen, is given in Figure 4.9. Here, a couple of things can be observed. For lower temperatures up to 30 °C, the gain measurements match very well and the Dynamic Histogram is able to pick up the unphysical instantaneous changes in gain without any problems. At higher temperatures, there is some tension between the gain measurements, and the gain measurement starts to become less precise. This can again be attributed to the higher dark count rates at higher temperatures, which essentially decreases the integration time, since there are more 1 p.e. pulses arriving in 2000 traces, after which a new gain snapshot is taken. The optimization to integrate according to the amount of arriving 1 p.e. pulses seems straightforward, but was not carried out in this work.

Overall, the Dynamic Histogram method manages to measure the gain successfully. Even more nuanced features in Figure 4.9 like the rise in gain for the two last plateaus due to a slight temperature dip can be made out, which are lost in the dark count spectrum gain values.

4.2 Gain Stabilization

For measurements with a SiPM, it is desirable to have stable gain, as it is key to good resolution for various applications. Therefore, the gain of a SiPM is often held at a certain level with some kind of stabilization scheme. Here, a standard way to stabilize the gain is discussed, which uses the control coefficient k_c measured above combined with a temperature measurement from an external sensor. Furthermore, the fine-grained gain measurement with the Dynamic Histogram is used to stabilize the gain by itself without external temperature measurements using a PID controller.

4.2.1 Direct Stabilization

One of the easiest ways to stabilize the gain of a SiPM is to measure the temperature at the SiPM and adjust the applied voltage with temperature

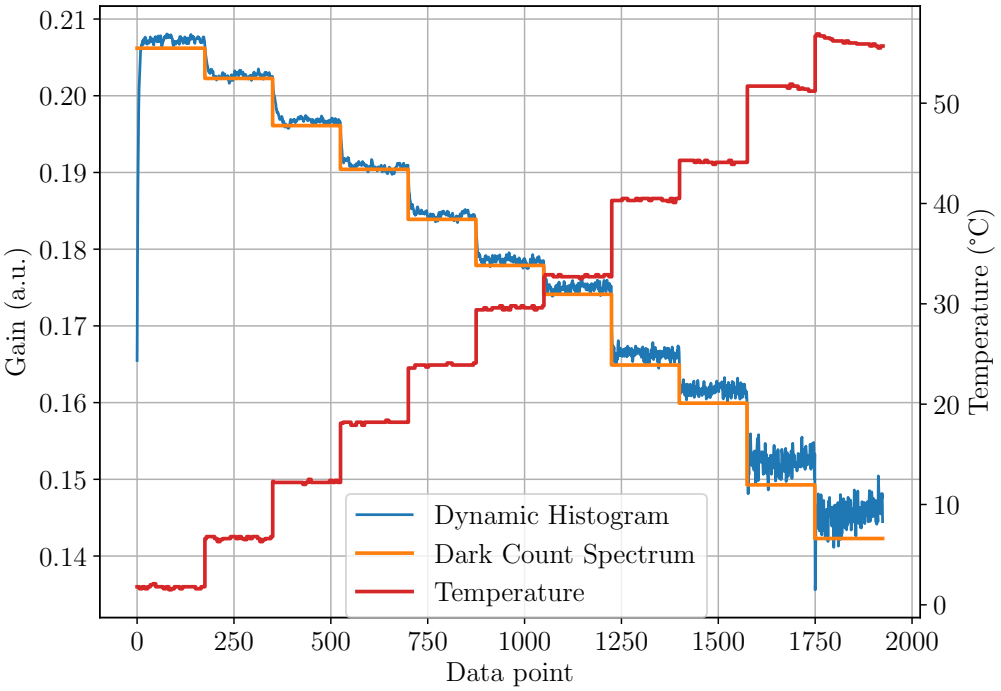


Figure 4.9: Comparison of gain as measured by the Dark Count Spectrum method and Dynamic Histogram method. The red line shows the temperature at each measurement point on the y-axis on the right hand side.

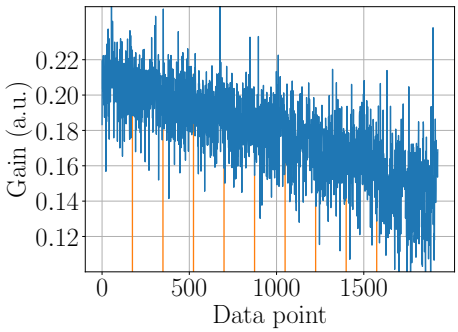
changes according to a correction coefficient. This is the "Direct Stabilization" explored in this section. In the temperature dependant measurements of a SensL C-Series SiPM above, this correction coefficient k_c was determined in Equation 4.6. The stabilization scheme used here works as follows:
At the beginning of a measurement a goal temperature T_g and starting voltage V_S were set. For convenience, the current temperature at the start was chosen. For later measurements, the then-current temperature T_{curr} is read and the change in temperature is calculated

$$\Delta T = T_g - T_{\text{curr}}. \tag{4.11}$$

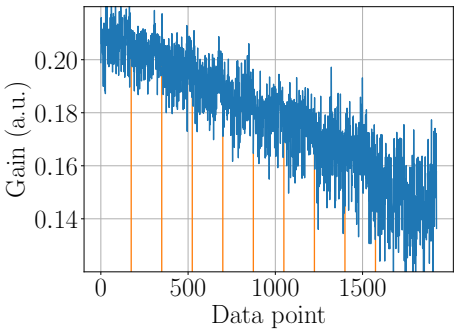
From this, a new voltage was calculated, which gave the SiPM the same gain as in the starting conditions T_g and V_S, according to k_c

$$V_{\text{new}} = V_S + k_c \cdot \Delta T \tag{4.12}$$

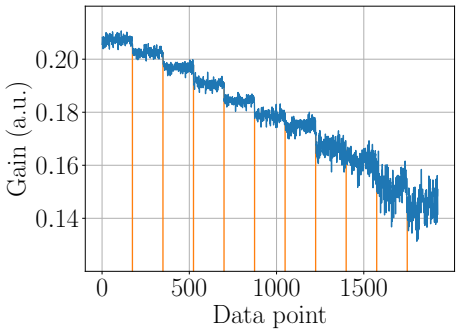
and the new voltage is applied to the SiPM. To shield against incorrect temperature readings, the voltage is set only if the new voltage falls into a safety



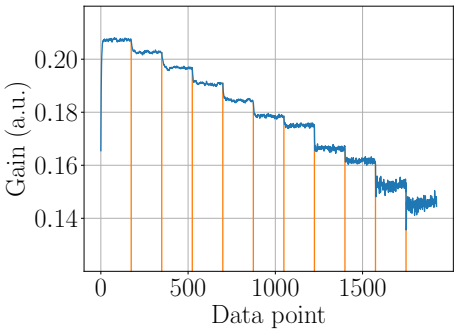
(a) k = 0.5.



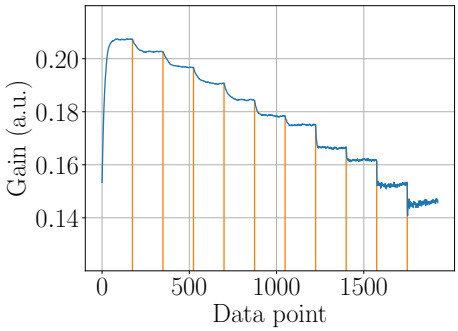
(b) k = 0.1.



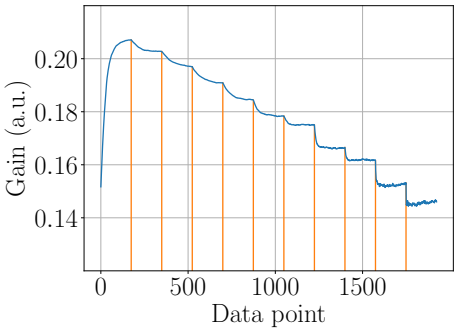
(c) k = 0.005.



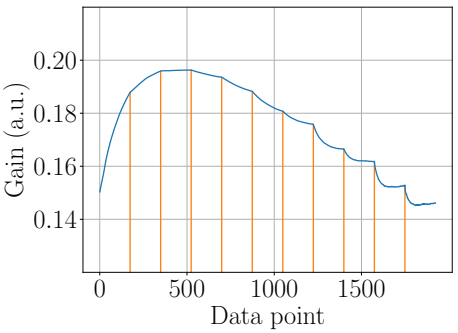
(d) k = 0.0005, weight of choice.



(e) k = 0.0001.



(f) k = 5×10^{-5}.



(g) k = 1×10^{-5}.

Figure 4.10: Gain measured by the Dynamic Histogram method with the data used for Figure 4.6. The orange lines indicate a discontinuity in time, where the temperature was stabilizing for ~ 30 min. Note the different scale on (a).

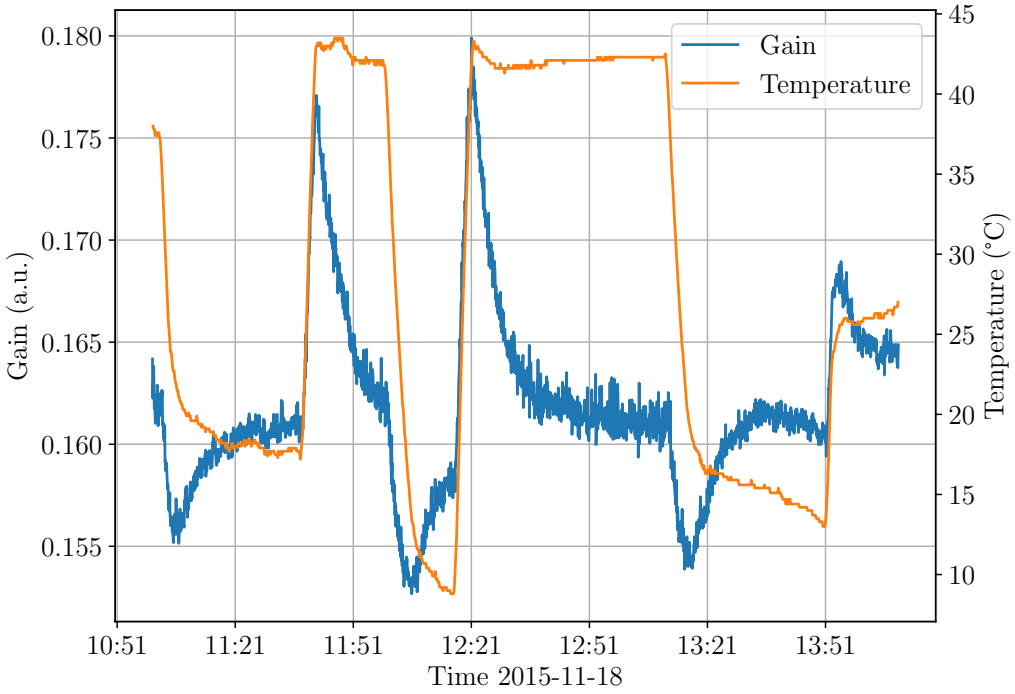
interval $26\,\mathrm{V} < V < 30\,\mathrm{V}$. This procedure is carried out every 10 s.

For the gain stabilization the same experimental setup as above was used, which can be seen schematically in Figure 4.1. Again, only dark count data was considered and the climate chamber was kept in darkness during the stabilization runs. Additionally, the SiPM was located inside a light-tight metal box. Of the two temperature sensors, the one inside the metal box, right next to the SiPM, was used to stabilize the gain. During the stabilization runs, a random trigger was used to record ~ 500 traces/s, temperature readings and applied voltage were logged every 10 s.

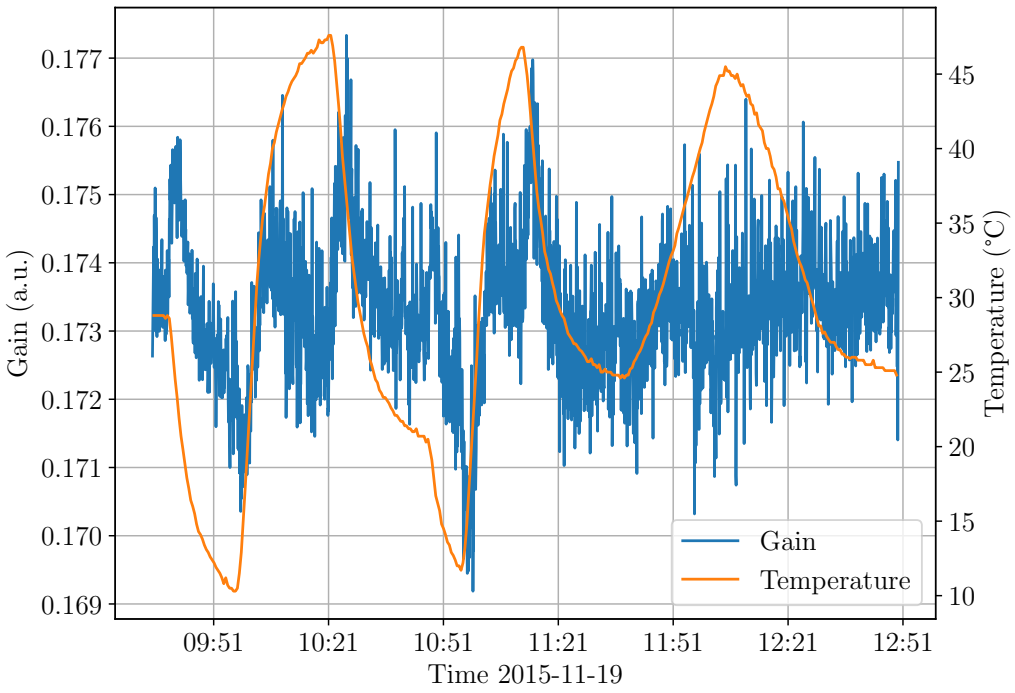
In total, two stabilization runs were performed for the direct stabilization, lasting $\sim 1\,\mathrm{h}$ and $\sim 3\,\mathrm{h}$, respectively. During these runs, the temperature inside the climate chamber was set to values between 50 °C and 10 °C while the gain was kept stable by the scheme described above. Since only temperature readings were used in the stabilization, the gain was measured offline after the run. For the gain measurement, the Dynamic Histogram method was used, since it allows a more fine-grained tracking of the gain and is more robust[3] for the large amount of data than fitting dark count spectra. The gain measurements over the course of the two stabilization runs can be seen in Figure 4.11. These plots show the gain measurement in blue and the temperature readings of the inner temperature sensor in orange. It can be said that the gain stabilization works as intended for periods where the temperature was stable, as the gain returned to its starting value (~ 0.163 in Figure 4.11a), regardless of the absolute temperature. However, during periods, where the temperature changes, the gain does not remain at a constant value and changes as well. This effect is particularly pronounced in Figure 4.11a, and demonstrates one of the main flaws of this method of gain stabilization. For this particular run, the temperature sensor was in contact with the metal box, which has different heat capacity and thermal conductivity than the air, which surrounds the SiPM. Thus the gain is stabilized with a faulty temperature measurement, which results in wrong gain values. In the second run (see Figure 4.11b), care was taken to prevent the contact between temperature sensor and metal box, and thus the gain is much more stable in the presence of temperature changes.

In a real experiment, the problem with temperature readings is not as vital as here, since both the temperature sensor and SiPM would be mounted to the

[3] In general, good starting values for the fit of the peaks in dark count spectra are not available due to shifting baselines and changing gain, which also influences the ratio of peaks. This results in many faulty gain measurements for the large number of spectra fitted.



(a) Direct stabilization run on 2015–11–18.



(b) Direct stabilization run on 2015–11–19.

Figure 4.11: Stabilization runs with direct gain stabilization. On the first run, the temperature sensor was in contact with the metal box, therefore the measured temperature and the SiPM temperature differ, which results in over- or undershooting in the gain. On the second run, care was taken to prevent this contact.

same printed circuit board, and the temperature gradients are arguably not as high as in this experiment, but it is still a source for potential problems in an experiment.

It should be noted that the direct stabilization, as used here, presumes a linear relation between gain and temperature, as does the manufacturer ([85]) and as is commonly used in literature (e.g. [98, 99]). However, the measurements done to determine the correction coefficient k_c (see Figure 4.6) show a slight deviation from this. Furthermore, a commercial voltage supply for SiPM from manufacturer Hamamatsu [100], which is used in chapter 5, uses a second order polynomial instead of a first order line. This means there is still room for improvement for a direct gain stabilization and the implementation shown here is merely typical, to compare with the PID stabilization, which is described in the next section.

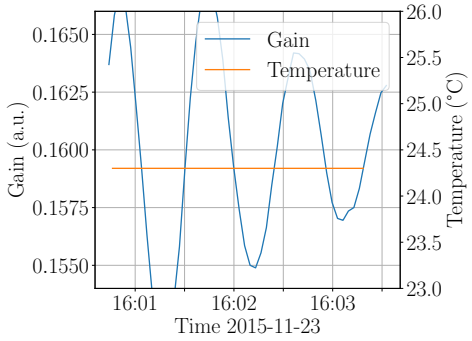
4.2.2 PID Stabilization

So far, the Dynamic Histogram method was used to measure the gain "offline", on data, which was recorded beforehand. However, it is also possible to measure the gain in real time by feeding new traces to the Dynamic Histogram as they are recorded. In order for such a scheme to be feasible, processing of a trace must not take longer than recording the next trace, i.e. event extraction (see subsection 4.1.2) and insertion into the Dynamic Histogram (see subsection 4.1.4) must happen quickly. In the current implementation, this takes $\sim 210\,\mu s$. For comparison, one trace lasts $\sim 500\,ns$, which is considerably faster, but since only $\sim 500\,\text{traces/s}$ can be recorded, real time gain tracking is possible with this setup.

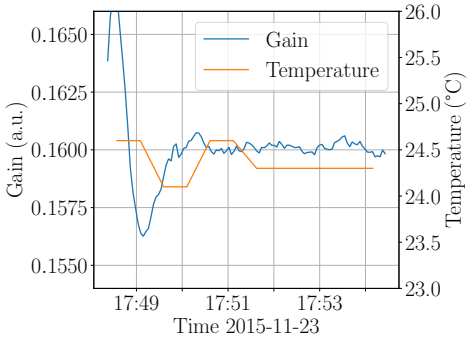
As real time gain measurements are available, one can use the gain as the process variable for the feedback loop instead of the temperature. To stabilize the gain to a certain value G_g, a proportional-integral-derivative controller (PID controller) can be used. As the name suggests, a PID controller uses a weighted sum of three terms to determine the new control value (which in this case is the applied voltage)

$$V(t) = K_p \Delta G(t) + K_i \int_0^t \Delta G(t')dt' + K_d \frac{d\Delta G(t)}{dt}, \tag{4.13}$$

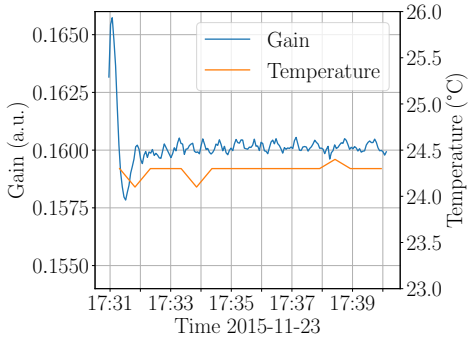
with the weights K_p, K_i and K_d for the proportional term, integral term and derivative term, respectively. The error value $\Delta G(t)$ is equivalent to the change in temperature in Equation 4.11 for the temperature control case and given as



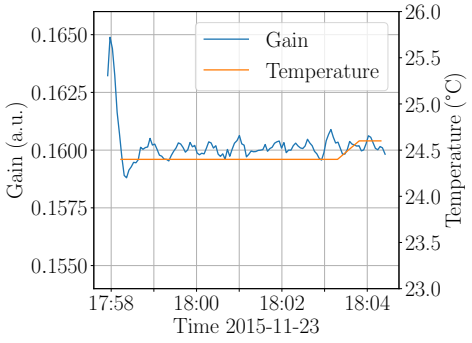
(a) $K_p = 10$, $K_i = 2$, $K_d = 0$, duration: 168 s.



(b) $K_p = 6.5$, $K_i = 0.5$, $K_d = 6.5$, duration: 363 s.



(c) $K_p = 13$, $K_i = 0.5$, $K_d = 10$, duration: 549 s.



(d) $K_p = 18$, $K_i = 0.5$, $K_d = 18$, duration: 389 s.

Figure 4.12: Evolutions of the PID stabilized gain for different PID weights. The jumps in temperature are due to the limited resolution of the temperature readout.

the difference between the set gain G_g and the currently measured gain G_{curr}

$$\Delta G(t) = G_g - G_{\text{curr}}. \tag{4.14}$$

In order to reliably stabilize the gain, values for K_p, K_i and K_d have to be found, which maintain the set gain G_g while minimizing oscillations around G_g and overshooting and still providing a reasonable response time. To find suitable values, a PID scheme was built into the live gain measurement with the Dynamic Histogram. For every 2000 traces, a new error value ΔG was generated, from which a new voltage was calculated using Equation 4.13. Various values for the weights were tested by setting the goal gain G_g to a certain value and observing the evolution of the system. Examples of this can be found in Figure 4.12. In practice, it was difficult to find parameters which provide both a fast response time and show only minimum overshooting. The over-

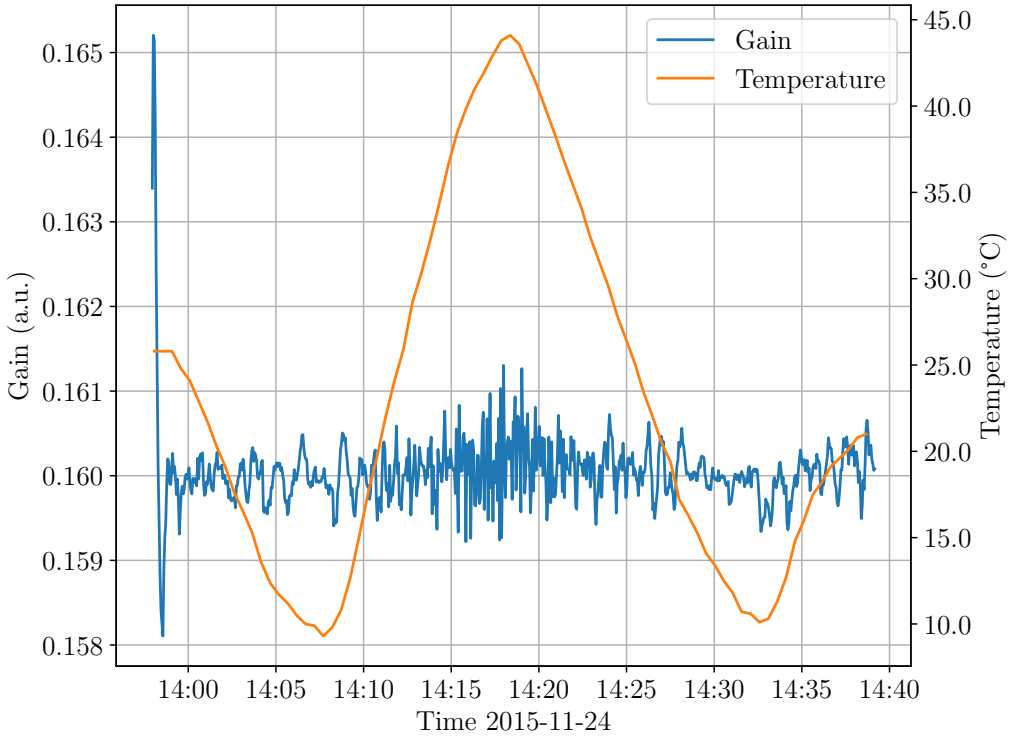


Figure 4.13: Gain stabilization run on 2015–11–24 using the PID stabilization. The initial set voltage was 28.0 V.

shooting is mostly driven by the integral term (compare Figure 4.12a), which is useful to eliminate accumulating errors, that cannot be corrected using the proportional term. For the gain stabilization, this turned out to not be an issue, resulting in a small K_i value, where the overshooting only happens for drastic gain changes. Furthermore, a bigger K_d value prevents overshooting (compare Figure 4.12c, 4.12d), at the expense of less stability for the most likely case, where the temperature does not change much. Since the drastic gain changes in the data are artificial, the overshooting does not happen with continuous data and causes no further problems in the stability. In the end, values between the ones used in Figure 4.12c and Figure 4.12d were chosen $K_p = 13$, $K_i = 0.5$, $K_d = 13$.

Two different stabilization runs were performed for the PID stabilization. In the first one, the typical use case for gain stabilization was tested; holding the gain at a constant value while the surrounding temperature changes. For this test, the goal gain was set to $G_g = 0.16$ a.u and the initial voltage applied to the SiPM was 28.0 V. Over the time of $\sim$ 40 min, the temperature was changed from room temperature to $\sim$ 10 °C and $\sim$ 45 °C, and back to room

temperature in reverse order. The results from this run can be seen in Figure 4.13. Is it apparent that gain stabilization works very well, as the gain stays at the set gain G_g regardless of absolute temperature or temperature gradient. Additionally, the gain fluctuations are smaller in amplitude than their counterparts in the direct gain stabilization case (see Figure 4.11b). To quantify this, the relative standard deviations of both data can be considered, which are $\sigma_{\text{direct}} = 5.9\,\%$ for the direct stabilization, and $\sigma_{\text{PID}} = 1.8\,\%$ for the PID stabilization. Thus the PID stabilization is more accurate by a factor of ~ 3.
Like in Figure 4.9, the time constant for the mean values of $\mu_{0\,\text{p.e.}}$ and $\mu_{1\,\text{p.e.}}$ are temperature dependent, which makes the frequency of gain updates seem temperature dependant, e.g. at the lowest temperatures, there are less 1 p.e. pulses from thermal excitation, thus a single gain measurement with the Dynamic Histogram takes into account less data points than it does for higher temperatures. This makes the measurement at high temperatures seem more volatile. Regardless, the measurement shows no major contribution from this effect on the overall stability of the gain.

The second stabilization run demonstrates the capability of the PID stabilization to set the SiPM to specific gains. To test this, the SiPM was held at a constant temperature of $\sim 26\,^{\circ}\text{C}$ at a initial voltage of 28.0 V. The initial goal gain was set to $G_g = 0.18\,\text{a.u.}$, and throughout the run, it was set to $G_g = 0.15\,\text{a.u.}$ and subsequently $G_g = 0.14\,\text{a.u.}$, while the temperature stayed constant. The results from this run are shown in Figure 4.14. In this plot, it can be seen that the SiPM gain is held stable regardless of G_g, although there is some slight overshooting when settling on a new gain value, which was also present in earlier measurements. It is however noticeable, that the last $G_g = 0.14\,\text{a.u.}$ is not reached exactly, but a gain value slightly below $0.14\,\text{a.u.}$ is held. The reasons for this are not clear, but since the difference between the goal gain and the actual gain remains constant, this behaviour could possibly be fixed with more optimum PID weights. In particular the integration weight K_i could be adjusted to eliminate the residual error.

So far, gain measurements from the Dynamic Histogram were used to show the stability of the PID stabilization. While the gain values used in the plots were produced offline on recorded data, gain measurements from the live Dynamic Histogram were used in the PID controller to stabilize the gain. Thus the quantity, which is stabilized, is also used as a metric for the stabilization. As was shown above, the gain measurement with the Dynamic Histogram matches quite well (and is arguably superior) to the gain measurement done

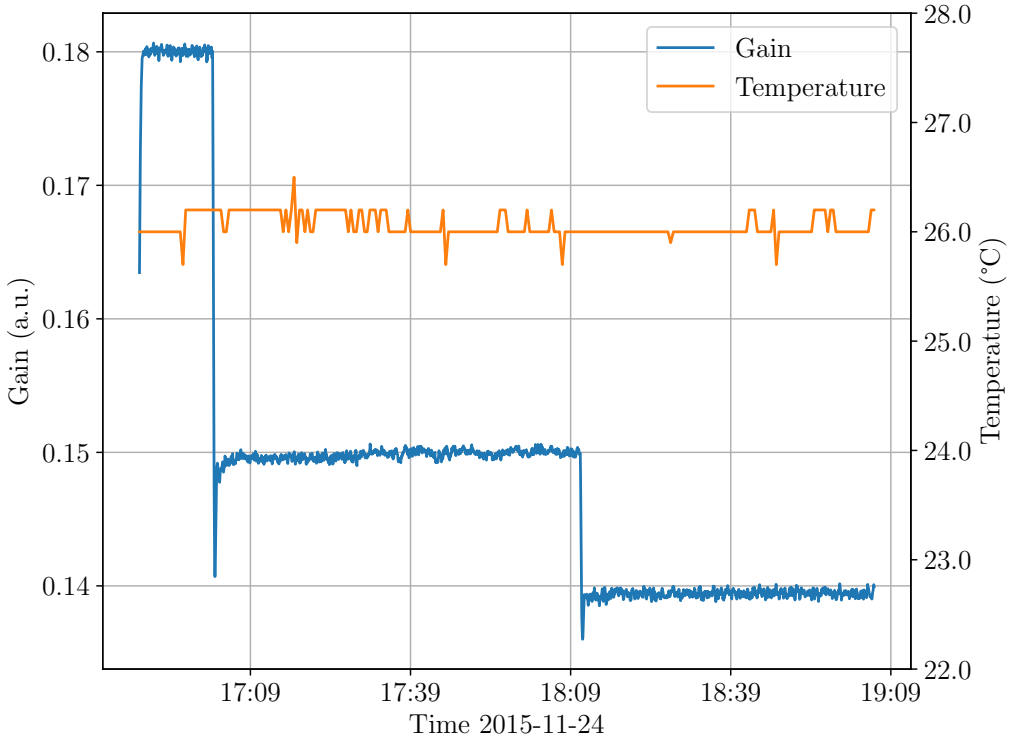


Figure 4.14: Gain stabilization run on 2015–11–24 using the PID stabilization. Throughout the run, the goal gain G_g is set to three different values. The apparent jumps in temperature are due to the limited resolution of the temperature readout.

with dark count spectra, therefore these measurements strongly suggest that the SiPM gain is in fact being stabilized. As a further verification of a successful gain stabilization, spectra with both kinds of gain stabilization and without gain stabilization are compared. As only dark count data were used, it is possible to generate a dark count spectrum as described in subsection 4.1.3 from the data taken in Figure 4.13. Unfortunately, no data was taken with the same temperature curve used in Figure 4.13 with direct stabilization or without gain stabilization at all. However, data of the system for different temperatures without gain stabilization are available from the measurement for Figure 4.6. Thus, it is possible to construct an equivalent spectrum by combining the data from different temperatures in a way, which mimics the temperature curve in Figure 4.13. Additionally, the same temperature range was covered in the direct stabilization run shown in Figure 4.11b, therefore a comparable dark count spectrum is available for the direct stabilization case as well. As the three data sets differ in length, the length of the smallest data set (PID stabilization) was chosen for every spectrum, i.e. 1 500 000 traces, which

correspond to a duration of $\sim 41\,\mathrm{min}$.

The resulting spectra are shown in Figure 4.15. The spectra confirm what is also seen in the Dynamic Histogram gain plots, that the PID stabilization provides a better stabilization than the direct stabilization based on temperature measurements. It is interesting to note that even in the non-stabilized case, there is a clear distinction between 0 p.e. peak and 1 p.e., but the distinction of higher p.e. peaks becomes increasingly hard.

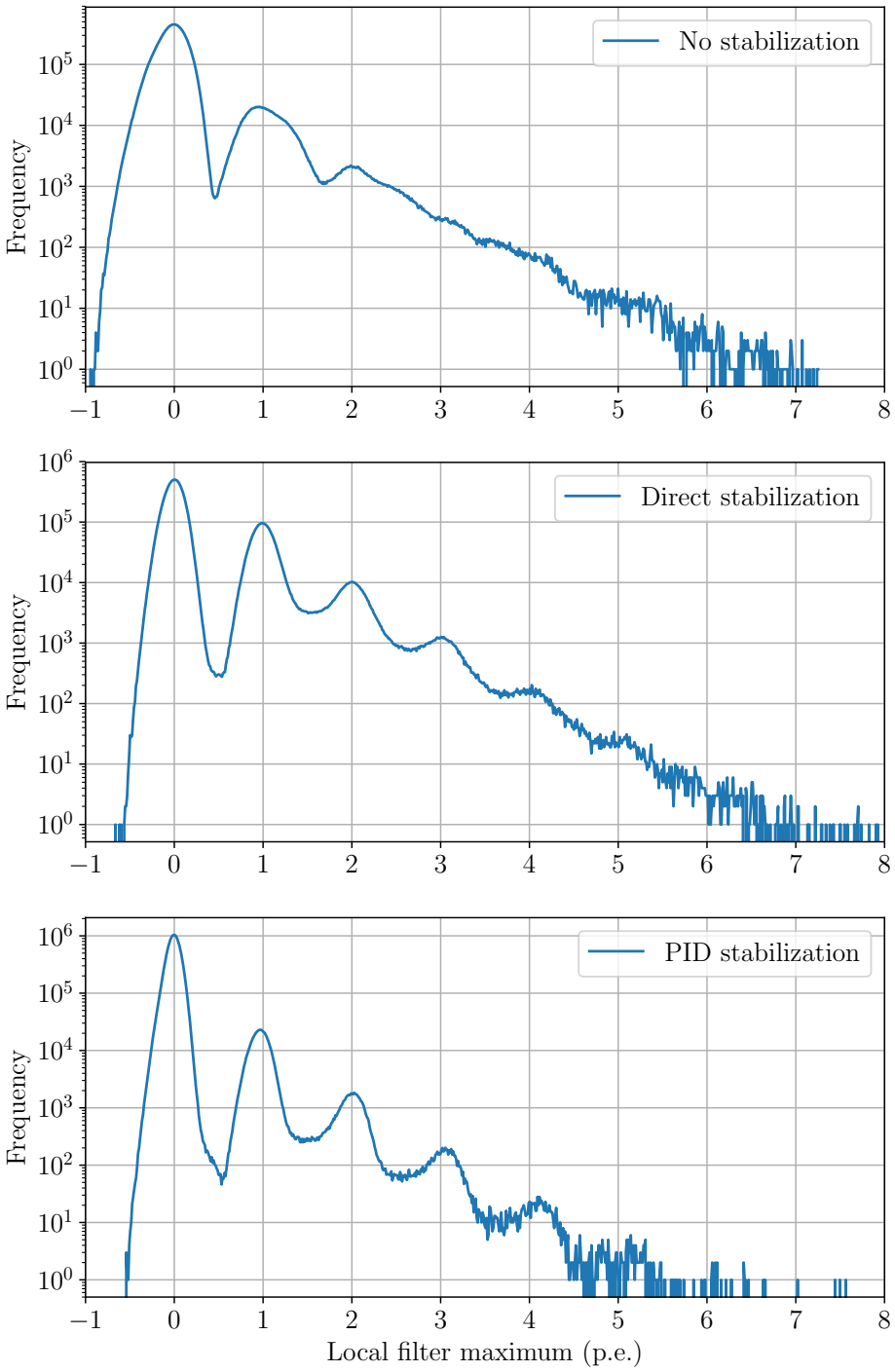


Figure 4.15: Dark count spectra with a temperature curve similar to the one used in Figure 4.13. The PID stabilization produces the best separation between p.e. peaks. The broadening of the peaks comes from different gain due to changing temperatures and is most pronounced at higher p.e. peaks.

4.3 Viability of PID Gain Stabilization

This section aims to explore the viability of the PID gain stabilization as described above in the context of an EUSO-like experiment. Thus, particularly the SiPM Elementary Cell Add-on (SiECA) of the EUSO Super Pressure Balloon mission will be taken as reference. It is thoroughly discussed in chapter 5 below.

In order for the PID gain stabilization to work, two conditions have to be met. First, the raw traces of the single pixels need to be available to categorize the pulses. In SiECA, only a high level view of the data (i.e. trigger counts per pixel) is written out. Secondly, the rate of traces and 1 p.e. signals needs to be high to measure the gain precisely in a short amount of time. As SiECA was triggered externally every 15 s, and only then data was processed, both conditions are not met, thus the method is not applicable to SiECA. Furthermore, the required power consumption for the digital processing of every detector pixel is too high for the relatively small power budget of a balloon flight mission. Nevertheless, one can think of similar experiments, e.g. on a space station, which do not have such a tight power budget and where the Dynamic Histogram method is in principle available. Therefore, the data of SiECA can be used to estimate if the Dynamic Histogram and the PID gain stabilization can work in a space-born experiment.

Apart from the data acquisition, the main difference between SiECA and the measurements in this chapter is that SiECA uses a more recent SiPM array by Hamamatsu, whereas in this chapter, a single pixel SensL C-Series SiPM was used. On a single pixel level, they are comparable as both have the same $3 \times 3\,\mathrm{mm}^2$ dimensions and both are sensitive in the UV, but the newer Hamamatsu SiPM of SiECA has decreased dark count rates. Since the Dynamic Histogram and thus the gain measurement depends on dark counts, lower dark count rates mean more data has to be sampled in order to measure the gain precisely. Thus, the temperature gradients, which can be handled by the gain stabilization, cannot be as steep for lower dark counts rates. Furthermore, the atmosphere at altitudes in which SiECA was operated is considerably colder than the laboratory environment in which the gain stabilization was tested. This further decreases the dark count rate and adds to this effect.

For the SensL SiPM, the dark count rate can be measured with the data used in Figure 4.9 by counting the 1 p.e. peaks, which are categorized by the Dynamic Histogram. This leads to a dark count rate of $\lambda_{\mathrm{dc,\,SensL}} \sim 400\,\mathrm{kHz}$

at 10 °C, the lowest temperature for which the gain stabilization was tested. In EUSO-SPB, SiECA was operated at ~ -30 °C and showed a dark count rate of ~ 20 kHz (see chapter 5), a reduction by a factor of ~ 20. Thus, the temperature gradients, which can be stabilized against, are ~ 20 times less steep than in the experiment above at 10 °C. The steepest gradient, which was tested in the above experiment was a change of 34.6 K in 11 min, which results in an average temperature gradient of $189\,\frac{\mathrm{K}}{\mathrm{h}}$. However, this change was handled without any problems by the PID stabilization using the Dynamic Histogram and steeper changes in temperature were not possible with the climate chamber. Therefore, this gives a conservative estimate of the capabilities of the method. The steepest temperature gradient of SiECA in EUSO-SPB is $< 3\,\frac{\mathrm{K}}{\mathrm{h}}$, as will be seen in the next chapter. With the low dark count rate of SiECA, the PID stabilization is able to handle a temperature gradient of at least $189\,\frac{\mathrm{K}}{\mathrm{h}}/20 \approx 9.5\,\frac{\mathrm{K}}{\mathrm{h}}$, which is far greater than the gradients experienced by SiECA. Thus, the gain stabilization as presented here can work for real experiments, given the necessary DAQ systems.

4.4 Conclusions and Outlook

In this chapter, the change in SiPM gain due to temperature effects was discussed and measured experimentally. Furthermore, means to correct for the change in gain by altering the applied voltage were presented including a novel approach, which does not rely on external temperature measurements.
First, the experimental setup was described, which allows recording dark count data from SiPM while controlling the ambient temperature. With these dark count data, a matched filter was used to extract events from traces by picking the local maxima of a filtered trace. From said events, dark count spectra were created, which allow measurements of the SiPM gain by comparing the position of the 0 p.e. peak and 1 p.e. peak. Using the gain measurement from dark count spectra, the dependency of SiPM gain on temperature and applied voltage were measured, and a correction factor was determined, which allows adjustments of the applied voltage based on a measured temperature difference to keep a stable gain.
Then, a novel approach was presented to construct simplified dark count spectra by keeping track of the positions of the 0 p.e. and 1 p.e. peaks and sorting incoming events into the correct category. This allows for an effective gain measurement in real time. Suitable parameters for the Dynamic Histogram were found and the gain measurements of the new method were validated by comparison to dark count spectra measurements.
In the end, two gain stabilization methods were tested. The first one uses

temperature measurements from an external sensor and the correction factor, which was determined beforehand to stabilize the SiPM gain. This approach is simple to implement and works quite well and can be considered state of the art. In the experiment, care had to be taken to keep the temperature sensor from contact with other materials, especially for steep temperature gradients. In the second scheme, a PID controller was used to stabilize the gain by itself with online gain measurements from the Dynamic Histogram, which enables this novel approach. This worked very well and the system was able to keep a stable gain even for high temperature gradients. To further verify successful gain stabilization, dark count spectra from data using both gain stabilization schemes and no stabilization were compared.

While the PID stabilization with the Dynamic Histogram works well in the laboratory, it has a few drawbacks compared to the direct stabilization. In order to measure the gain online, the traces have to be available, i.e. a fully digital approach to data acquisition is needed. For experiments with a strict power budget, e.g. balloon missions, this is not feasible. Adding to this, the extraction of events from the traces is fairly expensive computationally, which makes it prohibitive for some experiments. However, the implementation used in this work was done purely on the CPU. Significant performance improvements should be possible by using GPU or FPGA for this task.
Furthermore, in the current implementation, a new gain value is generated every 2000 traces. This is not ideal, since it does not take into account the rate of 1 p.e. events, which can become quite low for extreme cooling of the SiPM or bright illumination. Before being used in a real experiment, these effects have to be taken into account and moreover, optimum PID parameters need to be found.

Chapter 5

EUSO-SPB

In this chapter, data recorded by the EUSO super pressure balloon mission (EUSO-SPB) is evaluated, in particular data by the silicon photomultiplier elementary cell add-on (SiECA). After a brief introduction of the experiment, the SiECA data set is explored. This includes the categorization of event types and the development of cuts to distinguish physical from unphysical data. A model of the data is developed and fitted to the data, which leads to estimates of the UV night sky background rates and SiPM dark count rates. Furthermore, transient light events recorded by SiECA are identified. In the end, the SiECA results are briefly compared to the data of the main camera of EUSO-SPB.

5.1 EUSO-SPB

The EUSO super pressure balloon mission (EUSO-SPB) is a pathfinder for the detection of Ultra High Energy Cosmic Rays (UHECR) from space. It was launched on April 24th 2017 from Wanaka, New Zealand under the balloon program of the NASA Columbia Scientific Balloon Facility. EUSO-SPB was the third SPB test flight from Wanaka, with a projected duration of up to 100 days at a floating altitude of $\sim 33\,\mathrm{km}$. Its main scientific goal was the measurement of UHECR from sub-orbital space by observing the fluorescence of nitrogen molecules in the atmosphere [17] (see subsection 2.2.2), and the measurement of background UV light over the ocean and of UV light from clouds. Due to an anomaly in the super pressure balloon, EUSO-SPB could not stay at its floating altitude and the flight was terminated prematurely after 12 days on May 6th 2017. In the following, a brief overview of the EUSO-SPB instrument and its systems will be given. For a more detailed description, the reader is referred to [101, 17] and references therein.

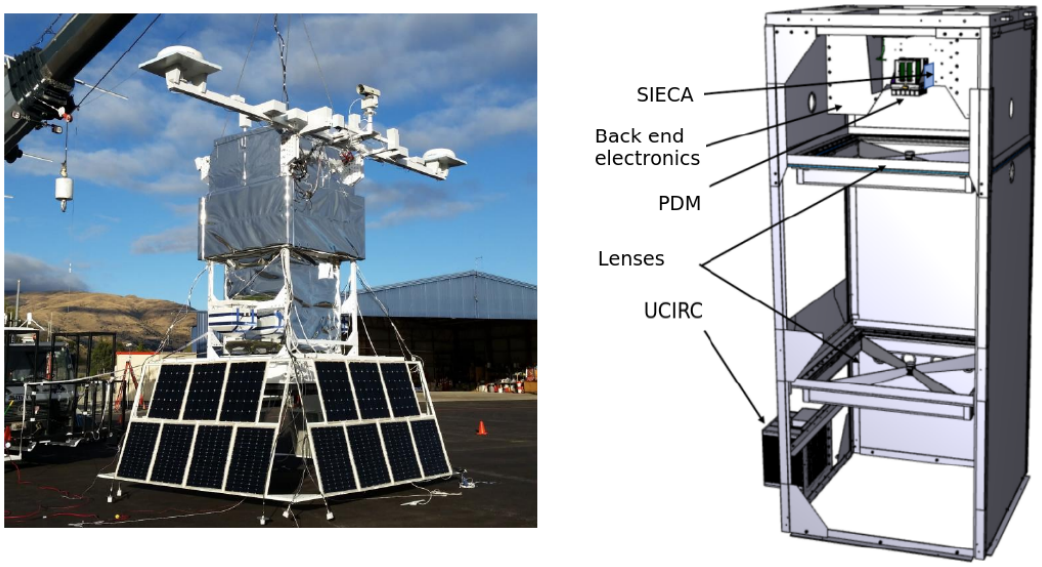


Figure 5.1: Left: Picture of the complete EUSO-SPB instrument. Right: Schematic view of EUSO-SPB instrument. Taken from [101].

5.1.1 Main Instrument

EUSO-SPB is the second balloon flight with a EUSO-like detector, succeeding the EUSO balloon flight from Timmins, Canada in 2014 [102, 103]. As a pathfinder for a JEM-EUSO-like mission, the design of the main camera, the Photo Detection Module (PDM), follows the original JEM-EUSO design [72]. A picture of the complete instrument as well as a schematic is shown in Figure 5.1.

The focal surface of the PDM consists of 36 multi-anode photomultipliers (PMT), with 64 pixels each, resulting in 2304 pixels. The 36 multi anode PMT are grouped into 9 Elementary Cells (EC), with 2×2 multi anode PMT each. Each pixel has a size of $3 \times 3\,\mathrm{mm}^2$. In order to minimize light pollution outside the desired UV waveband, Schott BG3 filters [104] are glued to the multi anode PMT. In addition to the focal surface, the PDM contains ASIC based readout [105], a high voltage power supply and the PDM board, which hosts the FPGA providing the logic for the PDM. In contrast to the first EUSO balloon flight, where the camera was triggered externally, the PDM board is capable of self triggering the readout of the PDM using a trigger algorithm described in [106].

The optics of EUSO consist of two PMMA fresnel lenses with a diameter of 1 m. Before the flight, the optics were calibrated [107] at Colorado School

of Mines. The calibration included measurement of the point spread function with a result of 10 mm. This corresponds to 3 pixels in the PDM. Additionally, the optical throughput of the system was measured, with a resulting efficiency of 30 %. The field of view was measured with laser pulses and by observing stars in the night sky to $(11.10 \pm 0.15)^\circ$. At the nominal flight altitude of 33 km, this corresponds to a field of view on ground of ~ 6 km. The energy threshold for CR from trigger and optics is 10^{18} eV.

To facilitate communication with the instrument and downloading flight data, the gondola was equipped with two Iridium Pilot antennas. In addition to the data runs, a small 40 s portion of the data were downloaded directly after recording to monitor the system. Therefore a UV LED, which fired every 16 s, was mounted near the first lens to illuminate the PDM. Power to the instrument was provided by a battery pack, which was charged during daytime by solar panels, which were mounted alongside the gondola (see Figure 5.1).

Beside the PDM, EUSO-SPB was equipped with two auxiliary cameras. The University of Chicago Infrared Camera (UCIRC) [108] and the silicon photomultiplier elementary cell add-on (SiECA). The purpose of UCIRC was the measurement of clouds and monitoring of atmospheric conditions by observing clouds in two different bands with a field of view larger than the PDM's. Therefore, it was mounted outside the optical path as can be seen in Figure 5.1. However, due to low temperatures, UCIRC was shut off for most of the flight and did not provide much meaningful data.

5.1.2 SiECA

SiECA is a single EC research and development camera to study the use of SiPM as light detectors in a EUSO-like experiment. It is designed to function as an additional EC to the PDM, which increases the focal surface of the detector and shares the optics of the PDM. SiECA uses 2×2 Hamamatsu S13361–3050AS–08 SiPM arrays, an 8×8 array of $3 \times 3\,\mathrm{mm}^2$ pixels. Instead of the normal epoxy window, silicone resin windows are used, which increases the sensitivity in the UV band (see Figure 3.6). To minimize light pollution, the same BG3 filters, which are used for the PDM, are glued to the SiPM pixels. Each SiPM array is voltage biased by two Hamamatsu C11204–02 [100], one for the lower half of the array and one for the upper half, adding up to a total of eight voltage supplies for SiECA. Through four temperature sensors mounted on the PCB next to the SiPM, the voltage supplies also provide gain stabilization (compare subsection 4.2.1). Eight CITIROC ASICs by Weeroc [109] digitize the signals of the four SiPM arrays, again dividing each SiPM into

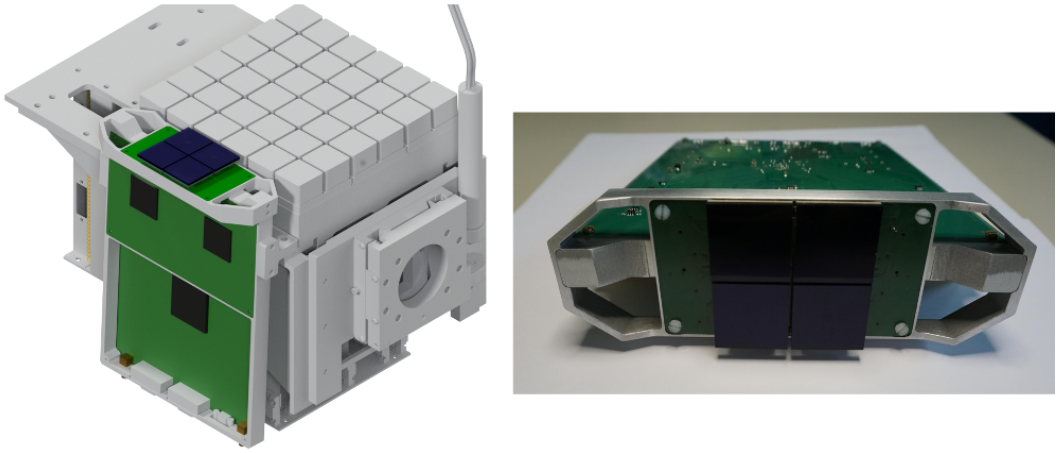


Figure 5.2: Left: Schematic of SiECA mounted next to the PDM. Right: Picture of SiECA. Taken from [18].

an upper and lower half. In order to use the same optics, SiECA is mounted alongside the PDM, which is shown in Figure 5.2.

5.2 Data Quality

As discussed in section 5.1 the main scientific goals of the EUSO-SPB mission were the measurement of UHECR and the UV background over water and clouds, with the SiECA prototype being an ancillary research and development test. During the flight, complications concerning electrical interfere emerged between SiECA and the PDM, which led to instability in the PDM whenever SiECA was turned on. Because of this, after the anomaly in the super pressure balloon was detected, taking data with the PDM was prioritized to reach the scientific goals of the mission. Thus, only few data were recorded by SiECA in comparison to the PDM. This section will give an overview of the data and discuss the SiECA data for use in reaching the EUSO-SPB scientific goals.

5.2.1 Data Overview

Due to its design as an add-on to the PDM, SiECA does not have any trigger logic on its own. In its intended operation, the triggers of the PDM would be passed down to SiECA which allows for synchronous triggering and the use of SiECA as an extension of the PDM. Due to technical problems during commissioning and obvious time constraints, this scheme could not be implemented and instead a random trigger scheme was used for SiECA. In this scheme, SiECA was triggered externally every 15 s from the data processor stack of EUSO-SPB. However, since it was designed with a synchronous

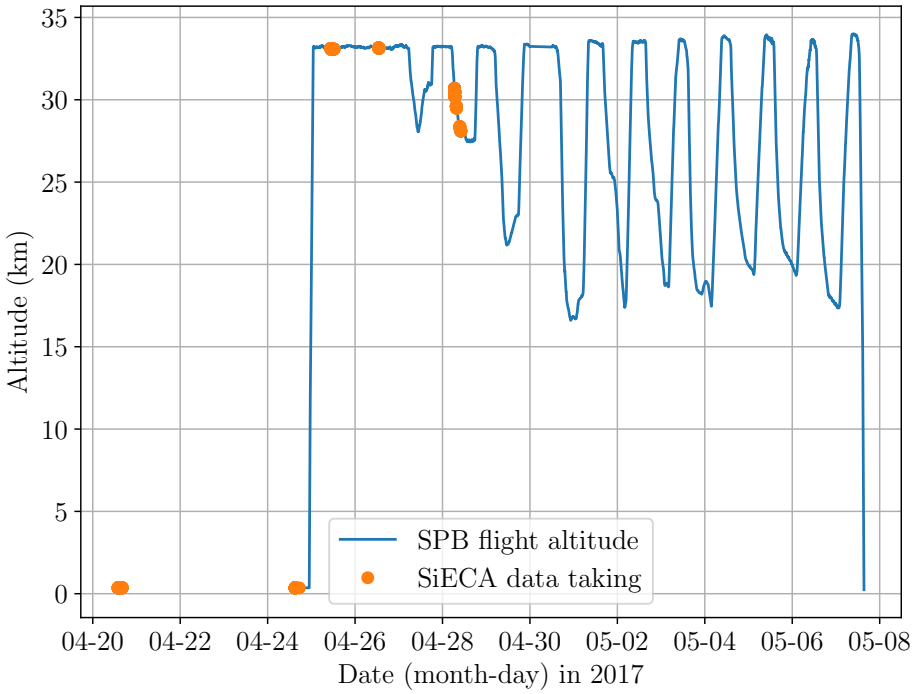


Figure 5.3: Flight curve of EUSO-SPB showing the flight altitude over time. The anomaly with the balloon can be seen in the steep drops during night time. Orange dots indicate the time when SiECA data was recorded.

trigger in mind, no precision clock was provided, which means the data was timestamped by the CPU of the data processor, with a precision of only 1 s.

Each trigger produced a trace, which consists of 128 data points with each data point lasting 1 Gate Time Unit (GTU, 2.5 µs). This amounts to a total length of 320 µs per trace. During the time of the balloon flight 284 traces were recorded by SiECA, thus only $\sim$ 90 ms of flight-data is available. In addition to this, there are 94 traces from on-site tests in Wanaka, but most of them are empty or do not contain valid data. See subsection 5.2.4 for a more detailed description.

After each trigger, the recorded data is transmitted from SiECA to the data processor of EUSO-SPB and can later be downloaded along with the data from other systems. During the flight, two Iridium [101] satellite antennas were used to transmit the data from the balloon down to earth. For the analyses, the downloaded raw binary data was converted with a script supplied with the data to an easily explorable ROOT [110] tree file format. In this format, each

Table 5.1: Overview of completeness values in SiECA data. For any incomplete traces, only the last four GTUs are missing.

	incomplete	complete
number of traces	231	147
number of traces in-flight	176	108
	missing	recorded
number of GTUs	924	47460
number of GTUs in-flight	704	35648

trace contains count data for each of the 256 pixels and 128 GTUs. In addition to the pixel data, there are two types of metadata: a Unix timestamp (with a precision of 1 s) and the completeness of the trace. Due to the nature of the download and the priorities shifting to UHECR detection with the PDM, not all traces contain the full 128 GTUs. The completeness value denotes how many GTUs were downloaded for each trace. In practice this does not present as much of a problem as may seem, because there are only two types of completeness values in the data: 124 and 128. A value of 128 means the trace is complete, 124 means the last four bins are missing. Since traces miss at most the last four bins, all traces are used for the analyses, although for incomplete traces, the missing four bins (which still contain some arbitrary data in the *.root* files) are discarded. An overview of the completeness of the data can be seen in Table 5.1.

In addition to the pixelated photo-detectors, there were multiple auxiliary systems onboard. In particular, NASA provided a differential GPS receiver called Compass, which was able to precisely measure time, position and orientation of the balloon. Using this data, the flight curve of the balloon is plotted in Figure 5.3. It can be seen that there are two records of data at the target altitude of ~ 33 km, as well as a few records in the fourth night, on April 28th. Because of interference between SiECA and the PDM, no more data was taken after the anomaly became apparent. A thorough look at the records can be found in subsection 5.2.4.

5.2.2 Trigger Counting

To understand the images SiECA produces, it is important to understand how the data is generated. SiPM signals in SiECA are digitized in the Citiroc ASIC by Weeroc [109]. In this ASIC the raw SiPM signals are amplified, and if some

trigger condition is met, digital signals are produced to indicate which channels triggered in a set period of time. The Citiroc ASIC uses one of two different amplifiers, one with high gain, one with lower gain, to allow for a higher dynamic range. For SiECA only the high gain amplifier is used. Whenever an amplified signals exceeds a set threshold, a digital logic signal is emitted until the signal falls below the threshold again, which happens typically in $\sim 10\,\mathrm{ns}$ [98]. For each channel, the amount of triggers are counted within 1 GTU, and this count is then written into a buffer to be read out. While this approach can lead to an underestimation of the real photon count e.g. when multiple photons hit one pixel within the time window of 10 ns, for low counts this gives a good approximation of the real photon rate. Since SiECA is simply counting triggers, the probability of missing a photon can be estimated using Poisson statistics as

$$p_{\mathrm{random}} = 1 - e^{-\Delta T \cdot R}\,, \tag{5.1}$$

where ΔT is the length of the coincidence time window and R is the rate of incoming photons (over threshold). For 10 counts per GTU and a coincidence time window of 10 ns, this yields a coincidence probability of 0.04 %, which can be neglected. As will be seen in subsection 5.2.5, 10 counts per GTU is a very conservative estimate, therefore even less photons are missed due to such coincidence.

The reason for the low count rates is the way SiECA was calibrated before the flight. Since SiPM have a (compared to PMT) high dark count rate of $\sim 1\,\mathrm{MHz\,mm^{-2}}$ at room temperature, triggering on a single p.e. can lead to pile-up and thus an overestimation of the recorded light, as triggers would be dominated by dark counts. To avoid this issue, the amount of triggered events in 1 GTU has to be lowered. While this could be achieved by applying a lower supply voltage to the detector, gain and PDE would be sacrificed, which both depend on the applied voltage, see Figure 4.6, [87]. To keep the trigger rate manageable while still having high gain and PDE, the trigger threshold of the detector was increased. Since the dark counts happen only at the 1 p.e.-level and higher p.e. dark counts only happen because of coincidence as described in Equation 5.1 (neglecting a few % from optical crosstalk), increasing the trigger threshold drastically lowers the amount of triggers due to dark counts [111]. This decrease can be described by a Borel-distribution [112]. To find the optimal configuration for the flight, various combinations of thresholds and gains were tested in the Single Photon Calibration Stand at KIT [69] by the Karlsruhe group. For each setting, the *excess photons* were determined, which are defined as the difference between the trigger count in one illuminated pixel and the average trigger count of all other dark pixels for one trace (128 GTUs). By maximizing the *excess photons*, optimal settings were found and one uniform

threshold was set for the whole detector. More details can be found in [18]. This works because the pixels were flat fielded to show uniform gain across the detector and the gain was stabilized against temperature changes as described in subsection 5.1.2. The temperature stabilization was provided by the bias voltage supplies (Hamamatsu C11204–02 [100]), relying on four temperature sensors on the detector surface. The threshold in the actual flight configuration is not known in absolute units of p.e. A method to estimate the threshold using the flight data is presented in subsection 5.3.6, which yields a threshold of $\sim 1.5\,\mathrm{p.e.}$

5.2.3 Camera Images

To convert the flat 256 pixel array from the *.root* data files to the pixel positions on the detector, a pixel mapping is required. In this thesis, a mapping from [113] was used, which was supplied alongside the raw data. This mapping places the observer behind the camera looking through the optics down to earth. A visualization of the mapping can be seen in Figure 5.4, and an example image from the camera is shown in Figure 5.5. In this chapter, the coordinates of rows and columns in the camera images will be denoted by an y- and x-coordinate, respectively.

In this pixel mapping, it is easy to see how the detector is constructed: The whole detector can be divided into four parts, one for each 64 pixel SiPM array. Each SiPM array is itself vertically divided into two parts. This division comes from the fact that the Citiroc ASIC only has 32 input channels and therefore each SiPM array is read out by 2 ASICs. The same division is also found for the bias voltage supplies, with each voltage supply serving the same 32 pixels as are processed by one ASIC. The example image in Figure 5.5 shows one trace containing 128 GTUs. Since the count rates are typically low (see subsection 5.2.5), and each pixel contains at the most 1–2 counts per GTU, count rates are averaged over the whole trace, which lasts $128 \times 2.5\,\mu\mathrm{s} = 320\,\mu\mathrm{s}$ and is generated from one trigger. In Figure 5.5 the mean count rate per GTU and pixel is shown. It can immediately be seen that each pixel of the detector worked. However, the trigger rate on the top right of the detector is evidently lower than for the remaining pixels. This feature is present in all the data and can be traced back to the division of each detector into two regions. For this particular region, either the trigger threshold was higher than for the remaining pixels, therefore yielding less triggers or the gain was lower, with an unchanged trigger threshold. This particular event shows a cloud moving through the field of view of the detector, which is further discussed in section 5.4.

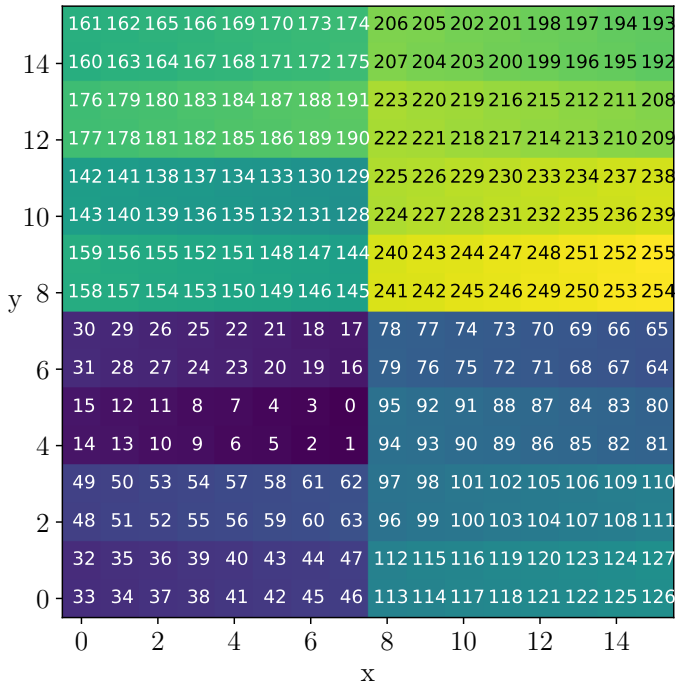


Figure 5.4: Pixel mapping used in this thesis, starting at (7, 5) for pixel # 0. The color gradient follows the increasing channel numbers. The division of the detector into 4 SiPM arrays can easily be seen from this. Each SiPM array is furthermore divided vertically, e.g. for the bottom left SiPM array, the divisions are from 0 to 31 and from 32 to 63.

5.2.4 Voltage Biased Data

As can be seen in Figure 5.3, SiECA data is spread across a few time periods of data taking with pauses in between, which will be called "records" in the following. In this section, the data quality of the various records will be analysed and it will be discussed which record can be considered for the analysis. From the flight curve seven records of data can be made out, two before the start of the flight, and during the flight one on April 25th, one on April 26th and three on April 28th. To evaluate the data, the spectrum of each record is constructed by averaging the counts of all pixels for each GTU in the record. From these spectra, it can be inferred if the detector was turned on. The spectra for the different records are shown in Figure 5.6. It can be seen that each record shows events between 5 and 10 counts, which will later be classified as "spikes". They appear even if the detector was turned off. Apart from that, only three records show any data beside the first bin: The one on April 25th as well as

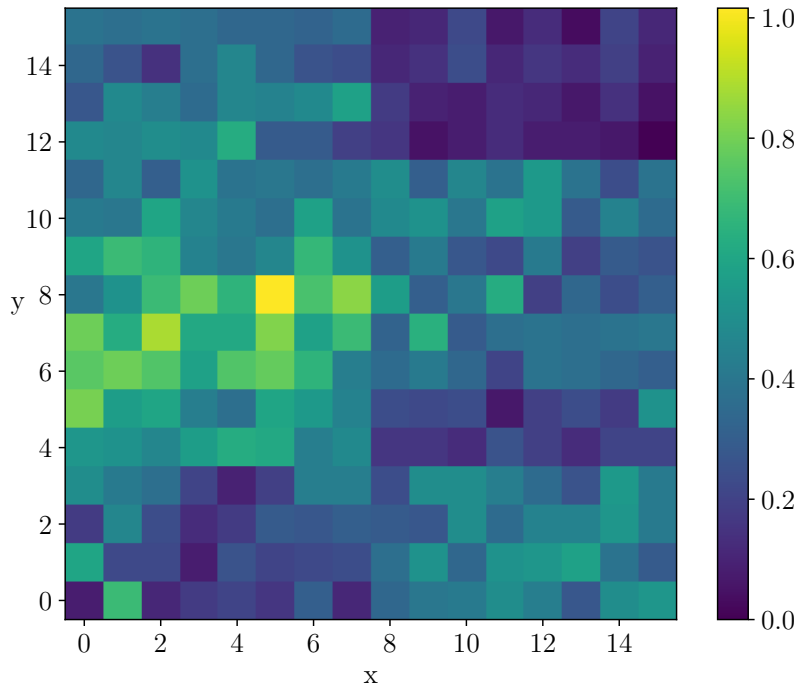
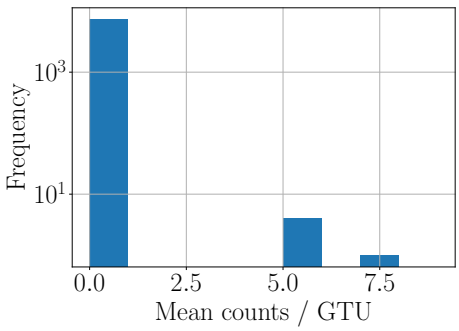
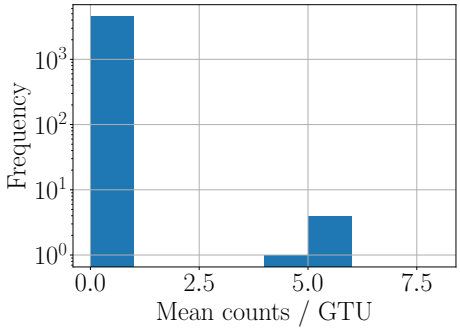


Figure 5.5: Example camera image from SiECA, each pixel is averaged over one trace, i.e. 128 GTUs. Each averaged pixel trigger rate is given in $\frac{\text{counts}}{\text{GTU}}$. This image shows a cloud moving through the field of view on 2017–04–28 07:38:32.

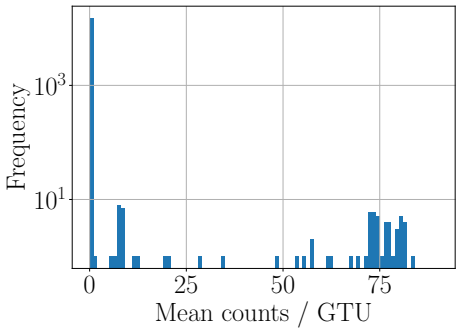
the last two records on April 28th. When magnifying the first bin of these spectra (see Figure 5.7), it becomes apparent that only for those three records data was recorded, the other records show a flat zero, except for spikes. This yields three records of voltage biased data on two days: the 25th and the 28th. Since the timestamps of each event are known, a simple timestamp cut on the data can be used in order to filter voltage biased data. This will be called the unbiased data cut in the following. From the 47460 complete GTUs, 21604 can be considered valid data. Table 5.2 gives an overview over the records.



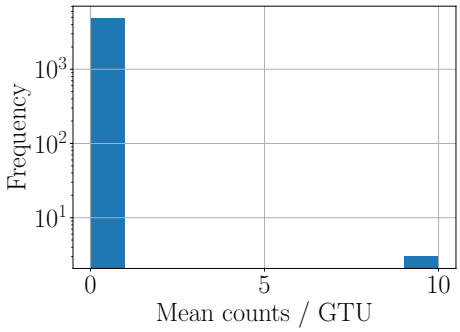
(a) Record April 20th.



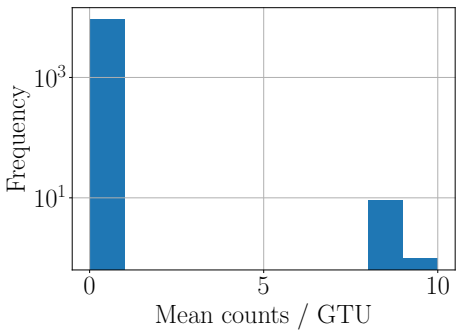
(b) Record April 24th.



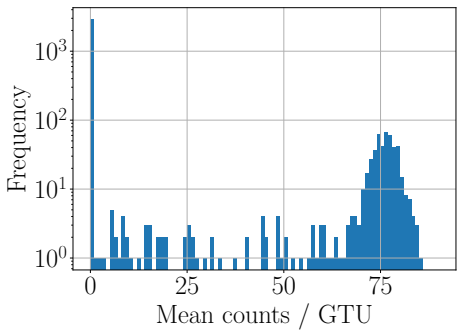
(c) Record April 25th.



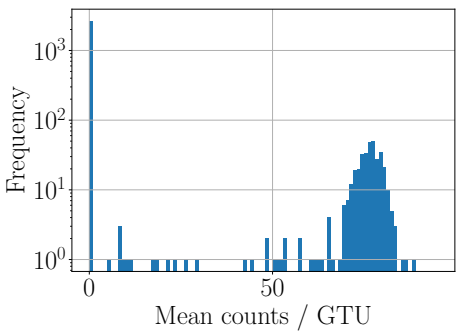
(d) Record April 26th.



(e) Record April 28th 1.

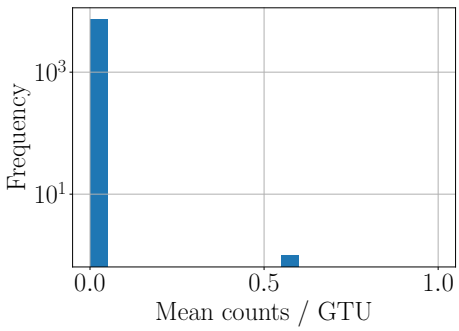


(f) Record April 28th 2.

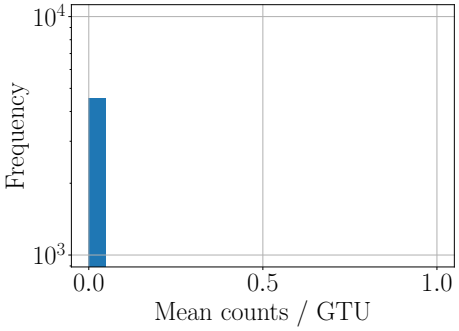


(g) Record April 28th 3.

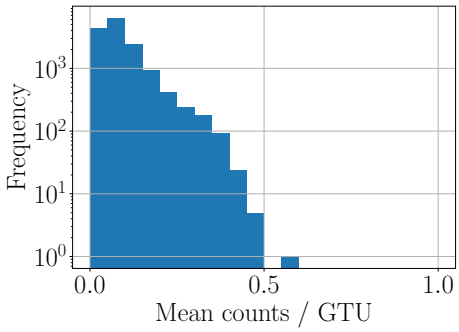
Figure 5.6: Averaged spectra of the 7 records of data.



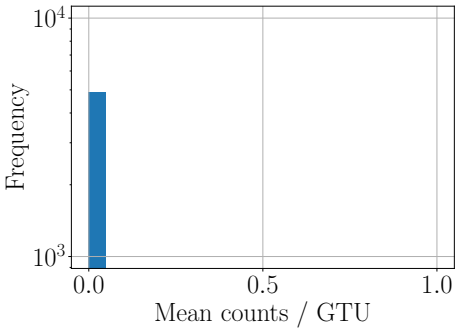
(a) Record April 20th.



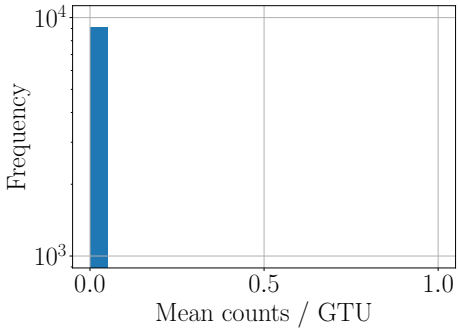
(b) Record April 24th.



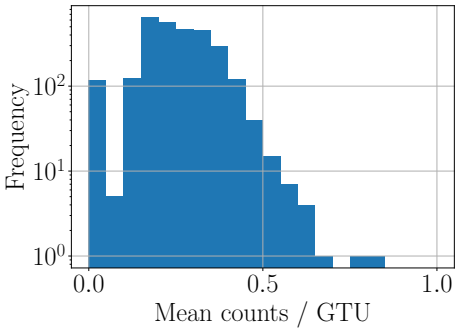
(c) Record April 25th.



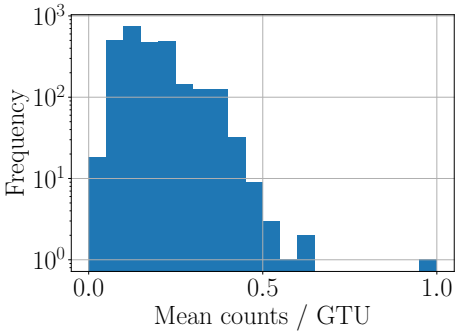
(d) Record April 26th.



(e) Record April 28th 1.



(f) Record April 28th 2.



(g) Record April 28th 3.

Figure 5.7: Magnification of the first bin of Figure 5.6. Only the records April 25th, 28th 2 and 28th 3 contain physical data.

Table 5.2: Overview of voltage biased GTUs in SiECA data.

Record	GTUs	is voltage biased
April 20th 2017	7284	no
April 24th 2017	4528	no
April 25th 2017	15188	yes
April 26th 2017	9156	no
April 28th 2017 1	9156	no
April 28th 2017 2	3408	yes
April 28th 2017 3	3008	yes

5.2.5 Event Classification

Up until now, two different cuts were used, the completeness cut described in subsection 5.2.1, which works on GTU level and removes the GTUs, which were not written completely (this only affects the last four bins of any trace). The second cut was the unbiased data cut described in subsection 5.2.4, which works on a trace level and removes any trace, which was recorded during a time when the detector was not voltage biased. In addition to these two cuts, there are several more cuts needed to remove unphysical data from the data set for the analysis. All the following cuts will work on the GTU level and only remove the GTUs in question, while the remaining GTUs of the trace remain in the data set. An overview of all cuts is shown in Figure 5.8. In this histogram, the average counts of all pixels in one GTU are plotted. The colors indicate the remaining data after a certain cut. From this histogram alone, data can be distinguished into four broad categories:

- **Very low count rates in the first bin**. Due to the relatively high trigger threshold, the first bin includes background counts from the ultraviolet night sky background as well as dark counts from the SiPM themselves. This data will be discussed thoroughly in section 5.3.

- There is a **clustering of GTUs between 5 and 9 average counts**. These GTUs are called "spikes", which are an anomaly in the detector with a very distinct signature both in time and in the images they produce. These items will be discussed in subsection 5.2.6.

- The **peak at the highest counts** is one of the most obvious features in the histogram, as well as the sharp cutoff after this peak. These GTUs are called "pattern GTUs", because they produce a distinct image pattern, but unlike the spikes happen at seemingly random times. Both the high counts and the cutoff indicate a saturation in the detector. These events will be discussed in subsection 5.2.7.

- **The continuum between the spikes and the pattern GTUs** is the last category in the histogram. These are either pattern-like GTUs where the pattern occurred only a fraction 1 GTU, resulting in a lower count rate than the usual 80 counts, or events where single pixels exhibit unusually high counts.

As will be seen below, physical events predominantly appear in the first bin of Figure 5.8. In the following, the unphysical events will be categorized, and cuts will be defined to distinguish the unphysical data from the physical data.

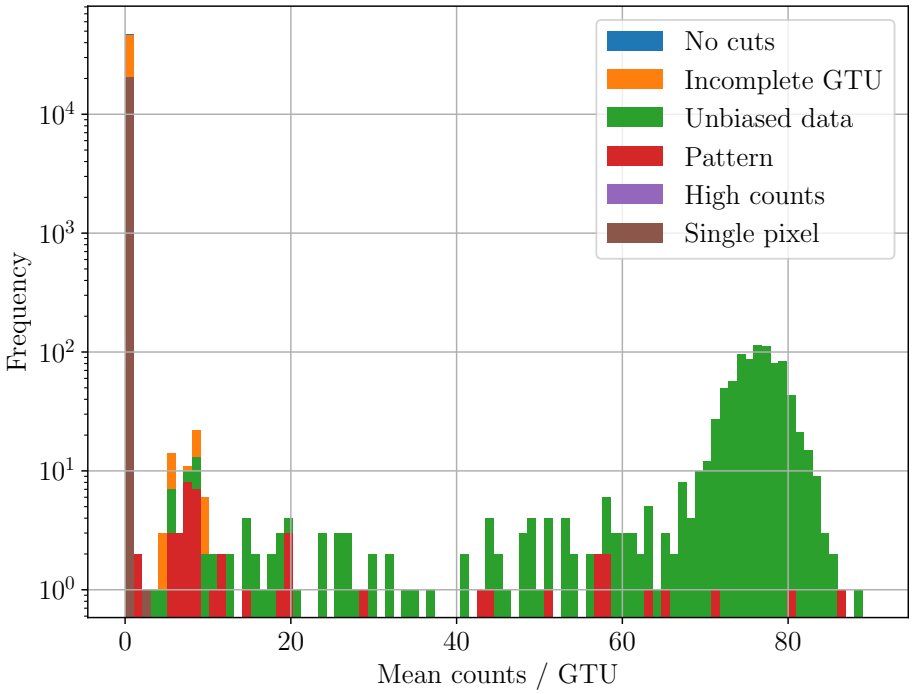
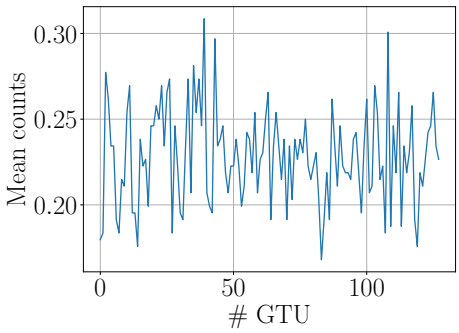


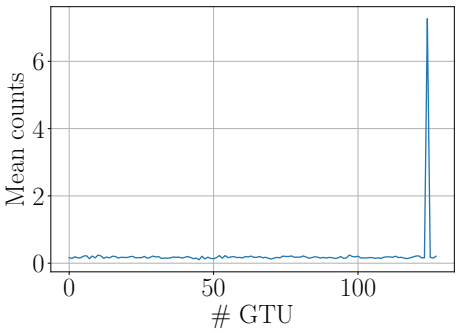
Figure 5.8: Average counts in SiECA per GTU shown for a bin width of 1. The colored area indicates the remaining data after a certain cut. Almost all physical data lies in the first bin. The incomplete GTU cut and unbiased data cut only remove some data in the first bin. The majority of the spectrum does not survive the pattern cut, the remaining data with higher mean counts are removed by the high count cut.

5.2.6 Spikes

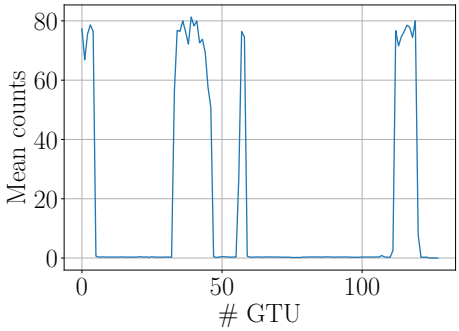
Spike GTUs can easily be recognized when looking at timelines of traces as can be seen in Figure 5.9. In these timelines the average counts of all pixels per GTU are plotted over time. For spikes, the 124th GTU shows between 5 and 9 counts while the neighbouring GTUs have much lower count rates, see Figure 5.9b. When looking at the camera images of such GTUs, spikes produce a distinct pattern, which can be seen in Figure 5.10. It is worth noting that while the average count for the GTUs is only around 8 $\frac{\text{cts}}{\text{GTU}}$, single pixels which contribute to the pattern exhibit up to 200 counts in one GTU. Since this pattern only shows up in the 124th GTU, a connection to incomplete traces seems plausible, however the spikes happen in both cases, for complete as well as incomplete traces. The reasons for the spike events are unclear, but the pattern hints at malfunctions in the electronics: When comparing the channel numbers



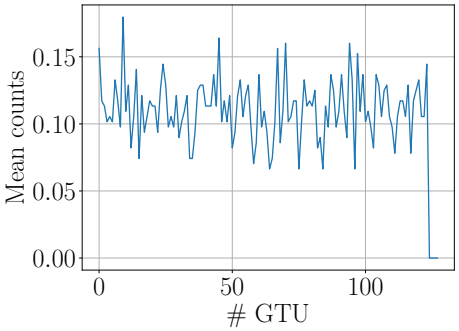
(a) Normal event which will pass all cuts.



(b) Spike event with characteristic spike on the 124th GTU.



(c) Timeline containing four pattern events which produce the highest average counts in SiECA data.



(d) Incomplete trace where the last 4 bins are missing and filled with zeroes.

Figure 5.9: Timelines for different types of events in SiECA data. Mean pixel count for the detector is plotted over corresponding GTU.

from the used channel mapping (see Figure 5.4) of pixels involved, it can be seen that the pattern is made up of two spots of consecutive pixel numbers, from 14–19 and from 141–147. This suggests some form of electrical crosstalk in the wiring of those pixels.

For the actual analyses, these pixels do not pose a problem, since the high count rates of up to 200 counts per GTU only appear in the spikes events and these GTUs can therefore be cut from the analysis data. For this cut, a GTU is classified as a spike if it occurs on the 124th GTU and the maximum count of a single pixel exceeds 40.

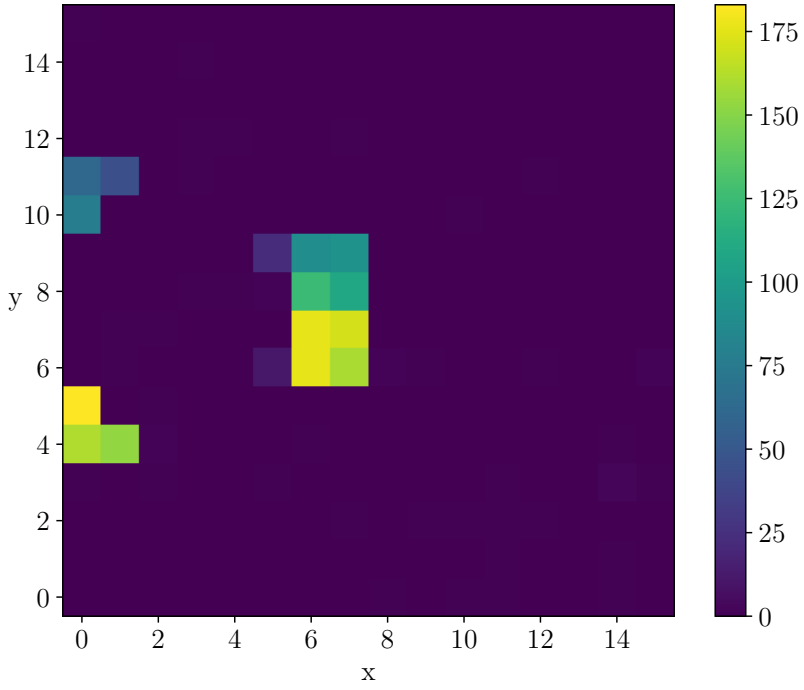
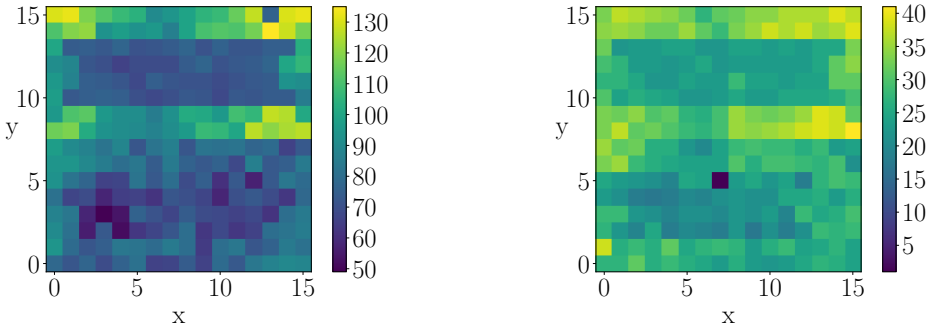


Figure 5.10: Camera image showing spike event on 124th GTU. These events produce the highest count rates in single pixels up to 200 $\frac{\text{cts}}{\text{GTU}}$.



(a) Example of a pattern GTU @ 2017–04–28 07:34:18. Match score: 0.92, Mean counts / GTU: 83.52 $\frac{\text{cts}}{\text{GTU}}$.

(b) Example of a pattern GTU @ 2017–04–28 07:36:47. Match score: 0.84, Mean counts / GTU: 26.90 $\frac{\text{cts}}{\text{GTU}}$.

Figure 5.11: Examples of GTUs showing pattern events.

5.2.7 Pattern Events

While spike GTUs produce the highest count rates in single pixels, pattern GTUs have the highest count rates across the whole detector. When looking at these high count GTUs, it can be seen that even if the actual GTUs look different, they all share one underlying pattern, see Figure 5.11. To extract this pattern, an average is taken over all GTUs with a high average count rate. From Figure 5.8, an average count rate of 65 counts can be chosen as a threshold for this cut (this may seem arbitrary, but the pattern will later be used directly as a cut, so the exact threshold value is not vital here). The pattern extracted from this data can be seen in Figure 5.12. From the high difference in count rate of neighbouring pixels and the distribution of count rates, it is unlikely that this pattern was caused by a physical event on the detector. It is also interesting to note that in comparison to the spike GTUs the whole detector has a high count rate and no consecutive channels are showing special features. Moreover, it is surprising how the borders between high and lower counts do not follow the bias voltage and ASIC division, and even cut through this division as can be seen in the upper part of the pattern image.

As with the spikes, the reasons for the pattern GTUs are not clear. Since for most pattern GTUs, the whole detector shows high counts and there are no higher count events (see Figure 5.8), the detector seems to saturate. In order to remove these GTUs reliably, a better method to distinguish pattern GTUs from GTUs with valid data is needed. Since the average pattern is known (see Figure 5.12), a template matching approach can be used to classify GTUs as pattern GTUs. The algorithm used is described in [114] and implemented

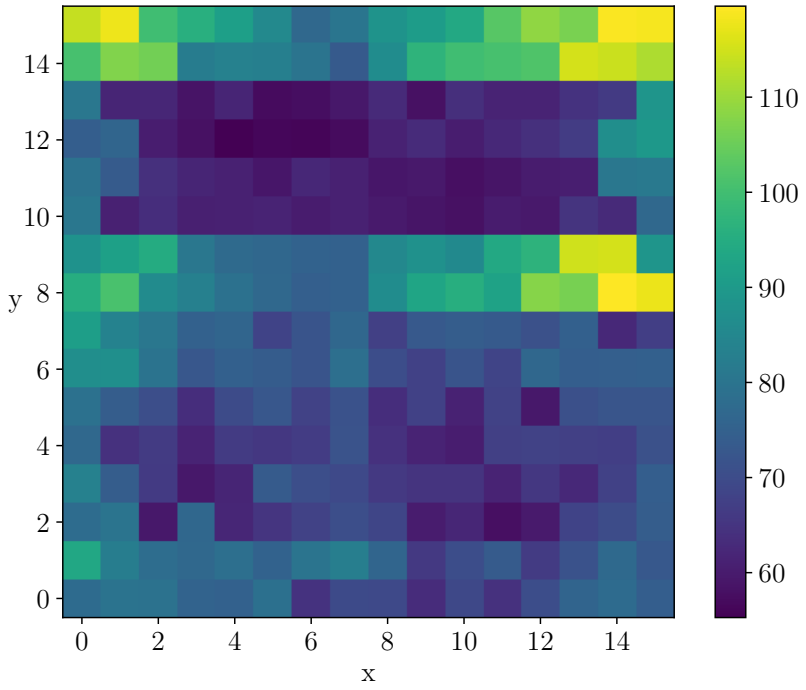


Figure 5.12: Camera image of average pattern in $\frac{\text{cts}}{\text{GTU}}$. This image is used as a feature to match pattern events against.

in [115]. It is based on a normalized cross-correlation of an image to a specified feature image and outputs a correlation coefficient between 1 (perfect match) and -1 (perfect anti-match). In this case the feature image is the average pattern image which has the same dimensions as every other camera image, therefore potential limitations of the normalized cross-correlation correlation such as image scaling, rotation and perspective distortion are not an issue here. Since the method is normalized, even GTUs, in which the pattern lasted only for a fraction of 1 GTU and will therefore have a lower amplitude will produce a high match score. This can easily be seen when plotting the match scores of a camera image against its average count rate, as shown in Figure 5.13. For this plot, only the incomplete GTU cut and unbiased data cut were used, the spike events are still present. Several populations of events can be made out. The biggest population clusters at very low average count rates and match scores between -0.2 and 0.2. These traces include the majority of physical events and should not be removed by the cut. On the other hand, the events with high match score are pattern events, which cluster at high count rates. Besides these two clusters, two branches can be made out. The first branch consists of events in the upper left area of the plot, which have high match

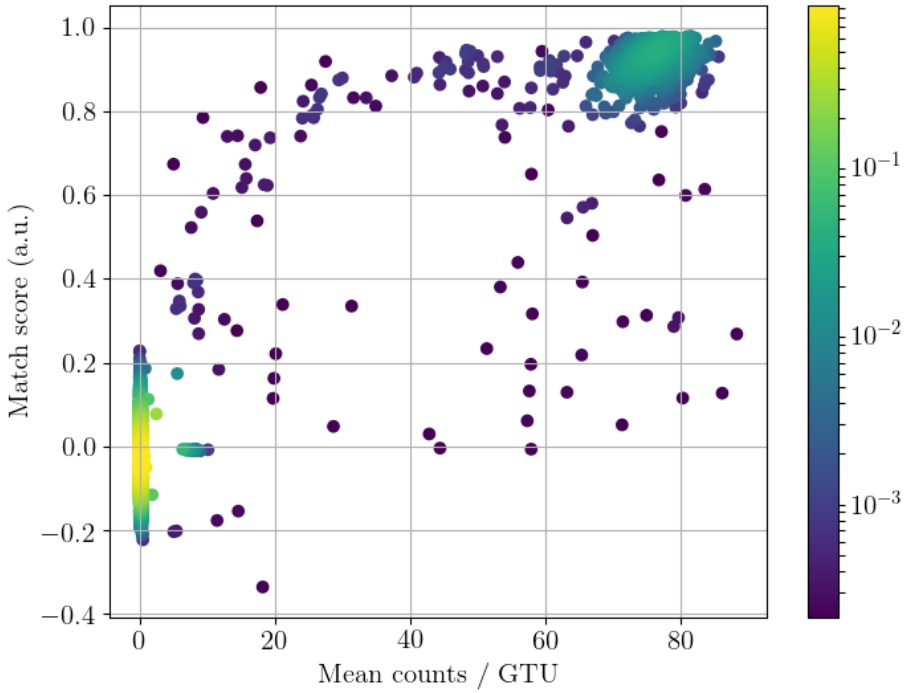
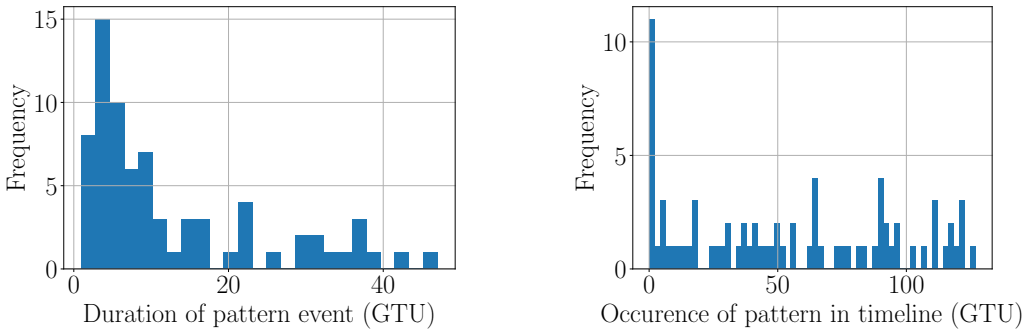


Figure 5.13: Match score of GTUs plotted against the average count rate. The colorbar indicates the density of GTUs on the plot. A match score of 0.25 is used as a threshold for the pattern cut.

scores. These are events, in which the pattern can clearly be made out. The example pattern GTUs in Figure 5.11 fall into this category. It is interesting to note that there exist pattern events with very low average count rates. The second branch consists of events on the bottom right of the plot with moderate match scores, but high average count rates. These events often show the pattern only in a fraction of the pixels, such as the GTUs in Figure 5.15. Additionally, the spike events are visible as a clustering at a match score of 0 and average count rates between 5 and 9. From these data, it is not obvious what the best values for a cut on pattern events is. In the end, a match score of 0.25 was defined as the threshold for pattern GTUs. This removes most of the pattern events and preserves the clustering of non-pattern events at low count rates completely. Some pattern events still remain after the cut, but will be removed by a following cut later on.

When looking at timelines of traces containing pattern GTUs, it can be seen that unlike spikes, pattern events can last longer than one single GTU (see Figure 5.9c). This is further explored in Figure 5.14a, which shows the



(a) Distribution of duration of pattern events. Duration is defined as consecutive GTUs which pass the pattern cut.

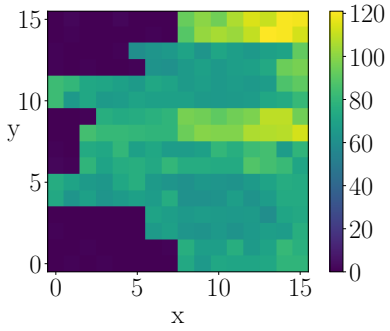
(b) Distribution of occurrence of first pattern event after the trigger.

Figure 5.14: Properties of pattern events.

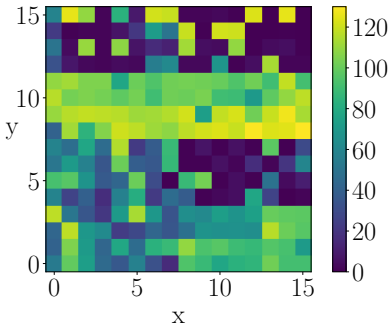
distribution of durations of pattern events. The duration of the event is defined as the amount of consecutive GTUs which pass the pattern cut. While there is a peak at 5 GTUs, it is difficult to infer anything about the nature of the patterns from the duration alone. Figure 5.14b shows the distribution of time between the first GTU of a trace and the first occurrence of a pattern event. There is a slight enhancement in the first GTU, but apart from that the distribution appears to be flat, which hints at an external source for the pattern, rather than the trigger itself.

One apparent feature of the pattern are two uppermost lines of pixel (line 14 and 15 in Figure 5.12). When examining the way SiECA is mounted next to the PDM, pixels with big y-components are expected to be closest to the PDM. Since there was electrical interference between SiECA and the PDM (which caused SiECA to be shut down after the anomaly was discovered), the pattern might be the effect of this interference on SiECA. Since the precision of SiECA timestamps is only 1 s, it is not possible to correlate the pattern events with effects in the PDM. Due to the template matching cut it is however possible to exclude pattern GTUs from the analysis, and since the rise and fall time of pattern events seems to be less than 1 GTU, it can be assumed that they do not produce artifacts in neighbouring GTUs.

In subsection 5.2.5, a continuum of events was present between spikes and the peak of pattern events at 80 couts. Some of these events survive even the template matching cut. A collection of typical events can be seen in Figure 5.15. In these images, some pixels show the same behaviour as in pattern



(a) Half pattern event @ 2017–04–28 07:38:47. Match score: 0.11, Mean counts / GTU: 19.76 $\frac{\text{cts}}{\text{GTU}}$.



(b) Half pattern event @ 2017–04–28 09:45:06. Match score: -0.004, Mean counts / GTU: 44.42 $\frac{\text{cts}}{\text{GTU}}$.

Figure 5.15: Examples of camera images where only some pixel follow the pattern. These events are not caught by the template match cut.

GTUs while the other pixels have low counts. Due to this, the GTUs are not detected by the template matching cut. However these are only 17 GTUs and can thus be eliminated by cutting any non-pattern pixels with an average count rate above 15 counts.

There is still one class of GTUs left that appear to show unphysical data: When looking at the remaining GTUs with the highest count rates, there are some GTUs in which one single pixel (with coordinates $x : 7$, $y : 5$) shows high counts while the remaining detector stays at low counts. This is problematic because from the point spread function of the optics [107], such a high count rate should be spread across multiple neighbouring pixels instead of being focused in one single pixel. To cut these GTUs, every GTU is cut where this particular pixel has > 5 counts and is the highest value of the GTU. This will be called the single pixel cut. Additionally, high counts (above 40 $\frac{\text{cts}}{\text{GTU}}$) only appear in pattern events, spike events and the single pixel events described here. Therefore one can simplify the spike cut and the continuum cut by eliminating all events where the highest count is above 40 $\frac{\text{cts}}{\text{GTU}}$. This cut will be called high counts cut. It also cuts off some few single pixel events, but these will be eliminated by a later cut anyway.

By using all the defined cuts, the subset of valid GTUs, which will later be used for the analysis, can be extracted. Table 5.3 gives an overview of the cuts in the order they are applied and of the remaining GTUs after each cut in the data set. As can be seen, more than half of the GTUs in the data set

Table 5.3: Overview of cuts for SiECA data.

cut	GTUs
No cuts	48384
Incomplete GTU	47460
Unbiased data	21604
Pattern	20672
High counts	20596
Single pixel	20572

are recorded when SiECA was not voltage biased. From the remaining 21604 GTUs, which represent ∼54 ms of data, only 1032 GTUs (∼2.5 ms) are cut out.

5.2.8 Estimation of Cosmic Ray Detection Likelihood

The main scientific goal of the EUSO-SPB mission was the detection of cosmic rays traversing the atmosphere. Due to its short recording time, it is unlikely to find such an event in the SiECA data. This section will give a rough estimation of the likelihood to detect such an event in SiECA data. From [116], around one detected CR is expected for the duration of measurement with the PDM. This value is based on simulations using an expected detector performance before that flight and expected cloud coverage. For this estimation a value of 1 is used. As discussed in the above section, the SiECA data set after the unbiased data cut consists of 21604 GTUs, which represent $\sim$ 54 ms of data taking. For the sake of simplicity, further GTU cuts are not used, as they remove only a small fraction of GTUs. For the PDM, the duration of the flight, for which data was downloaded, was $\sim$ 30 h. During this duration, the PDM was recording using an optimized trigger [106, 117]. In the value of expected CR for the PDM, the trigger is already included. In contrast to the PDM, SiECA is triggered externally every 15 s. Thus, the probability of recording a CR with SiECA is given by the fraction of the durations both detectors were recording

$$p_{\mathrm{CR}} = \frac{54\,\mathrm{ms}}{30\,\mathrm{h}} = 5 \times 10^{-7}. \tag{5.2}$$

Since this number is very low, no further effects such as geometric dimensions, optics point spread function or detection efficiencies are taken into account. Even this rough estimate illustrates that the likelihood of detecting a CR with SiECA data is very low, therefore no in-depth searches for cosmic rays were attempted.

5.3 UV Background Estimation

After having defined the cuts to distinguish the physical data from unphysical data, one can look into the content of the remaining data to better understand the detector and measure the UV night sky background.

5.3.1 Pixel Spectra Content

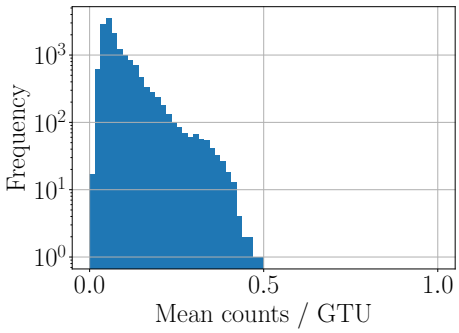
To evaluate the remaining data, the spectra of the pixels can again be studied. For simplicity, an average of all detector pixels is taken instead of examining 256 spectra separately. Since the remaining data consists of three records (see Table 5.2) and data was taken under different conditions for each record (e.g. altitude, compare Figure 5.3), the spectra have to be viewed separately and cannot be combined. The pixel spectra for the three records are shown in Figure 5.16. Since the measurements from the ASIC are trigger counts, Poisson-like distributions are expected, assuming a constant rate of triggers[1]. What is seen instead are distributions with several features. Compared to the first record, the spectrum of the third record looks much more irregular, even having three different peaks. From this it can be inferred that for the duration of data taking, the third record saw (at least) three periods were a specific trigger rate was predominant. In the same way, it can be assumed that for the first record, the rate was much less variable.

However, the trigger rate seen in the spectra are not to be taken as trigger rates from real photons. Since SiPM have rather high dark count rates, the spectra contain not only triggers from photons, but also from dark counts. This means that three different kinds of contributions are expected in the spectra:

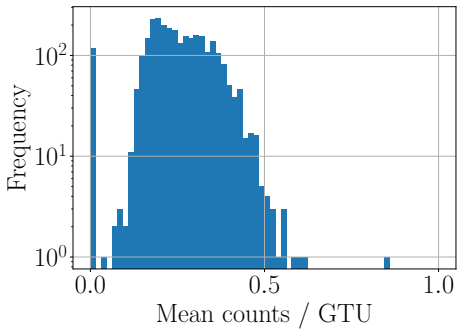
Dark counts which happen in every pixel because of thermal excitation of electron-hole pairs. Due to the gain stabilization of the voltage supply, the overvoltage for each pixel is kept constant throughout the flight and the dark count rate depends only on the probability of thermal excitation, i.e. on the temperature.

UV background light describes the residual light in the UV band, where the detector is sensitive. These events are a background for CR search. In the field of view of the camera, UV night sky background can be assumed to be diffuse and therefore uniform across the detector.

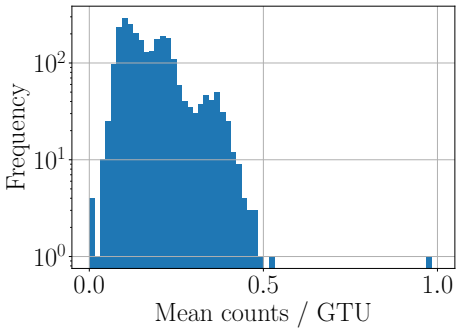
[1] For single pixels (and a constant rate), a Poisson spectrum is present. Due to averaging, non-integer values are produced, and the distribution cannot be Poisson anymore, even though the shape of the distribution is similar.



(a) Record 25th.



(b) Record 28th 2.



(c) Record 28th 3.

Figure 5.16: Mean counts per GTU for data which passes all cuts. The presence of peaks hints at different trigger rates for different periods of recording.

Transient light events which only appear for a short duration of time. These can be cosmic rays as well as high clouds moving through the field of view or lights from planes or ships. These events typically will only appear in a few pixels of the detector and can cause an enhancement in the spectra.

Of these signals, UV background and transient light events are of physical interest. Since this section is concerned with UV background estimation, the transient light events appear mainly as nuisance and data is selected in a way which minimizes transient light event contribution. In section 5.4 however, transient light events will be interesting as such events can make it possible to directly compare SiECA and the PDM.

From this, the irregular spectrum of the third record (see Figure 5.16c) can be explained with transient light events. Since transient light events dominate the trigger rate for a short period of time, they manifest in the spectrum as enhancements at higher rates. To get strong enhancements, the light events

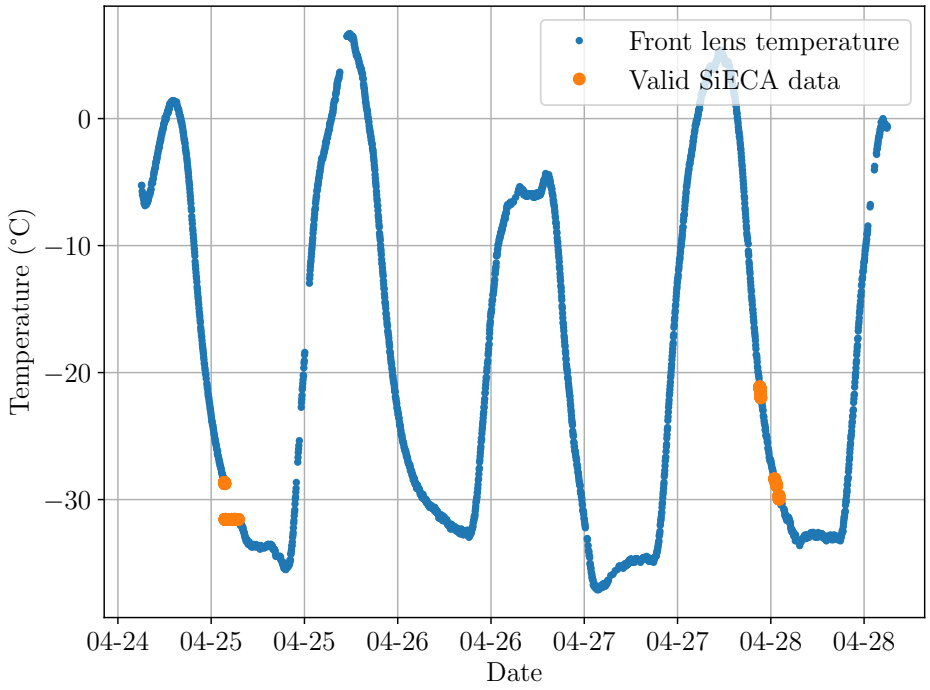


Figure 5.17: Temperature of the sensor closest to SiECA. The steep drop in the first record comes from missing temperature sensor data, a smooth decline is to be expected.

need to stay in the field of view for a very long time, e.g. slow moving clouds. Since only trigger counts are measured, there is no easy way to separate the dark counts from the UV background in the spectra. Both trigger rates λ_i can be described as Poisson distributed random variables with distributions $\mathcal{P}(\lambda_{\mathrm{dc}}, k)$ and $\mathcal{P}(\lambda_{\mathrm{bg}}, k)$ for dark count rate and night sky background rate, respectively. Because these two events happen independently, their sum is also Poisson distributed $\mathcal{P}(\lambda_{\mathrm{dc}} + \lambda_{\mathrm{bg}}, k)$. From one spectrum alone, it is impossible to make out which fraction of the counts come from dark counts and which from background, since they both happen at the same time with a constant rate (assuming no transient light events and a constant temperature). However, the dark count rate depends on the temperature and the night sky background is expected to differ slightly from night to night (e.g. due to moonlight and static cloud coverage). In the following, a method to disentangle both rates is presented, which uses this coherence of rates either in time (dark counts per pixel) or in space (background across the detector).

To estimate the dark count rates of the different records, the temperatures

of the detector have to be determined. Although there were four temperature sensors mounted on the detector to enable the voltage supply stabilization of gain [18], they were only used internally and not read out. Due to an issue in the voltage supply of the PDM not being able to turn on in low temperatures, both the PDM voltage supply and the PDM were heated. Because of this, SiECA is thermally isolated from the PDM and the temperature readings of the PDM's temperature sensors are not applicable to SiECA. Therefore, the temperature sensor on the front lens was chosen to measure the temperature of SiECA. The temperature of the detector for the records of valid data can be seen in Figure 5.17[2]. Immediately it can be seen that the first and the third record are recorded at the same temperature. The dark count rate of SiPM depends on the temperature and overvoltage (and the gain is stabilized), thus the same dark count rate can be assumed for data recorded in the first and in the third record in a given pixel. Naturally, the dark count rate differs from pixel to pixel, but for each given pixel, it is the same in both records, and as the temperature is constant, so is the dark count rate for each pixel. Therefore the pixel spectra, i.e. the probability to observe k counts per GTU, can be described in the following way:

$$\begin{aligned} \mathcal{P}_{i,1}(\lambda, k) &= \mathcal{P}_i(\lambda_{\mathrm{dc}_i}, k) + \mathcal{P}_i(\lambda_{\mathrm{bg1}}, k) \\ \mathcal{P}_{i,2}(\lambda, k) &= \mathcal{P}_i(\lambda_{\mathrm{dc}_i}, k) + \mathcal{P}_i(\lambda_{\mathrm{bg2}}, k) \\ &\vdots \end{aligned} \tag{5.3}$$

Each Poisson pixel spectrum $\mathcal{P}$ of detector i (one of the 256 pixels) consists of two components: A dark count rate λ_{dc_i} which stays the same for every pixel i (due to constant temperature) and a background rate $\lambda_{\mathrm{bg}n}$ for the nth night. In this context, the term "night" will be used for any extended period of data taking with constant conditions, guaranteeing the same uniform background rate across the detector. Since there are 256 pixels, this gives $256 \times N$ coupled equations for N given nights. Before the discussion of means to extract the rates from the pixel spectra and solve the set of equations, it has first to be determined how many records qualify as "nights".

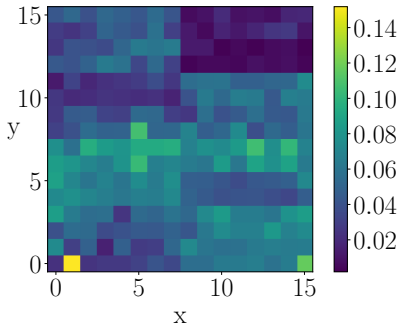
[2]The SiECA data points were inserted into the plot by matching the nearest timestamp of data taking with timestamps of the front lens temperature time series. For some of the first record of SiECA data, no data for the front lens temperature was available, which looks like a steep drop in temperature for SiECA. This drop can be expected to have happened smoothly over the time of data taking.

5.3.2 Data Selection

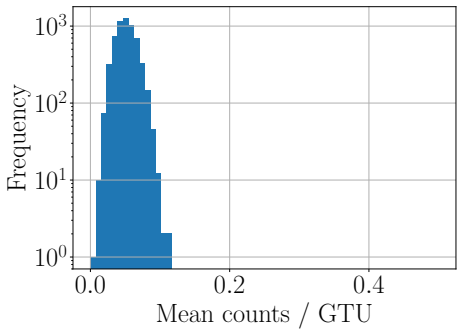
Due to higher temperature, the data in the second record shows a higher dark count rate than in the first and third records. Since the model in Equation 5.3 assumes a constant dark count rate for each pixel, the addition of a temperature dependence and thus varying dark counts per pixel is not trivial and introduces additional parameters into the model. Thus, the second record was discarded for this analysis. For the first and third record, data was not recorded continuously at all times, but with short breaks between patches of data. These patches can be used as basis for the "nights" and one can define patches as continuous durations of recording with a gap of at least 300 seconds to the next patch. From the data in the first and third record, which survive all cuts, 7 such patches are present, an overview can be seen in in Table 5.4.

To use the model described in Equation 5.3, dark count rates have to be constant for each pixel, and the night sky background rate has to be constant for each night. Since a constant dark count rate is given by the constant temperature, patches have to be found with minimal transient light event contribution to ensure constant night sky background rates across the detector. In other words, a cut on transient light events has to be found.

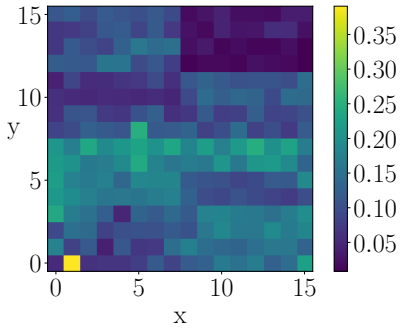
For the patches, both the average pixel spectra per patch and the average camera image per patch can be analysed. Examples of this can be seen in Figure 5.18. Both the images of the first two patches and the spectra are very similar, in counts as well as shape. Other patches show varying deviations from this, e.g. patch 5 (see Figure 5.18e, 5.18f) has a similar image, but a spectrum with multiple peaks. This implies that the background rate changed during recording due to transient light events, but in a diffuse way across the detector. Patch 7 (see Figure 5.18e, 5.18f) only has 1 peak in the spectrum, but clearly a higher rate on the right side of the detector (in section 5.4, this will be linked to a cloud moving slowly through the field of view). Since the first two patches have the lowest counts and appear to be very clean, the average image of these two patches can be used as a template for the background, i.e patches without transient light events, for template matching similar to the pattern cut in subsection 5.2.7. By demanding a match score > 0.9 and uniform background rates, i.e. only one peak in the spectrum, which was determined by eye, only four patches qualify: 1, 2, 3 and 6. These patches can be used as nights with independent background rates to solve the system described in Equation 5.3. It should be noted that the fourth night (patch 6) has much lower statistics compared to the other three, which makes it unclear if inclusion of this night to the data set improves the result. This question is resolved in subsection 5.3.4



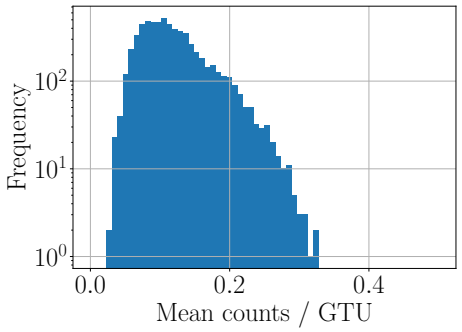
(a) Patch 1 image, match score 0.99.



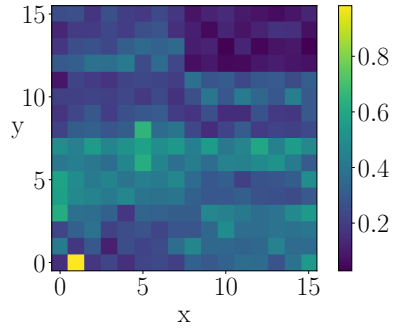
(b) Patch 1 spectrum.



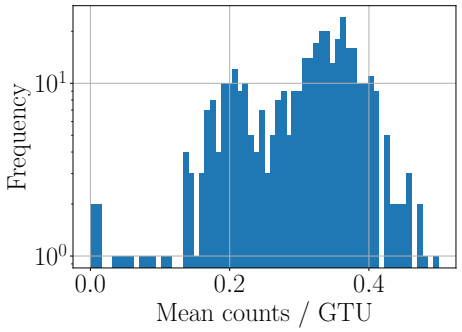
(c) Patch 3 image, match score 0.98.



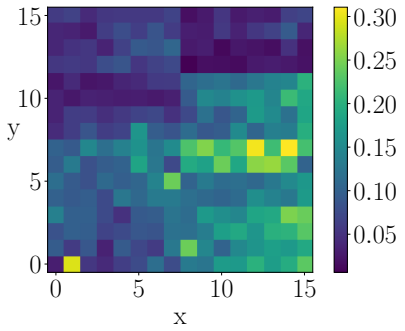
(d) Patch 3 spectrum.



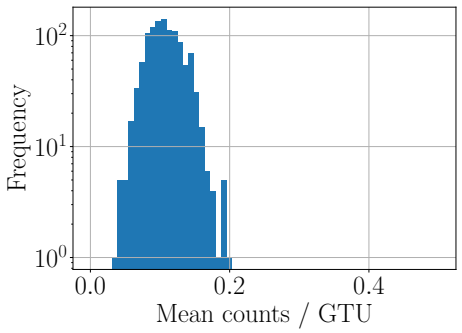
(e) Patch 5 image, match score 0.93.



(f) Patch 5 spectrum.



(g) Patch 7 image, match score 0.79.



(h) Patch 7 spectrum.

Figure 5.18: Examples of average camera images and corresponding spectra for some patches. Patch 5 and patch 7 contain transient events as seen by the multiple peaks in the spectrum (5.18f) and the low match score (5.18g).

with toy simulations.

Table 5.4: Overview of "nights" in the first and third record.

Patch	Start	Duration in GTUs
1	2017–04–25 10:41:31	5776
2	2017–04–25 10:53:16	1999
3	2017–04–25 11:36:58	6392
4	2017–04–25 12:09:11	317
5	2017–04–28 09:28:38	389
6	2017–04–28 09:35:42	676
7	2017–04–28 09:45:36	1120

5.3.3 Poisson Model Fits

In subsection 5.3.1, the pixel spectra were modeled as sums of Poisson distributed random variables described by certain parameters (in this case, the rates λ_{dc} and λ_{bg}), which can be written as $P(x|\theta)$, where P describes the distribution, x describes the data and θ describes the parameters of the distribution. To obtain the rates λ_i, the parameters θ which best describe the data need to be found. To infer these parameters from the data, *estimators* can be used, which are functions of the data used to estimate the value of the parameter θ [118]. For the analysis of SiECA data, two commonly adopted estimators were used: The maximum likelihood estimator and the least squares estimator, which are both frequentist methods (see [118, 119] for further reading). In this section, a short outline will be given of their differences and a toy simulation will be used to estimate their properties estimating Poisson distributed spectra.

The **maximum likelihood** estimator works by finding the parameter θ which maximizes the likelihood function $L(\theta)$. The likelihood can be obtained as the probability of the data under assumption of the parameters $L(\theta) = P(x|\theta)$ [118]. As the data consists of counts for single GTUs and pixels, the counts can be viewed as independent and identically distributed (per pixel/night), thus the likelihood can be written as

$$L(\theta) = \prod_{i}^{n} f(x_i; \theta), \tag{5.4}$$

where f is the probability distribution function describing the process (Poisson in this case), x_i are the data points (counts) and θ are the parameters of

the Poisson distributions. Since only the maximum of the likelihood function is relevant and the logarithm ln is strictly increasing, one can also use the logarithm of the likelihood $\ln L(\theta)$ instead of the likelihood $L(\theta)$. In this case, the product becomes a simple sum, which is easier to handle computationally. This is sometimes called the *log-likelihood*. The number of events n in the product in Equation 5.4 is considered fixed. If the number of observed events n is also randomly distributed (e.g. through a Poisson distribution with mean μ), it can be included into the likelihood as

$$L(\theta) = \frac{\mu^n}{n!} e^{-\mu} \prod_i^n f(x_i; \theta), \tag{5.5}$$

where μ is the mean of the Poisson distribution. This is called the *extended likelihood* estimator [120]. In this chapter both the extended maximum likelihood estimator and the usual maximum likelihood estimator are used to infer rates from Poisson spectra. For both versions of the maximum likelihood estimator as described above, the data is used in unbinned form, since certain binning can lead to aliasing and therefore loss of information for the data set. However, for the SiECA data set, each data point for every pixel consists of integer values (counts). Because of this, it is possible to construct the spectra in a way that no information is lost, e.g. by ensuring integer bins and a number of bins $n_{\text{bins}} > \max(\text{counts})$ for each spectrum. In this case, the unbinned data can be reconstructed from the binned histogram. For the following analyses, care was taken not to alias the spectra, therefore binning does not pose a problem, even when using binned data.

In physics, it is often possible to describe measurements y_i as sets of independent Gaussian distributed random variables x_i with mean values $\mu(x_i, \theta)$ and variances σ_i. To find the optimal parameters θ, the log-likelihood can be used as described above

$$\ln L(\theta) = -\frac{1}{2} \sum_i \frac{\left(y_i - \mu(x_i, \theta)\right)^2}{\sigma_i^2}. \tag{5.6}$$

To maximize the right-hand side, the sum has to be minimized. This is the basis of the **least squares** estimator

$$\chi^2 = \sum_i \frac{\left(y_i - \mu(x_i, \theta)\right)^2}{\sigma_i^2}. \tag{5.7}$$

Instead of maximizing the likelihood $L(\theta)$, the quadratic sum of the differences between data and assumed model values weighted by the variances is

minimized. This quantity is called χ^2 [3]. Although Gaussian variables were used to define χ^2 (and the likelihood $L(\theta)$ and χ^2 coincide for Gaussian variables), the least squares method can be used even if the measurements y_i are not Gaussian as long as they are independent [119]. In order to calculate χ^2, the variances σ_i of the data points have to be known. These are often not known a priori and must be estimated from the data. For Poisson variables x_i, the variances are given by $\sqrt{x_i}$ [121]. However, this poses a problem for binned data spectra with very low count rates (as in this case), as the variances become poorly estimated or even undefined in case of bins with 0 counts. In this chapter, this problem was mitigated by estimating the variances from the Poisson model $\mu(x_i, \theta)$ instead of the data. In the following, the two different types of estimators, maximum likelihood and least squares, will be treated as two independent methods for extracting parameters from Poisson spectra, as both can give valid results and have their strengths and weaknesses.

The SiECA data set consists of binned Poisson spectra with low mean values μ. In order to test the different estimators' ability to infer the mean values μ from the spectra, toy simulations were performed. For this, Poisson distributed data was randomly ([122]) generated for 45 different mean values between 0.001 and 0.9. Each distribution contained 5000 data points, which is similar to the statistics found in SiECA data usable for physics analysis (see Table 5.4). For each distribution, the optimal mean values were found using the estimators described above. To minimize the effect of statistical fluctuations, this procedure of generating data and fitting was carried out 20000 times per mean value μ. For every data set and estimator, the same start value of $\mu = 0.1$ was used for the optimization.
The χ^2 minimization was implemented using [123], where the raw data was first transformed into a histogram and χ^2 was defined as

$$\chi^2 = \sum_i \frac{(y_i - \mathcal{P}(x_i, \lambda))^2}{\mathcal{P}(x_i, \lambda)}, \tag{5.8}$$

where λ denotes the optimization parameter and estimate of μ, and the sum runs over every integer bin in the histogram.
For both usual and extended maximum likelihood estimation, a freely available implementation [124] was used, due to the increased numerical complexity compared to the least squares minimization, which proved to be too slow for the complete fit in subsection 5.3.4 when implemented in [123]. In order to find the error in the estimation for increasing mean values μ, for each distribution

[3] Under certain conditions, the minimum follows the χ^2-distribution, and is in general called χ^2 because of this, even when the particular distribution differs [119].

a relative error was calculated as

$$\delta\mu_i = \frac{|\mu - \lambda_i|}{\mu}, \tag{5.9}$$

where μ is the true value, λ is the optimal parameter from the estimation and the index i takes into account the 20000 optimal values per estimator and μ. From the distribution of $\delta\mu_i$, the error in estimation for a particular estimator is given as the mean value of this distribution

$$\Delta\mu = \overline{\delta\mu} = \frac{\overline{|\mu - \lambda|}}{\mu}. \tag{5.10}$$

The errors in estimation $\Delta\mu$ for 45 different values of μ can be seen in Figure 5.19. As expected, the errors $\Delta\mu$ get smaller for higher values of μ, since the data is more evenly distributed in the histogram and the statistical fluctuations of the 1st bin become less important. In the same way, the uncertainties $\sigma_{\Delta\mu}$ get smaller since outliers become less likely. The errors of estimation are reasonably low, reaching above 5 % only for the lowest values of μ. No significant difference can be seen in the error of estimation $\Delta\mu$ for any method, however the maximum likelihood methods tend to have marginally smaller errors as well as uncertainties $\sigma_{\Delta\mu}$. When fitting single Poisson spectra, the usual maximum likelihood estimation gives the same values as the extended one, they are not distinguishable in Figure 5.19. It is worth noting that the distribution of errors is not the same for the χ^2 method and the maximum likelihood methods. When taking the signed difference instead of the absolute value in Equation 5.9, the mean value $\Delta\mu$ should be almost zero if the distribution is symmetric. For the χ^2 case however, negative values are obtained which indicate that said method tends to overestimate the true value, while in the maximum likelihood case there is no preference in either direction. As the relative errors $\Delta\mu$ are similar however and the effect is small, this does not cause further problems.

5.3.4 Linear Approach

As described in Equation 5.3, each pixel spectrum can be considered a sum of two background spectra

$$\mathcal{P}(\lambda, k) = \mathcal{P}(\lambda_{\mathrm{dc}}, k) + \mathcal{P}(\lambda_{\mathrm{bg}}, k), \tag{5.11}$$

one from the SiPM dark counts λ_{dc} and one from night sky background λ_{bg}. From single spectra the two rates are impossible to disentangle however, since both background events happen at the same time with constant rates. By selecting the data for diffuse night sky background rates, the night sky rate

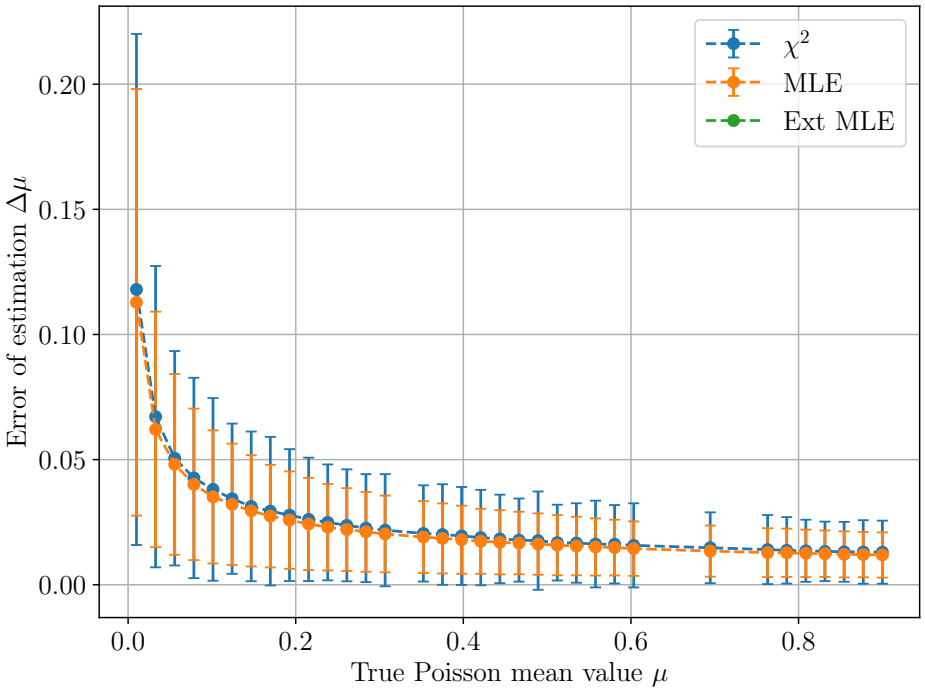


Figure 5.19: Errors of estimation for single Poisson spectra with low counts for three different estimators. For single spectra, extended and normal maximum likelihood estimators give the same result. The least squares estimator performs only marginally worse.

λ_{bg} can be constrained to be the same across the detector for every pixel in each night. At the same time, due to constant temperature, each pixel has the same dark count rate λ_{dc} for every night. Using this, a model can be build with $256 \times n$ parameters $f(x, \theta)$ and the methods described in subsection 5.3.3 can be used to fit the parameters θ to the data x. Here, x denotes the whole data set described in subsection 5.3.1 and θ are the $256 \times n$ parameters with the following constraints:

$$\mathcal{P}_{i,n}(\lambda) = \mathcal{P}_{i,n}(\lambda_{\text{dc}_i}) + \mathcal{P}_{i,n}(\lambda_{\text{bg}_n}) \; \forall i \in \{\text{pixels}\}, \forall n \in \{\text{nights}\}, \tag{5.12}$$

where $\mathcal{P}(\lambda)$ is the Poisson distribution $\mathcal{P}(\lambda, k)$, $\{\text{pixels}\}$ denotes the set of 256 pixels and $\{\text{nights}\}$ denotes the n nights. Because of the high dimensionality of the model and the amount of constraints, finding the best fit parameter θ is computationally intensive and can take a lot of time. Before estimating the accuracy of the method with toy data, a computationally simpler approach is presented:

Hybrid Linear System

For this problem, sums of Poisson distributions have to be considered. It can be shown that given two Poisson distributed random variables $X \sim \mathcal{P}(\lambda_1)$ and $Y \sim \mathcal{P}(\lambda_2)$, the sum of the variables $X + Y$ is Poisson distributed again $X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$ [125]. Using this, a system of linear equations for the rates can be formulated

$$\begin{aligned}
\lambda_{\mathrm{dc},1} + \lambda_{\mathrm{bg},1} &= \lambda_1 \\
\lambda_{\mathrm{dc},2} + \lambda_{\mathrm{bg},1} &= \lambda_2 \\
&\vdots \\
\lambda_{\mathrm{dc},1} + \lambda_{\mathrm{bg},2} &= \lambda_{257} \\
\lambda_{\mathrm{dc},2} + \lambda_{\mathrm{bg},2} &= \lambda_{258} \\
&\vdots \\
\lambda_{\mathrm{dc},255} + \lambda_{\mathrm{bg},n} &= \lambda_{256\cdot n-1} \\
\lambda_{\mathrm{dc},256} + \lambda_{\mathrm{bg},n} &= \lambda_{256\cdot n},
\end{aligned} \tag{5.13}$$

which consists of $256 + n$ unknowns and $256 \cdot n$ equations. Each equation corresponds to one single pixel spectrum in the data. Since the rates of single Poisson spectra can be extracted with good accuracy (see subsection 5.3.3), the rates λ_i can be considered known values. This system can be expressed in vector form as

$$r_i(x) = \sum_j a_{ij} x_j - \lambda_i, \tag{5.14}$$

where $i \in \{0..256 \cdot n\}$ and $j \in \{0..256 + n\}$. The vector x_j consists of the desired rates $\lambda_{\mathrm{dc/bg},j}$. For a solution $\|r_i(x)\|$ has to be minimized. Since the matrix a_{ij} has more rows than columns, the system is called overdetermined and has in general no solution [126]. Is is however possible to find a vector x_j which minimizes $\|r_i(x)\|$ using the least squares method [127][4]. This x_j does not satisfy every equation in the system, but gives values that come closest to solving the system exactly. The maximum likelihood method and the least squares method can be combined to solve this system using a hybrid approach: The maximum likelihood estimator gives slightly more accurate results (see subsection 5.3.3) while being computationally more expensive. Here, each spectrum only has to be fitted once, to get the right hand side values λ_i in Equation 5.13. The resulting system of equation can be solved using the faster least squares methods.

[4] Is it of course also possible to solve this system with a maximum likelihood estimator. However, there are a plethora of least squares methods and optimizations for this problem, which speed up the procedure, which is why this family of methods was preferred here.

In order to assess the viability of obtaining both SiPM dark counts rates $\lambda_{\mathrm{dc},i}$ and as night sky background rates $\lambda_{\mathrm{bg},n}$ as well as to estimate the accuracy such a fit, toy simulations were carried out.

Toy Simulations

For such a simulation to work, toy data are needed which resemble the real data as close as possible. The actual rates are not known a priori, but the rates can be constrained from the spectra and camera images. The data set consists of three or four nights, which have different average rates. Due to the constant temperature across all nights, the dark count rates $\lambda_{\mathrm{dc},i}$ do not change and an upper bound can be placed on the dark count rate by selecting the night with the lowest average rate $\overline{\lambda_{\mathrm{dc}}}$ which is $\sim 0.04\,\frac{\mathrm{cts}}{\mathrm{GTU}}$ in the second night. In this night, the night sky rate λ_{bg} is the lowest (since $\lambda_{\mathrm{dc}} = \mathrm{const}$) and λ_{dc} cannot be higher than this rate. For the upper bound, the rate of the brightest pixel $\max(\lambda_{\mathrm{dc}}) \sim 0.14\,\frac{\mathrm{cts}}{\mathrm{GTU}}$ is selected. In the same way, an upper bound for the background counts is given by the highest average night count ($\overline{\lambda_{\mathrm{bg}}} \sim 0.11\,\frac{\mathrm{cts}}{\mathrm{GTU}}$ from the third night), with a maximum of $\max(\lambda_{\mathrm{bg}}) \sim 0.39\,\frac{\mathrm{cts}}{\mathrm{GTU}}$. Using these values, data can be generated for a detector with p pixels in n nights by randomly generating $\lambda_{\mathrm{dc},i}$ with $i \in \{0..p\}$ using a uniform distribution from 0.001 to 0.15. The background rates $\lambda_{\mathrm{bg},j}$ with $j \in \{0..n\}$ are picked from a uniform distribution from 0.001 to 0.4. For similar statistics to the real data set, each spectrum contains 5000 counts. Using these rates, the rate for each pixel spectrum can be calculated according to Equation 5.13 and the spectra can be generated from Poisson distributions as in subsection 5.3.3.

With the generated data sets, best fit values of the rates can be produced and the different methods can be compared to the true values for the rates and to each other. For this comparison, all four methods described above are used, namely χ^2 minimization, usual and extended maximum likelihood approaches and the hybrid approach of extracting the rates first with a maximum likelihood extractor before solving the system of equations with the least squares method. For each method, the same starting values of $\lambda_{\mathrm{dc,\ start}} = 0.005$ and $\lambda_{\mathrm{bg,\ start}} = 0.2$ were used. These values were chosen with the maximum likelihood estimator in mind, which needs relatively specific starting values in order to converge successfully in comparison to the χ^2 minimization which is much more lenient with regards to its starting values. Therefore the starting value $\lambda_{\mathrm{bg,\ start}}$ was chosen as the mean value of the interval of its uniform distribution and $\lambda_{\mathrm{dc,\ start}}$ was chosen to be significantly smaller, as experience showed this to be necessary.
Similar to the toy simulation in subsection 5.3.3, it is important to understand

how well the methods work on any data set. Because of this, toy simulations were not only performed for the SiECA data set with 256 pixels and 3 nights, but for setups with smaller detector size as well. While in subsection 5.3.3, a single Poisson spectrum fit was done in less than a second, regardless of method, a fit to the complete data set can take minutes with the hybrid approach or the χ^2 minimization and up to hours with the maximum likelihood approaches (especially when using the extended version). This makes it very hard to collect large statistics, and limits the simulation to setups with up to 256 pixels and 50 fits per setup. Simulations were run for setups with 3 nights and varying amount of pixels from 5 to 256. Since the true values of the rates μ_i are known, the relative error can again be used as a measure for each fit

$$\delta\mu_i = \frac{|\mu - \lambda_i|}{\mu}, \tag{5.15}$$

where λ_i are the best fit values. Since it can be cumbersome to give the relative error of every single pixel and night rate especially for high amounts of pixels, two mean values will be given instead: One for the relative error of estimation for the night rates $\Delta\lambda_{\mathrm{bg}} = \overline{\delta\mu_{\mathrm{bg}}}$ and one for the dark counts $\Delta\lambda_{\mathrm{dc}} = \overline{\delta\mu_{\mathrm{dc}}}$. For each setup, 50 data sets were simulated and fitted. Due to the asymmetric nature of the distribution of mean values $\Delta\lambda_i$, the median value is given instead of the mean value. The error is calculated as the 15.87th and 84.13th percentile, which corresponds to $1\,\sigma$ for a Gaussian distribution.

Simulation Results

The result of the simulations for varying setups with 3 nights can be seen in Figure 5.20. At first glance, it can be seen that for both nights and dark counts, the χ^2 minimization gives the smallest error of estimation $\Delta\lambda$, with $\Delta\lambda_{\mathrm{dc}} < 20\,\%$ for the dark count rate and $\Delta\lambda_{\mathrm{bg}} < 10\,\%$ for a SiECA-like setup with 256 pixels. The fact that night rates are estimated with a smaller error is understandable due to having much more pixels than nights. For low pixel counts, the χ^2 minimization is beaten by the maximum likelihood methods, which perform similar to each other in this regime. For increasing pixel counts the usual maximum likelihood error increases while the extended maximum likelihood error follows the χ^2 method and stays competitive. Even though it sounds attractive in theory, the hybrid approach fails to give a solid estimation of the parameters for the data set. For the dark count rates, it never manages to converge and for the night rates, it completely stops working at ~ 125 pixels. Due to only having 50 fits per setup, the results are subject to statistical fluctuations, which manifest in erratic curves, which can be expected

to be smooth for high numbers of fits. As this is much more pronounced in the maximum likelihood estimators, the χ^2 minimization seems more robust with regards to small changes in the data set, which is further emphasized by the smaller errors bars in the χ^2-case. This was also experienced in the methods' dependency on the starting values for the fits, where the χ^2 minimization would arrive at the optimal values over a large set of starting values. Therefore, the χ^2 minimization was chosen as the best method to fit the model to the SiECA data set as it showed to have the smallest error of estimation, highest robustness and is computationally simpler than maximum likelihood methods and therefore much faster.

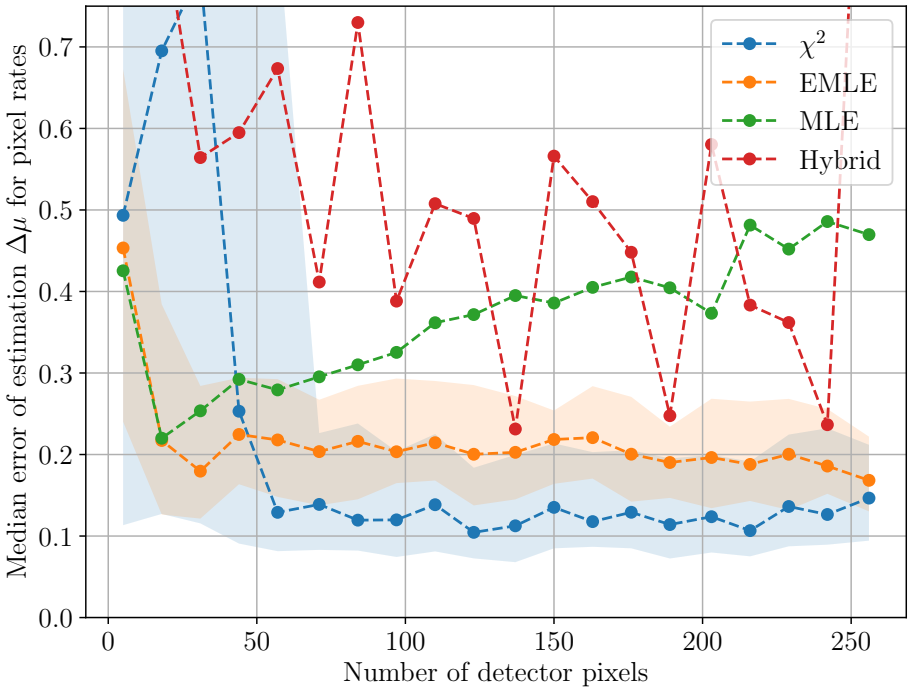
For simplicity, the spectra for the toy simulation were generated with 5000 counts per spectrum. In the real SiECA data set, the counts per spectrum vary per night, which can be seen in Table 5.4. The fourth qualifying "night" (patch 6) in particular contains only 676 counts. Therefore, another simulation was run using only the χ^2 estimator to predict if the nonuniform distribution of counts alters the error of estimation and to decide if it is worth including the fourth night with low statistics. The simulation was setup in the same way as above with the difference of using [5776, 1999, 6392, 676] counts for the spectra in consecutive nights, with one simulation using 3, the other using 4 nights. Instead of 50 fits per setup, 2000 fits were used since the expensive maximum likelihood fits could be excluded. The result of the simulations are summarized in Table 5.5. As can be seen, the inclusion of the fourth night increases the error of estimation slightly due to large statistical uncertainties in the fourth night pixel spectra, but the difference is small and well within any reasonable fluctuations. Additionally, the fourth night data are the only data in the set from the later records (on April 28th, see subsection 5.3.1), therefore it was included in the fit.

Table 5.5: Summary of the error of estimation $\Delta\lambda$ for excluding and including the fourth night.

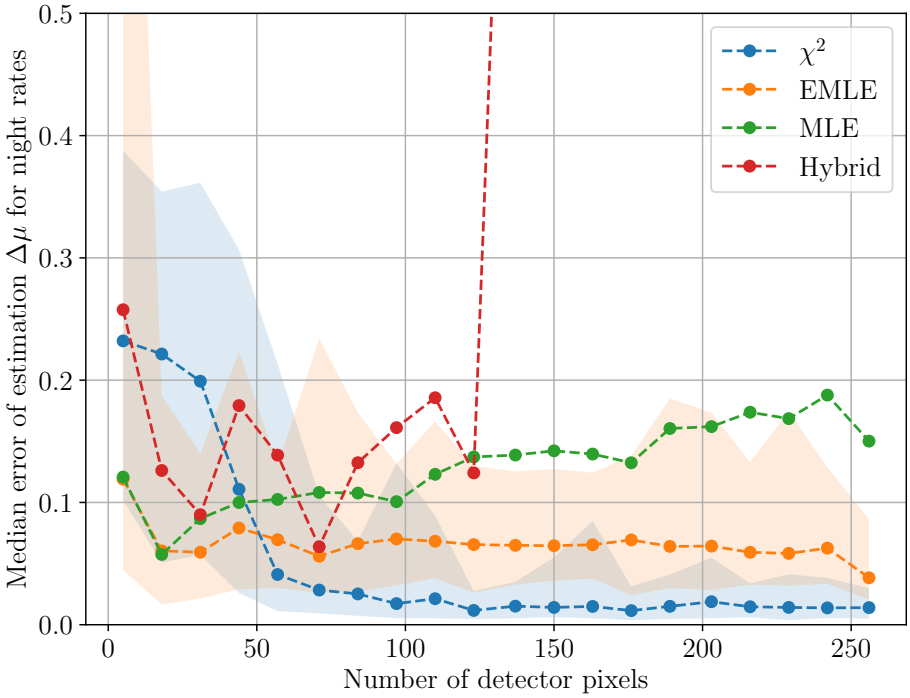
	Three nights	Four nights
$\Delta\lambda_{\mathrm{bg}}$	$0.016^{+0.024}_{-0.009}$	$0.019^{+0.032}_{-0.008}$
$\Delta\lambda_{\mathrm{dc}}$	$0.132^{+0.091}_{-0.051}$	$0.132^{+0.101}_{-0.049}$

SiECA Data Fit

The actual fit was performed with both the χ^2 minimization and the extended maximum likelihood estimator (EMLE). The results for the night sky back-



(a) Median error of estimation $\Delta\mu$ for dark count rates over number of detector pixels.



(b) Median error of estimation $\Delta\mu$ for night rates over number of detector pixels.

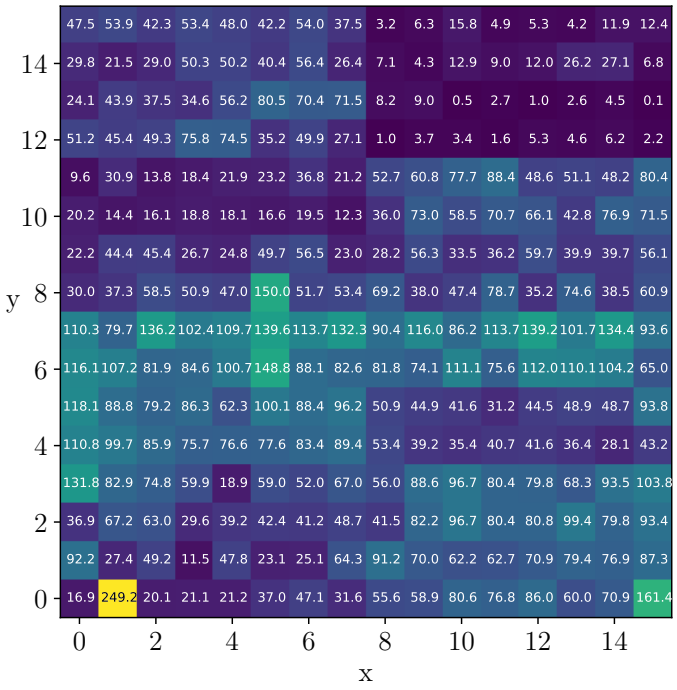
Figure 5.20: Results for toy simulation with 3 nights and varying pixel numbers. For the sake of readability, errors are given as shaded regions only for the two best methods.

ground rates can be seen in Table 5.6. It can be seen that the values for the background rates differ much from night to night and span over two orders of magnitude. The first two nights have the lowest rates, which was already known from subsection 5.3.1, as they were used as a definition for a clean average image when searching for qualifying nights. The third and fourth night show much higher rates, but it is hard to infer if this is due to cloud coverage or the lower flying height (see Figure 5.3). While it is reassuring to see both estimators arrive at similar rates λ_{bg}, the discrepancy between the actual values is bigger than the 2 % error of estimation derived above. This is a hint that the linear model used for this fit describes the data insufficiently. The following subsections will improve on this simple model.

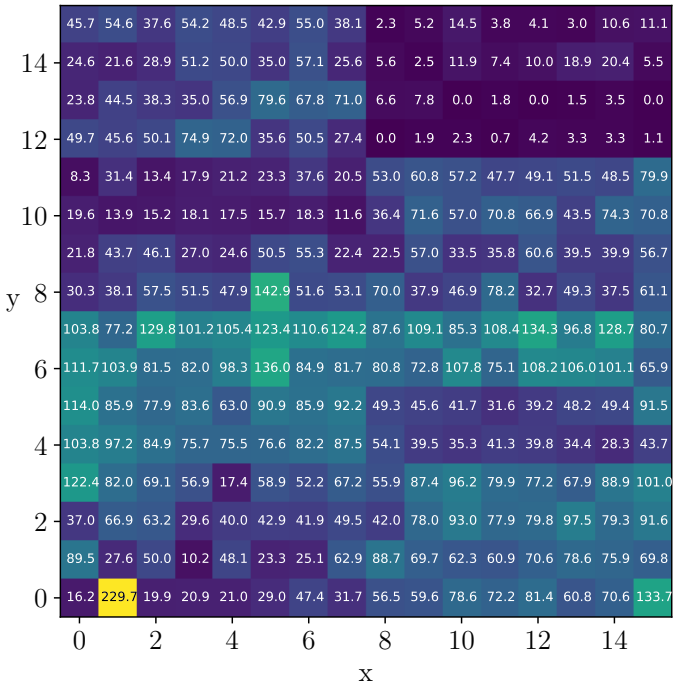
Table 5.6: Counting rates of the night sky background for different "nights" using the linear model. Nights correspond to the patches defined in Table 5.4. The last row shows the average dark count rate.

Night	λ_{bg} with χ^2 ($\frac{\mathrm{cts}}{\mathrm{GTU}}$)	λ_{bg} with EMLE ($\frac{\mathrm{cts}}{\mathrm{GTU}}$)
1	1.44×10^{-3}	1.80×10^{-3}
2	9.97×10^{-4}	1.06×10^{-3}
3	5.96×10^{-2}	5.58×10^{-2}
4	1.63×10^{-1}	1.46×10^{-1}
λ_{dc}	5.61×10^{-2}	5.41×10^{-2}

The resulting dark count values λ_{dc} with the χ^2 minimization and extended maximum likelihood estimation are shown in Figure 5.21. As expected, the distribution of rates follows the clean average image (similar to Figure 5.18a) for both methods. In these images, two features stand out: The area in the top right with 4×8 pixels shows lower rates than the remaining pixels. Since these 32 pixels share one voltage supply, it is plausible that they received a supply voltage lower than the rest. The second pixel on the bottom row stands out with a much higher count rate than the remaining pixels. Like with the night sky background rates, there is little difference in the rates estimated by the two different methods, however the χ^2 minimization estimates slightly higher rates for most pixels. This behaviour could already be seen in subsection 5.3.1, when looking at single Poisson spectra. The average dark count rate by χ^2 minimization is $\overline{\lambda_{\mathrm{dc}}} = 0.056\,\frac{\mathrm{cts}}{\mathrm{GTU}} = 22.5\,\mathrm{kHz}$, which is a surprisingly high value. A characterization of the same SiPM used for SiECA by [98] showed a dark count rate of $\lambda_{\mathrm{dc}} \sim 150\,\mathrm{kHz}$ at 0 °C. From the temperatures of the SiECA SiPM at time of recording ($\sim -30\,^{\circ}\mathrm{C}$), this rate is expected to be $\sim 10\times$ lower [99],



(a) Dark count rates obtained by χ^2 method in $10^{-3}\ \frac{\text{cts}}{\text{GTU}}$. Mean $\lambda_{\text{dc}} = 0.056\ \frac{\text{cts}}{\text{GTU}}$



(b) Dark count rates obtained by extended maximum likelihood method in $10^{-3}\ \frac{\text{cts}}{\text{GTU}}$. Mean $\lambda_{\text{dc}} = 0.054\ \frac{\text{cts}}{\text{GTU}}$

Figure 5.21: Resulting dark count rates distribution using the linear model. Both methods are in good agreement.

which leads to an expected rate of $\lambda_{\mathrm{dc,\,exp}} \sim 15\,\mathrm{kHz}$, comparable to the measured rate. However, due to the calibration procedure (see subsection 5.2.2), a threshold higher than 0.5 p.e. is expected and thus a lower trigger rate. In principle, results with a higher dark count rate and a lower background rate could be explained if the fit is unable to distinguish the rates properly and the lowest background rate is set to a near 0 value, only giving the differences in rates between the nights and adding the missing offset to the dark counts. However, the toy simulations were unable to reproduce such a scenario and the fit was able to separate the rates properly, as shown above. Still, the linear model is a simplification where every detector has the same properties and perfect detection efficiency. The next section tries to improve on these flaws by introducing the photon detection efficiency as a fit parameter.

5.3.5 Photon Detection Efficiency

When studying the distribution of dark count rates (see Figure 5.21), it is apparent that not all pixels have the exact same rate. Since the dark count rate as well as gain and other SiPM parameters depend on the overvoltage, the different pixels can be expected to have different overvoltages. This however also impacts the pixels' ability to detect incoming photons, which is parametrized as photon detection efficiency (PDE). Due to differences in PDE the apparent rates from the night sky background as seen in specific pixels are expected to be different, even if the actual rate λ_{bg} stays the same across the detector. To account for this effect, a new class of parameters ϵ_i is introduced to the model, which describe the detector efficiency for each pixel. This changes Equation 5.12 to

$$\mathcal{P}_{i,n}(\lambda) = \mathcal{P}_{i,n}(\lambda_{\mathrm{dc}_i}) + \mathcal{P}_{i,n}(\epsilon_i \cdot \lambda_{\mathrm{bg}_n})\ \forall i \in \{\text{pixels}\}, \forall n \in \{\text{nights}\}, \qquad (5.16)$$

where ϵ_i is the PDE for every pixel. For the model, this nearly doubles the amount of free parameters from $260 = 256\,\lambda_{\mathrm{dc}} + 4\,\lambda_{\mathrm{bg}}$ to $516 = 256\,\lambda_{\mathrm{dc}} + 256\,\epsilon + 4\,\lambda_{\mathrm{bg}}$. To estimate the ability of the fit to deal with the increased amount of parameters, toy simulations were performed with the updated model. Both dark counts rates λ_{dc} and background rates λ_{bg} are handled in the same way as for the linear model. The additional efficiencies ϵ_i are taken randomly from a uniform distribution of values between 0.15 and 0.35, which are typical values for PDEs (see e.g. [87]). For the Poisson spectra generation, the background rates were multiplied with the detector efficiency of the respective pixel. Each data set was fitted with the linear model described in subsection 5.3.4 and the model including the detector efficiencies to compare both the efficiency model with the true values and to estimate the linear model's error when dealing

with data in which pixels have nonuniform efficiencies. The simulations were run for one setup with 256 pixels and 4 nights for 200 times using the χ^2 minimization method. An overview of the results are shown in Table 5.7. It is

Table 5.7: Overview of fit results for toy simulation with varying detector efficiencies. The results are presented as error of estimation $\Delta\lambda$ as described in Equation 5.9.

	Linear model	Efficiency model
$\Delta\lambda_{\mathrm{bg}}$	$0.77^{+0.03}_{-0.05}$	$0.33^{+0.13}_{-0.12}$
$\Delta\lambda_{\mathrm{dc}}$	$0.60^{+0.50}_{-0.29}$	$0.48^{+0.69}_{-0.35}$
$\Delta\lambda_{\epsilon}$	—	$0.30^{+0.09}_{-0.06}$

no surprise that the model with the efficiency parameter has smaller errors of estimation, since the data was constructed using this exact model. However, is it notable that both models have similar errors regarding the dark count rate λ_{dc}. When estimating the background rate λ_{bg}, the linear model collapses completely with an error above 70 %. Even the model which takes efficiency into account has a much larger error when compared to the results of the linear model in Table 5.5. There are several reason for this. The most obvious one is the increased amount of fit parameters, which allow for more local minima in the χ^2 minimization and therefore a more complicated fit. Additionally, in the model dark count rate and PDE are treated as two completely independent quantities, which they clearly are not. Both depend on the overvoltage of the SiPM and are therefore correlated, but for a more realistic model even more parameters would be needed, which is unfeasible. For the real data, the assumption of independence between the parameters is stronger, since the efficiency can not only be PDE but also include efficiencies of the optics. The dark count rate in the data can also be enhanced by electric noise which can cause extra triggers and further weakens the correlation.

The fit to the data was performed using the χ^2 minimization, the results for the background rates are summarized in Table 5.8 and the dark count rates and PDE values are shown in Figure 5.22. When comparing the night sky background rates to the results of the linear model (see Table 5.6), it can be seen that except for the second patch, all rates are approximately one order of magnitude higher, but the ratios of rates in different patches stays about the same. From the model, higher background rates make perfect sense, since the rates are multiplied by an efficiency < 1, and the difference in count rate can be distributed not only to the dark count rates but also to the efficiencies.

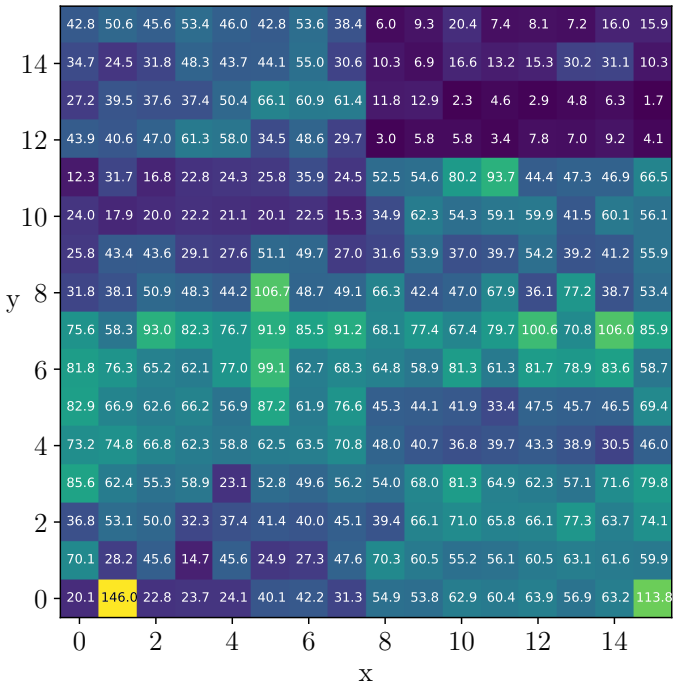
In the linear model, the pixel with the lowest count rate essentially gives an upper bound to the background rate, since pixels cannot have negative dark count rates. The addition of an efficiency mitigates this effect, since said pixel can be given lower efficiency, which results in overall higher background rates.

The same argument can be used to understand the lower dark count rates (see Figure 5.22a) compared to Figure 5.21a. The PDE distribution (see Figure 5.22b) mostly follows the dark count rate distribution. As expected, the hot pixel at coordinate (1, 0) has the highest PDE with 63.6 % and the top right area has the lowest PDE with 5.0 %. Both of these values seem unphysical. The high value of 63.6 % is outright impossible as there are no commercially available SiPM with PDEs in this region. The low values are possible in principle, but only for very low overvoltages, which make it hard to detect any signal at all. To ensure that these extreme values do not alter the best fit values for the other pixels, two subsets of the data were fitted without the extreme pixels. For the first subset (subset_1), the hot pixel at (1, 0) was cut from the data, for the second subset (subset_2) both the hot pixel and the top right area were cut. Results from these fits can be seen in Table 5.8. The only value that changes significantly is the background rate for the second patch, where the subsets produce higher values, although the absolute difference is still small. The other parameters, particularly the PDE values stay the same, which confirms the model's ability to describe the data, even if the interpretation of the efficiency parameter as the true PDE is questionable as it describes a combination of efficiencies, not only the PDE.

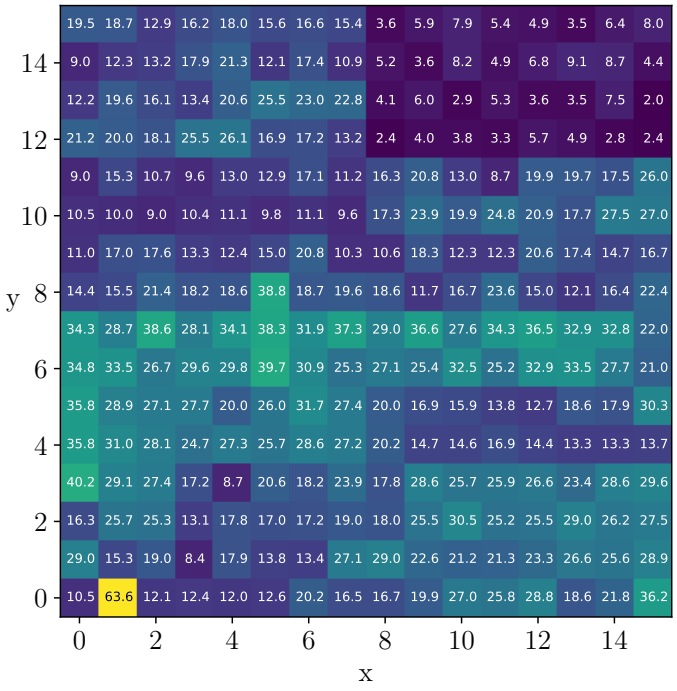
Table 5.8: Counting rates of the night sky background for different "nights" using efficiency models. Nights correspond to the patches defined in Table 5.4. The last row shows the average dark count rate.

Night	λ_{bg} ($\frac{\text{cts}}{\text{GTU}}$) full	λ_{bg} ($\frac{\text{cts}}{\text{GTU}}$) subset_1	λ_{bg} ($\frac{\text{cts}}{\text{GTU}}$) subset_2
1	1.28×10^{-2}	1.26×10^{-2}	1.50×10^{-2}
2	1.62×10^{-3}	1.90×10^{-3}	3.53×10^{-3}
3	3.69×10^{-1}	3.54×10^{-1}	3.64×10^{-1}
4	8.79×10^{-1}	8.44×10^{-1}	8.50×10^{-1}
λ_{dc}	4.80×10^{-2}	4.76×10^{-2}	5.25×10^{-2}

A interesting feature can be seen in some pixels when comparing the dark count rates with the efficiencies (see Figure 5.22). For most pixels, the relative values are similar, which means a pixel with high dark count rate will also have a high efficiency. For some pixels however, this does not hold true, e.g. pixels



(a) Dark count rates in 10^{-3} $\frac{\text{cts}}{\text{GTU}}$. Mean $\lambda_{\text{dc}} = 0.048$ $\frac{\text{cts}}{\text{GTU}}$



(b) PDEs from the data obtained by using the efficiency model in %.

Figure 5.22: Resulting dark count rates and PDE distribution using the efficiency model.

with (x, y) coordinates (10, 11) and (11, 11). Both pixels have a relatively high dark count rate, but show very low efficiencies. This hints at a time-dependent rate, which was only present in these pixels and not in the rest of the detector, i.e. transient light events, leading to a higher dark count rate.

As pixel (1, 0) and the top right region differ much in count rate from the rest of the detector and cannot be satisfyingly explained due to PDE, other possibilities have to be considered. Since the threshold of the detector is not known precisely, the next chapter will discuss the influence of the detector threshold to the data.

5.3.6 Thresholds

So far the detector threshold has not entered into the discussion of the data and the reconstruction of the rates. Implicitly it was assumed that whenever an electron-hole pair (and therefore a charge carrier avalanche) was created, the SiPM would generate a signal. Since in a ideal finger-spectrum of a SiPM, there is a distinct valley between the pedestal and the 1 p.e. peak, a threshold of 0.5 p.e. was used implicitly for the above discussion. In subsection 5.2.2 a time window of 10 ns was used to estimate the loss of photons due to random coincidence. The same principle can be used to describe the expected rate under the influence of a (integer[5]) threshold. The following estimations are based on timing coincidence, the contribution of optical crosstalk is not included.

Due to the constant temperature, the dark count rate can be modeled as a Poisson process with a constant rate λ_{dc}. For the night sky background the real source spectrum is unknown, but the "nights" used for the analysis were constructed in a way which guarantees an almost constant rate (see subsection 5.3.1), therefore the background rate can be modeled as constant as well. Whenever a single 1 p.e. signal is created, an interval $T = 10\,\mathrm{ns}$ exists for coincidence to reach a particular threshold. Since for a 0.5 p.e. threshold, every signal is above threshold, this is trivial. Considering two constant rates r_1 and r_2, the trigger rate R becomes

$$\begin{aligned} R_{0.5} &= r_1 \cdot P(\geq 0 \text{ in } T) + r_2 \cdot P(\geq 0 \text{ in } T) \\ &= r_1 + r_2, \end{aligned} \tag{5.17}$$

where $P(\geq 0 \text{ in } T)$ is the probability to detect 0 or more signals in T, which of course is 1. Since r_1 and r_2 are rates, a single count, which starts the

[5]Integer in this case means that the difference between thresholds can only be 1, although the particular thresholds are applied in between two integer bins in the spectrum, i.e. 0.5 and 1.5.

coincidence interval, happens $(r_1 + r_2)$ times per unit of time. For a threshold of 1.5 p.e., the trigger rate R can be expressed in a similar way

$$\begin{aligned} R_{1.5} &= r_1 \cdot P(\geq 1 \text{ in } T) + r_2 \cdot P(\geq 1 \text{ in } T), \\ &= (r_1 + r_2) \cdot P(\geq 1 \text{ in } T) \\ &= (r_1 + r_2) \cdot \big(1 - \mathcal{P}_{r_1 \cdot T}(0) \cdot \mathcal{P}_{r_2 \cdot T}(0)\big), \end{aligned} \tag{5.18}$$

where $P(\geq 1 \text{ in } T)$ is the probability to detect 1 or more signals in T and $\mathcal{P}_r(k)$ denotes the probability to measure 0 events for a Poisson variable with mean value r. For each single signal, a second signal is needed to reach the trigger threshold. This signal can either come from r_1 or r_2, and can be expressed by $1 -$ *probability of measuring nothing at all from both variables in* T, which is $1 - \mathcal{P}_{r_1 \cdot T}(0) \cdot \mathcal{P}_{r_2 \cdot T}(0)$. Extended to a 2.5 p.e. threshold, the rate becomes

$$\begin{aligned} R_{2.5} &= (r_1 + r_2) \cdot P(\geq 2 \text{ in } T) \\ &= (r_1 + r_2) \cdot \Big(1 - \Big(\mathcal{P}_{r_1 \cdot T}(0) \cdot \mathcal{P}_{r_2 \cdot T}(0) \\ &\qquad + \mathcal{P}_{r_1 \cdot T}(0) \cdot \mathcal{P}_{r_2 \cdot T}(1) + \mathcal{P}_{r_1 \cdot T}(1) \cdot \mathcal{P}_{r_2 \cdot T}(0)\Big)\Big), \end{aligned} \tag{5.19}$$

because two additional signals are needed to reach the threshold, which is the inverse of the probability to measure either 0 from both rates, exactly 1 from r_1 and 0 from r_2, or vice versa. This leads to an expression for a $k.5$ p.e. threshold

$$R_{k.5} = (r_1 + r_2) \cdot \left(1 - \sum_{i=0}^{k-1} \sum_{j=0}^{k-1-i} \mathcal{P}_{r_1 \cdot T}(i) \cdot \mathcal{P}_{r_2 \cdot T}(j)\right). \tag{5.20}$$

With this expression, it is possible to examine changes in integer threshold to the observed rate by the detector using expected values. From the PDM, $\sim 2 \frac{\text{cts}}{\text{pixel GTU}}$ are expected for the background rate λ_{bg}. For the dark counts, an expected rate $\lambda_{\text{dc}} \sim 0.05 \frac{\text{cts}}{\text{GTU}}$ was constructed for a trigger threshold of 0.5 p.e. in subsection 5.3.4. Although the fitted values seem to be higher, this value can be used as an estimate of the trigger threshold. The trigger rates at certain thresholds are shown in Figure 5.23. Due to the nature of timing coincidence and the narrow coincidence window of $T = 10$ ns, the trigger rates drop rapidly with increasing threshold, more than 10^{-2} per threshold step. From this and the measured trigger rates ($\sim 10^{-2}$, see subsection 5.3.4), a threshold of 1.5 p.e. seems the most likely for SiECA.

One apparent feature of the rate distribution over the detector is the dark area in the top right and the hot pixel at (1, 0). The strong deviation from the mean trigger rate could be explained by introducing different trigger thresholds

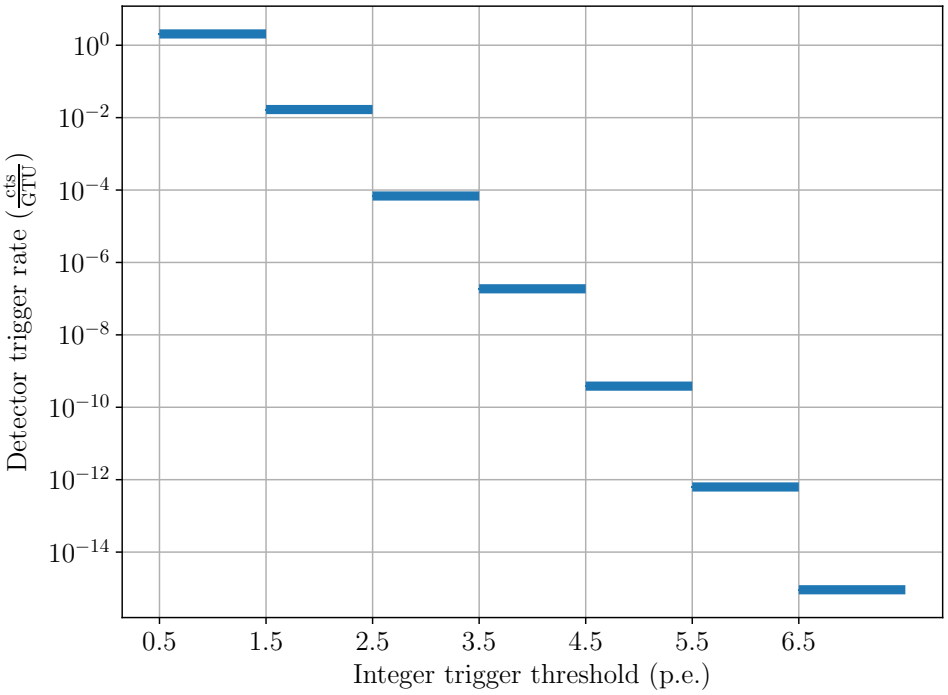


Figure 5.23: Detector trigger rates for varying thresholds with $\lambda_{\text{bg}} = 2\,\frac{\text{cts}}{\text{GTU}}$ and $\lambda_{\text{dc}} = 0.05\,\frac{\text{cts}}{\text{GTU}}$. The short coincidence time window of $T = 10\,\text{ns}$ causes the rate to drop rapidly per threshold.

$\mathcal{T}$ to the model. For a mean threshold $\overline{\mathcal{T}} = n.5\,\text{p.e.}$, the top right area has a higher threshold $\mathcal{T}_{\text{top}} = n.5 + 1\,\text{p.e.}$ resulting in a lower trigger rate and the hot pixel at (1, 0) a lower threshold $\mathcal{T}_{(1,\,0)} = n.5 - 1\,\text{p.e.}$. This threshold effect can be introduced into both the linear model and the efficiency model. However, instead of describing the measured spectrum as a sum of Poisson spectra like in Equation 5.12, each spectrum is described by a single Poisson spectrum with altered rate

$$\mathcal{P}(\lambda_{\text{i, n}}) = \mathcal{P}\big(R_{\mathcal{T}_i}(\lambda_{\text{dc}_i}, \lambda_{\text{bg}_n})\big)\ \forall i \in \{\text{pixels}\}, \forall n \in \{\text{nights}\}, \tag{5.21}$$

where $R_{\mathcal{T}_i}(\lambda_{\text{dc}_i}, \lambda_{\text{bg}_n})$ denotes the trigger rate as described in Equation 5.20 with rates $(r_1, r_2) = (\lambda_{\text{dc}_i}, \lambda_{\text{bg}_n})$. For the efficiency model, r_2 is additionally multiplied with the efficiency parameter $r_2 = \epsilon_i \cdot \lambda_{\text{bg}_n}$. Since the efficiency parameters were introduced to explain the difference in trigger rates between the regions, for the fit they serve a similar purpose.

Fits of the modified model without efficiencies to the data were carried out for two different average thresholds, $\overline{\mathcal{T}} = 1.5\,\text{p.e.}$ and $\overline{\mathcal{T}} = 2.5\,\text{p.e.}$ For

comparison of the resulting rates, it is not convenient to give every dark count rate like in Figure 5.21, therefore the given values are summarized to 7 rates: The night sky background rates of the four nights λ_{bg_i}, one rate $\lambda_{(1,\,0)}$ for the hot pixel at (1, 0), one average value for the top right area λ_{top} and one value for the mean of the remaining pixels λ_{dc}. This summary can be seen in Table 5.9. As expected, the model with higher thresholds produces higher values for the rates. The values for the night sky background rates follow the same trends as in the linear model or the efficiency model, in that the ratio of the fourth to the third patch is ~ 2, even if the second patch rate in the $\overline{\mathcal{T}} = 2.5\,\mathrm{p.e.}$ case is surprisingly low. For the dark count rates the introduction of thresholds reverts count rates qualitatively. The hot pixel at (1, 0) which has the highest trigger rates in the models without thresholds now shows the lowest value because of its low threshold, and the top right area needs high rates due to high threshold.

The same trend can be seen when adding efficiencies to the threshold model, see Table 5.10. The hot pixel, which was fitted with a 63.6 % PDE above receives the lowest PDE with 3.1 % and the top right area goes from 5.0 % to 58.0 % in the $\overline{\mathcal{T}} = 1.5\,\mathrm{p.e.}$ case. These values seem unphysical since such extreme PDEs are not likely, in particular not for the top right area where the upper half of the pixels of the same SiPM needed to show much higher PDE than the lower pixels on the same device. It is worth remembering that all camera pixels share one threshold value and the different thresholds in the model are created by different gains. To have a higher threshold, the top right area needed to have low gain, which is in tension with the high dark count rates and PDEs which are fitted, especially with regards to the lower half of the same device. This leads to the conclusion that the threshold model does

Table 5.9: Resulting rates for the SiECA data set fitted with the threshold model without efficiencies. The rates are given in $\frac{\mathrm{cts}}{\mathrm{GTU}}$.

	$\overline{\mathcal{T}} = 1.5\,\mathrm{p.e.}$	$\overline{\mathcal{T}} = 2.5\,\mathrm{p.e.}$
λ_{bg_1}	0.039	0.232
λ_{bg_2}	0.028	0.0016
λ_{bg_3}	1.41	6.66
λ_{bg_4}	3.01	12.7
$\lambda_{(1,\,0)}$	0.048	4.93
λ_{top}	12.9	33.2
λ_{dc}	4.0	18.9

Table 5.10: Resulting rates for the SiECA data set fitted with the threshold model including efficiencies. The rates are given in ($\frac{\text{cts}}{\text{GTU}}$).

	$\overline{\mathcal{T}} = 1.5\,\text{p.e.}$	$\overline{\mathcal{T}} = 2.5\,\text{p.e.}$
λ_{bg_1}	0.52	4.62
λ_{bg_2}	0.20	3.81
λ_{bg_3}	8.19	22.7
λ_{bg_4}	15.9	38.8
$\lambda_{(1,\,0)}$	0.137	5.32
λ_{top}	10.2	28.4
λ_{dc}	3.55	17.5
$\epsilon_{(1,\,0)}$	3.1	20.2
ϵ_{top}	58.0	53.9
ϵ_{dc}	25.9	35.9

not describe the data properly. This conclusion may not be surprising, since the model uses integer thresholds, which are a simplification valid only for extremely thin peaks in the finger spectrum. In reality, the single peaks in the finger spectrum are Gaussian distributed with a certain width σ, which can lead to overlap in the tails of different peaks depending on the noise of the detector. In such a case, the threshold can not only be set to integer values, but on a continuum between peaks. In a perfect model, each pixel's gain, dark count rate and detector noise would be modelled, and the threshold would be determined based on one fixed value by gain and noise alone. However, this would result in even more parameters to fit and is therefore not feasible with the available data set and was thus not attempted.

5.3.7 Subset Fit

All of the three presented models, namely the linear model, efficiency model and the threshold model suffer from the same problem. They are unable to properly deal with the strong differences in rate of the average pixel, the hot pixel and the top right area. Both the efficiency model and the threshold model try to solve this problem, but ultimately fail as they result in unphysical values. However, there is a common pattern in the results of all the models. The ratio of $\lambda_{\text{bg},4}/\lambda_{\text{bg},3}$ has a value of ~ 2.5 in every model. This suggests that there is indeed a significant change in triggers between the third and fourth night. As the values for the first two nights are dominated by the extreme pixels, one can look at the "average" subset of the data, excluding the hot pixel

and top right area. This was already done for the efficiency model as a test of robustness. Since the efficiency was introduced to bridge the gap between the extreme regions and the average pixels, it makes sense to use this subset also for the linear model. As the addition of the efficiency parameters weakens the model's ability to separate background and dark count rates (compare Table 5.7), using the simpler linear model is preferred. Additionally, the average subset most closely resembles the ideal data set used for verification of the method in subsection 5.3.4. The fit to the average subset of the linear model is shown in Table 5.11 as well as the fit to the full data set carried out in subsection 5.3.4 for comparison.

Table 5.11: Resulting rates of the linear model fitted to only the average subset of the data. For comparison, the fit to the full data is given in the first column.

Night	λ_{bg} ($\frac{\mathrm{cts}}{\mathrm{GTU}}$) full	λ_{bg} ($\frac{\mathrm{cts}}{\mathrm{GTU}}$) subset
1	1.44×10^{-3}	1.07×10^{-2}
2	9.97×10^{-4}	9.25×10^{-3}
3	5.96×10^{-2}	7.96×10^{-2}
4	1.63×10^{-1}	1.88×10^{-1}
λ_{dc}	5.61×10^{-2}	4.95×10^{-2}

From the results in Table 5.11 the fit to the subset finds a slightly lower dark count rate, but increased background rates. In the case of the first two nights, the values are increased by one order of magnitude. This behaviour was also seen in the efficiency fit and is most likely caused by the exclusion of the top right area, which artificially increases the dark count rate due to lower photon detection efficiency. The result suggests that for the first two nights, the count rates are still dominated by dark counts and even in the third and fourth night, the dark counts contribute to the total count rate in a significant manner. Since the subset most closely resembles the data set of the verification (see section 5.3.4), these values can be considered the most trustworthy and will be used for the comparison to the PDM in the following section. Overall, the dark count rates seem to be quite large, as a rate of $\sim 5 \times 10^{-2}\ \frac{\mathrm{cts}}{\mathrm{GTU}}$ is expected for the SiPM at 0.5 p.e. threshold. Since the actual threshold is clearly higher, a decreased dark count rate is expected. A contribution of counts due to interference (see section 5.4) could explain the increased rate. In the final subset fit, the average dark count rate is ~ 5 times higher than the lowest UV background rate. As SiPM dark count rates are likely to decrease even more in the following years (see section 3.4), and the dark count rate is arguably

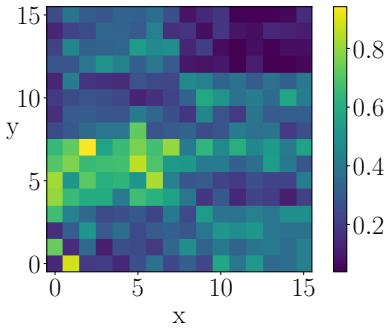
inflated, one can be optimistic that the dark count is not prohibitive for CR experiments using SiPM.

5.4 Comparison of SiECA and PDM data

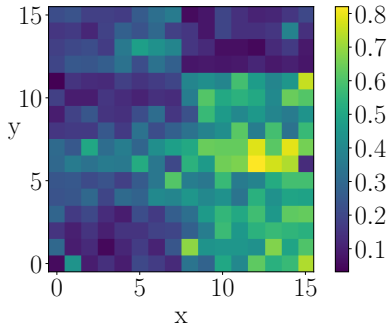
Due to the low statistics of the SiECA data set, it is difficult to obtain with certainty absolute values for the measured rates and efficiencies. Due to its nature as an auxiliary device however, the performance of SiECA can be measured with respect to the PDM. This can be achieved by studying events, which happen in both detectors at the same time, or transition from one detector to the other. Ideally, such an event would be a CR which passes through the PDM and goes through the SiECA field of view. Apart from cosmic rays, other events which do not change in brightness from one detector to the other also qualify. These events were classified as *transient light events* in subsection 5.2.1 and can be high clouds or light from a plane passing through the field of view. Apart from transient light events, there is a light source intrinsic to EUSO-SPB, which might qualify. The PDM was monitored using a bright health LED, which fired every 16 s. Since it provides a very bright signal, it makes an ideal tool for comparison. Due to the SiECA random trigger however, no signal of the health LED could be found in the data.

5.4.1 Transient Light Events

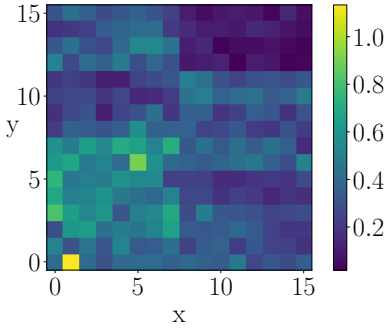
In subsection 5.3.2, a criterion was developed to find transient light events as a means to cut them and focus on clean background data. Using this cut, three patches of data recording can be made out, which contained significant transient light events, one in the first record, on April 25th, and one in each record on April 28th. An example of a spectrum which contains transient light events can be seen in Figure 5.18f. To resolve these transient light events, camera images have to be averaged over the duration of one trace, 128 GTUs. Both the first transient light event on April 25th and the last transient light event on April 28th seem to be static in camera space with decreasing trigger intensity. However, the lack of movement could be due to the small amount of recorded traces per transient light event. An example of such a transient light event can be seen in Figure 5.24a. For the remaining transient light event on April 28th, more data is available and a blob of light can be seen moving from the center toward the bottom right out of the field of view and a second blob can be seen emerging from the bottom left and moving towards the top right in consecutive camera image averages. The transient light event consists of



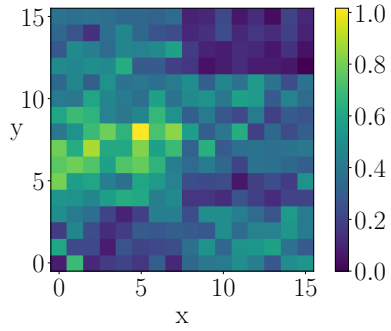
(a) Transient light event on 2017–04–25 12:27:54.



(b) Transient light event on 2017–04–28 07:35:18.



(c) Transient light event on 2017–04–25 07:37:47.



(d) Transient light event on 2017–04–25 07:38:32.

Figure 5.24: Examples of transient light events. The third and fourth image show the same event, ~ 1 min apart.

25 traces of 128 GTUs and lasts about 10 min, several snapshots can be seen in Figure 5.24b to Figure 5.24d. The event itself is separated into two parts of consecutive triggers with 15 s between traces, separating the first blob of light and the second blob of light by ~ 2 min. Since clouds recorded by the PDM showed a similar velocity ($\sim 0.5 \frac{\text{EC}}{\text{GTU}}$) and distribution of light [17], it is plausible that this transient light event shows clouds as well.

To search for the corresponding event in the PDM, the timestamps of the SiECA transient light event can be matched with the closest timestamps in the PDM data, and the same 128 GTU averages can be taken, which results in 25 camera images from the PDM. Due to the slow velocity of the clouds, the small time resolution of ~ 1 s of SiECA does not pose a problem. An example of such an camera image can be seen in Figure 5.23 During the time of this transient light event, only 2 of the 9 ECs of the PDM were turned on, which

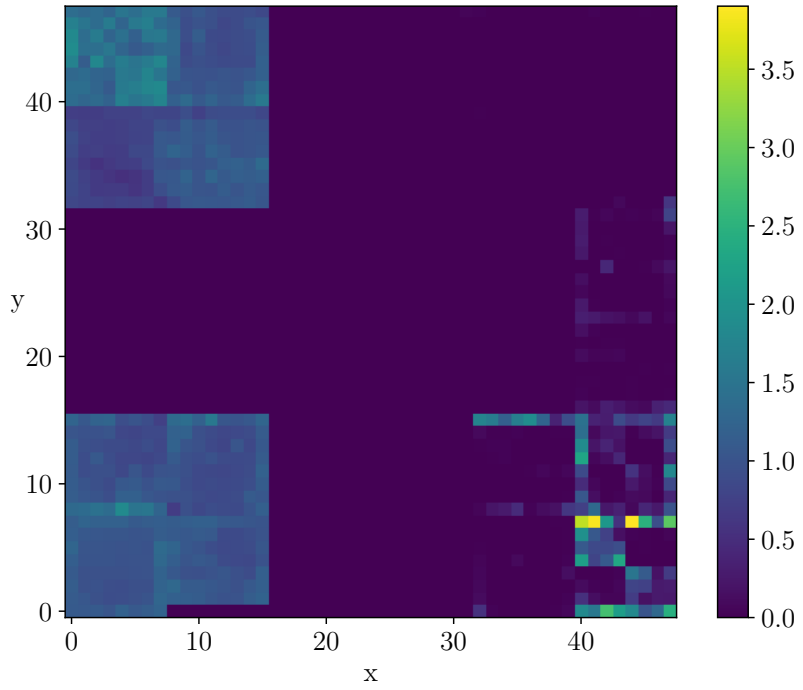


Figure 5.25: PDM image during the time of potential cloud event in SiECA.

makes matching the blobs of light near impossible. However, the patterns of movement seen in the two working ECs of the PDM support the assumption of moving clouds in the field of view. This proves that SiECA is indeed able to measure physical events.

5.4.2 Night Sky Background

Another way to compare the performance of SiECA and the PDM is by looking at the night sky background itself. The night sky background light is uniform on a per-night-basis not only for all SiECA pixels, but for the PDM as well. By comparing the trigger rates from the night sky background in SiECA with the overall trigger rates in the PDM, which does not suffer from dark counts, one can compare both devices' performance in measuring UV light.

In the above section, cuts were defined on the data to obtain a data set which contains as little as possible *transient light events* and four patches to use for the night sky background analysis (see Table 5.4). Since some PDM ECs were shut off during the time (see Figure 5.25), care has to be taken

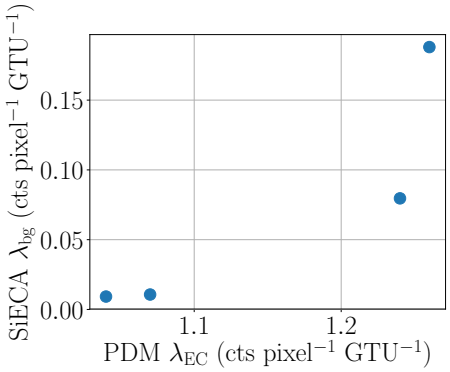
Table 5.12: Comparison of SiECA λ_{bg} with top left PDM EC trigger rate λ_{EC}. Additionally, the fraction of active pixels of the PDM φ_{px} is given.

Patch	λ_{bg} $(\frac{\text{cts}}{\text{GTU}})$	λ_{EC} $(\frac{\text{cts}}{\text{GTU}})$	φ_{px}
1	1.07×10^{-2}	1.07	0.23
2	9.25×10^{-3}	1.04	0.23
3	7.96×10^{-2}	1.24	0.57
4	1.88×10^{-1}	1.26	0.88

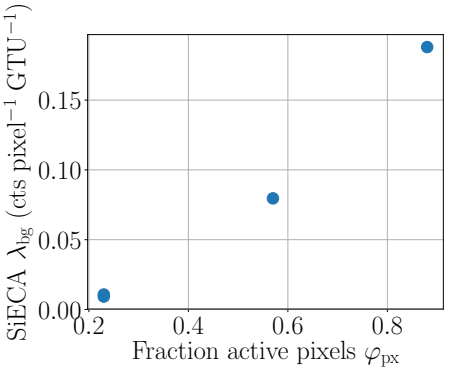
to only use ECs, which were active for every patch of the SiECA data set. The amount of active ECs varies over the time of recording, from two ECs as in Figure 5.25, to every EC working for later timestamps. Since the top left EC is the only one which is active for whole duration, data from this EC is chosen as a reference to compare SiECA against. For SiECA, the results from the linear model to the average subset were chosen (see Table 5.11), since the problematic differences in rate of the extreme pixels are not present here. The comparison of the SiECA background rates with the trigger rates of the top left EC in the PDM can be seen in Table 5.12.

In an ideal case with a perfect threshold of 0.5 p.e., these values should correlate heavily and change in a similar way. Qualitatively, this is the case, as the ordering of the values is the same for both detectors. Quantitatively however, the values differ dramatically, the SiECA rates λ_{bg} span multiple orders of magnitude while the PDM values λ_{EC} only change for $\sim 20\,\%$. A plot showing the values' correlation can be seen in Figure 5.26a. The Pearson correlation efficient between both value sets is $\rho_{\text{bg,EC}} = 0.89$, which hints at a connection between the values. Due to the difference in magnitude of the values it is clear however, that the values obtained by the efficiency model λ_{bg} are not easily compatible with the PDM values λ_{EC}. The strong deviations in magnitude between λ_{bg} and λ_{EC} could possibly be explained with a more comprehensive threshold model, which includes Gaussian distributed p.e. peaks, since the simpler integer threshold model in subsection 5.3.6 showed signs of being very sensitive to small changes in one of the two rates, i.e. the background rate. However, this is not in the scope of this work and was not investigated further.

A remarkable feature in Table 5.12 is that the rate λ_{bg} increases as a higher fraction of PDM pixels were active. Since there was interference between the two cameras, it is possible that the increasing amounts of triggers in SiECA is due to increasing interference with the PDM when more pixels are active. A plot showing this correlation can be seen in Figure 5.26b. These two value sets



(a) Correlation plot between the PDM trigger rate in the top left EC and the fitted SiECA background rate. The two sets have a correlation coefficient of $\rho_{bg,EC} = 0.89$.



(b) Correlation plot between the fraction of active pixels and the fitted SiECA background rate. The two sets have a correlation coefficient of $\rho_{bg,\varphi_{px}} = 0.99$.

Figure 5.26: Correlation plots for SiECA background rate λ_{bg}.

are strongly correlated with a Pearson correlation coefficient of $\rho_{bg,\varphi_{px}} = 0.99$. This makes it likely that the rates λ_{bg} are affected by the interference. Due to the low statistics of only four data points, it is very difficult to ascribe the change in rates to one particular of the two effects or to estimate the relative strengths of both effects. It should be noted that the PDM rate λ_{EC} also shows signs of this correlation with the active pixel fraction, as the Pearson correlation coefficient is quite high $\rho_{EC,\varphi_{px}} = 0.93$.

5.5 Conclusions and Outlook

In this chapter, a thorough look into the SiECA data set of the EUSO-SPB flight was presented including classification of pathological events, estimation of the UV night sky background, the search for interesting time-dependent phenomena and a brief comparison to the PDM data.
At first, an overview was given of the SiECA data set by means of trigger rate histograms for single GTUs. From this, different classes of unphysical data were identified and methods were developed to reliably cut these events from the physical data. Most notably, these include spike events which show a distinct signature in time and in the images they produce, as well as pattern events which happen at seemingly random times and show a distinct pattern in the images. Using the developed cuts, a subset of valid data was constructed, which was used as a basis for physics analysis.

A model describing the camera spectra through Poisson distributions was developed, which separates the counts into temperature dependent dark counts and time dependent UV background counts. Cuts were developed to identify data subsets with and without the presence of time dependent light sources. Using different subsets, a method was developed to disentangle the dark count rates from the background rates in the spectra. The feasibility of this method was shown in toy simulations. With this method, rates were extracted from the valid data subset and evaluated. A few adjustments of the model were proposed and carried out in order to better describe the specific features of the data, including the introduction of an efficiency parameter motivated by the SiPM PDE and the addition of detector thresholds. A threshold model was developed and a mean threshold in the order of 1.5 p.e. was estimated. By using a subset of the pixels, which eliminates threshold effects, rates could be extracted for the dark count rate and the UV background rate. While the dark count rate seems unusually high, it is still only ~ 5 times higher than the lowest UV background rate.

A comparison of the performance of SiECA with the PDM was conducted by utilizing time dependent light sources such as clouds. At least two events were found in the SiECA data set with features similar to clouds which were found in the PDM data. Due to the high amount of inactive PDM pixels it was however not possible to directly match a single cloud in both detectors. The constructed UV background light rate in SiECA was compared to the PDM trigger rates and correlations were found between the rates as expected, but also between the amount of active PDM pixels and observed UV background

rates, which hints at noise as a source of the increasing counts. As SiPM dark count rates and UV background rate are comparable even with this inflated dark count rate due to noise, SiPM dark counts are not prohibitive for CR experiments in near space

While the model worked well in the simulations, it had problems to describe the SiECA data set perfectly. This is mostly due to the relatively small data set as fits with more complex models, which describe the intricacies of the detector, were not feasible. A model, which takes into account the Gaussian nature of the p.e. peaks should cope better with the small differences in gain and thus threshold, which are present in the detector. With a larger data set, different values of dark counts for different temperatures could be used, even the suspected correlation between active PDM pixels and count rates could be utilized. Another fact, which the model used in this work does not consider, is the presence of electric noise in the system. Instead of describing a dark count rate and a background rate, the current models describe a static rate and a time dependent rate, both of which could also partially be caused by noise. This would explain e.g. the high dark count rate in presence of a 1.5 p.e. threshold.
With larger statistics and an improved model, it could become possible to benefit from the intrinsic SiPM dark counts and characterize the detector under live conditions.

Chapter 6

Summary and Outlook

Silicon photomultipliers (SiPM) are promising alternatives to conventional photomultipliers for the search for cosmic rays in space or near-space environments. The main goal of this thesis was to evaluate if SiPM are suitable candidates for CR detection from above. Despite their many advantages like lower weight and lower supply voltage, there are in particular two drawbacks of SiPM investigated in this thesis, which could prohibit the use of SiPM for CR detection from space: The strong dependence of SiPM parameters on the ambient temperature and the comparably high dark count rate.

Chapter 4 describes the study of temperature dependence of the SiPM gain, for which a setup was built to measure the gain of one SiPM pixel at different temperatures using a commercial climate chamber. With a fast, digital data acquisition based on a DRS4 chip, the dependence of the temperature on the SiPM gain was measured. To correct for the changes in gain, the supply voltage needs to be adjusted according to the temperature. For this, a correction factor was determined from the data to stabilize the gain by measuring the current temperature. An algorithm was developed to reliably and efficiently measure the SiPM gain in real-time over a large temperature range, based on an abstraction of the standard practice to measure histograms of SiPM signals to determine the gain. It exploits the presence of dark counts by using them for continuous gain measurement. The results of the algorithm were verified with a conventional method to measure the gain and were in good agreement. The availability of real-time gain measurements allows a novel way of gain stabilization, which does not rely on external (and potentially error-prone) temperature measurements: A PID controller uses the real-time gain measurements to stabilize the gain to a given value. Using the climate chamber, the novel gain stabilization scheme was shown to work well, especially for high temperature changes. A comparison between the different methods was car-

ried out, where the novel method provided a better stabilization ($\sim 3\times$ smaller gain fluctuations) than the conventional method using external temperature measurements.
To investigate the viability of the novel method in a near-space experiment, the temperature changes in the climate chamber were compared to the temperature changes measured in the EUSO super pressure balloon (EUSO-SPB) flight. Considering the different dark count rates of the different SiPM and lower overall ambient temperature, the temperature changes in the climate chamber were still $\sim 3\times$ higher than in the EUSO-SPB experiment. Since the gain stabilization worked without any problems in the climate chamber, the novel method can handle the temperature changes in EUSO-like experiments.

So far, the novel method of gain stabilization can be considered as a proof of concept. In order to work reliably in an experiment with many pixels, there are still some optimizations to be done, e.g. the current implementation does not take into account the temperature dependence of the dark count rate, which leads to more uncertainty in the gain measurement at high temperatures. Furthermore, the method relies on a purely digital approach to data acquisition, which can be prohibitive for experiments with a tight power budget, such as balloon experiments.

In chapter 5, data from the SiPM elementary cell addon (SiECA) of the EUSO-SPB pathfinder were analysed, which are the first data of a SiPM camera for UHECR search in near-space. It was distributed over 7 separate data sets. Cuts were defined to discard the unphysical data, e.g. data where the SiPM were not voltage biased or certain pixels showed unphysical high counts. This included the construction of a noise pattern, which could be eliminated from the data with a normalized cross-correlation method. The remaining data showed no cosmic-ray-like event. An algorithm was developed to distinguish the 7 data sets as either static (4) or including transient events (3), based on the relative brightness across the detector and the spectral shape of the data set.
The counts of the 4 static data sets consist of SiPM dark counts and UV background. To disentangle the two rates, a model was developed, which uses the fact that the counts are Poisson distributed. It exploits the consistency of per-pixel dark count rates due to constant ambient temperature throughout recording, and the uniformity of the UV background across the detector for each data set. Toy simulations were carried out to evaluate different fit methods, and both a *least squares* and *maximum likelihood* approaches worked for extracting the different rates. The rates extracted from the data show that

dark count rates and UV background rates are in the same order of magnitude ($\sim 10^{-2}$ cts/2.5 µs/pixel). Furthermore, the inclusion of threshold effects to the model allows estimation of the unknown SiECA pixel threshold to ~ 1.5 p.e. For the 3 data sets including transient events, two moving light dots were identified by averaging the recorded traces, which resemble the movement of clouds seen in the EUSO-SPB main camera. A comparison with the EUSO-SPB main camera confirms correlations to the SiECA count rates and hints at electrical crosstalk of some sort as an additional source of counts in both cameras.
From the available data, it can be concluded that the SiECA prototype mainly worked as intended. All pixels were functional and the camera was able to record atmospheric UV background light and moving light with features characteristic for clouds. A comparison of the dark count rates and UV background rate demonstrates that SiPM dark counts are not prohibitive for the detection of CR from near-space using SiPM.

There is still huge potential for CR observation from above. A second balloon flight is planned as a direct successor to EUSO-SPB, called EUSO-SPB2. It will have an improved camera with Schmidt design optics and observe in tilt mode instead of nadir mode. This allows the observation of nearly horizontal extensive air showers. Furthermore, EUSO-SPB2 will feature an increased time resolution to detect Cherenkov light in addition to fluorescence light. EUSO-SPB2 will work as a pathfinder for the ambitious probe of extreme multi-messenger astrophysics (POEMMA) mission, which will consist of two twin satellites flying in formation using both a fluorescence camera and a Cherenkov camera, which will use SiPM as light detectors. Beside the observation of charged CR with increased exposure, it will detect γ-rays and the Cherenkov radiation of up-going τ decays from cosmogenic ν_τ neutrinos.

Bibliography

[1] L. Bergström and A. Goobar, *Cosmology and Particle Astrophysics (Springer Praxis Books)*. Springer, 2006.

[2] **Planck** Collaboration, Y. Akrami *et al.*, "Planck 2018 results. I. Overview and the cosmological legacy of Planck," arXiv:1807.06205 [astro-ph.CO].

[3] S. M. Faber and J. S. Gallagher, "Masses and mass-to-light ratios of galaxies," *Annual Review of Astronomy and Astrophysics* **17** no. 1, (1979) 135–187, https://doi.org/10.1146/annurev.aa.17.090179.001031.

[4] T. Banks, D. B. Kaplan, and A. E. Nelson, "Cosmological implications of dynamical supersymmetry breaking," *Phys. Rev. D* **49** (1994) 779–787. https://link.aps.org/doi/10.1103/PhysRevD.49.779.

[5] M. Takeda *et al.*, "Energy determination in the Akeno Giant Air Shower Array experiment," *Astropart. Phys.* **19** (2003) 447–462, arXiv:astro-ph/0209422 [astro-ph].

[6] C. B. A. McCusker and M. M. Winn, "A new method of recording large cosmic-ray air showers," *Il Nuovo Cimento* **28** (1963) 175–178.

[7] C. J. Bell *et al.*, "The upper end of the observed cosmic ray energy spectrum," *Journal of Physics A: Mathematical, Nuclear and General* **7** no. 8, (1974) 990. http://stacks.iop.org/0301-0015/7/i=8/a=007.

[8] R. Baltrusaitis *et al.*, "The utah fly's eye detector," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **240** no. 2, (1985) 410 – 428. http://www.sciencedirect.com/science/article/pii/0168900285906588.

[9] P. Sokolsky and G. B. Thomson, "Highest Energy Cosmic Rays and results from the HiRes Experiment," *J. Phys.* **G34** (2007) R401, arXiv:0706.1248 [astro-ph].

[10] J. Blümer, R. Engel, and J. R. Hörandel, "Cosmic rays from the knee to the highest energies," *Progress in Particle and Nuclear Physics* **63** no. 2, (2009) 293–338. https://doi.org/10.1016/j.ppnp.2009.05.002.

[11] B. R. Dawson, M. Fukushima, and P. Sokolsky, "Past, Present and Future of UHECR Observations," *PTEP* **2017** no. 12, (2017) 12A101, arXiv:1703.07897 [astro-ph.HE].

[12] **Pierre Auger** Collaboration, J. A. Chinellato and the Auger Collaboration, "Reviewing recent results from pierre auger observatory," *Journal of Physics: Conference Series* **485** no. 1, (2014) 012033. http://stacks.iop.org/1742-6596/485/i=1/a=012033.

[13] P. Tinyakov, "Latest results from the telescope array," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **742** (2014) 29 – 34. http://www.sciencedirect.com/science/article/pii/S0168900213014587. 4th Roma International Conference on Astroparticle Physics.

[14] A. Haungs, G. Medina-Tanco, and A. Santangelo, "Special issue on the jem-euso mission," *Experimental Astronomy* **40** no. 1, (2015) 1–2. https://doi.org/10.1007/s10686-015-9483-9.

[15] H. E. Bergeson *et al.*, "Measurement of light emission from remote cosmic-ray air showers," *Physical Review Letters - PHYS REV LETT* **39** (1977) 847–849.

[16] D. Renker and E. Lorenz, "Advances in solid state photon detectors," *Journal of Instrumentation* **4** no. 04, (2009) P04004. http://stacks.iop.org/1748-0221/4/i=04/a=P04004.

[17] **JEM-EUSO** Collaboration, L. Wiencke and A. O. and, "EUSO-SPB1 mission and science," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Oct, 2017. https://doi.org/10.22323/1.301.1097.

[18] **JEM-EUSO** Collaboration, A. Haungs *et al.*, "SiECA: Silicon photomultiplier prototype for flight with EUSO-SPB," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Oct, 2017. https://doi.org/10.22323/1.301.0442.

[19] **HiRes** Collaboration, D. J. Bird *et al.*, "The cosmic-ray energy spectrum observed by the fly's eye," *The Astrophysical Journal* **424** (1994) 491. https://doi.org/10.1086/173906.

[20] **Ice Cube** Collaboration, R. Abbasi, Y. Abdou, and T. A.-Z. and, "Measurement of the anisotropy of cosmic-ray arrival directions with icecube," *The Astrophysical Journal Letters* **718** no. 2, (2010) L194. http://stacks.iop.org/2041-8205/718/i=2/a=L194.

[21] **KASCADE-Grande** Collaboration, W. D. Apel *et al.*, "Kneelike structure in the spectrum of the heavy component of cosmic rays observed with kascade-grande," *Phys. Rev. Lett.* **107** (2011) 171104. https://link.aps.org/doi/10.1103/PhysRevLett.107.171104.

[22] **Particle Data Group** Collaboration, M. Tanabashi *et al.*, "Review of particle physics," *Phys. Rev. D* **98** (2018) 030001. https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[23] **PAMELA** Collaboration, O. Adriani *et al.*, "Pamela measurements of cosmic-ray proton and helium spectra," *Science* **332** no. 6025, (2011) 69–72, http://science.sciencemag.org/content/332/6025/69.full.pdf.

[24] K. Kobayakawa, Y. Sato, and T. Samura, "Acceleration of particles by oblique shocks and cosmic ray spectra around the knee region," *Phys. Rev.* **D66** (2002) 083004, arXiv:astro-ph/0008209 [astro-ph].

[25] P. L. Biermann, N. Langer, E.-S. Seo, and T. Stanev, "Cosmic rays IX," *Astronomy & Astrophysics* **369** no. 1, (2001) 269–277. https://doi.org/10.1051/0004-6361:20010083.

[26] R. Plaga, "A possible universal origin of hadronic cosmic rays from ultrarelativistic ejecta of bipolar supernovae," *New Astronomy* **7** no. 6, (2002) 317 – 336. http://www.sciencedirect.com/science/article/pii/S1384107602001355.

[27] V. S. Ptuskin, "Transport of high energy cosmic rays," *Advances in Space Research* **19** no. 5, (1997) 697–705. https://doi.org/10.1016/s0273-1177(97)00390-6.

[28] G. Giacinti, M. Kachelrieß, and D. V. Semikoz, "The escape model for galactic cosmic rays," *Journal of Physics: Conference Series* **632** (2015) 012094. https://doi.org/10.1088/1742-6596/632/1/012094.

[29] K. Greisen, "End to the cosmic-ray spectrum?," *Physical Review Letters* **16** no. 17, (1966) 748–750. https://doi.org/10.1103/physrevlett.16.748.

[30] G. T. Zatsepin and V. A. Kuzmin, "Upper limit of the spectrum of cosmic rays," *JETP Lett.* **4** (1966) 78–80. [Pisma Zh. Eksp. Teor. Fiz.4,114(1966)].

[31] **HiRes** Collaboration, R. U. Abbasi *et al.*, "First observation of the greisen-zatsepin-kuzmin suppression," *Physical Review Letters* **100** no. 10, (2008) . https://doi.org/10.1103/physrevlett.100.101101.

[32] **Pierre Auger** Collaboration, J. Abraham *et al.*, "Observation of the suppression of the flux of cosmic rays above 4×10^{19}eV," *Phys. Rev. Lett.* **101** (2008) 061101, arXiv:0806.4302 [astro-ph].

[33] **Telescope Array** Collaboration, D. Ivanov, "TA spectrum summary," in *Proceedings of The 34th International Cosmic Ray Conference — PoS(ICRC2015)*. Sissa Medialab, Aug, 2016. https://doi.org/10.22323/1.236.0349.

[34] **Pierre Auger** Collaboration, F. F. and, "The cosmic ray energy spectrum measured using the pierre auger observatory," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Aug, 2017. https://doi.org/10.22323/1.301.0486.

[35] **HiRes** Collaboration, High Resolution Fly'S Eye Collaboration *et al.*, "Measurement of the flux of ultra high energy cosmic rays by the stereo technique," *Astroparticle Physics* **32** (2009) 53–60, arXiv:0904.4500 [astro-ph.HE].

[36] D. Allard, E. Parizot, E. Khan, S. Goriely, and A. V. Olinto, "UHE nuclei propagation and the interpretation of the ankle in the cosmic-ray spectrum," *Astron. Astrophys.* **443** (2005) L29–L32, arXiv:astro-ph/0505566 [astro-ph].

[37] V. de Souza and and, "Testing the agreement between the x_max distributions measured by the pierre auger and telescope array observatories," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Aug, 2017. https://doi.org/10.22323/1.301.0522.

[38] **Pierre Auger Collaboration*** Collaboration, A. Aab and P. Abreu, "Depth of maximum of air-shower profiles at the pierre auger

observatory. ii. composition implications," *Phys. Rev. D* **90** (2014) 122006. https://link.aps.org/doi/10.1103/PhysRevD.90.122006.

[39] R. U. Abbasi *et al.*, "Study of Ultra-High Energy Cosmic Ray composition using Telescope Array's Middle Drum detector and surface array in hybrid mode," *Astroparticle Physics* **64** (2015) 49–62, arXiv:1408.1726 [astro-ph.HE].

[40] **Telescope Array** Collaboration, R. U. Abbasi *et al.*, "Indications of Intermediate-Scale Anisotropy of Cosmic Rays with Energy Greater Than 57 EeV in the Northern Sky Measured with the Surface Detector of the Telescope Array Experiment," *Astrophys. J.* **790** (2014) L21, arXiv:1404.5890 [astro-ph.HE].

[41] A. Haungs, H. Rebel, and M. Roth, "Energy spectrum and mass composition of high-energy cosmic rays," *Reports on Progress in Physics* **66** no. 7, (2003) 1145. http://stacks.iop.org/0034-4885/66/i=7/a=202.

[42] J. F. Carlson and J. R. Oppenheimer, "On multiplicative showers," *Phys. Rev.* **51** (1937) 220–231. https://link.aps.org/doi/10.1103/PhysRev.51.220.

[43] B. Rossi and K. Greisen, "Cosmic-ray theory," *Rev. Mod. Phys.* **13** (1941) 240–309. https://link.aps.org/doi/10.1103/RevModPhys.13.240.

[44] K. Kamata and J. Nishimura, "The Lateral and the Angular Structure Functions of Electron Showers," *Progress of Theoretical Physics Supplement* **6** (1958) 93–155.

[45] H. P. Vankov, N. Inoue, and K. Shinozaki, "Ultrahigh energy gamma rays in the geomagnetic field and atmosphere," *Phys. Rev. D* **67** (2003) 043002. https://link.aps.org/doi/10.1103/PhysRevD.67.043002.

[46] J. Matthews, "A Heitler model of extensive air showers," *Astroparticle Physics* **22** (2005) 387–397.

[47] J. Engel, T. K. Gaisser, P. Lipari, and T. Stanev, "Nucleus-nucleus collisions and interpretation of cosmic-ray cascades," *Phys. Rev. D* **46** (1992) 5013–5025. https://link.aps.org/doi/10.1103/PhysRevD.46.5013.

[48] R. Engel, D. Heck, and T. Pierog, "Extensive air showers and hadronic interactions at high energy," *Annual Review of Nuclear and Particle Science* **61** no. 1, (2011) 467–489. https://doi.org/10.1146/annurev.nucl.012809.104544.

[49] V. F. Hess, "Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten," *Phys. Z.* **13** (1912) 1084–1091.

[50] **LOPES** Collaboration, H. Falcke, "Radio Detection of Ultra-High Energy Cosmic Rays," arXiv:0804.0548 [astro-ph].

[51] P. A. Cherenkov, "Visible luminescence of pure liquids under the influence of γ-radiation," *Dokl. Akad. Nauk SSSR* **2** no. 8, (1934) 451–454. [Usp. Fiz. Nauk93,no.2,385(1967)].

[52] **Pierre Auger** Collaboration, J. Abraham and M. Aglietta, "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** no. 1, (2004) 50 – 95. http://www.sciencedirect.com/science/article/pii/S0168900203033497.

[53] K. Greisen, "Cosmic ray showers," *Annual Review of Nuclear Science* **10** no. 1, (1960) 63–108, https://doi.org/10.1146/annurev.ns.10.120160.000431.

[54] **KASCADE** Collaboration, T. Antoni and W. D. Apel, "Test of high-energy interaction models using the hadronic core of eas," *Journal of Physics G: Nuclear and Particle Physics* **25** no. 10, (1999) 2161. http://stacks.iop.org/0954-3899/25/i=10/a=313.

[55] M. A. K. Glasmacher *et al.*, "The cosmic ray energy spectrum between 10**14-eV and 10**16-eV," *Astropart. Phys.* **10** (1999) 291–302.

[56] I. Frank and I. Tamm, "Coherent visible radiation of fast electrons passing through matter," in *Selected Papers*, pp. 29–35. Springer Berlin Heidelberg, 1991 (1937).

[57] M. Giller, G. Wieczorek, A. Kacperczyk, H. Stojek, and W. Tkaczyk, "Energy spectra of electrons in the extensive air showers of ultra-high energy," *J. Phys.* **G30** (2004) 97–105.

[58] J. W. Fowler *et al.*, "A Measurement of the cosmic ray spectrum and composition at the knee," *Astropart. Phys.* **15** (2001) 49–64, arXiv:astro-ph/0003190 [astro-ph].

[59] **TUNKA** Collaboration, B. K. Lubsandorzhiev, "TUNKA-EAS Cherenkov experiment in the Tunka Valley," *Nucl. Instrum. Meth.* **A595** (2008) 73–76.

[60] **H.E.S.S.** Collaboration, J. A. Hinton, "The Status of the H.E.S.S. project," *New Astron. Rev.* **48** (2004) 331–337, arXiv:astro-ph/0403052 [astro-ph].

[61] D. Ferenc, "The magic gamma-ray observatory," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **553** no. 1, (2005) 274 – 281. http://www.sciencedirect.com/science/article/pii/S0168900205016098. Proceedings of the fifth International Workshop on Ring Imaging Detectors.

[62] T. C. Weekes *et al.*, "VERITAS: The Very energetic radiation imaging telescope array system," *Astropart. Phys.* **17** (2002) 221–243, arXiv:astro-ph/0108478 [astro-ph].

[63] M. Nagano, K. Kobayakawa, N. Sakaki, and K. Ando, "Photon yields from nitrogen gas and dry air excited by electrons," *Astropart. Phys.* **20** (2003) 293–309, arXiv:astro-ph/0303193 [astro-ph].

[64] M. Ave, M. Bohacova, and B. Buonomo, "Spectrally resolved pressure dependence measurements of air fluorescence emission with airfly," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **597** no. 1, (2008) 41 – 45. http://www.sciencedirect.com/science/article/pii/S016890020801276X. Proceedings of the 5th Fluorescence Workshop.

[65] **HiRes** Collaboration, R. U. Abbasi *et al.*, "Search for point-like sources of cosmic rays with energies above 10** 18.5 eV in the HiRes. 1. Monocular data-set," *Astropart. Phys.* **27** (2007) 512–520, arXiv:astro-ph/0507663 [astro-ph].

[66] R. Benson and J. Linsley, "Satellite observation of cosmic ray air showers," *International Cosmic Ray Conference* **8** (1981) 145–148.

[67] **JEM-EUSO** Collaboration, J. H. Adams, "The jem-euso mission: An introduction," *Experimental Astronomy* **40** no. 1, (2015) 3–17. https://doi.org/10.1007/s10686-015-9482-x.

[68] *Pierre Auger Observatory and Telescope Array: Joint Contributions to the 35th International Cosmic Ray Conference (ICRC 2017)*. 2018. arXiv:1801.01018 [astro-ph.HE].

[69] M. Karus, *Development of a Calibration Stand for Photosensors for Extremely High-Energy Cosmic Ray Research*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2016. 51.03.04; LK 01.

[70] S. E. Falk, *Atmospheric Influences on Space-Based Observations of Extremely High-Energy Cosmic Rays*. PhD thesis, 2014. 51.04.02; LK 01.

[71] **JEM-EUSO** Collaboration, J. Jr *et al.*, "An evaluation of the exposure in nadir observation of the jem-euso mission,".

[72] **JEM-EUSO** Collaboration, J. H. Adams *et al.*, "JEM-EUSO observational technique and exposure," *Experimental Astronomy* **40** no. 1, (2014) 117–134. https://doi.org/10.1007/s10686-014-9376-3.

[73] L. M. Barbier *et al.*, "Nightglow: an instrument to measure the earthâs nighttime ultraviolet glowâresults from the first engineering flight," *Astroparticle Physics* **22** no. 5, (2005) 439 – 449. http://www.sciencedirect.com/science/article/pii/S0927650504001720.

[74] **JEM-EUSO** Collaboration, J. H. Adams, "Science of atmospheric phenomena with JEM-EUSO," *Exper. Astron.* **40** no. 1, (2015) 239–251.

[75] A. Guzman, G. Saez-Cano, and K. Shinozaki, "The jem-euso observation in cloudy conditions," *Experimental Astronomy* **40** (2014) .

[76] **JEM-EUSO** Collaboration, P. von Ballmoos *et al.*, "A balloon-borne prototype for demonstrating the concept of jem-euso," *Advances in Space Research* **53** no. 10, (2014) 1544 – 1550. http://www.sciencedirect.com/science/article/pii/S0273117713007527. Cosmic Ray Origins: Viktor Hess Centennial Anniversary.

[77] **JEM-EUSO** Collaboration, F. Bisconti, "EUSO-TA prototype telescope," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **824** (2016) 603–605. https://doi.org/10.1016/j.nima.2015.09.105.

[78] **JEM-EUSO** Collaboration, F. Capel, A. Belov, M. Casolino, and P. Klimov, "Mini-EUSO: A high resolution detector for the study of terrestrial and cosmic UV emission from the International Space Station," *Advances in Space Research* **62** (2018) 2954–2965, arXiv:1709.00405 [astro-ph.IM].

[79] R. J. Hodgkinson, "Impact ionization and quantum efficiency in silicon," *Proceedings of the Physical Society* **82** no. 1, (1963) 58. http://stacks.iop.org/0370-1328/82/i=1/a=308.

[80] R. J. McIntyre, "A new look at impact ionization-part i: A theory of gain, noise, breakdown probability, and frequency response," *IEEE Transactions on Electron Devices* **46** no. 8, (1999) 1623–1631.

[81] T. Kirn *et al.*, "Wavelength dependence of avalanche photodiode (apd) parameters," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **387** no. 1, (1997) 202 – 204. http://www.sciencedirect.com/science/article/pii/S0168900296009904. New Developments in Photodetection.

[82] V. Golovin, "Avalanche photodetector." Russian agency for patents and trademarks, 1998. # 2142175.

[83] Z. Sadigov, "Avalanche detector." Russian agency for patents and trademarks, 1998. # 2102820.

[84] O. Mineev *et al.*, "Scintillator counters with multi-pixel avalanche photodiode readout for the nd280 detector of the t2k experiment," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **577** no. 3, (2007) 540 – 551. http://www.sciencedirect.com/science/article/pii/S0168900207007607.

[85] SensL, *C-Series Low Noise, Blue-Sensitive Silicon Photomultipliers*, 4, 2018. Rev. 3.0.

[86] Hamamatsu, *MPPC (Multi-Pixel Photon Counter) arrays S13361-3050 series*, 8, 2017.

[87] T. Jammer, *Characterization of Silicon Photomultipliers for use in a Camera for Cosmic Rays*. Diploma Thesis, Universität Tübingen, 2014.

[88] W. G. Oldham, R. R. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *IEEE Transactions on Electron Devices* **19** no. 9, (1972) 1056–1060.

[89] A. N. Otte, "SiPM's a very brief review," in *Proceedings of International Conference on New Photo-detectors — PoS(PhotoDet2015)*. Sissa Medialab, Jun, 2016. https://doi.org/10.22323/1.252.0001.

[90] G. Hurkx, H. de Graaff, W. Kloosterman, and M. Knuvers, "A new analytical diode model including tunneling and avalanche breakdown," *IEEE Transactions on Electron Devices* **39** no. 9, (1992) 2090–2098. https://doi.org/10.1109/16.155882.

[91] A. L. Lacaita, F. Zappa, S. Bigliardi, and M. Manfredi, "On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices,".

[92] J. H. Swoger and S. J. Kovacic, "Enhanced luminescence due to impact ionization in photodiodes," *Journal of Applied Physics* **74** (1993) 2565–2571.

[93] Hamamatsu, *Silicone layered TSV MPPC array S13361-3050xx*, 5, 2015.

[94] "Mppc array." https://www.hamamatsu.com/eu/en/product/optical-sensors/mppc/index.html. Accessed: 2018-09-17.

[95] S. Ritt, R. Dinapoli, and U. Hartmann, "Application of the drs chip for fast waveform digitizing," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **623** no. 1, (2010) 486 – 488. http://www.sciencedirect.com/science/article/pii/S0168900210006091. 1st International Conference on Technology and Instrumentation in Particle Physics.

[96] Keithley, *Series 2400 SourceMeter SMU Instruments*, 11, 2017. Rev. 1.

[97] G. Turin, "An introduction to matched filters," *IRE Transactions on Information Theory* **6** no. 3, (1960) 311–329.

[98] M. Renschler *et al.*, "Characterization of Hamamatsu 64-channel TSV SiPMs," *Nucl. Instrum. Meth.* **A888** (2018) 257–267, arXiv:1804.00897 [astro-ph.IM].

[99] A. N. Otte, D. Garcia, T. Nguyen, and D. Purushotham, "Characterization of Three High Efficiency and Blue Sensitive Silicon Photomultipliers," *Nucl. Instrum. Meth.* **A846** (2017) 106–125, arXiv:1606.05186 [physics.ins-det].

[100] Hamamatsu, *C11204-02 Power supply for MPPC Operation Manual*, 4, 2016. Rev. B.

[101] **JEM-EUSO** Collaboration, S. Bacholle, "The EUSO-SPB instrument," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Aug, 2017. https://doi.org/10.22323/1.301.0384.

[102] **JEM-EUSO** Collaboration, P. von Ballmoos, "The EUSO-BALLOON mission," in *Proceedings of The 34th International Cosmic Ray Conference — PoS(ICRC2015)*. Sissa Medialab, Aug, 2016. https://doi.org/10.22323/1.236.0322.

[103] **JEM-EUSO** Collaboration, V. Scotti and G. Osteria, "EUSO-balloon: The first flight," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **824** (2016) 655–657. https://doi.org/10.1016/j.nima.2015.10.070.

[104] Advanced Optics SCHOTT AG, *Optical Filter Glass*, 06, 2017.

[105] S. Ahmad *et al.*, "Spaciroc: a rad-hard front-end readout chip for the jem-euso telescope," *Journal of Instrumentation* **5** no. 12, (2010) C12012. http://stacks.iop.org/1748-0221/5/i=12/a=C12012.

[106] **JEM-EUSO** Collaboration, M. E. Bertaina *et al.*, "The trigger logic of EUSO-SPB and its performance," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Aug, 2017. https://doi.org/10.22323/1.301.0443.

[107] **JEM-EUSO** Collaboration, J. B. Eser *et al.*, "Preflight calibration and testing of EUSO-SPB in the lab and the desert," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Aug, 2017. https://doi.org/10.22323/1.301.0457.

[108] **JEM-EUSO** Collaboration, A. V. Olinto, L. Allen, M. Rezazadeh, and S. M. and, "UCIRC: Infrared cloud monitor for EUSO-SPB," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Aug, 2017. https://doi.org/10.22323/1.301.0436.

[109] J. Fleury *et al.*, "Petiroc and citiroc: front-end asics for sipm read-out and tof applications," *Journal of Instrumentation* **9** no. 01, (2014) C01049. http://stacks.iop.org/1748-0221/9/i=01/a=C01049.

[110] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," *Nucl. Instrum. Meth.* **A389** (1997) 81–86.

[111] R. Falkenstein, *Characterization of Silicon Photomultipliers and scintillator light yield measurements*. Diploma Thesis, Universität Tübingen, 2011.

[112] S. Vinogradov, "Analytical models of probability distribution and excess noise factor of solid state photomultiplier signals with crosstalk," *Nuclear Instruments and Methods in Physics Research A* **695** (2012) 247–251, arXiv:1109.2014 [physics.ins-det].

[113] W. Painter. Personal communication, 2017.

[114] J. P. Lewis, "Fast normalized cross-correlation," in *Vision interface*, vol. 10, pp. 120–123. 1995.

[115] S. van der Walt *et al.*, "scikit-image: image processing in Python," *PeerJ* **2** (2014) e453. http://dx.doi.org/10.7717/peerj.453.

[116] **JEM-EUSO** Collaboration, F. Fenu *et al.*, "Expected number of extensive air showers observable by EUSO-SPB," in *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, Aug, 2017. https://doi.org/10.22323/1.301.0426.

[117] **JEM-EUSO** Collaboration, G. Abdellaoui *et al.*, "Cosmic ray oriented performance studies for the JEM-EUSO first level trigger," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **866** (2017) 150–163. https://doi.org/10.1016/j.nima.2017.05.043.

[118] **Particle Data Group** Collaboration, K. A. Olive *et al.*, "Review of Particle Physics," *Chin. Phys.* **C38** (2014) 090001.

[119] G. Cowan, *Statistical data analysis*. Oxford University Press, USA, 1998.

[120] R. Barlow, "Extended maximum likelihood," *Nuclear Instruments and Methods in Physics Research A* **297** (1990) 496–506.

[121] I. N. Bronstejn, *Taschenbuch der Mathematik*. Harri Deutsch, Frankfurt am Main, 2012.

[122] M. Matsumoto and T. Nishimura, "Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Trans. Model. Comput. Simul.* **8** no. 1, (1998) 3–30. http://doi.acm.org/10.1145/272991.272995.

[123] M. Newville, T. Stensitzki, D. B. Allen, and A. Ingargiola, "LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python," Sept., 2014. https://doi.org/10.5281/zenodo.11813.

[124] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," *eConf* **C0303241** (2003) MOLT007, arXiv:physics/0306116 [physics]. [,186(2003)].

[125] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*. Springer Texts in Statistics. Springer, New York, third ed., 2005.

[126] T. J. Rivlin, "Overdetermined systems of linear equations," *SIAM Review* **5** no. 1, (1963) 52–66, https://doi.org/10.1137/1005005. https://doi.org/10.1137/1005005.

[127] G. Williams, "Overdetermined systems of linear equations," *The American Mathematical Monthly* **97** no. 6, (1990) 511–513.

Danksagung

Zuallererst gilt mein Dank Prof. Josef Jochum, für die Möglichkeit, die Arbeit an seinem Lehrstuhl anzufertigen und dass er mir als Ansprechpartner jederzeit zur Seite stand. Vielen Dank auch an Prof. Tobias Lachenmaier für zahlreiche hilfreiche Ratschläge und Anmerkungen.
Vielen Dank an Jörg Bayer vom IAAT, der mir sowohl bei den Experimenten mit der Klimakammer als auch bei Fragen zur Ballonmission jederzeit helfen konnte.
I want to thank Will Painter from KIT for discussions regarding SiECA and its data.
Ein besonderer Dank geht an Alex, Christian und Alex für ausgiebige, fruchtbare Diskussionen und nicht zuletzt für die Hilfe beim Korrekturlesen der Arbeit sowie an Ann-Kathrin für ein offenes Ohr bei Statistikfragen.
Ein allgemeiner Dank geht an die Mitglieder der beiden Arbeitsgruppen, für einen freundlichen und offenen Arbeitsalltag und für interessante Gespräche bei Kaffee, Kuchen und Bier.
Ich danke Lisa für ihre fortlaufende Unterstützung, auf die ich mich jederzeit verlassen kann.
Zuletzt vielen Dank meiner Familie, die mich herzlich und geduldig auf meinem Weg unterstützt.