Basic characterization of Silicon Photomultipliers for Imaging Atmospheric Cherenkov Technique

by
Sin-iu Ho

supervised by Prof. Dott. Andrea Santangelo
Dr. Gerd Pühlhofer
Dr. Baiyang Bi

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Master of Science
in Physics

Institut für Astronomie und Astrophysik (IAAT)
Eberhard Karls Universität Tübingen
Tübingen, Germany

May 21, 2024

Abstract

The imaging atmospheric Cherenkov technique (IACT) holds a pivotal role in ground-based gamma-ray astronomy. It captures Cherenkov radiation directly induced by very high-energy gamma rays (VHEGRs) from various intriguing astrophysical sources. Traditionally reliant on photomultiplier tubes (PMTs) for light detection, recent advancements in IACT camera designs have introduced the adoption of silicon photomultipliers (SiPMs) as promising alternatives. Institut für Astronomie und Astrophysik Tübingen (IAAT), actively engaged in the Cherenkov Telescope Array (CTA) project, has developed a keen interest in exploring the application of SiPM technology. The objective is to couple SiPMs modules to the well-developed IACT camera, FlashCam, in forthcoming endeavors.

The first half of this study provides a comprehensive review of key concepts pertaining to IACT objectives and technology. It elucidates the operating principles and basic characteristics of SiPMs, highlighting their potential as viable photodetectors in gamma-ray astronomy alongside PMTs.

The latter part of the study involves a preliminary characterization of a sample from the Hamamatsu MPPC S13360 series, covering static, dynamic, and optical aspects. Various performance and noise parameters are assessed, demonstrating close alignment with the reference values.

Despite unfulfilled achievement to develop the a robust and integrated analysis software and experimental setups, this study represents a meaningful step towards a thorough understanding of SiPM characteristics, laying the groundwork for future advancements in VHEGR astronomy through enhanced detector performance and evaluation methodologies.

Contents

List of Figures

List of Tables

List of Acronyms

ADC analog-to-digital converter . . . 27
AGN active galactic nucleus . . . 7
APD avalanche photodiode . . . 46
AQC active quenching circuit . . . 51
ASIC application specific integrated circuit . . . 36
ASTRI Astrofisica con Specchi a Tecnologia Replicante Italiana . . . 35
CCD charge-coupled device . . . 22
CHEC Compact High Energy Camera . . . 35
CGRO Compton Gamma Ray Observatory . . . 11
CMOS complementary metal-oxide-semiconductor . . . 51
CR cosmic ray . . . 1
CTA Cherenkov Telescope Array . . . i
CTAO Cherenkov Telescope Array Observatory . . . 31
DAQ data acquisition . . . 28
DC Davies–Cotton . . . 23
DCR dark count rate . . . 64
DGRB diffuse gamma-ray background . . . 7
DLED Differential Leading Edge Discriminator . . . 103
EAS extensive air shower . . . 2
EBL extragalactic background light . . . 35
EGRET Energetic Gamma Ray Experiment Telescope . . . 11
ENF excess noise factor . . . 57
FACT First G-APD Cherenkov Telescope . . . 23
FBK Fondazione Bruno Kessler . . . 37
FEB front-end board . . . 36
FEE front-end electronics . . . 37

FF fill factor . . . 60
FoV field of view . . . 22
FPM focal plane module . . . 36
FPGA field-programmable gate array . . . 37
FSRQ flat-spectrum radio quasar . . . 38
GCT Gamma-ray Cherenkov Telescope . . . 36
GM-APD Geiger-mode avalanche photodiode . . . 48
GRB gamma-ray burst . . . 7
HAWC High-Altitude Water Cherenkov . . . 15
H.E.S.S. High Energy Stereoscopic System . . . 3
HEGRA High Energy Gamma Ray Astronomy . . . 23
IAAT Institut für Astronomie und Astrophysik Tübingen . . . i
IACT imaging atmospheric Cherenkov technique . . . i
ISM interstellar medium . . . 5
KM2A Kilometre Square Array . . . 15
LAT Large Area Telescope . . . 11
LHAASO Large High Altitude Air Shower Observatory . . . 1
LSB least significant bits . . . 27
LST Large-Sized Telescope . . . 22
MAGIC Major Atmospheric Gamma Imaging Cherenkov . . . 3
MAPM multi-anode photomultiplier . . . 35
MEPhI Moscow Engineering Physics Institute . . . 53
MCA multichannel analyzer . . . 86
MPPC Multi-Pixel Photon Counter . . . 37
MST Medium-Sized Telescope . . . 3
NIF near-infrared . . . 53
NSB night sky background . . . 3
PDE photon detection efficiency . . . 26
PDP Photon Detection Plane . . . 36
PDL picosecond diode laser . . . 106
PHS pulse height spectrum . . . 104
PMT photomultiplier tube . . . i
PQC passive quenching circuit . . . 51
pSCT Prototype Schwardschild–Couder Telescope . . . 37
PSF point spread function . . . 26

Chapter 1

Introduction

Gamma rays are at the highest energy end of the electromagnetic spectrum, with energy generally having a lower limit of several hundred keV[1] and no clear upper limit. In 2021, China's Large High Altitude Air Shower Observatory (LHAASO) observed a gamma ray source with an energy of 1.4 PeV, called a PeVatron[1]. This means that the gamma-ray energy spectrum already spans more than 10 orders of magnitude. Table 1.1 is a widely accepted demarcation of gamma-ray energy regimes in recent years[2].

Table 1.1: Energy regimes in gamma-ray astronomy

Energy regime	Acronym	Energy range
Low-energy	LE	< 100 MeV
High-enegy	HE	100 MeV–100 GeV
Very-high-energy	VHE	100 GeV–100 TeV
Ultra-high-energy	UHE	> 100 TeV

Among these regimes, very high-energy gamma ray (VHEGR) astronomy has developed rapidly in the past one or two decades, and has become one of the main pillars of high-energy astrophysics[3] together with ultra-high-energy cosmic ray (UHECR) astronomy, ultra-high-energy neutrino astronomy and other fields. VHEGR astronomy carries an important task, thas is, to solve the mystery of the origins and acceleration mechanisms of cosmic rays.

Cosmic rays (CRs) were discovered in 1912 by Viktor Hess during a high-altitude balloon

[1]Metric unit prefixes: k = kilo- = 10^3, M = mega- = 10^6, G = giga- = 10^9, T = tera- = 10^{12}, P = peta- = 10^{15}.

experiment. The study of CRs led to the development of particle physics from the 1930s to the 1950s. Among them, positron, muon, pion, and kaon were all discovered through experiments or observations related to CRs[4]. However, 99% of primary CRs are atomic nuclei, and the remaining 1% are made up of electrons and positrons. In other words, all of CRs are charged particles. When they travel through the universe, the magnetic field permeating space will deflect their trajectories through the Lorentz force. This makes it impossible for people to infer the origins of CRs based on their arrival directions. Nevertheless, according to the hadron model of the gamma ray production mechanism (Section 2.2.2), protons in CRs can interact with the interstellar medium where consequently neutral pions (π^0) produce gamma rays, thereby preserving the information about the sites from which CRs could be accelerated. Neutrinos, which are also electrically neutral, play the same role.

How do people detect VHEGRs? In fact, neither gamma rays nor primary CRs can directly reach the Earth's surface because they interact with atomic nuclei in the atmosphere, creating what is known as extensive air showers (EASs). Due to this atmospheric barrier, gamma rays were once the last "unopened window" for astronomical observations in the entire electromagnetic spectrum, forcing people to send instruments carried by satellites or high-altitude balloons to explore gamma rays. However, the energy spectrum of gamma-ray differential flux usually obeys the power law $\mathrm{d}N/\mathrm{d}E \propto E^{-\alpha}$, where $\alpha \approx 2$–3 is the spectral index. This suggests, when the energy is 10 times higher, the flux lowers by 1% to 0.1%. With an eye on this, space-based instruments are limited by the detector's sensitive volume, so extending flux sensitivity to TeV energies would not be an easy task. This gains favor for the indirect detection techniques that use EAS as a mediator to detect VHEGRs.

For VHEGRs, the sole process in a EAS is the electromagnetic cascade, in which high-energy photons generate electron-positron pairs through pair production, and electrons generate high-energy photons through bremsstrahlung. The number of secondary particles continues to multiply, and their energy continues to decay until they reach the critical energy of $\sim 85\,\mathrm{MeV}$. At such relativistic energies, electrons move faster than the speed of light in the air, so they emit Cherenkov radiation in a cone-shaped volume with small half-vertex angle ($\sim 1^\circ$). The spectrum of this light peaks near UV light and it can illuminate a circle with a radius of 100 meters on the ground for several nanoseconds. By collecting and analyzing the Cherenkov flashes with ground-based observatories, the energy and arrival direction of primary gamma rays can be reconstructed, and valuable information about the sources can be inferred. This is the basis of the atmospheric Cherenkov technique. The development of this technology had been fraught with difficulties for decades. The culprit was that primary

CRs, which mainly trigger hadronic cascades, also produce Cherenkov flash events in a similar way as gamma rays do, but with an overwhelming amount of counts.

This problem is gradually overcome by imaging atmospheric Cherenkov technique (IACT) which strives to use the difference in images of Cherenkov flash to discriminate gamma-ray-induced showers and cosmic-ray-induced showers. Since the 2000s, the sensitivity and angular resolution of the third-generation (also the current-generation) IACTs such as High Energy Stereoscopic System (H.E.S.S.)[5], Major Atmospheric Gamma Imaging Cherenkov (MAGIC)[6], and Very Energetic Radiation Imaging Telescope Array System (VERITAS)[7] at sub-TeV energies have been comparable to space-based instruments such as Fermi-LAT. What is particularly admirable is that the sensitivity is extended to TeV energies, becoming the beginning of TeV astronomy. This huge detection potential has led to an international collaboration for construction of Cherenkov Telescope Array (CTA), which aims to raise the sensitivity by an order of magnitude more than current generation within energies ranging from 20 GeV to 300 TeV, potentially increasing the number of TeV sources[2] into the thousands[8].

Traditionally an IACT uses photomultiplier tubes (PMTs) as photodetectors for its focal plane cameras. PMTs have excellent sensitivity to Cherenkov light and high gain ($\sim 10^6$), both of which are what an IACT requires. However, PMTs have the problem of aging, which prevents PMT-based IACT cameras from operating under strong-light conditions (such as moonlight), thus shortening the telescope's duty cycle. Furthermore, considering long-term operation, the gain of PMT is asked be maintained at a relatively low level ($\sim 4 \times 10^4$)[9]. Due to these limitations, while letting a portion of telescopes continue to employ PMTs, CTA consortium is also actively looking for alternatives to PMTs, such as silicon photomultipliers (SiPMs). For example, although the Medium-Sized Telescope (MST) of CTA (Fig. 1.1) initially adopted PMTs, it aims to couple SiPMs to its currect camera during a future upgrade. The concept of this semiconductor detector was proposed in the 1990s, and its performance has continued to improve along with semiconductor technology. Now, not only are SiPMs comparable to PMTs in terms of photon counting and photon timing, but it even becomes the preferred option in many application fields (e.g., high-energy physics, medical imaging) that traditionally employ PMTs[10]. However, it is known that SiPMs still have some shortcomings that PMTs do not have, such as a longer temporal signal tail, a spectral response that is more susceptible to the influence of night sky background (NSB), and

[2] *cf.* As of July 2023, there are 272 TeV sources listed in TeVCat (TeV gamma-ray source catalog, http://tevcat2.uchicago.edu/).

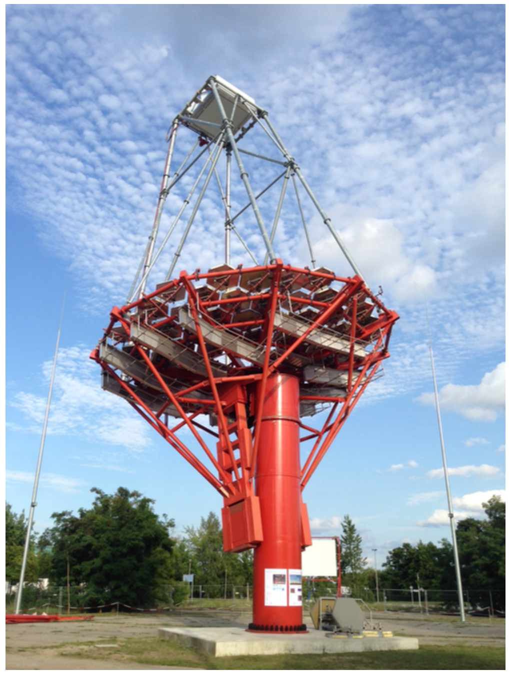


Figure 1.1: Medium-Sized Telescope prototype installed in Berlin, Germany. Photo credit: CTAO website.

correlated noise such as optical crosstalk. These are all issues that people need to address when trying to apply SiPMs on IACT.

In order to better understand the characteristics of SiPM for its future application on IACTs, this thesis will focus on the characterization of SiPMs. Chapters 2 and 3 constitutes the elaboration of motivation behind this study. Chapter 2 will briefly introduce VHEGR sources, VHEGR emission mechanisms, and list some of the techniques for VHEGR detection. Chapter 3 will explain the principles of IACTs and also present the current status and vision of IACTs, including an overview of CTA. As background knowledge for this study, Chapter 4 turns the focus to the working principle and performance characteristics, noise characteristics and applications of SiPMs and its basic unit—single photon avalanche diode (SPAD). The following two chapters, Chapters 5 and 6, are the experimental aspects of this study, in which the experimental methods and results are divided in to three categories and discussed separately. Chapter 7 gives the experimental summary and prospects for future research work.

Chapter 2

Very-High-Energy Gamma Rays

2.1 VHEGR sources

The number of celestial objects that can be studied by very high-energy gamma ray (VHEGR) astronomy is currently quite limited, but in the long run, it has been approximately increasing in an exponential fashion thanks to the advancement of technology. Referring to the Kifune plot (Fig. 2.1), it can be found that this growth trend also appears for X-ray and sub-TeV gamma-ray sources. A detailed introduction to each type of objects is beyond the scope of this thesis. For more comprehensive understanding in this subject, one may refer to [2]. [11] is another monograph focusing on objects that will be explored by CTA. A more concise overview of galactic and extragalactic VHEGR sources can be found in [12].

Supernova remnant

Supernova remnants (SNRs) are structures formed by the interaction between the material thrown out during a supernova explosion and the interstellar medium (ISM). Due to the existence of shock waves, supernova remnants are also places where particles are accelerated. Accelerated charged particles can produce high-energy gamma rays following the leptonic model or the hadronic model (see Section 2.2 for details). It is not yet certain which model dominates the gamma-ray emission from supernova remnants, perhaps both have a certain proportion. Parameterization studies are still in progress[14, 15].

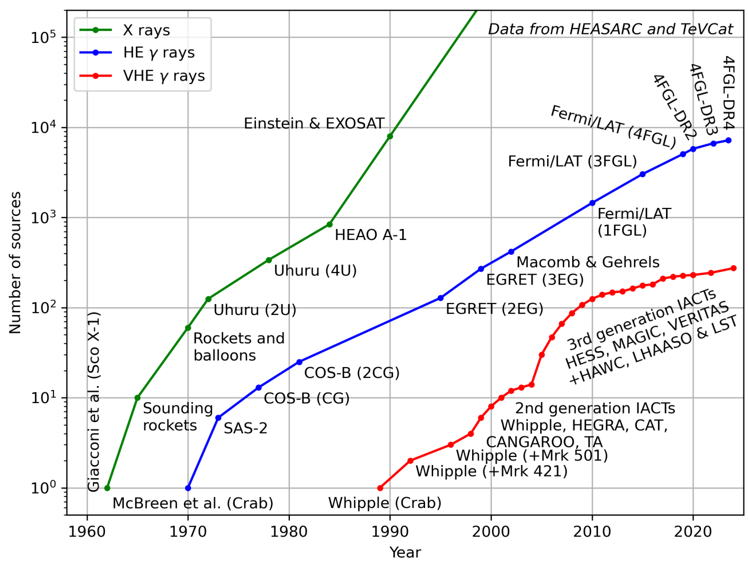


Figure 2.1: Kifune plot is a concept proposed by Tadashi Kifune in the 1990s[13]. What is presented here is a lastest version created by Stephen Fegan (https://github.com/sfegan/kifune-plot).

Pulsar and pulsar wind nebula

Pulsars are rapidly rotating neutron stars. A neutron star is a compact object left behind after a supernova explosion. Its magnetosphere has an extremely strong magnetic field. ($\sim 10^{12}$ gauss), which can accelerate the electron-positron plasma to form a cold ultra-relativistic outflow, which is the pulsar wind. The termination shock between the pulsar wind and the interstellar medium forms a pulsar wind nebula (PWN); this is where the pulsar wind emits high-energy gamma rays through non-thermal processes[2]. The Crab Nebula and the Vela Nebula are also typical examples of this type of systems.

Microquasar

Microquasars are essentially X-ray binaries with relativistic jets. In addition to the thermal emission of the accretion disk (X-ray band), they may also posses the conditions to form gamma rays, but so far only sporadic or insignificant cases are found in the GeV to TeV bands. A few examples include SS 433, Cygnus X-1, and Cygnus X-3[16].

Active galactic nucleus

According to the unified model of active galactic nucleus (AGN), these bright and distant objects are essentially galactic nuclei of supermassive accreting black holes. Seyfert galaxies, quasars, radio galaxies, blazars and other spectroscopically classified celestial bodies are all included in the explanation of the AGN model.[17]

Gamma-ray burst

Gamma-ray bursts (GRBs) are described as the most powerful explosion event in the universe; it can emit gamma rays lasting from a few seconds to tens of seconds, with an energy of up to 10^{54} erg. It is usually accompanied by an afterglow in different bands. Bursts lasting less than 2 seconds are called short GRBs, and bursts longer than 2 seconds are called long GRBs; according to currently accepted models, the former are believed to be produced when compact binary stars (such as binary neutron stars or neutron star-black hole systems) merge; the latter are thought to originate from jets produced when the core of a massive star collapses.

Diffuse gamma-ray background

The diffuse gamma-ray background (DGRB) consists of a combination of galactic emission and extragalactic isotropic emission with energies between 0.1 GeV and 0.1 TeV. The isotropic gamma-ray background can be explained by star-forming galaxies[18].

2.2 Mechanisms of VHEGR emission

In VHEGR astronomy, the mechanisms of gamma-ray production can be mainly divided into two types: *leptonic* and *hadronic*. The leptonic mechanism is mainly caused by the electromagnetic interaction of high-energy electrons, while the hadronic mechanism mainly produces neutral pions (π^0) through the strong interaction between high-energy protons and ISM gases, and then the neutral pion decays into gamma rays. Both electrons and protons are accelerated by some astrophysical source such that they have power-law flux spectra of the form

$$\phi_e \propto E_e^{-p}, \tag{2.1}$$

where p is the power-law index for charged particles. And this characteristic will also be transferred to the photon's flux spectrum through the mechanisms of VHEGR emission.

2.2.1 Leptonic origin

Non-thermal bremsstrahlung

The bremsstrahlung is the radiation emitted by a charged particle due to deceleration of electric field posed by another charged particle. Thermal bremsstrahlung takes place when the emitting particles (typically electrons) are in thermal equilibrium with their surroundings (typically hot plasma), and the distribution of electron velocities is Maxwellian. This contributes to X-ray spectrum more than to gamma-ray spectrum. And its spectrum is featured by a flat plateau up to a cutoff frequency. While the electrons are too relativistic to establish thermal equilibrium with surroundings, they emit *non-thermal* bremsstrahlung, through which the photon spectrum obeys a power law,

$$\phi_\gamma \propto E_\gamma^{-p} \tag{2.2}$$

Inverse Compton scattering

The inverse Compton (IC) scattering is the process in which low-energy photons are scattered by a high-energy electron. Its cross-section is determined by the photon energy in the rest frame of electron compared to the rest mass of the electron. Let Γ be the Lorentz factor of the electron, ϵ the photon energy before scattering, m_e the rest mass of electron and c the speed of light in vacuum. If $\Gamma\epsilon \ll m_e c^2$, the cross-section of photon–electron scattering is described by the Thomson scattering, of which the cross-section is electron or photon energy and is given by

$$\sigma_{\mathrm{T}} = \frac{8}{3}\pi r_0^2 \simeq 6.65 \times 10^{-25}\,\mathrm{cm}^2, \tag{2.3}$$

where r_0 is the classical electron radius. In this case the average energy of the upscattered photon is approximately proportional to the electron energy squared, i.e. $E_\gamma \propto E_{\mathrm{e}}^2$. For a power-law distribution of electron energy spectrum (Eq. (2.1)) will also result in a power-law energy spectrum of upscattered photons:[12, 19, 20]

$$\phi_\gamma \propto E_\gamma^{-(p+1)/2}. \tag{2.4}$$

If $\Gamma\varepsilon \gg m_{\mathrm{e}}c^2$, the scattering is described by the Klein–Nishina formula, with one simplified form being

$$\sigma_{\mathrm{KN}} = \frac{3}{8}\sigma_{\mathrm{T}} x^{-1}\ln(4x), \qquad \text{with } x = \frac{E_{\mathrm{e}}\epsilon}{(m_{\mathrm{e}}c^2)^2}. \tag{2.5}$$

In this case the resulting energy spectrum of upscattered photons will be described by a steeper power law[12]:

$$\phi_\gamma \propto E_\gamma^{-(p+1)}. \tag{2.6}$$

Synchrotron radiation

When relativistic charged particles are accelerated perpendicular to their velocity in a magnetic field, synchrotron radiation can be produced in the tangential direction and produce a continuum. If the electrons have a power-law spectrum like Eq. (2.1), then the energy spectrum of synchrotron photon will also follows a power law[12, 19]:

$$\phi_\gamma \propto E_\gamma^{-(p+1)/2}. \tag{2.7}$$

However, since the energy of the synchrotron photon is much smaller than that of the parent particles, and there is a high-energy cutoff in the synchrotron radiation continuum, the synchrotron radiation usually make up the low-energy component in the leptonic model of VHEGR emission. However, the synchrotron photons can be upscattered by the very same electron population that produced them through the inverse Compton process. This is the synchrotron self-Compton (SSC) process[21, 22]. Furthermore, if the magnetic field is strong enough, synchrotron radiation can be tightly coupled to pair generation, leading to efficient cascade development[20].

2.2.2 Hadronic origin

The inelastic collision of relativistic protons and atomic nuclei with the surrounding gas will produce unstable secondary pions, kaons and hyperons. Some of the decays modes of these unstable particles can produce VHEGRs. The most important channel is that a neutral pion (π^0) decays into two beams of gamma rays:

$$\pi^0 \longrightarrow 2\,\gamma \tag{2.8}$$

with a mean lifetime of only $8.4 \times 10^{-17}\,\mathrm{s}$. According to the kinematics, the production of π^0 requires that the kinetic energy of the proton should exceed $E_{\mathrm{th}} = 2m_\pi c^2(1 + m_\pi/4m_\mathrm{p}) \simeq 280\,\mathrm{MeV}$, where $m_\pi \simeq 135\,\mathrm{MeV}/c^2$ is the rest mass of π^0 and $m_\mathrm{p} \simeq 938\,\mathrm{MeV}/c^2$ is the rest mass of proton[20]. Contrastingly, the charged pions (π^+ and π^-) decay into muons, which in turn decay into electrons. These decays involve the production of neutrinos:

$$\begin{aligned}
\pi^+ &\rightarrow \mu^+ + \nu_\mu \\
\pi^- &\rightarrow \mu^- + \overline{\nu}_\mu \\
\mu^+ &\rightarrow e^+ + \nu_e + \overline{\nu}_\mu \\
\mu^- &\rightarrow e^- + \overline{\nu}_e + \nu_\mu
\end{aligned} \tag{2.9}$$

The energy of π^0 is evenly distributed between the resulting photons during the decay process. Thus, if the energy flux of π^0 adheres to the power law $\phi_{\pi^0} \propto E_{\pi^0}^{-p}$, a power law with the identical spectral index p exists for the energy flux of gamma rays, i.e.,

$$\phi_\gamma \propto E_\gamma^{-p}. \tag{2.10}$$

2.3 Detection techniques of VHEGRs

The detection techniques of astrophysical gamma-ray sources can be divided into *direct detection* and *indirect detection*. Direct detection is a technique used to detect gamma rays outside the atmosphere, as outlined in Section 2.3.1. Conversely, indirect detection involves observing the extensive air shower (EAS) produced by the gamma rays interacting with the atmosphere, as further explained in Section 2.3.2.

2.3.1 Space-based techniques for lower-energy gamma rays

Since the Earth's atmosphere is opaque to gamma rays, direct detection must be conducted outside the atmosphere, using detectors mounted on satellites or high-altitude balloons. This technique mainly utilizes the electron-positron pair production that is induced by gamma rays in detector materials. For example, the EGRET detector[1] mounted on the Compton Gamma Ray Observatory (CGRO) used a spark chamber as its tracker. The spark chamber itself is a gas ionization chamber in which multiple layers of thin metal plates are subjected to high voltage. When gamma rays pass through the metal plate, the generated electrons and positrons that trace a trajectory, providing information on the arrival direction, and then the charged particles deposit energy in the scintillator calorimeter. CGRO as a whole could detect gamma rays from 30 MeV to 30 GeV[23]. This kind of space-based detectors needs to be equipped with an anti-coincidence dome to reject other ionization radiation events. Employed in 2008, the LAT[2] on the Fermi Gamma-ray Space Telescope also uses a similar principle, but it uses conversion foils with silicon-strip detectors (SSDs) to replace the function of the spark chamber[24] and push the detectable energy to 300 GeV.

Space detectors have made many contributions to sub-TeV astronomy, such as discovering the diffuse gamma-ray emission on the Galactic plane, performing gamma-ray survey missions, establishing gamma-ray source catalog (e.g. EGRET catalogs[25, 26]), etc. Nevertheless, the flux of TeV gamma rays is so low that the typical collection area (on the order of square meters) of a space detector does not provide a sufficient event rate to study VHEGRs.

[1] The Energetic Gamma Ray Experiment Telescope (EGRET). https://heasarc.gsfc.nasa.gov/docs/cgro/egret/.
[2] The Large Area Telescope (LAT). https://fermi.gsfc.nasa.gov/science/instruments/lat.html.

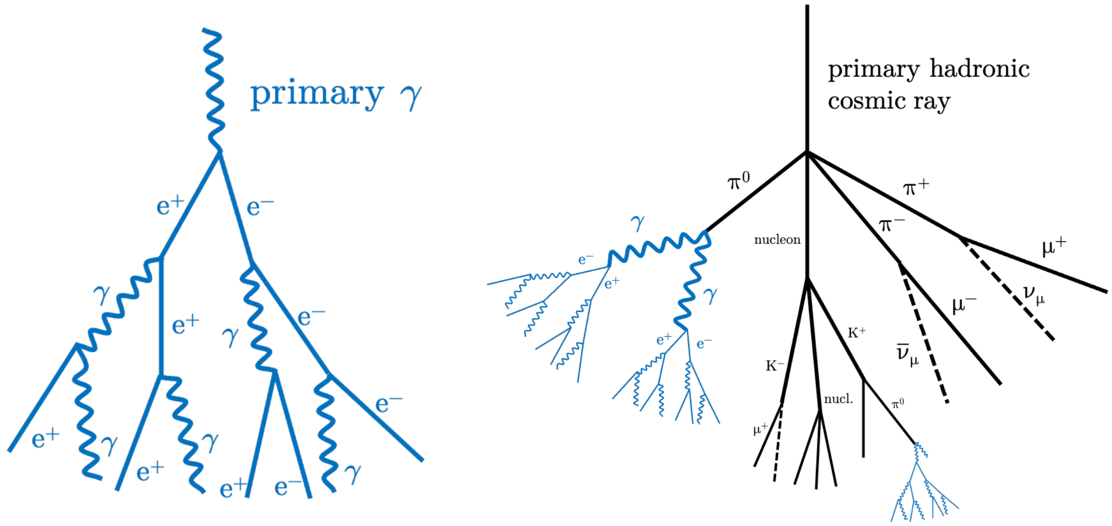
2.3.2 Ground-based techniques for VHEGRs

The upper limit of sensitivity of space-based detectors prompted the search for ground-based gamma-ray observatories, i.e. the technique known as indirect detection, although it had begun developing about 20 years earlier than space-based detectors. Indirect observations of VHEGR rely on extensive air shower (EAS). In fact, following the discovery of EAS by Pierre Auger in 1938[27], many of the characteristics of EAS were well studied in the particle detector array experiments. For instance, the fact that EAS emits Cherenkov radiation was confirmed by simple photomultiplier tube experiments in 1953[28]. In the early 1960s, secondary charged particles in the EAS and induced Cherenkov light were proposed to explore gamma-ray sources. This can be considered the beginning of *ground-based gamma-ray astronomy*[3]. Later, the Whipple Observatory and the Crimean Astrophysical Observatory, each attempted to build reflecting telescopes to search for potential point sources. But no convincing TeV source was found until the TeV gamma rays from the Crab Nebula were detected by the Whipple Observatory in 1989 with a confidence level of 9σ[29]. For a more detailed development history during this period, one may refer to [30].

Extensive air shower

An extensive air shower (EAS) is a cascade of particles initiated by high-energy cosmic rays (including gamma rays) in the atmosphere. The so-called cascade means that the number of particles increases step by step in the process, as depicted in Fig. 2.2. One can distinguish between electromagnetic cascades and hadronic cascades.

Electromagnetic (EM) cascades are typically initiated by high-energy electrons, positrons and photons (γ). Electrons or positrons can produce photons through bremsstrahlung, and photons can produce electron–positron pairs through pair production. In both processes, the number of particles in the cascade double, whereas the average energy per particle reduces to one half. Secondary electrons lose energy mainly through ionization, and secondary photons lose energy primarily through the Compton effect until they are absorbed by the atmosphere or reach the ground (as long as the primary energy is sufficient). When the energy of each particle is less than the critical energy $E_c \approx 85\,\text{MeV}$, energy loss begins to dominate the shower development, and the shower gradually dies out in the atmosphere. Although it is possible for electrons to produce hadrons and muons through quantum electrodynamic processes, their numbers are far less than those in hadronic cascades. Therefore, an EM



(a) A gamma-ray shower is purely composed of EM cascades.

(b) A hadronic shower is composed by hadronic cascades, EM cascades, and muonic components.

Figure 2.2: Schematic diagrams of gamma-ray and hadronic showers. The blue part represents the EM cascades.

cascade can be regarded as purely composed of electrons, positrons and photons. Fig. 2.2a shows a VHEGR induces an EAS with pure EM cascades.

Hadronic cascades are triggered by high-energy nuclei (including protons and heavy nuclei) and produce various types of baryons and mesons. The former are mainly nucleons, and the latter include π mesons and K mesons, which are unstable and decay into more particles. For example, a neutral pion (π^0) decays into two photons (Eq. (2.8)) and a charged pion decays ($\pi^{\pm}$) into a muon and a muon neutrino (Eq. (2.9)). Among these, the production of photons means that the hadronic cascade will also produce electromagnetic cascades. Also, the muons resulting from $\pi^{\pm}$ decay are minimum ionizing particles in air and their relativistic speed also allows them to reach the ground before decaying. The special thing is that, muons are products of hadronic showers and become one of the useful features to distinguish gamma-ray showers from hadronic showers. A hadronic shower is shown in Fig. 2.2b, where hadronic cascades, EM cascades, and muonic components can be seen.

The longitudinal development of EAS is closely related to the density profile of the atmosphere. Whether the EAS can reaches the ground depends on the energy of the primary particles. Generally, a TeV-range EAS develops to its maximum extent at $\sim 10\,\text{km}$ above sea level, after which it begins to decay. At sea level, EAS can be effectively detected with

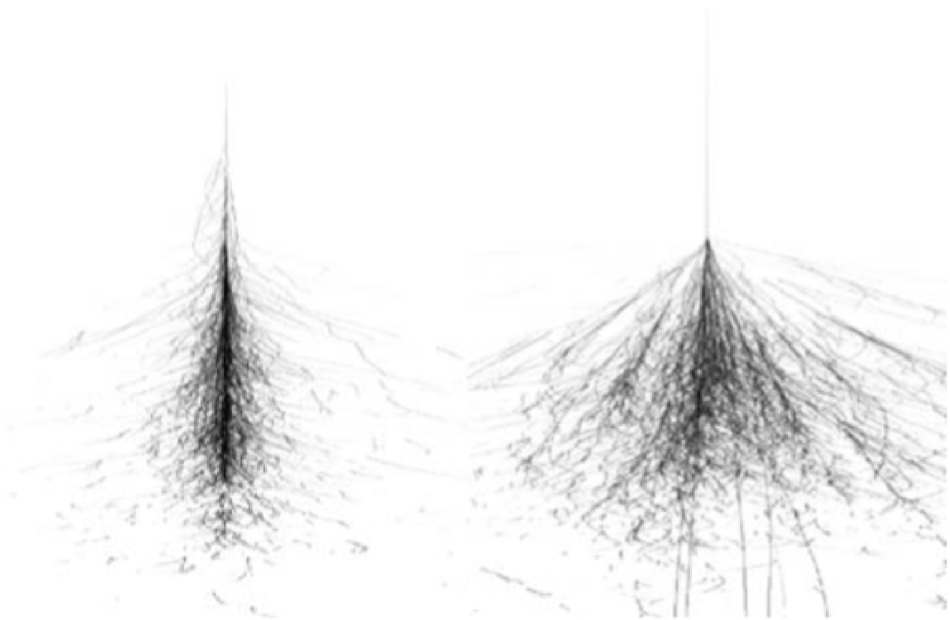


Figure 2.3: EAS development based on Monte Carlo simulations. *Left*: gamma-ray shower. *Right*: hadronic shower. Credit: [32].

a particle detector array only when the primary energy is $\gtrsim 1\,\mathrm{PeV}$. Hence, increasing the elevation of the detectors lowers the energy threshold of observation. In addition to the longitudinal development, the secondary particles in EM cascades and hadronic cascades receive transverse momentum due to Coulomb scattering and strong interaction, respectively, thus resulting in the lateral development of EAS. Compared with EM cascades, larger transverse momentum transfer is seen in the strong interaction. As a result, hadronic cascades develop more laterally than gamma-ray shower and are more irregular in shape, as shown in Fig. 2.3.

Monte Carlo simulations provide the most detailed analysis of EAS because they can take into account all particle interactions and tracking the evolution of all particles in the atmosphere. A commonly used example is the CORSIKA code,[3] with Fig. 2.3 being its results.

Only a qualitative description is given here; for a more comprehensive exposition about EAS, see [4, 31].

[3]https://iap.kit.edu/corsika/.

Particle detector array

An EAS is mainly composed of charged particles such as electrons and muons when it reach the ground. Therefore, the particles developed in an EAS can be "sampled" by setting up particle detector array in a vast collection area. Such detectors include scintillator detectors and water Cherenkov detectors (WCDs). A scintillator absorbs charged particles to produce scintillation light. This weak light signal can be detected by a single-photon detector (see Section 3.1.4) coupled to the scintillator. Because muons are more penetrative than electrons, the scintillator detector specially designed for muons must be covered under an absorbing layer such as lead plate or soil. Examples of scintillator detector arrays dedicated to VHEGRs include the Tibet ASγ experiment[33] and the Kilometre Square Array (KM2A) of LHAASO[34], both of which are located in the Northern Hemisphere. In its design stage, the Southern Wide-field Gamma-ray Observatory (SWGO) is positioned to become the next-generation particle detector array, partnering with CTA to jointly explore the TeV gamma-ray sky in the Southern Hemisphere[35].

On the other hand, a WCD uses a pool of water as a medium, in which the traversing relativistic charged particles can produce Cherenkov radiation (cf. Section 3.1.1), which is also detected by single photon detectors. Examples of WCD arrays include the Milagro experiment and its follow-up experiment called the High-Altitude Water Cherenkov (HAWC) Gamma-Ray Observatory[36].

Atmospheric fluorescence detector

An indirect method of detecting EAS is to use the fluorescence emission by nitrogen molecules in the atmosphere due to the excitation of EAS. Most of the fluorescent emission is UV between 300 and 400 nm and can be collected using mirrors and detected by photomultiplier tubes. Such an instrument is called an atmospheric florescence detector or, if can image the fluorescence with a photodetector array, an atmospheric fluorescence telescope. Although the fluorescence emission is isotropic, these instruments are primarily designed to observe the longitudinal development of the EAS from the side and is used to reconstruct the energy of primary cosmic ray. Moreover, the fluorescence light of EAS is extremely weak, so this method can only be used to detect the EAS with $E \gtrsim 100\,\mathrm{PeV}$[31], which belongs to the regime of UHECR. On the contray, the fluorescence emission of VHEGR, with energy usually $< 1\,\mathrm{PeV}$, is almost negligible[37]. Hence, atmospheric fluorescence detectors are not currently an available technique for VHEGR.

Atmospheric Cherenkov techniques

Collecting the Cherenkov radiation emitted by relativistic EAS secondary particles when they travel through the atmosphere is also an indirect method of detecting EAS. When the Cherenkov radiation is emitted, a cone is formed with the core of the cluster as its axis, and a Cherenkov light pool is formed on the ground. Regarding this, Section 3.1.1 will have a more detailed description.

There are at least two atmospheric Cherenkov techniques that can reconstruct the arrival direction and energy of primary cosmic rays. One technique is to set up a light collector array with a long baseline on the ground like a particle detector array, use the arrival time and density of photons and the spatial information of the array to reconstruct the wavefront of Cherenkov light, and then calculate the energy and arrival direction of EAS. This technology is called *wavefront sampling technique*. Since it is not required to obtain a detailed image of Cherenkov light, only one or more photomultiplier tubes are used in the focal plane of each collector, with relatively large pixels, to observe a wide field of view (from roughly larger than $10^\circ \times 10^\circ$ to $\sim 1\,\mathrm{sr}$), so it is also called a wide field-of-view or wide-angle Cherenkov telescope. TAIGA-HiSCORE is an example of wavefront sampling technique[38].

Another atmospheric Cherenkov technique is Imaging Atmospheric Cherenkov Technique (IACT),[4] of which an elaboration will be given in Chapter 3.

The techniques introduced in this section are summarized in Fig. 2.4.

[4]The letter “A” in IACT sometimes also stand for “air,” and the letter “T” also stands for “telescope.”

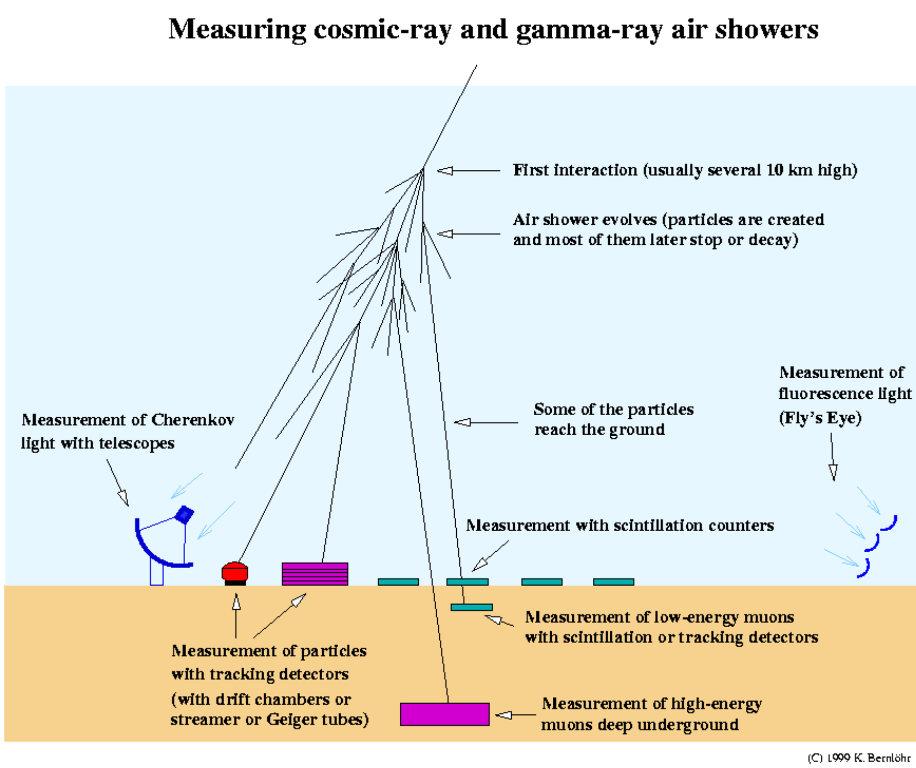


Figure 2.4: Ground-based EAS techniques. Credit: Wikimedia. Original author: K. Bernlöhr.

Chapter 3

Imaging Atmospheric Chererekov Technique

3.1 The concept of IACT

3.1.1 Cherenkov radiation

A charged particle at rest in a dielectric medium will cause isotropic electric polarization. Now, we consider a moving charged particle with a uniform speed v in a dielectric medium. Electromagnetic fields associated with the charged particle can polarize the medium. But the speed at which the polarization spreads out is the phase velocity of electromagnetic waves, c/n, which is slower than the speed of light in vacuum c, where n is the index of refraction ($n > 1$). When $v < c/n$, the electromagnetic field in the medium is perturbed, but the perturbation relaxes through destructive interference, and no light is emitted. However, when $v > c/n$, the perturbation of the field interferes constructively, and asymmetric polarization is seen along the trajectory of the charged particle.

According to Huygens' principle, the successive spherical wavelets moving outward form a conic wavefront, emitting anisotropic radiation. Considering the simple geometric configuration in Fig. 3.1, one can find the angle θ_{Ch} between the direction of Cherenkov emission and that of the particle motion satisfies

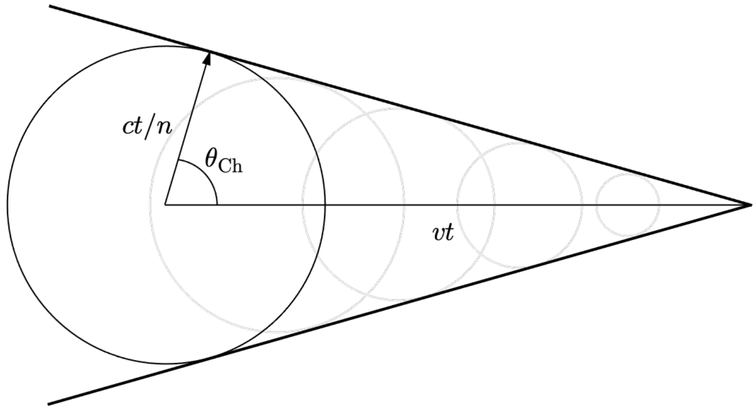


Figure 3.1: Geometry of Cherenkov radiation. In this figure, vt represents the distance traveled by a charged particle with velocity v during a time duration t, and ct/n denotes the distance over which polarization spreads during the same time duration.

$$\cos\theta_{\mathrm{Ch}} = \frac{c/n}{v} = \frac{1}{n\beta}, \tag{3.1}$$

where $\beta = v/c$. Setting $\theta_{\mathrm{Ch}} = 0$ in Eq. (3.1) yields the Cherenkov energy threshold

$$E_{\mathrm{Ch}} = \frac{mc^2}{\sqrt{1-\beta^2}} = \frac{mc^2}{\sqrt{1-n^{-2}}}, \tag{3.2}$$

where m is the particle's rest mass.

The Cherenkov radiation is named after P. A. Cherenkov, who discovered it during an experiment under S. I. Vavilov's supervision in 1934[39, 40]. The corresponding theory was given in 1937 by I. Frank and I. Tamm, who described the energy spectrum of Cherenkov radiation. An extension of the Frank–Tamm formula gives the differential number of photons emitted per unit path length dx in the wavelength range λ to $\lambda + \mathrm{d}\lambda$:[41, 42]

$$\frac{\partial^2 N}{\partial x \partial \lambda} = \frac{2\pi\alpha z^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right), \tag{3.3}$$

where z is the charge carried by the charged particle in the unit of the elementary charge e, and $\alpha = e^2/\hbar \approx 1/137$ is the fine structure constant.

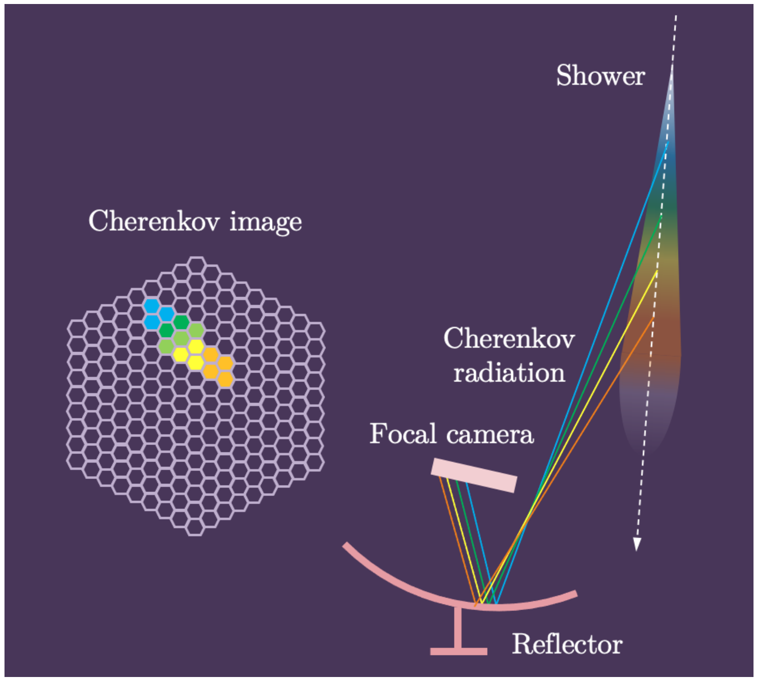


Figure 3.2: Principle of the IACT. Different colors represent different stages of shower development and the corresponding parts of an image from the camera's view.

The secondary particles of a high-energy EAS can reach relativistic speeds and thus meet the conditions for emitting Cherenkov light. Therefore, detecting the atmospheric Cherenkov radiation caused by an EAS is theoretically feasible. A crude and straightforward idea is to use a telescope to receive Cherenkov light, just like in optical astronomy. The earliest experiments in the 1950s placed one PMT at the focal plane of the optical system. These experiments were successfully detected Cherenkov light[28] but could not retrieve the information of the primary particle. In the following decades, two atmospheric Cherenkov techniques were developed. The *sampling* atmospheric Cherenkov technique, briefly described in Section 2.3, builds telescope arrays to sample the wavefront. The *imaging* atmospheric Cherenkov technique (IACT) takes another path—it builds telescopes with photodetector arrays to resolve the image of Cherenkov light. As depicted in Fig. 3.2, during the development of EAS in the Earth's atmosphere, Cherenkov photons reach the ground surface at various zenith angles. These photons can subsequently be collected by an optical system (e.g., a reflector, as described in Section 3.1.3) captured by a multi-pixel focal camera (as described in Section 3.1.4), facilitating the resolution of Cherenkov radiation.

3.1.2 Observing atmospheric Cherenkov radiation

The index of refraction of air, as well as atmospheric absorption and scattering, will affect the propagation of Cherenkov light. In addition, there are at least two important observational interferences: the night sky background and moonlight. These matters are explored below.

Index of refraction of air

As can be seen in Eq. (3.1), to observe the atmospheric Cherenkov radiation, the index of refraction of air n_{air} needs to be taken into consideration. An approximate estimation is given by[43]:

$$n_{\mathrm{air}} \simeq 1.0 + 0.000283 \cdot \frac{\rho(z)}{\rho_0} = 1.0 + 0.000283 \cdot e^{-z/H}, \tag{3.4}$$

where z is the altitude in km, $\rho(z)$ is the air density reasonably modeled by an exponential function $\rho(z) = \rho_0 e^{-z/H}$, $\rho_0 \simeq 1.225 \times 10^{-3}\,\mathrm{g\,cm^{-3}}$ is the air density at the ground level, and $H \simeq 8.4\,\mathrm{km}$ is the atmospheric scale height. Based on Eqs. (3.1) and (3.4), we know the θ_{Ch} changes as the EAS propagates downward in the atmosphere. This feature is shown schematically in Fig. 3.2.

Absorption and scattering of the atmosphere

The atmospheric Cherenkov radiation also suffers from extinction because of the absorption and scattering of air[44, 45], which results in the specrum shown in Fig. 3.3. The absorption below 340 nm is dominated by ozone (O_3), which creates the opacity of the ultraviolet band. The scattering and absorption of molecules and aerosols[1] also contribute to the extinction in the visible band. The molecular scattering is described by Rayleigh scattering theory, while the aerosol scattering is described by Mie scattering theory. The presence of clouds also has a complex impact on Cherenkov light, so IACTs also need observation stations to monitor real-time meteorological conditions[46].

[1] Aerosols are solid or liquid particles stably suspended in a gaseous medium.

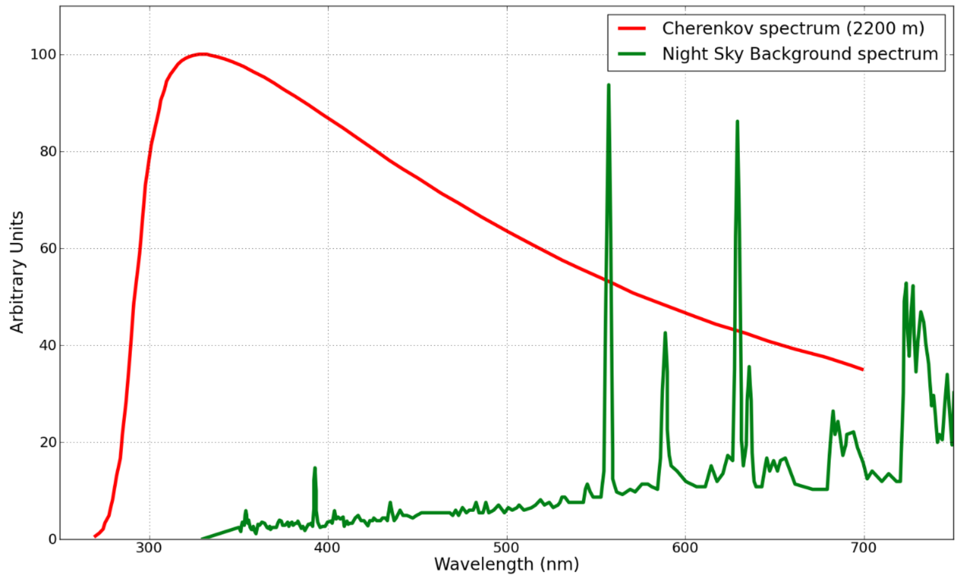


Figure 3.3: The spectrum of the atmospheric Cherenkov radiation (*red curve*) and the spectrum of the night sky background (NSB, *green curve*). Both were measured in La Palma at an altitude of 2200 m above sea level. Credit: [51]. Original data from [52, 53].

Night sky background

In addition to the above-mentioned atmospheric factors, there is also a weak diffuse light, called the night sky background (NSB), that will interfere with the observation of Cherenkov light. The NSB originates from multiple sources, including night airglow, zodiacal light, diffuse galactic light, and diffuse extragalactic light. These components vary with time as well as terrestrial and celestial coordinates, complicating the study of the NSB[47]. Fig. 3.3 shows the spectrum of NSB measured in La Palma (one of CTA's scheduled sites), where has been considered and some prominent airglow emission lines can be seen. The NSB brightness based on *in situ* direct measurement using charge-coupled device (CCD) can be expressed in magnitude. However, the NSB rate per pixel is often used for IACT camera pixels. A typical value for the NSB rate is 1.7×10^{12} NSB photons per second per steradian per square meter[48]. In terms of the pixel field of view (FoV) and pixel area designed for Large-Sized Telescopes (LSTs) in the CTA project, the NSB rate is hundreds of MHz[49]. An excessively high NSB rate has many adverse effects on photodetectors, especially on SiPMs. Because of the recovery time of the bias voltage (see, e.g., Section 4.1.7,) under high NSB rates, the incomplete recovery of the bias voltage will result in a lower gain and detection efficiency.[50].

Moonlight

Traditionally, an IACT can only operate under dark conditions, which excludes moonlit nights or twilight. The reason is that intense background light aggravates the NSB rate, reduces the sensitivity of PMTs to Cherenkov light, and causes faster aging or even damage of the PMTs. Under the strict limitation of observation time, traditional PMT-based IACTs can only operate under a duty cycle of about 10%. Operating the PMTs at a lower gain or higher anode current can relieve this limitation to some extent[42]. In contrast, using SiPMs for IACT observations will not be as restrictive as using PMTs. The First G-APD Cherenkov Telescope (FACT) has proven the observational stability of SiPMs during moonlit nights[48].

3.1.3 Optical system of IACT

The spectral maximum of Cherenkov light is in the UV band, so the optical system of IACT telescopes are tailored to focus UV light effectively. Besides, an IACT has a wider FoV (several degrees) to match the typical size of a Cherenkov image. To ensure that the extremely weak Cherenkov light is collected, an IACT is equipped with a reflector with an effective area of up to hundreds of m^2, focusing the light into the focal plane where the camera is mounted.

Most of the reflectors are not formed in one piece, but use hundreds of small mirrors (facets) with a diameter $\lesssim 1\,\mathrm{m}$ to form a large segmented or tessellated reflector. This can reduce cost and difficulty of optical design. The facets used to be circular, e.g., in CANGAROO, High Energy Gamma Ray Astronomy (HEGRA), H.E.S.S. 12 m telescope. Nowadays, to collage the facets as tighly as possible, the shape of IACT facets are usually square (e.g. MAGIC) or hexagon (e.g. H.E.S.S. 28 m telescope, and CTA's telescopes). The base material of mirrors can be glass or aluminum, with multi-layer coatings (such as SiO_2, Cr, Al) to maximize the reflectance of Cherenkov light[54]. In addition, each facet is coupled to an actuator. This is to adjuct the mirros initially and compensate either initial or long-term mechanical deformation, which undermines the alignment accuracy of the facets and eventually the quality of imaging[55].

Current generation IACTs often use Davies–Cotton (DC) optics. The original motivation for this design was for solar furnaces[56]. The off-axis focusing makes it relatively easy to achieve a large FoV[57]. The focal ratio of the facet in DC optics is the same as the focal ratio of the telescope as a whole, which is approximately unity[43]. However, the DC optics is subject to arrival time dispersion, where light reflected from different parts of the reflector arrives at different times. This effect scales with telescope size, and will lengthen the integration time of signals[3].

Therefore, larger telescopes may choose a parabolic design, letting the reflected light to reach the focal plane at the same time, such as the LST. In addition, according to CTA's lastest proposal, all of Small-Sized Telescopes (SSTs) will adopt an innovative double-mirror design—the Schwarzschild–Couder (SC) optics[58]. Although the SC optics has a more complex mechanical design, it can reduce the point dispersion function, which is conducive to the anglular resolution[59].

In recent years all large IACTs use alt-azimuth mounts for the positioning. Steering can be swift if the telescope is light-weighted; for example, LST's steering system can orient to any position in the sky within 20 seconds[60].

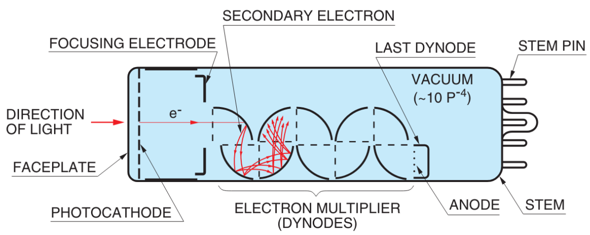
3.1.4 Cameras of IACT

The camera of IACT is a highly compact detector system that contains a pixelated photon-detection plane, readout electronics, data acquisition, different levels of triggering, calibration, monitoring, and other subsystems. The description here will focus more on photodetectors.

Photodetectors and their peripherals

Detecting Cherenkov flashes with high efficiency is the primary goal of an IACT camera. Cherenkov light produced by EAS is not only transient but also very faint. For a 1 TeV gamma-ray EAS, only a few hundred photons per square meter can be collected in an event, and the event only lasts for several nanoseconds[20]. Although common in optical telescopes, CCDs, with integration times ranging from 0.1 s to 100 s, are unsuitable for IACT due to their slower response time to capture events as rapid as Cherenkov light.

Instead, an IACT requires *single-photon detectors* for its photodetectors. In other words,



(a) Sketch of a PMT. Credit: [61].

(b) A 12-pixel PMT module used in FlashCam. Credit: [62].

Figure 3.4: Photomultiplier tube (PMT)

the performance IACT pursues is the fast response and the sensitivity to single photons.

The photomultiplier tube (PMT) was the earliest single-photon detector put into use by IACT and by many other applications that require low-light-level detection. It remains the mainstream option in the current-generation IACTs. A PMT converts incident photons into photoelectrons through photoelectric emission on the photocathode. Then, the electric field accelerates the photoelectrons, generating more electrons through secondary emission on multiple dynode stages, and finally, the PMT outputs macroscopic current at its anode. This process is visualized in Fig. 3.4a.

A silicon photomultiplier (SiPM) is a monolithic array of SPADs, which are photodiodes that operate at high reverse bias and utilize the avalanche breakdown of the semiconductor as the multiplication mechanism of carriers. Many properties of SiPMs have become competitive with those of PMTs. The introduction to the silicon photomultiplier (SiPM) will be the focus of Chapter 4.

The following are some of the basic performance parameters of photon counters:

1. **Gain** — the multiplication factor of the primary photoelectrons. The gain of different photon counters ranges from thousands to millions, depending on the application. The premise is to bring a good signal-to-noise ratio to the readout system, which is the key to single-photon detectors being able to detect faint light.

2. **Detection efficiency** — the probability that an incident photon will be detected and registered as a signal by the photodetector. This probability is a combination of multiple factors, such as quantum efficiency in photoelectric conversion and structure-related geometric efficiency. For example, the photon detection efficiency (PDE) of SiPM (Section 4.3.3) is its distinctive way of describing the detection efficiency.

3. **Photon number resolution** — the ability to resolve the number of detected photons. In practice, the discreteness of photon numbers is manifested by the well-separated photoelectron peaks observed on a single-photoelecton (SPE) spectrum. The photon number resolution can be quantified by measuring the relative peak width on the SPE spectrum. Concepts such as charge resolution, pulse amplitude resolution, and single photoelectron resolution are highly relevant to photon number resolution. The use of these terms depends on how the SPE spectrum is measured and how the peak discreteness is defined. Judging from the peak discreteness on the SPE spectrum, SiPM performs better than PMT in general.

4. **Photon timing resolution** — the ability to measure the arrival time of photons. The time elapses between the optical input and the electrical output is called timing latency. The fluctuation of this quantity is called timing jitter or *transit time spread* (TTS), which serves as a measure of timing resolution.

5. **Dynamic range** — the range in which the signal amplitude is proportional to the number of detected photons. Beyond this range, a saturation effect will occur; that is, the signal amplitude will no longer increase linearly with the number of photons. So, the dynamic range effectively sets the boundaries for the detectable photon numbers. For IACT cameras, the required dynamic range is at least a few thousand and below ten thousand photoelectrons[63].

A single-photon detector can only measure photon counts at a fixed point on the focal plane. To obtain an image of Cherenkov light, we need a detector array to “pixelate” the focal plane. The density of pixels in the focal plane determines the spatial resolution of the image, which should be optimized based on the field of view, the expected image properties, the point spread function (PSF) of the telescope, and the dynamic range of sigle pixel, as well as considerations such as readout system design and cost.

Dead space on the focal plane is the space between pixels where reflected light is missed. Also, an IACT usually does not have peripheral shielding to block stray light from large incident angles. To reduce the dead space, block the stray light, and thus maximize the light collection efficiency, most IACT cameras are equipped with non-imaging light concentrators in front of the photodetector[64]. These are sometimes referred to as light guides, light collectors, or light funnels. The Winston cone, with a parabolic profile, is a commonly used design for the light concentrator. For IACTs, the entrance and exit apertures are usually hexagonal. However, it is shown that other designs, e.g., quadratic or cubic Bézier curve, outperform the Winston cone[65]. Light concentrators with exit apertures adapted to square SiPMs have also been developed[50]. The inner surface of the light concentrator, like that on the facet, is also coated to optimize the reflection of near UV Cherenkov light. In addition, the dust protection glass in front of the camera window can also selectively reflect more light of wavelength $> 550\,\mathrm{nm}$[66].

Last but not least, the plane where the photodetector is located can be modularized. This practice allows for sharing components in circuit topology, saves costs, facilitates mechanical assembly, and allows the camera to adapt to different sizes and layouts. Fig. 3.4b portrays FlashCam's 12-pixel module, which is an example of modularization.

Readout and trigger system

Generally speaking, a detector's readout system includes a series of signal conditioning stages, such as amplification, filtering, shaping, and digitization. Due to the transient nature of Cherenkov light, the IACT camera specifically requires a waveform sampling rate close to the GHz level.

At the input of the readout circuit, the current signal from the single-photon detector is usually converted into a voltage signal through a load resistor. A preamplifier (pre-amp) immediately following the photodetector output can improve the signal-to-noise ratio (S/N). The signal might need to be shaped to eliminate stacking effects and obtain the correct pulse height. Shaping can also take place in specific amplifier topologies with integrator functions.

Due to the wide dynamic range and the single-photon resolution of the single-photon detector, when it comes to the selection of analog-to-digital converters (ADCs), the resolution (number of bits) and least significant bits (LSB) must also fit the detector correspondingly. Given this, when the cost of the ADC is also taken into consideration, a "dual-gain" method may be used to amplify the signal; that is, the signal is split into two paths and amplified using a low-gain and a high-gain channel, respectively, so as to provide appropriate gains to different amplitude[67].

In contrast, the FlashCam designed for CTA adopts a single gain design and a fully digital method in its readout system. The signal enters the digitalization stage after a single amplifier with non-linear gain characteristics. The pulse height is reconstructed in a sophisticated algorithm, which eliminates the requirement for additional channels[62, 68, 69].

In order to filter out the events of interest and initially reject NSB, the readout circuit should not only be responsible for sampling but also be included in the triggering system. The trigger system may involve multiple levels with different trigger rates, processing back and forth on samples temporarily stored in buffers. By properly designing the trigger strategy, accidental triggers by NSB are supposed to be well suppressed, while spatial homogeneity and symmetry can be achieved simultaneously. The design philosophy of FlashCam's trigger system can be learned in [62].

After being triggered, candidate events enter the data acquisition (DAQ) system, whose output provides the raw level of the data processing architecture (see next section). In DAQ systems that require communication between cameras and servers, using Gigabit Ethernet is a practical and convenient option.

IACT cameras have other subsystems, such as mechanical, cooling, power supply, safety, and calibration systems. Introducing them would be beyond the scope of this thesis.

3.1.5 Analysis methods of IACT

The data processing of IACTs can be divided into several data levels (DLs, as shown in Fig. 3.5[70]), and methods of data analysis are reviewed by [71].

DL0 is the raw data, which is the waveform from each DAQ channel, the pulse height of which has been sampled by the readout and trigger systems, and converted into a physical quantity that represents the number of photons using time integration, such as the charge expressed in photoelectrons (p.e.). Also included in this level are the further cleaning of NSB[47], electronic noise, and the homogenization of the signal according to flat-fielding.

The next level of data processing (DL1 to DL2) is to establish a Cherenkov image based on

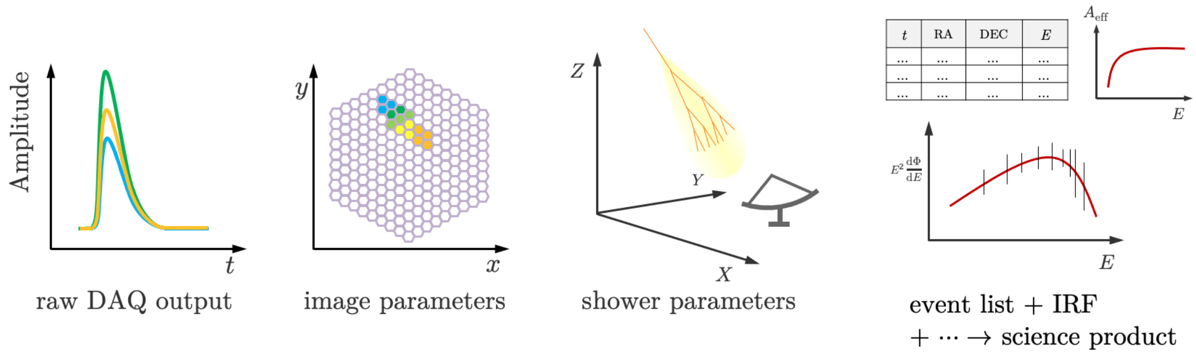


Figure 3.5: Data levels (DLs) of IACT. The definition of DLs is based on CTA. Adapted from [71].

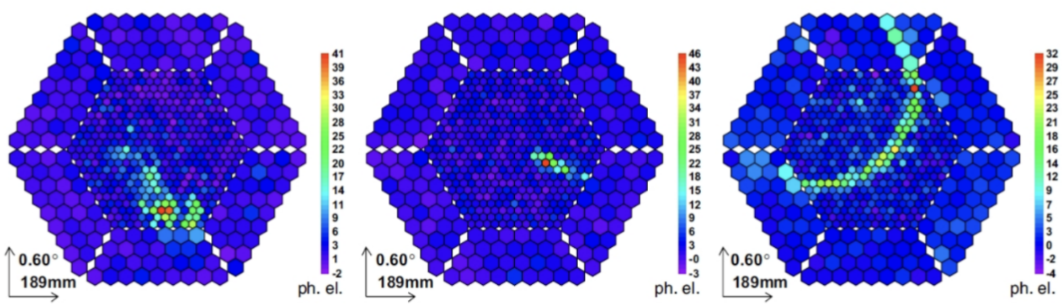


Figure 3.6: Typical Cherenkov images of a hadron-like event (*left*), a gamma-like event (*middle*), and a muon ring (*right*). Credit: [73].

the pixel layout and then reconstruct the EAS event. Effective suppression of background signals is crucial for the detection of gamma-ray signals. Such a background is mainly composed of hadronic-rich showers, and thus the process of separating gamma-ray showers from hadronic showers for IACTs is commonly referred to as gamma/hadron separation, which is also the key to the success of IACT. Fig. 3.6 shows the distinction of Cherenkov images between a typical gamma-ray event and a typical hadronic event. As part of the hadronic showers, the Cherenkov radiation of muons is featured by a ring-shaped image and is also a tool for gamma/hadron separation and calibration[72].

The first analysis of Cherenkov image based on image parameters to perform background suppression was developed by A. M. Hillas in mid 80s. By means of Monte Carlo simulations, Hillas demonstrated that parametrizing the Cherenkov images up to the second moments and applying suitable cuts on the parameters should be reliable for IACTs to dis-

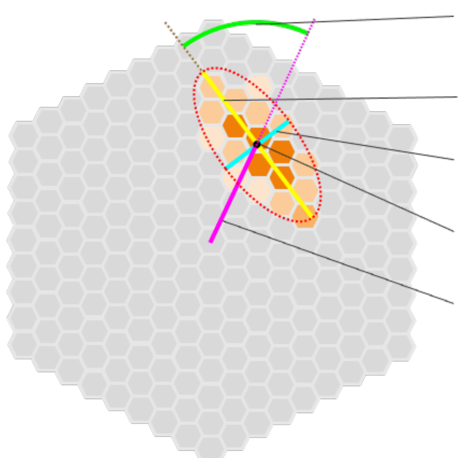


Figure 3.7: Hillas parameters. COG = center of gravity of the image. Adapted from [76].

criminate between the gamma-ray-induced and proton-induced showers[74]. The accuracy of this method has been demonstrated by many algorithms based on actual observations and simulations[75]. As shown in Fig. 3.7, the two second moments of the image along its minor axis and major axis are named WIDTH and LENGTH, respectively. If the shower is caused by hadrons, both the WIDTH and LENGTH of the image will be larger for a given image intensity. The energy of a shower can be estimated from the image SIZE (the first moment of the image) and the DISTANCE from the center of gravity and to the source position, which is inferred from the asymmetry of the distribution along both sides of the main axis.

However, symmetry along the major axis will result in degenerate solutions, where the source position cannot be estimated by observation with a single IACT. This problem can be solved by introducing higher-order moments of the image or using stereoscopic imaging (Fig. 3.8), which has become standard practice in the current-generation IACTs. Stereoscopic imaging not only solves the difficulty of a single IACT to distinguish local muon events (i.e. with the images that exceed the pixel range) from gamma-ray events, but also improves angular resolution, energy resolution, background suppression, and sensitivity[32, 72].

Alternatively, the model analysis method uses Monte Carlo simulation to establish a semi-analytical shower development model and many image templates and then performs a maximum likelihood fitting of all pixels available in the camera[77, 78]. Based on this concept, H.E.S.S. developed an event reconstruction algorithm known as ImPACT,[2] which improved the angular resolution[79]. Apart from the Hillas parameters and model analysis, event re-

[2] Image Pixel-wise fit for Atmospheric Cherenkov Telescopes

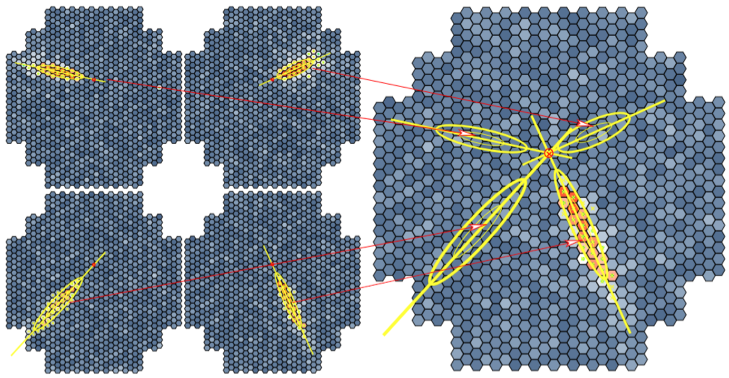


Figure 3.8: Four IACTs (here, H.E.S.S.) conduct stereoscopic observation to reconstruct a gamma-ray event. Credit: [32].

construction methods based on convolutional neural networks (CNNs) are also in use and under development for IACTs[80].

After obtaining the reconstructed data of the gamma-ray EAS event, we finally enter the high-level data analysis of IACT (from DL3 to DL5). At this stage, different gamma-like events are selected from the event list and matched, with the aid of instrument response functions (IRFs). After data binning, modeling, and fitting, scientific outputs such as source catalogs, sky maps, energy spectra, and light curves are generated. This can be completed by some open-source software, such as gammapy[81].

3.2 Cherenkov Telescope Array

3.2.1 Overview of CTA

The Cherenkov Telescope Array (CTA) is the next-generation IACT under construction. The array will consist of telescopes of three different diameters—Large-Sized Telescope (LST), Medium-Sized Telescope (MST), and Small-Sized Telescope (SST). Fig. 3.9 provides a graphical comparison of all telescope types in the CTA. With these telescopes and the optimized configuration of the array, the sensitivity of CTA is reported to be at least an order of magnitude better than existing VHEGR instruments[82]. The CTA is an international project run by a consortium involving 1500 members from more than 150 institutions in 25 countries. Once completed and opened, the Cherenkov Telescope Array Observatory (CTAO) will op-

Figure 3.9: Computer renderings of all types of telescopes in the CTAO: SST, Schwardschild–Couder Telescope (SCT), MST, and LST (from left to right). Credit: [83].

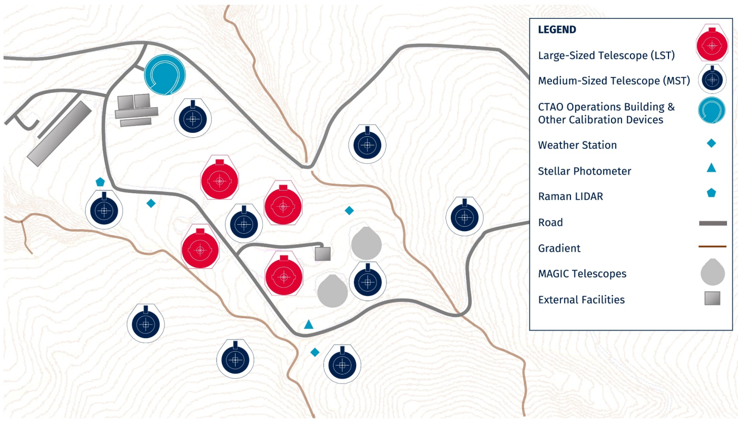
erate as an open observatory that accepts observation proposals from interested researchers and provides tools and support for data analysis[82].

Sites and layout optimization

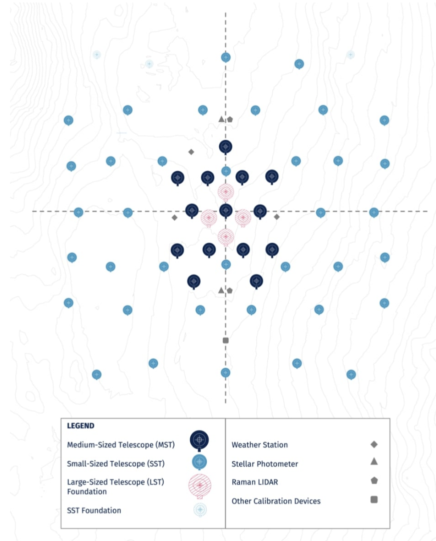
To provide all-sky coverage, the CTAO will consist of two arrays. The one in the Northern Hemisphere, called CTAO-North, is located on the island of La Palma in the Canary Islands (Spain), and the other in the Southern Hemisphere, called CTAO-South, is located in the Atacama Desert in Chile, close to Cerro Paranal. Both locations possess excellent seeing conditions that make them also the best places for optical and infrared observations.

A series of five large-scale Monte Carlo simulations were performed to optimize the layout of the telescope array, resulting in a full-scope array layout containing 118 telescopes[3] (called the Omega Configuration). Due to capiacty and funding constraints, a layout scenario called the Alpha Configuration was proposed and studied between 2016 and 2019, in which the northern and southern sites specialize temporarily[84]. According to the fine-tuned Alpha Configuration (Fig. 3.10) in 2021, the coordinates of each telescope are clearly specified such that the expected sensitivity of the array is guaranteed.

[3]The Omega Configuration fixed the number to 4 LSTs, 25 MSTs and 70 SSTs for the southern array, and 4 LSTs and 15 MSTs for the northern array.



(a) Layout of CTAO-North



(b) Layout of CTAO-South

Figure 3.10: Alpha Configuration. Credit: CTAO[85].

Table 3.1: Comparison between CTAO-North and CTAO-South.

	CTAO-North	**CTAO-South**
Altitude	2200 m	2100 m
Telescopes	4 LSTs & 9 MSTs (total: 13)	14 MSTs & 37 SSTs (total: 51)
Energy range	20 GeV – 5 TeV	150 GeV – 300 TeV
Specialized in	extragalactic γ-ray sources	galactic γ-ray sources

The Alpha Configuration for CTAO-North (Fig. 3.10a) focuses on the low- and mid-energy extragalactic sources with an energy range between 20 GeV and 5 TeV; 4 LSTs and 9 MSTs distributed over an area of about 0.5 km^2 will account for this observation goal. On the other hand, the Alpha Configuration for CTAO-South (Fig. 3.10b) is optimized in mid- and high-energy range (150 GeV up to 300 TeV), which specialize in the galactic gamma-ray sources. For this purpose, it will accommodate 37 SSTs and 14 MSTs, spread over an area of $\sim 3\,\text{km}^2$. In both arrays, foundations for additional telescopes are also included to meet the needs of future enlargement of the arrays. The comparison of the arrays is given in Table 3.1.

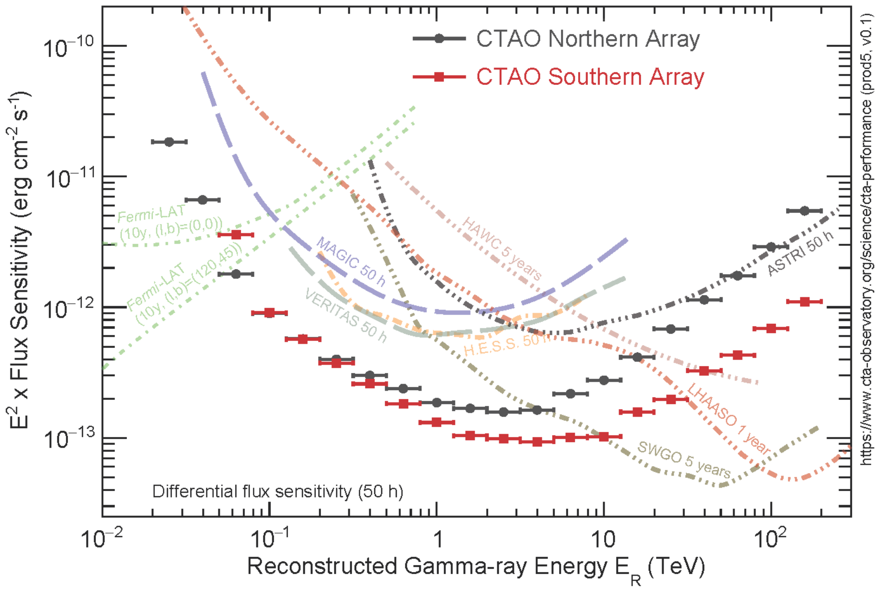


Figure 3.11: The differential sensitivity of CTA as a function of reconstructed energy, compared with other instruments. For the CTA arrays, the energy bins have a width of 0.2 decades. Credit: CTAO[87] with further references.

Performance

The criteria used to evaluate the performance of array layout includes differential sensitivity, effective collective area, angular resolution, energy resolution, and residual background rate[86].

The differential sensitivity is defined as the minimal flux to detect a point-like source with the significance $> 5\sigma$. Fig. 3.11 provides a comparison of the sensitivity of the different instruments. It is seen that sensitivity of CTA will reach $10^{-13}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$ and its energy coverage will reach approximately 100 TeV. Above this value, the particle detector arrays (HAWC, LHAASO, and SWGO) take avandages. In fact, because different reconstruction processes are used in different γ-ray detection techniques, this comparison is only a rough estimate.

Furthermore, CTA boasts an angular resolution ranging from 0.02° to 0.2° within its sensitive energy range, surpassing both current IACTs and air shower arrays in this regard. The angular resolution of CTA compared with other instruments is presented in Fig. 3.12.

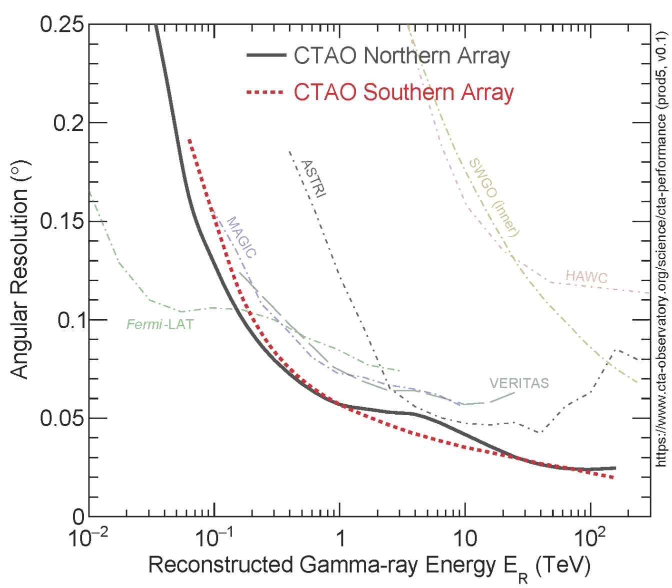


Figure 3.12: The angular resolution of CTA as a function of reconstructed energy, compared with other instruments. Credit: CTAO[87] with further references.

3.2.2 Small-Sized Telescope (SST)

The Small-Sized Telescope (SST) is a modified Schwarzschild–Couder (SC) telescope, with a primary mirror diameter of 4.3 m, a secondary mirror of 1.8 m, and a field of view of 8.8°. Optimized observation energies of SSTs are from 5 TeV to 300 TeV. Because extragalactic VHEGRs with energy higher than 10 TeV will be absorbed by the extragalactic background light (EBL)[88], the main observation objects are sources within the Milky Way. Therefore, all SSTs are located in CTAO-South, which is suitable for surveying the Milky Way. According to the Alpha configuration, CTAO-South will build at least 37 SSTs, making it the largest IACT array ever built. The final goal is to increase the number of telescopes to 70.

The optical design of SSTs is a modified version of the ASTRI-Horn, an dual-mirror SC telescope designed by ASTRI[4] group. And the camera of SSTs is based from the IACT camera, CHEC-S[5], in which 64 (8×8) SiPMs form a pixel tile, and 32 pixel tiles form a 2048-channel focal plane camera. The multi-coated entrance window of the camera functions as a

[4]Astrofisica con Specchi a Tecnologia Replicante Italiana (ASTRI), an Italian institute is engaged in building the ASTRI Mini-Array at the Teide Observatory on the island of Tenerife of the Canary Islands, aiming to detect γ rays with $E \sim 100$ TeV using only 9 IACTs[89].

[5]The Compact High Energy Camera (CHEC) is a camera design for SSTs. The prototype CHEC-M uses multi-anode photomultipliers (MAPMs), while the CHEC-S uses SiPMs[90].

bandpass filter (290 nm to 550 nm). The CHEC-S and the ASTRI telescope were combined to perform verification and on-sky test in 2019[91], and was later selected as the finalized version.

Other designs that have been proposed for SST include the single-mirror Small-Sized Telescope (SST-1M)[92, 93] and the dual-mirror Gamma-ray Cherenkov Telescope (GCT)[94]. Although the SST-1M was not selected as the finalized version of SST, its prototype is still under commissioning[95].

3.2.3 Medium-Sized Telescope (MST)

The sensitivity of the MSTs is optimized within the CTA's core energy range of 150 GeV to 5 TeV. The optical design of the MST is a modified version of DC optics with an diameter of 12 m, taking into account the PSF in the camera FoV and the isochronicity of the reflector. The support structure of the MST is made of steel to maintain rigidity, and the mirror part is inlaid with 86 small hexognal facets. In 2013 the MST prototype was built in Berlin.

Two versions of MST cameras were developed—NectarCAM and FlashCam. Given production capacity and funding constraints, MSTs of CTAO-North will employ the NectarCAM, while MSTs of CTAO-South will employ the FlashCam. Both of the camera versions offer nearly identical performance and have compatible interfaces[84]. Currently, both cameras use PMT as the photodetector, and the customization for SiPM is still under development.

NectarCAM comprises 265 seven-pixel PMT modules, resulting in an 1855-pixel camera. Within each pixel module, the signal is split into high-gain and low-gain channels. Both channel are amplified by the two linear preamplifiers on the focal plane module (FPM), and by the amplifier ACTA on the front-end board (FEB) amplifies, respectively. An additional trigger channel branch out in the FEB, as well. The events are sampled by "NECTAr" in the FEB, which is an analog pipeline application specific integrated circuit (ASIC) with a high sampling rate of 1 GSample/s. The signal is then digitized using a 12-bit ADC. In 2019, NectarCAM was installed in the MST prototype in Berlin[96]. Recent upgrades to FEB and NECTAr have improved the charge resolution and reduced the dead time[97].

FlashCam adopts a modular design concept of full digital readout and triggering within a horizontal architecture. Customized for the MSTs, FlashCam's Photon Detection Plane (PDP) comprises 147 twelve-pixel modules, arranged in a hexagonal array encompassing a total of 1764 PMTs with a side-to-side diameter of 2.1 m. Each module integrates essential functionalities including pre-amplification, high-voltage power supply, slow control, and monitoring. Utilizing a nonlinear preamplifier, each PMT pixel enables single-gain preamplification, with signal amplitude reconstruction achievable within a nonlinear range exceeding 200 photoelectrons (p.e.). Within the Readout System, the 12-bit Flash ADC facilitates continuous digitization with a sampling rate of 250 MSa/s, while the field-programmable gate array (FPGA) undertakes preprocessing, buffering, and triggering of the digitized signal. This combination of FPGA and the Flash ADC yields numerous advantages, including cost-effectiveness, low power consumption, high triggering performance, and superior amplitude and timing resolution across a wide dynamic range spanning from 0.2 p.e. to 3000 p.e. FlashCam underwent verification on H.E.S.S.'s fifth Cherenkov telescope (CT5) in 2019.[98].

Schwarzschild–Couder Telescope (SCT)

In addition to the MSTs based on DC optics, CTAO also has a proposal to build MSTs based on SC optics in the future, also known as the SCT. The primary mirror of the SCT will be 9.8 m in diameter, divided into 48 panels, and arranged in two concentric circles; the secondary mirror is 5.4 m in diameter. Like the DC-based MST, the FoV of SCT can reach 8°. For the photodetector, the SCT plans to use smaller and more compact SiPMs with respect to the traditional PMTs. Two SiPM models were tested in the prototype camera—Hamamatsu's Multi-Pixel Photon Counter (MPPC) and a high-density (HD) SiPM model developed by Fondazione Bruno Kessler (FBK). The full version of the camera will use 11 328 SiPM pixels, which are divided into 177 focal plane modules. Each module is coupled to a front-end electronics (FEE) module. An upgraded FEE module consists of a preamplifier stage (called SMART) and two independent ASICs—the TARGET-C ASIC for sampling and digitization, and the T5TEA ASIC for trigger[99]. The upgraded FEE modules has demonstrated an improved charge resolution[59].

The Prototype Schwardschild–Couder Telescope (pSCT) built at Whipple Observatory was inaugurated and commissioned in 2019. The pSCT has only 25 modules and a small 2.7° field of view, and yet it successfully detected the Crab Nebula with a significance of 8.6σ in 2020[100]. pSCT will continue to upgrade its focal plane module during the development phase[59].

3.2.4 Large-Sized Telescope (LST)

The sensitivity of the LSTs is specialized in the energy range of 20 GeV to 200 GeV, which can overlap with space-based instruments such as Fermi. The diameter of LST is comparable to the H.E.S.S. CT5, but the four telescopes allows the LST array to be more sensitive than Fermi and H.E.S.S., and also makes LSTs an ideal tool for studying transients and weak soft gamma ray sources[60].

The LST features a segmented parabolic mirror boasting a 23-meter diameter, coupled with a focal plane camera housing 1855 high quantum efficiency PMTs, offering a FoV spanning approximately 4.5°. Modularized into units of 7 channels, each sharing a readout circuit, the PMT undergoes dual-gain amplification stages, with the trigger signal derived from the high-gain output. To optimize power consumption, the Readout Board integrates a ASIC known as DRS4, which utilizes analog memory[6] for GHz sampling. The triggering mechanism involves the summation and comparison of signals between adjacent PMT modules' front-end and backplane boards, with digitization occurring between Level 0 and Level 1 trigger stages. Control over trigger signal distribution and clock synchronization is facilitated by the Trigger Interface Board (TIB)[9].

The first LST (LST-1) was completed in October 2018, and was first lighted and commissioned in December of the same year. Civil engineering works and installation of the other 3 telescopes have begun in 2023[9, 60]. During the commisioning, LST-1 has accumulated more than 800 hours of observations and reported couple of scientific results, including the Crab Nebula, PeVatron LHAASO J2108+5157[102], the outburst of the recurrent nova RS Ophiuchi, and the FSRQ[7] OP 313, which is the first detection of the VHEGR below 100 GeV[103].

Table 3.2 presents a detailed comparison of their respective technical specifications.

[6] MAGIC and FACT similarly employ DRS4 chips, where "DRS" stands for Domino Ring Sampler—an analog sampling based on a switched capacitor array (SCA). This method is also utilized by NECTAr and TARGET[101], distinguishing it from FlashCam's use of continuous digitization.

[7] flat-spectrum radio quasar (FSRQ), a type of AGN.

Table 3.2: Comparison of LST, MST, and SST. Source: [84, 104, 105].

	LST	MST	SST
Energy range	20 GeV – 150 GeV	150 Ge – 5 TeV	5 TeV – 300 TeV
Number of telescopes	4 (North)	9 (North) + 14 (South)	37 (South)
Mount	alt-azimuth	alt-azimuth	alt-azimuth
Optical design	parabolic	modified DC	SC
Tessellated mirror	198 facets, hexagonal	88 facets, hexagonal	18 facets (primary) monolithic (secondary)
Reflector diameter	23.0 m	12.0 m	4.3 m (primary) 1.8 m (secondary)
Focal length	28.0 m	16.0 m	2.15 m
Effective area	$\sim 400\,\mathrm{m}^2$	$88\,\mathrm{m}^2$	$\sim 8\,\mathrm{m}^2$
Field of view	4.5°	7.7° (NectarCAM) 7.5° (FlashCam)	9°
Pointing precision	$< 14''$	$< 7''$	$< 7''$
Number of pixels	1855	1855 (NectarCAM) 1764 (FlashCam)	2048
Pixel size (imaging)	0.1°	0.17°	0.16°
Photodetector type	PMT	PMT	SiPM
Readout data rate (readout of all pixels; before array trigger)	24 Gb/s	12 Gb/s	2.55 Gb/s

Chapter 4

Silicon Photomultiplier

4.1 Background knowledge about SiPM

4.1.1 Semiconductor

In solids, the periodic arrangement of the crystal lattice forms an energy band structure, of which the two most common energy bands are the valence band and the conduction band. The electrons in the valence band are bound by atoms and participate in bonding, while the electrons in the conduction band are not bound and can move freely. For semiconductors, the two energy bands separate by a region called the band gap in which no electron energy states exist. Typically, the conduction band has higher energy than the valence band, so the conductivity of the material depends on the size of the band gap. Typical values for the band gap of semiconductors range from 0.7 eV to 3 eV. This puts the conductivity of a semiconductor somewhere between that of a conductor (with overlapping conduction and valence bands) and an insulator (with a large band gap).

Doping an electron donor in an intrinsic semiconductor creates some impurity levels close to the conduction band, forming an n-type semiconductor in which the electrons raised to the conduction band appear as majority carriers. Conversely, doping with an electron acceptor creates impurity levels close to the valence band. This results in a p-type semiconductor where electron holes are left behind as electrons are raised to the those impurity levels. Electron holes can move in the valence band and behave as positively charged majority carriers in p-type semiconductors. The band gap of silicon (Si), a widely used semiconductor material, is approximately 1.1 eV at room temperature. For Si, the donor impurity is typically a group VA element (such as phosphorus), while the acceptor impurity is typically a group IIIA element (such as boron). Fig. 4.1 illustrates the band structure of n-type and p-type semiconductors.

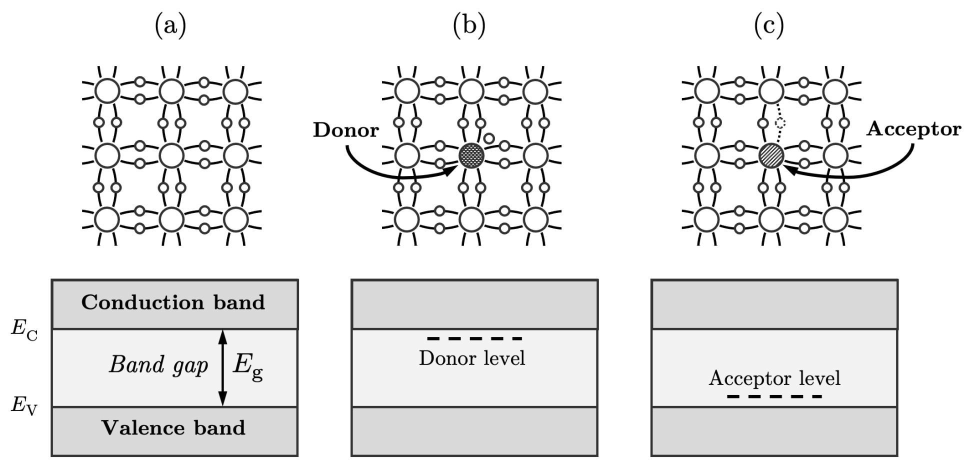


Figure 4.1: Covalent bondings (upper panels) and energy band diagrams (lower panels) of (a) intrinsic, (b) n-type, and (c) p-type semiconductors. Adapted from [106].

If an external field disturbs the thermal balance of carriers in a semiconductor, it will induce the generation and recombination of carriers, which are closely related to the behavior of optoelectronic devices. During carrier generation, electrons gain enough energy to enter the conduction band, correspondingly leaving holes in the valence band. The generation of carriers can occur through multiple pathways between band-to-band, shallow traps, and deep traps, where traps refer to impurities and lattice defects in the band gap. Three types of carrier production can be distinguished[107]:

- **Photoexcitation** induced by electromagnetic field (photons). Absorption of photons leads to excitation of electrons. This is the principle of photodiode detection of light.
- **Field ionization** induced by electic field. This can be caused by impact ionization, the Frenkel–Poole effect, or the tunneling effect. Impact ionization can lead to avalanche multiplication in the avalanche photodiode.

- **Thermal excitation** induced by phonon field. This is due to the random motion of atoms within the crystal lattice above the absolute zero. The thermal excitation assisted by deep traps is called the Shockley–Read–Hall (SRH) generation/recombination, which is the leading cause of dark noise in a SiPM (see Section 4.4.1). In addition, both the Frenkel–Poole and tunneling effects mentioned above can enhance the SRH generation in a highly reversed biased p–n junction[108].

4.1.2 p–n junction

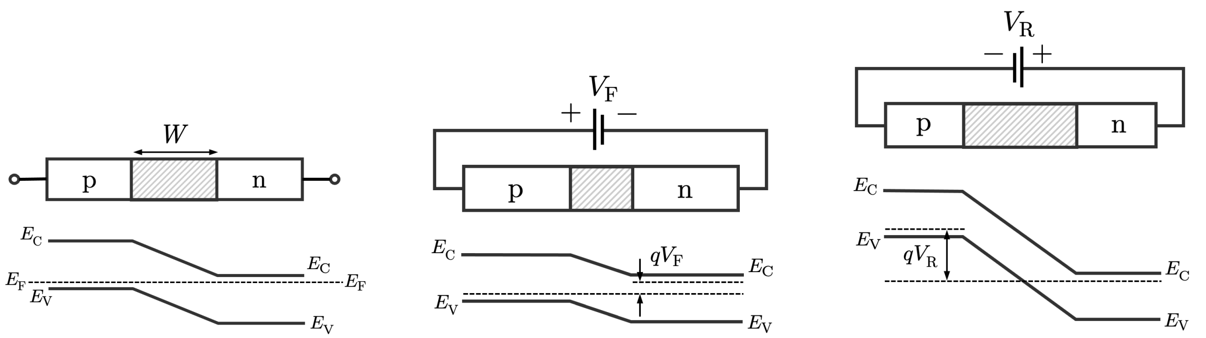
The p–n junction is the boundary of p-type and n-type semiconductors. During the formation of the p–n junction, the majority carriers of both sides diffuse towards the opposite side due to the concentration gradient. When thermal equilibrium is reached, some features can be observed:

- There is a built-in potential that blocks further diffusion of majority carriers across the junction.
- There is a depletion region at the junction, where the concentration of majority carriers becomes very low.
- There is a band bending—the n-type region bends upward, and the p-type region bends downward (Fig. 4.2a), which can be interpreted as the alignment of the Fermi levels[1].

Given the profile of doping concentration, one can calculate the depletion layer width and junction capacitance of the p–n junction, both of which has direct effect on the behaviors of SiPM. The width of the depletion layer W_{dep} determines the range within which avalanche breakdown can occur. Therefore, it also determines the avalanche triggering probability and breakdown voltage. Another important parameter is the junction capacitance[2] C_{J}, which determines the gain and the recovery time of a SPAD or a SiPM microcell (see Eqs. (4.6) and (4.11)). C_{J} is related to the geometry of the p–n junction and can be approximately described by the formula for a parallel-plate capacitor:

[1]The Fermi level represents the energy at which the occupation probability of an electron is 50%. Because of the presence of impurity levels, the Fermi level of p-type semiconductors is close to the valence band, and that of n-type semiconductors is close to the conduction band. The term “Fermi” originates from the Fermi–Dirac distribution, which describes the statistical distribution of electrons in an energy band.

[2]Also known as diode capacitance in general; also known as microcell capacitance in context of SiPM.



(a) Thermal equilibrium (zero bias). E_F is the Fermi energy. W is the width of the depletion region.

(b) Forward bias. V_F is the forward bias voltage.

(c) Reversed bias. V_R is the reverse bias voltage.

Figure 4.2: Biasing of p–n junctions, showing the bending of the energy bands. Adapted from [109].

$$C_J = \epsilon_0 \cdot \epsilon_{Si} \cdot \frac{A_J}{W_{dep}}, \tag{4.1}$$

where $\epsilon_0 = 8.854 \times 10^{-14}\,\mathrm{F\,cm^{-1}}$ is the vacuum permittivity, $\epsilon_{Si} = 11.9$ is the relative permittivity of silicon, A_J is the cross-section area of the junction. For SPADs, the typical value of C_J is $\lesssim 1\,\mathrm{pF}$[110, 111].

Forward bias and reverse bias

Applying a voltage polarity (i.e., bias) can change the width of the depletion layer and bring functionality to the p–n junction. Forward bias refers to the biasing where a positive voltage is applied to the p-side and negative to the n-side, and the reverse bias refers to the opposite polarity. The corresponding energy band changes are shown in Figs. 4.2b and 4.2c.

A forward bias narrows the width of depletion region, lower the potential barrier across the junction so that the majority carriers of both p- and n-sides are able to diffuse through the depletion region. This is how a diode conducts electricity under forward bias.

On the contrary, when a reverse bias is applied, the majority carriers are pushed further away from the junction, the depletion region becomes wider, and the built-in voltage also increases. These factors make it less likely that majority carriers will diffuse across the junction. However, there is still an extremely small current ($< 10^{-12}$ A, called the reverse saturation current) arising from the thermal generation and the diffusion of minority carriers[112]. So, the p–n junction exhibits very high resistance under reverse bias conditions.

Electrical breakdown

When the reverse voltage reaches specific critical points, the junction experiences an electrical breakdown, i.e., a sudden surge in reverse current. There are two main mechanisms for electrical breakdown of p–n junctions.

The first is band-to-band tunneling, as shown in Fig. 4.3a. Its formation conditions are usually a heavily doped PN junction and a narrow depletion region. When the external electric field is strong enough, the energy band bends more, and the depletion zone becomes narrower, which may cause *Zener breakdown*.

The second mechanism is avalanche multiplication, as shown in Fig. 4.3b. If carriers obtain enough energy while accelerated by an external electric field under reverse bias, they can ionize atoms by collision. This process is called impact ionization. As mentioned earlier, the ionization here actually refers to an electric field-induced carrier generation, where electrons enter the conduction band and leave holes in the valence band. To the first-order approximation, the minimum energy required for impact ionization is 1.5 times the band gap[113]. Detailed reasons can be found in [109, 114]. The electron-hole pairs produced by impact ionization can themselves be accelerated and continue to collide with atoms, triggering impact ionization once again, leading to a local chain reaction whereby the number of carriers multiplies like an avalanche.

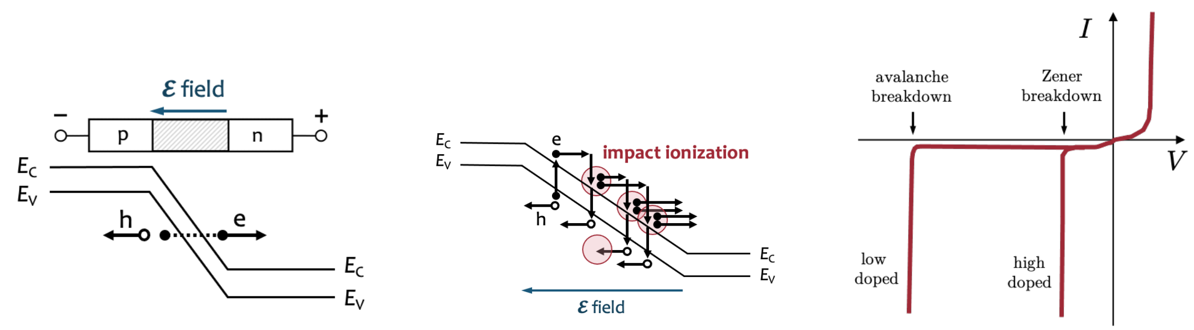
When the electric field is moderate, the localized high carrier concentration (*micro-plasma*) caused by multiplied carriers will effectively shield the external electric field, thereby preventing carrier acceleration and ultimately quenching avalanche multiplication. This scenario is described as self-quenched. Only if the reverse bias voltage is greater than the *breakdown voltage*[3] of the p–n junction is the avalanche multiplication self-sustained, bringing about an *avalanche breakdown*. In this case, avalanche multiplication does not stop on its own unless

[3] Zener breakdown, of course, has its own definition of "breakdown voltage." However, in the context of photodiode research, the breakdown voltage is understood as the one defined for avalanche breakdown. We follow this definition.

quenched by an external circuit, called a quenching circuit.

The avalanche breakdown of p–n junction can be well formulated by the concept of ionization integral. In fact, the ionization integral provides the condition of the avalanche breakdown and also theoretical calculation of breakdwon voltage.

Zener breakdown and avalanche breakdown generally do not occur on the same devices. If the doping conditions are favorable to electric field enhanced band-to-band tunneling (such as a Zener diode), then the device will undergo Zener breakdown when the reverse bias voltage is still low; otherwise, it will have a high breakdown voltage beyond which avalanche breakdown takes place, as shown in Fig. 4.3c.



(a) Band-to-band tunneling can cause Zener breakdown.

(b) Avalanche multiplication can cause avalanche breakdown.

(c) The I–V curve of the PN junction shows two different breakdown voltages.

Figure 4.3: Electrical breakdown of p–n junction

4.1.3 Operation modes of photodiodes

Photodiodes are semiconductor devices that convert light into photocurrent using a p–n junction. As mentioned in the Section 4.1.2, applying reverse bias can increase the depletion area on both sides and expand the light absorption volume. The intensified electric field in the depletion area provides faster and more linear response. Also, the lack of current flow also makes the photocurrent component easier to identify. Thus, photodiodes typically operate in reverse bias. Fig. 4.4 illustrates the photodiode's current-voltage (I–V) characteristics under reverse bias, from which different operation modes can be distinguished.

The so-called *linear mode* refers to the biasing below the breakdown point V_{BD}. The pho-

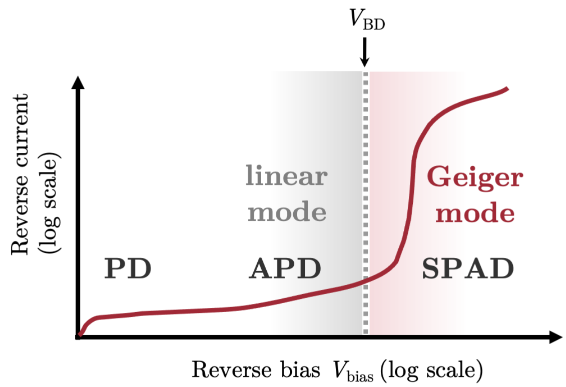


Figure 4.4: Operation modes of the photodiode. Referenced from [115]. PD = photodiode, e.g., PIN photodiode that will be introduced in Section 4.1.4.

todiode's output photocurrent $I_{\mathrm{PD}}^{\mathrm{linear}}$ in the linear mode is proportional to the incident light intensity and can be expressed by[116],

$$I_{\mathrm{PD}}^{\mathrm{linear}} = G \cdot R(\lambda) \cdot \Phi_E, \tag{4.2}$$

where G is the gain, which is the factor by which the carriers are multiplied in the p–n junction; R is the responsivity in the unit $\mathrm{A\,W^{-1}}$, which is dependent on wavelength λ; Φ_E is the radiation power incident on the photodiode, and the unit is W. The product of G and $R(\lambda)$ can be referred to as photosensitivity, and $(R(\lambda) \cdot \Phi_E)$ can be interpreted as the primary photocurrent.

In contrast, the *Geiger mode* is a region biased above V_{BD}, where the output current is *not* proportional to the incident light intensity. The name "Geiger" originates from the Geiger region of gaseous ionization chambers.[117]

Whether there is a clear demarcation between PIN photodiodes and avalanche photodiodes (APDs) is a question that is not easy to answer. The reason is that APDs often use different doping structures from PIN photodiodes due to functional requirements. The I–V curve of APDs under illumination has the feature of "punch-through voltage"[118], but the details are not discussed here.

The operation modes of photodiodes and their properties can be summarized in Table 4.1.

Table 4.1: Operation modes of photodiodes

Mode	Deivces	Avalanche multiplication	Reverser bias V_{bias}	Gain
Linear mode	PIN photodiode	almost none	$V_{\text{bias}} < V_{\text{BD}}$	~ 1
	APD	yes, self-quenched		$< 10^3$
Geiger mode	SPAD, SiPM microcell	yes, self-sustained	$V_{\text{bias}} > V_{\text{BD}}$	$\sim 10^6$

4.1.4 PIN photodiode

PIN photodiodes are photodiodes using a p-i-n structure that improves light absorption capabilities. In the p-i-n structure, a layer of intrinsic semiconductor (i-region) is sandwiched between p-type and n-type semiconductors. In practice, the p-i-n structure can achieved by letting a high ohmic p-layer (π layer) or a high ohmic n-layer (ν layer) sandwiched between a heavily doped p-layer (p^+ layer) and a heavily doped n-layer (n^+ layer)[119]. The purpose of heavily doping both sides is to form a good ohmic contact so that charges will not accumulate at the metal–semiconductor interface. As for the i-region, it becomes a wide depletion region with a uniform electric field, which acts as a photosensitive volume. This structure brings high quantum efficiency. However, such a design cannot provide internal gain and thus cannot achieve single photon detection.

4.1.5 Avalanche photodiode

Avalanche photodiodes (APDs) are photodiodes that are designed so that their biasing is slightly below the breakdown point V_{BD}. Its output current is still proportional to the incident light intensity (linear mode). The advantage of APDs compared to PIN photodiodes is their internal gain, allowing them to detect lower amounts of light. The disadvantage is that the gain also fluctuates due to the statistical fluctuation in the avalanche process. Given this limitation, the useful gain range for silicon APDs can only be set between 100 and 1000; the numbers are even lower for germanium and InGaAs APDs[116]. If an APD is required for single-photon detection, special measures regarding the readout circuitry (such as limited operating bandwidth and low-noise amplifiers) must be adopted[120].

4.1.6 Single-photon avalanche diode

Biased above the breakdown voltage, Geiger-mode avalanche photodiodes (GM-APDs) can attain much higher gain than ordinary PIN photodiodes and linear-mode APDs. Amplifying light signals with high gain is advantageous for single-photon detection. This is why GM-APDs are also called single-photon avalanche diodes (SPADs). The gain of SPAD is nominally infinite. However, in reality, even if no external quenching is implemented, the gain is limited by the space charge effect, contact resistance, diode heating, and circuit impedance[121], and it will not reach infinity.

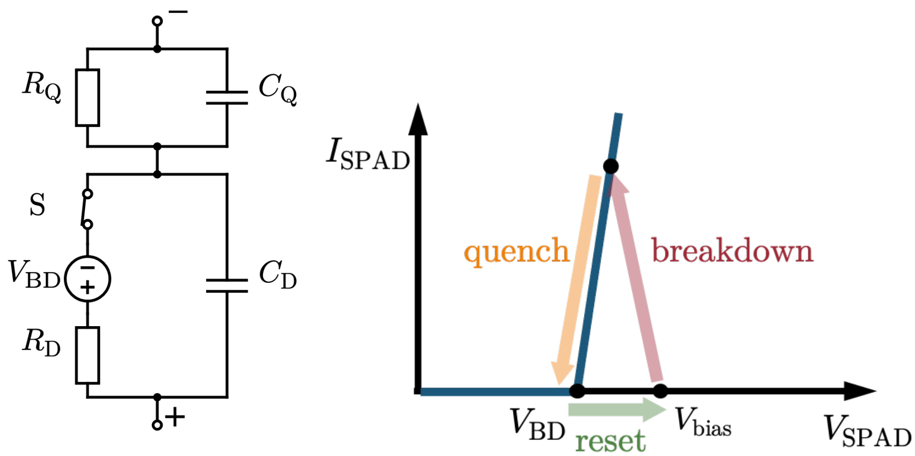
4.1.7 Passive quenching of SPADs

In this subsection, the passive quenching of SPADs will introduce, which is the key to understanding the long-tail response of SiPMs.

The avalanche breakdown of SPAD requires external quenching. The simplest way to quench the breakdown is to connect the SPAD in series with a high ohmic resistor with a resistance R_{Q}, known as the quenching resistor. The typical value of quenching resistance R_{Q} is $\gtrsim 100\,\mathrm{k\Omega}$. This method is called *passive quenching* because it only involves a passive component.

The equivalent electrical model of the passively quenched SPAD has been proposed by [122] and reformulated by, e.g., [110, 111, 121]. Let us consider a version depicted in Fig. 4.5a. Here, the quenching resistor is modeled as R_{Q} and C_{Q} in parallel, where C_{Q} is its parasitic capacitance. The SPAD is modeled as a constant voltage source V_{BD}, diode resistance R_{D}, and a switch S connected in series, together in parallel with C_{D}, the diode capacitance (i.e., junction capacitance, see Section 4.1.2).

Fig. 4.5b shows the dynamical evolution of the three phases. When $V_{\mathrm{SPAD}} = V_{\mathrm{bias}}$, the SPAD is in its idle state, waiting for the arrival of photons. Upon the triggering of avalanche breakdown, the SPAD goes through three stages—breakdown, quenching, and reset, and then returns to the idle state.



(a) Equivalent electrical model. (b) The evolution of the three stages on the I–V curve.

Figure 4.5: Passive quenching.

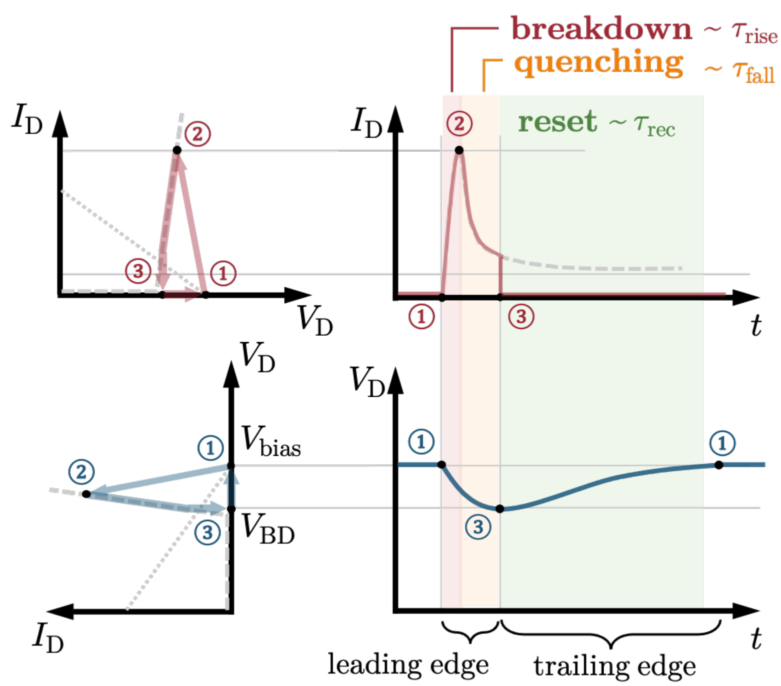


Figure 4.6: Output voltage V_D and current I_D of SPAD during breakdown phase, quenching phase, and reset phase. *Upper left*: I–V curve; *lower left*: I–V curve (rotated by 90°); *upper right*: current vs. time; *bottom right*: voltage vs. time. The numbers indicate the three phases in order. The V_D and I_D axes are not true to scale.

Breakdown stage Once the avalanche breakdown is triggered, the switch S is closed, and C_{D} discharges suddenly, making the SPAD current I_{D} exponentially rise to its maximum

$$I_{\max} = \frac{V_{\mathrm{bias}} - V_{\mathrm{BD}}}{R_{\mathrm{D}}} = \frac{V_{\mathrm{OV}}}{R_{\mathrm{D}}}, \tag{4.3}$$

with a rise time constant

$$\tau_{\mathrm{rise}} = R_{\mathrm{D}} \times C_{\mathrm{D}}, \tag{4.4}$$

where C_{Q} is neglected in the first approximation. An estimation of τ_{rise} is given by $R_{\mathrm{D}} \times C_{\mathrm{D}} = (100\,\Omega) \times (100\,\mathrm{fF}) = 10\,\mathrm{ps}$.

Quenching stage As the external current starts to flow through the quenching resistor, its high impedance leads to a considerable voltage drop, making the SPAD current decay exponentially. With an RC circuit model, the asymptotic steady-state minimum is found to be

$$I_{\mathrm{f}} = \frac{V_{\mathrm{bias}} - V_{\mathrm{BD}}}{R_{\mathrm{D}} + R_{\mathrm{Q}}} \simeq \frac{V_{\mathrm{OV}}}{R_{\mathrm{Q}}}, \tag{4.5}$$

where the approximation holds under the assumption that $R_{\mathrm{Q}} \gg R_{\mathrm{D}}$, which is almost always true. The fall time constant of the current decay is given by

$$\tau_{\mathrm{fall}} = R_{\mathrm{D}} \times (C_{\mathrm{D}} + C_{\mathrm{Q}}). \tag{4.6}$$

For a functional quenching circuit, $I_{\mathrm{D}}(t)$ will never reach I_{f}. Instead, when $I_{\mathrm{D}}(t)$ is less than a critical value, the statistical fluctuation in the multiplication factor is sufficient to cease the breakdown. The critical value of quenching is called the *latching current* I_{latch}. There are no defined values for I_{latch}, but it is considered somewhere between $10\,\mu\mathrm{A}$ and $20\,\mu\mathrm{A}$[111, 113, 121].

As can be inferred from the existence of the constant voltage source in the model, the voltage across the SPAD during quenching will remain above the breakdown point.

Reset (recovery) stage When the avalanche breakdown is fully quenched, the switch S is open again. Now, C_{Q} and C_{D} are recharged through R_{load} and R_{Q} up to the original level, V_{bias}. The recovery time constant[4] is given by

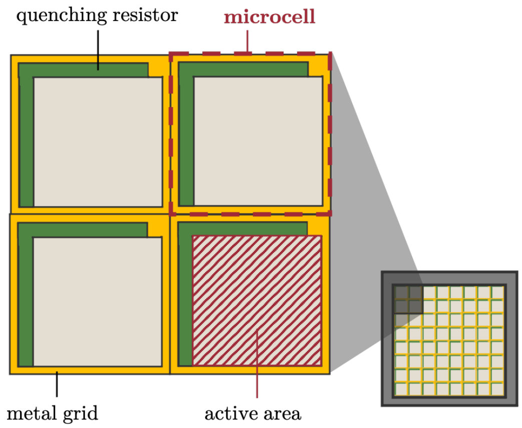
$$\tau_{\mathrm{rec}} = (R_{\mathrm{D}} + R_{\mathrm{load}}) \times (C_{\mathrm{D}} + C_{\mathrm{Q}}). \tag{4.7}$$

After approximately $5\tau_{\mathrm{rec}}$, the bias is restored to 99% of the intial level, and the SiPM is considered ready to acquire the correct pulse height from the arrival of a new photon[121]. A method to directly measure SiPM's recovery time constant can be found in, e.g., [50].

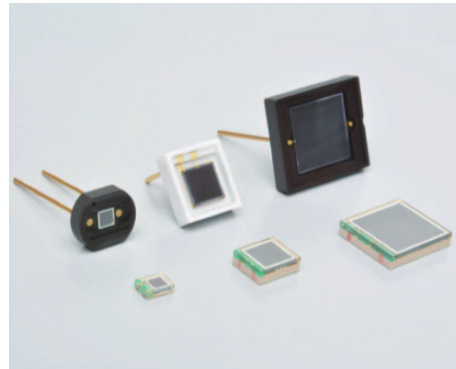
The temporal behaviors of voltage and current and their relations with the I–V curve is visualized in Fig. 4.6. The three time contants governing the three stages are also labeled. Basically, what we observe on the oscilloscope through through some readout circuits is the temporal behavior of voltage, which is featured by a long trailing tail. This is the nature of the RC circuit and also an inherent limitation of the passive quenching circuit (PQC). While the voltage across the SPAD is restored from $\sim V_{\mathrm{BD}}$ to V_{bias}, the ability to detect photons is also restored exponentially.

For single-cell SPADs or complementary metal-oxide-semiconductor (CMOS) SPADs, the passive quenching has been replaced by the active quenching circuit (AQC), which uses a transistor-based comparator to sense the avalanche breakdown and actively compensate a user-defined voltage at a user-defined hold-off time[110]. However, a AQC is space-consuming and will lower the geometric efficiency if the actively-quenched SPADs are to be integrated into a chip.

[4] Regarding the term "recovery time", there is ambiguity between different literatures. Some literature refers to "the time for the bias voltage to return to 99% of the initial level from the lowest point", which is similar to the definition of rise time; some literature refers to "the decay time constant of exponential fitting"[50]. In this study we use the latter definition.



(a) SiPM as an array of microcells.



(b) Hamamatsu MPPC® S13360 series in different packages. Photo credit: [124].

Figure 4.7: Silicon photomultipliers.

4.2 Concept and Realization of SiPMs

4.2.1 Concept

A SiPM can be defined as a *pixelated monolithic* array of SPADs. "Pixelated" means the SiPM is composed of *microcells*,[5] each of which is a SPAD in series with its quenching resistor. The number of microcells in a SiPM ranges between several hundred and several tens of thousands. The microcell pitch, i.e., the distance between two adjacent microcells, varies from 10 µm to 100 µm[123]. "Monolithic" means each SPAD is fabricated on the same substrate and connects in parallel to the same readout.

The simplified circuit diagram of a 4-microcell SiPM is shown in Fig. 4.8, from which one can quickly see how a SiPM works as a photon counter. All microcells are biased to the same voltage above the breakdown point and thus work in the Geiger mode. Therefore, triggering an avalanche breakdown in any single microcell will output an identical pulse response with the pulse height and same charge released. This is called the single-photoelectron response (SPR) in this thesis. If n different microcells are fired at the same time, then their outputs will be summed in the readout circuits n times and the result will be n times single-photoelectron

[5] Microcells are also known as "pixels" or simply "cells." Note than the name "pixel" does not imply SiPMs can resolve images like a camera. In fact, SiPMs do not have spatial resolution.

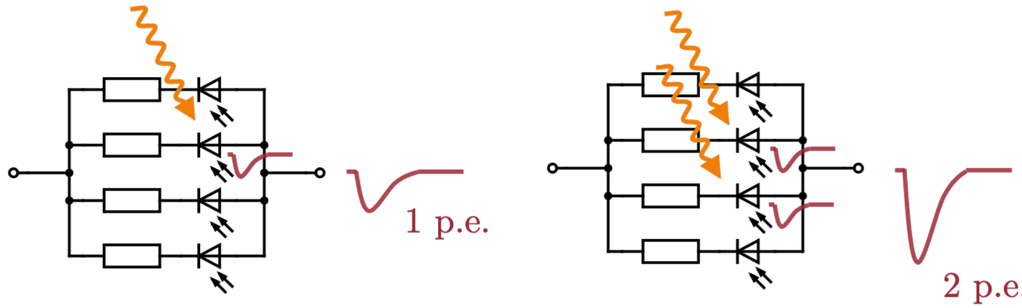


Figure 4.8: SiPM as a counter.

response (SPR), with n times the pulse height and n times the charge released. In other words, each microcell is a binary device whose output can be either 0 (OFF) or 1 (ON), while the combined output of N microcell in total provides N-bit resolution of output, making the SiPM an N-bit counter.

The concept of SiPM was proposed and patented in the late 1990s by Russian scientists Z. Sadygov and V. M. Golovin[121]. SiPMs overcame the lack of linearity of the single SPAD. Through continuous improvements in the 2000s, they finally achieved outstanding PDE, high internal gain, and low intrinsic noise.

Today, SiPMs are widely used in various applications, demonstrating their practicality and versatility. Well-known vendors of SiPM include Hamamatsu Photonics K.K., Onsemi (formerly SensL), and KETEK GmbH. Some academic institutes, e.g., Fondazione Bruno Kessler (FBK) and Moscow Engineering Physics Institute (MEPhI), are also engaged in the R&D of SiPMs, further expanding their potential applications.

Although SiPMs are mostly realized in silicon because of the well-established fabrication process of CMOS. The idea of realizing SiPMs with germanium and III/V materials is also envisaged especially for near-infrared (NIF) detection[110, 125].

4.2.2 Realization

Except for a few custom designs, SiPMs are usually manufactured using standard CMOS fabrication techniques, i.e., thermal oxidation, epitaxial growth, ion implantation, etc. Fig. 4.9 shows an exemplary structural design of SiPM based on the so-called p-on-n structure SPAD,

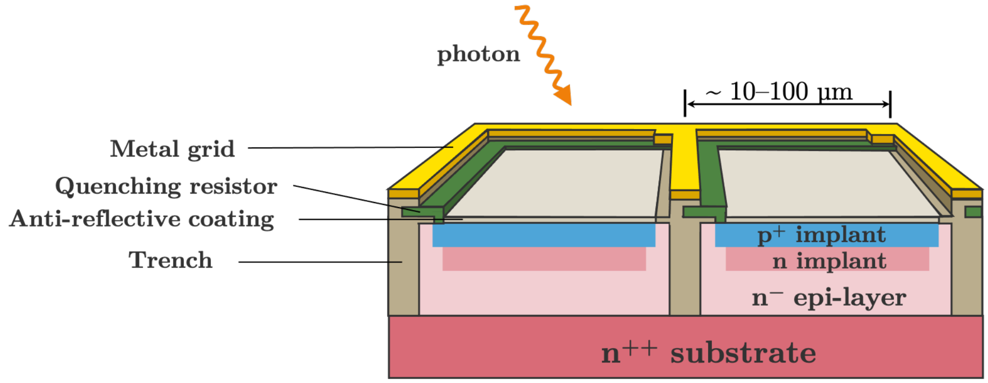


Figure 4.9: Exemplary structural design of a SiPM with p-on-n microcell.

where a shallow p-type Si is implanted on the top of n-type Si. By swapping the p-type and n-type, we have the n-on-p structure. The different orders of stacking p- and n-type layers is an approach to optimize SiPM's response to different wavelengths. This is due to the wavelength dependence of the absorption length and CMOS process considerations.

The crucial parts for avalanche breakdown are the *active region* and the *high-field region*. The active region is where carriers can be generated by photoexcitation and drift under the guidance of the electric field. This range is not always clearly defined but usually includes the lightly doped epitaxial layer, where the majority carrier is fully depleted, and a weak electric field is present. The typical depth of active region is a few µm.

Characterized by the maximal electric field and maximal ionization rate, the high-field region is where the carriers are most likely to trigger avalanche breakdown. Its typical depth is a few hundreds of nm.

Other parts of the SiPM structure include polysilicon quenching resistor to provide high resistance, metal grid to supply bias voltage to each microcell, anti-reflective coating to minimize reflection for certain wavelengths, and trenches to isolate inter-cell photon transmission so as to prevent optical crosstalk.

Note that Fig. 4.9 is only a conceptual sketch and may differ from real designs cater for different needs. For example, CMOS-compatible designs may have a feature called "guard rings". These are extra doping at the edges of active region to tune the electric field uniformity and prevent pre-mature breakdown[113]. This feature is not shown here.

4.3 Characteristics of SiPMs

4.3.1 Breakdown voltage and overvoltage

There are two dominating factors of breakdown voltage V_{BD}: the internal structure (especially refering to the doping profile) and temperature. The dependence of V_{BD} on the structure is embodied in the the ionization integral. Because the doping profile is designed by the manufacturer, the SiPM products has a nominally defined value of V_{BD}. However, due to the uncertainty in controlling doping in the actual fabrication process, the values of V_{BD} can spread out in a wide range for individual chips, especially if they are from different wafers or lots[126, 127]. Typical values of V_{BD} are tens of volts. For example, the SiPM used in our experiments has a breakdown voltage of about 52 V at room temperature.

Temperature dependence of breakdown voltage

The temperature dependence of the breakdown voltage can be understood qualitatively in the following way[113]. As the temperature increases, carriers collide with atoms or phonons in the crystal lattice more frequently, so they gain less average energy due to electric field acceleration during each collision. As a result, the critical electric field required to sustain avalanche breakdown is higher, so the breakdown voltage increases.

The temperature dependence of breakdown voltage can be linearly approximated near room temperature[128]:

$$V_{\mathrm{BD}}(T) = V_{\mathrm{BR},0}(1 + \beta(T - T_0)), \tag{4.8}$$

where the subscript 0 designates the reference value near room temperature T_0, and β is the temperature coefficient of breakdown voltage.

Overvoltage

Various parameters of SiPMs are dependent to difference between the reverse bias voltage V_{bias} and the breakdown voltage V_{BD}. So the definition of *overvoltage* V_{OV} is introduced:

$$V_{\text{OV}} := V_{\text{bias}} - V_{\text{BD}}. \tag{4.9}$$

The relative overvoltage $V_{\text{OV}}/V_{\text{BD}}$ can be defined as

$$\frac{V_{\text{OV}}}{V_{\text{BD}}} = \frac{V_{\text{bias}} - V_{\text{BD}}}{V_{\text{BD}}}. \tag{4.10}$$

Since the breakdown voltage V_{BD} vaires from different manufacturer and different models, when evaluating the dependence of performance parameters (e.g., as the PDE or the crosstalk probability) on overvoltage, it is best to use relative overvoltage as a reference parameter[129].

4.3.2 Gain

The gain G of photodiodes can be understood as the number of secondary carriers produced by each primary carrier. This is, however, not based on measurable quantities and is more suitable for theoretical considerations. Instead, a common operational definition for G is

$$G = \frac{Q}{e} = \frac{C_{\mu\text{cell}} \cdot V_{\text{OV}}}{e} = \frac{C_{\mu\text{cell}} \cdot (V_{\text{bias}} - V_{\text{BD}})}{e}, \tag{4.11}$$

where Q is the charge developed in one avalanche event and collected by the readout circuit, and q_{e} is the elementary charge. The charge Q is equal to the product of the overvoltage V_{OV} and the total capacitance $C_{\mu\text{cell}}$ of the microcell: $Q = C_{\mu\text{cell}} V_{\text{OV}}$; $C_{\mu\text{cell}}$ is the capacitance within one microcell and can be expressed as the sum of junciton capacitance and the capacitance of quenching resistor ($C_{\mu\text{cell}} = C_{\text{D}} + C_{\text{Q}}$). It is clearly seen that G is dimenionless.

From experimental point of view, Eq. (4.11) can be recasted in the following form:

$$G = \frac{Q}{e} = \frac{\int I_{\text{out}}(t)\,\mathrm{d}t}{e} = \frac{\int V_{\text{out}}(t)\,\mathrm{d}t}{R_{\text{load}} \cdot G_{\text{amp}} \cdot e}, \tag{4.12}$$

where $I_{\text{out}}(t)$ and $V_{\text{out}}(t)$ are SiPM's output current and output voltage, respectively. The integration of $V_{\text{out}}(t)$ over time amounts to the area between the signal and the baseline seen at the oscilloscope. Besides, the coversion I_{out} from to V_{out} relies on G_{amp}, the amplifier gain, and R_{load}, the resistance of the load circuit. These two factors can merge into one in certain cases, e.g., in our readout circuit with transimpedince amplifier (see Section 6.2). In addition, if a charge-to-digital converter (QDC) is adopted in the readout circuit, then the integration of I_{out} can be completed at the hardware level, provided that a proper integration time is chosen.

The fluctuation in gain arises from the uncertainty in the number of carriers developed in the avalanche multiplication process. Excess noise factor (ENF) is a way to quantify the fluctuation of gain. It is defined as

$$\text{ENF} = 1 + \frac{\sigma_G^2}{\langle G \rangle^2}, \tag{4.13}$$

where $\langle G \rangle$ and σ_G represent the mean and standard deviation of the gain G, respectively. For a SiPM, σ_G is small, so its ENF is almost equal to 1. The more important noise for SiPMs is the correlated noise (see Section 4.4.2), not the gain fluctuation. To adapt to this feature, the definition of ENF for SiPMs can be modified[113, 130].

4.3.3 Photon detection efficiency

The photon detection efficiency (PDE) of a SiPM is defined as the ratio between the number of the photons *detected* by the detector and the number of the photons *arriving* at the detector:

$$\text{PDE} = \frac{\text{Number of detected photons}}{\text{Number of arriving photons}}. \tag{4.14}$$

The PDE is generally factorized in three parameters: geometrical efficiency, quantum efficiency, and triggering probability[113, 121, 131, 132], i.e.,

$$\mathrm{PDE}(V_{\mathrm{OV}}, \lambda) = \mathrm{QE}(\lambda) \cdot P_{\mathrm{trig}}(V_{\mathrm{OV}}, \lambda) \cdot \mathrm{FF}(V_{\mathrm{OV}}, \lambda), \tag{4.15}$$

where QE is the quantum efficiency, P_{trig} is the triggering probability, FF is the fill factor (a measure of geometrical efficiency), V_{OV} is the overvoltage, and λ is the wavelength of incident photon. The definitions are explained below.

Avalanche triggering probability

The avalanche triggering probability[6] P_{trig} is the probability that a carrier triggers an avalanche breakdown. This concept seems intuitive, but different carriers, as well as different scenarios of carrier generation and transport, may have different avalanche triggering probabilities. We cannot assume *a priori* that they have the same value. For example, [108] defines different avalanche triggering probabilities when treating events such as dark pulse, optical crosstalk, afterpulse.

In general, avalanche triggering probabilities is obtained by derivation of measurable parameters and it is known experimentally that P_{trig} depends on voltage and wavelength. The dependence of overvoltage can be described by an empirical formula[108]:

$$P_{\mathrm{trig}} = 1 - e^{p \cdot (V_{\mathrm{BD}} - V_{\mathrm{bias}})}, \tag{4.16}$$

where p is a parameter obtained from fitting. It is worth mentioning that the avalanche triggering probability is also a factor of the photon detection efficiency (PDE, see Section 4.3.3), which also introduce voltage and wavelength dependence on the PDE.

To quantatively describe P_{trig}, we can first introduce the differential triggering probabilities for electrons and holes. Consider the one-dimensional model of depletion region. Let $P_{\mathrm{e}}(x)\ \mathrm{d}x$ be the probability that an eletron triggers an avalanche between positions x and $x + \mathrm{d}x$, and $P_{\mathrm{h}}(x)\ \mathrm{d}x$ be the counterpart for a hole, then the following differential equations

[6] also called triggering probability, Geiger triggering probability, Geiger probability, breakdown probability, turn-on probability P_{01}. The last alias is to contrast with the turn-off probability P_{10} in the context of hysteresis phenomenon of avalanche mechanism.

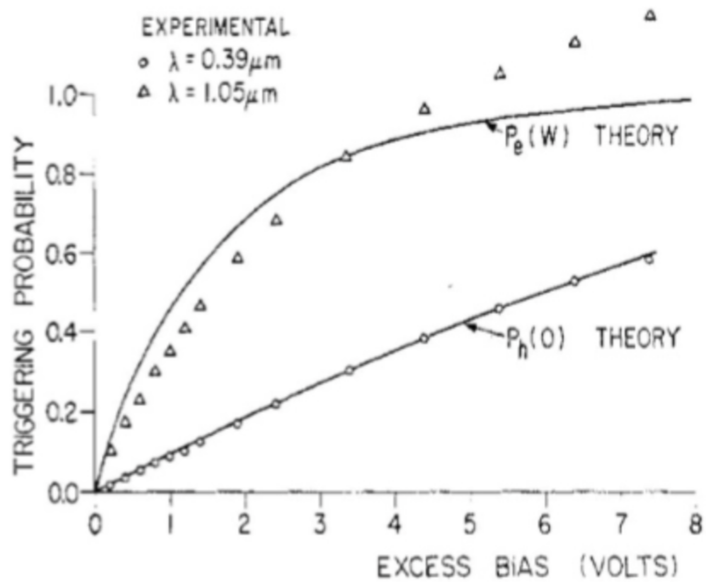


Figure 4.10: The avalanche triggering probability of electrons and holes in a SPAD, as a function of excess bias (=overvoltage). Credit: [133].

hold:[134]

$$
\begin{aligned}
\frac{\partial P_{\mathrm{e}}}{\partial x} &= (1 - P_{\mathrm{e}})\alpha_{\mathrm{p}}(P_{\mathrm{e}} + P_{\mathrm{h}} - P_{\mathrm{e}}P_{\mathrm{h}}), \\
\frac{\partial P_{\mathrm{h}}}{\partial x} &= -(1 - P_{\mathrm{h}})\alpha_{\mathrm{n}}(P_{\mathrm{e}} + P_{\mathrm{h}} - P_{\mathrm{e}}P_{\mathrm{h}}).
\end{aligned}
\tag{4.17}
$$

If we know the electric field profile $E(x)$ and the ionization coefficients α_{p} and α_{n}, we can integrate Eq. (4.17) with the boundary conditions in which carriers leaving the high field region have zero probability to trigger an avalanche, and obtain the total triggering probability[121]

$$
P_{\mathrm{trig}} = P_{\mathrm{e}} + P_{\mathrm{h}} - P_{\mathrm{e}}P_{\mathrm{h}}. \tag{4.18}
$$

If the overvoltage is not too high, the P_{trig} is proportional to electric field strength in the junction and thus overvoltage. But as P_{trig} approaches 100%, the proportionality loses. This can be seen in Fig. 4.10.

Besides, P_{trig} is affected when the microcell is still in its recovery state. P_{trig} can be formulated as an exponential function of time t using the recovery time constant τ_{rec},

$$P_{\text{trig}}(t) = P_{\text{trig},0}\left[1 - \exp\left(-\frac{t}{\tau_{\text{rec}}}\right)\right], \tag{4.19}$$

where $P_{\text{trig},0} = P_{\text{trig}}(t = 0)$ is the triggering probability in a fully recharged microcell and τ_{rec} is the recovery time constant of the microcell.

Fill factor

Not all photons striking the SiPM will enter the active (photosensitive) region of the p–n junction. The photon can be absorbed by the non-depleted region (such as the substrate), by the dead area between microcells (such as the metal grid), or by the shallow layer above the p–n junction. The photon can be reflected by the upper surface at entrance window. Penetrating throughout the device is possible, as well. Of course, photons of such high energy are not what the detector is targeting.

Despite the possibilities of all these scenarios, various techniques are used to minimize the loss of incoming photons. These are not too tricky thanks to the mature CMOS process in the semiconductor industry. For example, optically pure materials and appropriate nonreflective structures such as an intermediate Si_3N_4 layer can be used to prevent losses at the entrance window[120].

As a result, the *fill factor* (FF) is generally used to assess the geometrical efficiency of SiPMs. FF is defined as the ratio of the active (photosensitive) area A_{active} to the total area of the sensor A_{total}, i.e.,

$$\text{FF} = \frac{A_{\text{active}}}{A_{\text{total}}}. \tag{4.20}$$

The active area A_{active} is illustrated as the diagonal stripe fill in Fig. 4.7a.

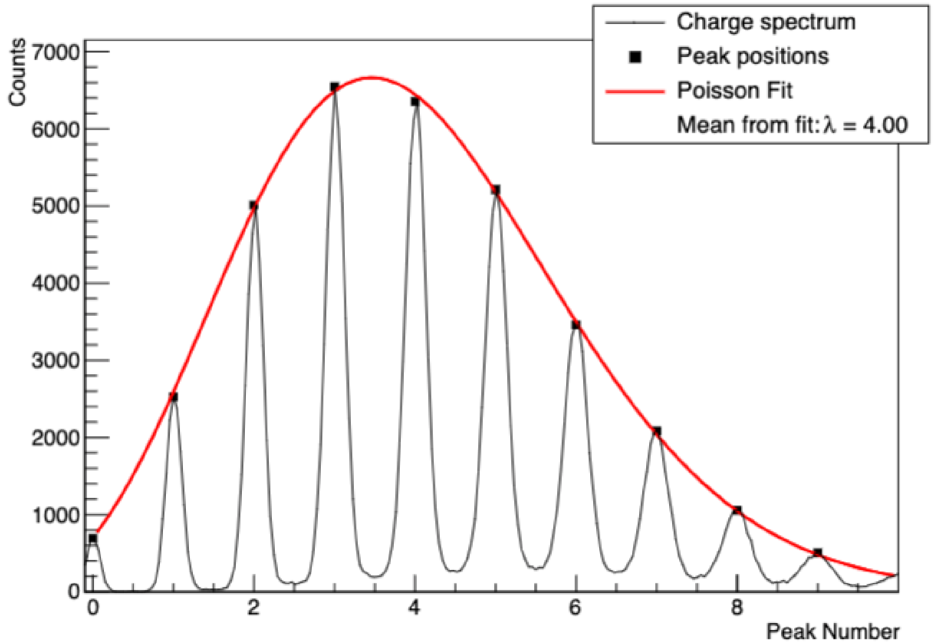


Figure 4.11: Example of single photoelectron spectrum. Credit: [135].

4.3.4 Single-photoelectron spectrum

In photon counting, the correspondence between the measured pulse height/charge and the detected number of photons can be displayed in single-photoelecton (SPE) spectra, as shown in Fig. 4.11.

The measured output of SiPM is a statistical distribution of pulse height or charge. By extracting the pulse heights from the SiPM's output waveform, one can plot the pulse height spectrum or amplitude spectrum. Alternatively, by performing time integration between the waveform and the baseline in a specified integrating gate (window), the charge can be calculated and be plotted as the charge spectrum. If a digitizer is used in the readout circuit, the number of photoelectrons (p.e.) can also be expressed in terms of ADC channel numbers through a conversion factor.

The amplitude spectrum is more susceptible to electrical noise and peak amplitude fluctuations in the readout. A lower readout bandwidth is beneficial to reducing the peak-to-valley ratio of the amplitude spectrum. As for the charge spectrum, it is more susceptible to the integration gate width and the correlated noise of SiPM[113]. Hence, optimizing the gate width is an indispensable step preceding plotting the charge spectrum.

At low light level, the number of photons is very small. A well-resolved SPE spectrum should allow each peak to correspond to an number of photoelectrons, highlighting the discrete nature of photons. These peaks are sometimes called photoelectron peaks or photopeaks[131].

In particular, the photoelectron peak corresponding to 0 p.e. is called *pedestal*. Its interpretation is that the integrared charge in the given window is equal to zero, which is equivalent to the situation where no photons are detected. The base can only be pedestal if the SiPM is read out with an external trigger system (e.g. synchronized with a pulsed light source). Fig. 4.11 is one example of this situation.

The peak-to-peak distance is proportional to the charge released by SiPM in a one-photon event, which is equivalent to Q in Eq. (4.12). Therefore, measuring the distance between photon peaks is also a method of obtaining gain[132].

The heights of the photoelectron peaks reflect the characteristics of Poissonian photon statistics[93, 136]. The probability density function of the Poisson distribution is

$$\text{Poisson}(x;\mu) = \frac{\mu^x}{x!}e^{-\mu}, \tag{4.21}$$

where x is a discrete random variable, μ is the mean of the distribution. The average number of photons detected by the SiPM can be extracted by fitting the Poisson distribution to the apices of all normalized photoelectric peaks. Combined the features of discrete photoelectron peaks and the Poisson distribution, the SPE spectrum resembles human fingers, and is thus nicknamed the "finger plot[137]."

The widths of the photoelectron peaks contain multiple components of noise. The pedestal width reflects the electrical noise. The widths of photoelectron peaks except pedestal reflects the effect of gain fluctuation. Since each photoelectron peak energy is approximately fitted using a Gaussian distribution, the standard deviation σ of the distribution provides a quantitative description of the photoelectron peak width. Assume σ_{ped} is the standard deviation associated with the pedestal width, and σ_G is the standard deviation associated with the gain fluctuation (see Eq. (4.13)), then the standard deviation σ_i associated with the i-th photoelectron peak can be described by[93, 138]

$$\sigma_i^2 = \sigma_{\text{ped}}^2 + i \cdot \sigma_G^2. \tag{4.22}$$

It can be seen that the photoelectron peak broadens with i. However, since σ_G is very small for SiPMs, this broadening is not necessarily obvious.

The width of the photoelectron peak also determines the photon number resolution (mentioned in Section 3.1.4), which can be expressed[7] by σ_i/i, the relative width of the photoelectron peak. The better the photoelectron peaks are separated, the lower σ_i/i is, and the better the quality of SPE spectrum. For IACT photodetectors, the charge resolution derived from the SPE spectrum is one of the key parameters for camera performance verification. The CTA also sets specific requirements for charge resolution within a specific p.e. range.

In addition, SPE spectrum are important in the camera calibration strategy because they can be used to extract single photodetector gain and overall optical detection efficiency for the flat-fielding of the camera[93].

In this study, SPE spectra are measured in dark conditions and at low light level. These will be presented in Sections 6.2 and 6.3, respectively.

4.4 Noise of SiPMs

For SiPMs, noise refers to avalanche events that are not caused by external photons. These types of noise limit the performance of timing and counting resolutions of SiPMs. The following distinctions can be made based on the causes of noise. The *primary noise* refers to the noise events that are causally independent of each other; and the *correlated noise* refers to the noise events that are causally dependent on other pulses.

The primary avalanche breakdowns can be either photon-generated or thermally-generated; while the former is considered signal, the latter is considered dark noise.

[7] Alternatively, if the photon number resolution is extracted from a pulse height spectrum, it is called the pulse height resolution $\sigma_{\text{PH}}/\text{PH}$[61]; if it is extracted from an amplitude spectrum, it is called the amplitude resolution σ_A/A; if extracted from a charge spectrum, it is called the charge resolution σ_Q/Q.

4.4.1 Dark noise

In certain types of photodetectors, such as SPAD and SiPM, avalanche breakdown can be triggered by non-photo-generated carriers. This phenomenon, known as *dark noise*, can generate pulse events even under dark conditions. These pulse events are referred to as dark counts, dark pulses, or dark events. The frequency at which dark counts occur is known as the *dark count rate* (DCR), which is measured in hertz (dark counts per second). When the photodetector produces a large number of dark counts on a macroscopic scale, the resulting current is called a *dark current*. Typically, the dark pulse rate is expressed in $\mathrm{MHz\,mm^{-2}}$ as it is proportional to the device area for a given SiPM design[121].

Dark pulses may be caused by thermal excitation or field-induced ionization (recall Section 4.1.1). The major contribution is the thermal excitation assisted by deep traps, which is the SRH generation/recombination. In addition, the Frenkel–Poole will lower the activation energy and enhance the generation rate of carriers[108, 113, 134].

An empirical formula can describe the dependence of DCR on the bias voltage V_{bias}[108]:

$$\mathrm{DCR} = N_{\mathrm{gen}} \cdot P_{\mathrm{trig}}^{\mathrm{DCR}} \cdot \exp(\beta \cdot V_{\mathrm{bias}}), \tag{4.23}$$

where N_{gen} is the generation rate of thermal excitation, $P_{\mathrm{trig}}^{\mathrm{DCR}}$ is the avalanche triggering probability of the thermally generated carriers, β is the fitting parameter that describes the increase of DCR with V_{bias} due to electrical-field-related effects. Many parameters in Eq. (4.23) are temperature-dependent. The temperature dependence of DCR is found to be[9, 132]:

$$\mathrm{DCR}(T) \propto T^{3/2} \exp\left(\frac{E_{\mathrm{g}}}{2k_{\mathrm{B}}T}\right), \tag{4.24}$$

where E_{g} is the band gap energy, k_{B} is the Boltzmann constant, and T is the absolute temperature.

Because of the nature of the Geiger mode, the shape of any dark pulse is precisely the same as that of a photo-induced pulse. As a coarse estimation, the number of pulses > 0.5 p.e. in the SPE spectrum divided by the time interval can be taken as the DCR. However, this estimation method cannot isolate the impact of correlated noise on the SPE spectrum unless statistical methods are used.

4.4.2 Afterpulsing and optical crosstalk

SiPM's correlated noise can be classified into two types based on its location. When correlated noise events happen in the initially triggered microcell, it is referred to as *afterpulsing*. When these events occur in the adjacent microcells, it is called *optical crosstalk*.

There are further classifications depending on the propagation paths of carriers and *secondary photons*. The so-called secondary photons are caused by hot-carrier luminescence of the primary avalanche breakdown in the initially triggered microcell. Each avalanche carrier is estimated to produce $\sim 10^{-5}$ secondary photons[113, 139]. These photon emissions can be photographed with sensitive, low-noise CCD cameras[129].

At least Five subtypes have been identified, four of which are related to the emission of secondary photons. They are listed below and are illustrated in Fig. 4.12.

(a) Afterpulsing due to diffusion. Some secondary photons transmit to the substrate of the very same microcell, and then some generated carriers diffuse back to its own active region and trigger an avalanche breakdown with a time delay.

(b) Afterpulsing due to trap release. Some primary avalanche carriers are trapped by deep or shallow energy levels in the band gap; some trapped carriers may be released after some ns or µs and may trigger an avalanche breakdown in the same microcell again.

(c) Prompt crosstalk. Some secondary photons directly transmit to the active regions of the adjacent microcells and trigger an extra avalanche breakdown there. Since the photons travel at the speed of light along this path, the time lapse between the primary and the secondary breakdowns is only a few picoseconds and the secondary event can be considered "prompt." So this phenomenon is also known as *direct crosstalk*.

(d) Delayed crosstalk. Some secondary photons are absorbed by the substrate of the adjacent microcells and then some generated carriers diffuse to its active region and

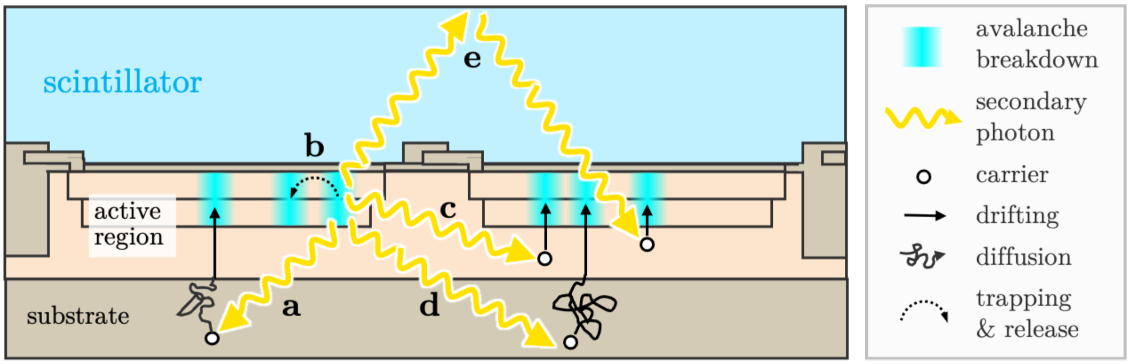


Figure 4.12: Different subtypes of SiPM's correlated noise. (a) afterpulsing due to diffusion, (b) afterpulsing due to trap release, (c) prompt crosstalk, (d) delayed crosstalk, and (e) external crosstalk.

trigger an extra avalanche breakdown there. Since this process involves the diffusion of carriers, it is also called *crosstalk due to diffusion*.

(e) External crosstalk. In some applications, the SiPM is optically coupled to a scintillator or applied to a protective layer. Some secondary photons can be reflected back to adjacent microcells through total internal reflection.[113, 115, 140]. This process is sometimes not explicitly distinguished but is incorporated into prompt crosstalk.

Because the delayed crosstalk and the afterpulsing involve delayed secondary avalanche breakdowns, they are collectively called *delayed correlated noise* and often compared[141].

Crosstalk probability

Any component of relevant noise can be quantified as a probability. The definition depends largely on the measurement method and is not necessarily universal. For example, the *(optical) crosstalk probability* P_{XT} measures the probability that the avalanche event in any microcell triggers another avalanche event in the neighborhood via direct transmission of secondary photons. It is usually is defined as[131, 132]

$$P_{\mathrm{XT}} = \frac{N_{>1.5\,\mathrm{p.e.}}}{N_{>0.5\,\mathrm{p.e.}}}, \tag{4.25}$$

where $N_{>k\,\mathrm{p.e.}}$ is the dark count with pulse height larger than k p.e. These counts can be obtained from the photopeaks on the SPE spectrum. $N_{>0.5\,\mathrm{p.e.}}$ is the dark count on the SPE spectrum except for the pedestal, and $N_{>1.5\,\mathrm{p.e.}}$ is the dark count except for the pedestal and the 1 p.e. photoelectron peak. Counts with amplitudes between 0.5 p.e. and 1.5 p.e. are considered to belong to the 1 p.e. photoelectron peak.

Strictly speaking, the crosstalk probability is a discrete probability distribution $P_{\mathrm{XT}}(k)$, where $(k = 0, 1, 2, 3, \dots)$ is the number of triggered microcells. Analytical probability models have been developed to describe $P_{\mathrm{XT}}(k)$. Some experimental results support the Borel distribution[130], and some support the "four nearest neighbors" model[142, 143].

Afterpulsing probability

Unlike crosstalk, which can be intuitively defined by setting a threshold for the photoelectron peaks on the SPEs spectrum, afterpulses are not defined using an threshold on any inter-time[8] distribution of pulses. After all, the temporal proximity of pulses does not imply causal correlation. In other words, any adjacent pulses on the waveform can either be causally independent or correlated.

The time-dependent *afterpulsing probability* $P_{\mathrm{AP}}(t)$ can be formulated based on two simple probability models—one describing the release of carriers from the trap and the other describing the re-triggering of the avalanche breakdown[141].

The first probability model, i.e. the probability P_{rel} that a carrier is released after being trapped for time t can be described by[144]

$$P_{\mathrm{rel}}(t) = \frac{1}{\tau_{\mathrm{rel}}} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{rel}}}\right), \tag{4.26}$$

where τ_{rel} is the characteristic time constant of releasing and the factor $1/\tau_{\mathrm{rel}}$ is the normalizing constant.

The second probability model, i.e. the probability that a released carrier re-trigger an avalanche breakdown, is essentially an avalanche triggering probability, which has been given

[8]The inter-time is defined as the interval between timestamps of adjacent pulses

in Eq. (4.19).

Combining the temporal behavior of Eqs. (4.19) and (4.26), we can model afterpulsing probability $P_{\mathrm{AP}}(t)$ as[141]

$$P_{\mathrm{AP}}(t) = \frac{\tau_{\mathrm{rec}} + \tau_{\mathrm{rel}}}{\tau_{\mathrm{rel}}^2} \cdot \left[1 - \exp\left(-\frac{t}{\tau_{\mathrm{rec}}}\right)\right] \cdot \exp\left(-\frac{t}{\tau_{\mathrm{rel}}}\right). \tag{4.27}$$

with $(\tau_{\mathrm{rec}} + \tau_{\mathrm{rel}})/\tau_{\mathrm{rel}}^2$ being the normalizing constant.

Amplitude–time distribution and the pile-up effects

The pile-up effects of dark pulses, which can be removed by proper algorithms, provide a simple method to distinguish different types of correlated noise. By correcting and gathering the pulse heights (amplitude) $\tilde{A}_i$, timestamp t_i and other information of each pulse identified on a waveform, one can create a data set with tuples of form $(A, \Delta t_i)$, where $\Delta t_i = t_{i+1} - t_i$ is the inter-time between the pulse i and the pulse that follows it. The following statements can then be deduced.

- The optical crosstalk can produce pulse heights of 2 p.e., 3 p.e., etc., because it triggers multiple microcells almost at the same time.
- The afterpulsing can cause a delayed secondary pulse with an correlated amplitude $\tilde{A} < 1$ p.e. because the microcell has not been restored from the primary avalanche breakdown yet.
- The delayed crosstalk can also cause a delayed secondary pulse, but its correlated amplitude is equal to 1 p.e. because the adjacent microcell is fully charged before the secondary avalanche breakdown is triggered.

These features are summarized in Fig. 4.13 and Table 4.2.

By plotting the corrected amplitude data $\tilde{A}_i$ against the corresponding inter-time data Δt_i, one obtain a scatter plot like Fig. 4.14, from which one can identify the clouds of data points representing different types of correlated noise[113, 145]. However, this method may not always work.

Table 4.2: Pulse features of different types of correlated noise under the influence of pile-up effects.

Types	Corrected amplitude $\tilde{A}$	Time delay
Prompt/external crosstalk	k p.e. ($k = 2, 3, \ldots$)	$\simeq 0$
Afterpulsing	< 1 p.e.	> 0
Delayed crosstalk	$= 1$ p.e.	> 0

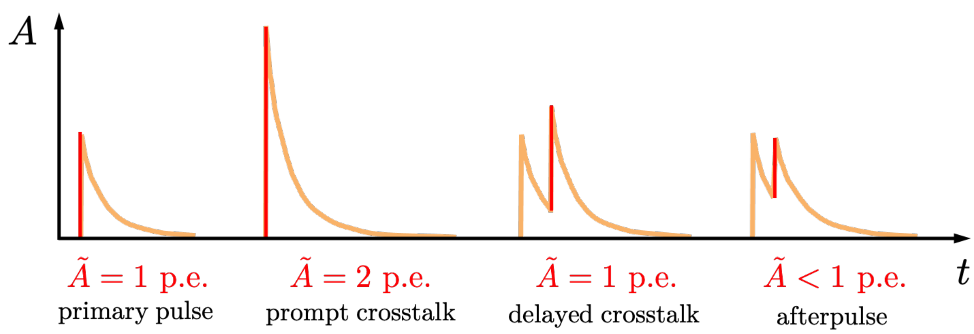


Figure 4.13: Pile-up effects of different types of correlated noise. The red lines next to the rising edges indicate the corrected amplitude $\tilde{A}$. Adapted from [141].

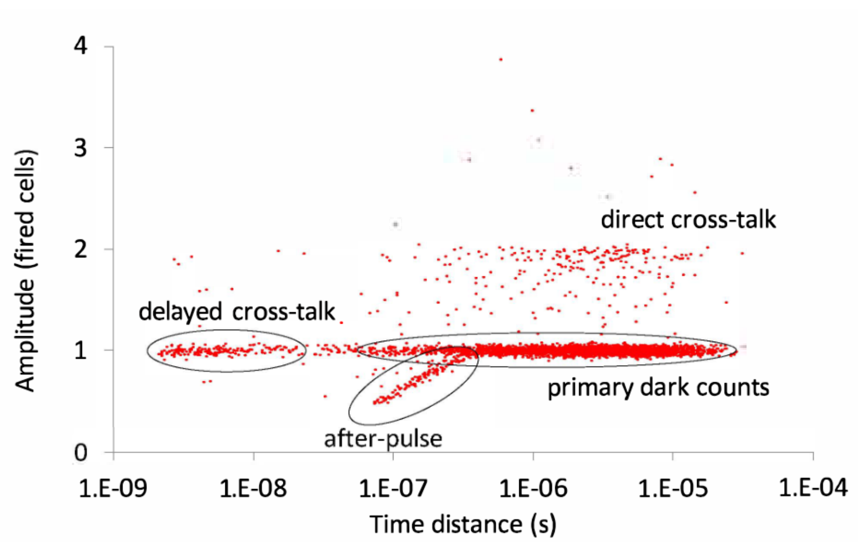


Figure 4.14: The scatter plot of amplitude A vs. inter-time Δt. Credit: [145].

Chapter 5

Static characterization

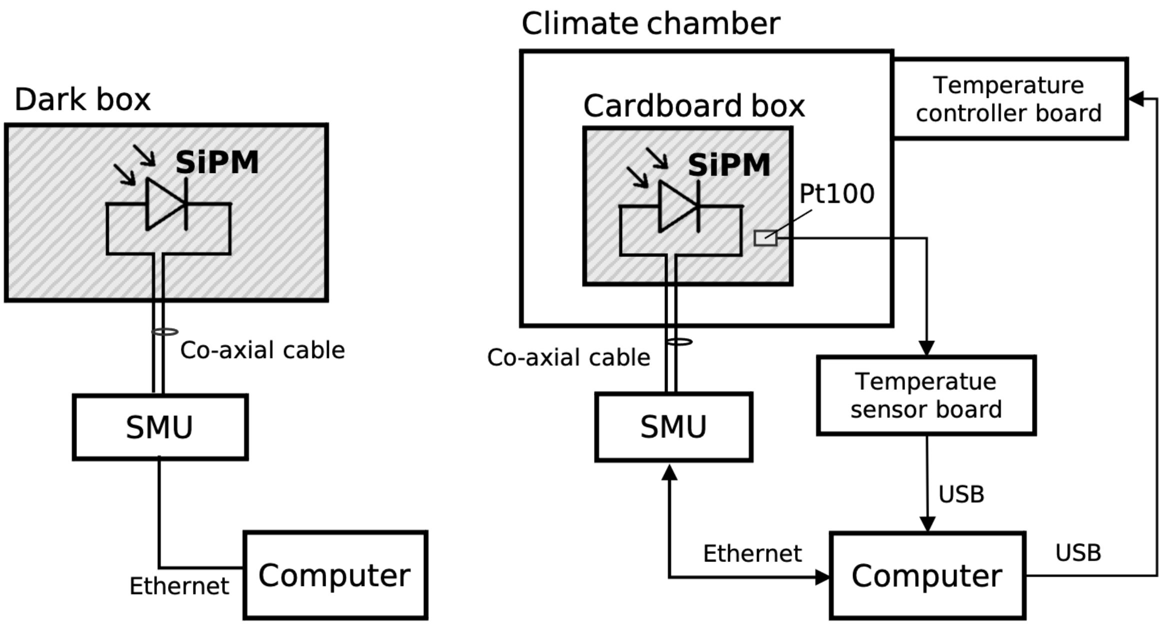
The properties of a SiPM strongly depend on intrinsic factors such as its size and structure, as well as external factors such as operating voltage and temperature. The process of determining the characteristics of SiPM is called *characterization*.

For a specific model of SiPM, understanding its current-voltage characteristics (i.e. I–V curve or I–V characteristics) has priority in the electrical characterization. This determines SiPM's breakdown voltage at a given temperature and thus determines the overvoltage that actually controls the effect of avalanche breakdown when the device is in its Geiger mode. The I–V characteristic involves measuring some time-independent behavior of device in some steady state (e.g. constant-voltage supply); it is one of the *static characteristics*.

Two important parameters of SiPM can be extracted from the static characteristics—the breakdown voltage V_{BD} and the quenching resistance R_{Q}. The former is crucial in determining the overvoltage, on which many other parameters depend; and although the latter has been defined during fabrication of the device, calculating its value is still helpful for understanding the pulse shape and for establishing an electrical circuit model.

5.1 Experimental setup

In principle, only a voltmeter and an ammeter are needed for static characterization of SiPMs. But in practice, the current measuring range before and after avalanche breakdown may span



(a) Setup for measurments in dark conditions.

(b) Setup for measurements in dark conditions and controlled temperature.

Figure 5.1: Experimental setups for the static characterization.

4 or 5 orders of magnitude. A highly accurate Source Measurement Unit (SMU) can simultaneously supply a constant voltage and perform current measurements from 10^{-9} ampere to 10^{-4} ampere, adjusting the range automatically to cover the 4 or 5 orders of magnitude. Therefore, a SMU is especially useful for static characterization. For example, Keithley's Model 2657A High Power System SourceMeter[1] is used in our experiments.

Fig. 5.1 shows the experimental setups for the static characterization. During measurements at room temperature, the SiPM is placed in a dark box. In the temperature control measurement, SiPM is placed in the climate chamber. Although the observation window of the climate chamber has been blocked, the SiPM is placed in a cardboard box to minimize the ambient light so as to ensure the SiPM is measured in a dark environment.

The climate chamber is controlled using a dedicated circuit board developed by the colleagues in IAAT that can communicate with a computer through a USB cable. Since the climate chamber display may not accurately reflect the temperature near the SiPM, a Pt100 resistive temperature sensor was placed near the SiPM, and its resistance was converted into temperature through another dedicated circuit board and read by the computer.

[1] https://www.tek.com/en/datasheet/2657a-high-power-system-sourcemeter-smu-instrument

In fact, for the most stringent standards regarding dark conditions and ambient temperature, a better solution would be to install a calibrating photodiode and a thermometer (in direct thermal contact with the device) on the test board.

The communication between the SMU and the computer is via Ethernet, and the instrument control is via the Test Script Processing® (TSP) scripting language, which is a hardware/software architecture developed by Keithley and based on Lua version 5.0[146]. The TSP script program can be written on a computer and then used to send and receive between the computer and the instrument through the open-source Virtual Instrument Software Architecture (VISA) library of Python.

In the measurements of both forward and reverse bias, linear staircase sweeps of voltage are used. For example, the bias voltage gradually increases from 51 V to 53.5 V in steps of 0.02 V, and the measurement time delay of each step is 0.3 s.

5.2 Forward I–V characteristic

The forward I–V curve of an ideal diode is described by the Shockley diode equation in the form of an exponential function:[112, 147]

$$I = I_{\mathrm{S}} \left[\exp\left(\frac{V}{\eta V_{\mathrm{T}}} \right) - 1 \right], \tag{5.1}$$

where I_{S} is the reverse-bias saturation current, V_{T} is the thermal voltage, and η is sometimes called the ideality factor.

However, when the diode is connected with a resistor in series, its I–V curve can be dominated by the behavior of the resistor and thus can be approximated by a piecewise linear model composed of a horizontal line at the zero current and a slanted line with non-zero x-intercept, with the slope being $1/R$. This is the case for the SiPMs with passive quenching circuits to quench the avalanche breakdown. Each microcell in the SiPM is connected in series with a polysilicon quenching resistor (recall Fig. 4.8). Given the quenching resistance R_{Q} and the number N_{cell} of microcells connected in parallel, the slope of the slanted linear function is $N_{\mathrm{cell}}/R_{\mathrm{Q}}$. Hence, the quenching resistance can be extracted by performing linear fitting on

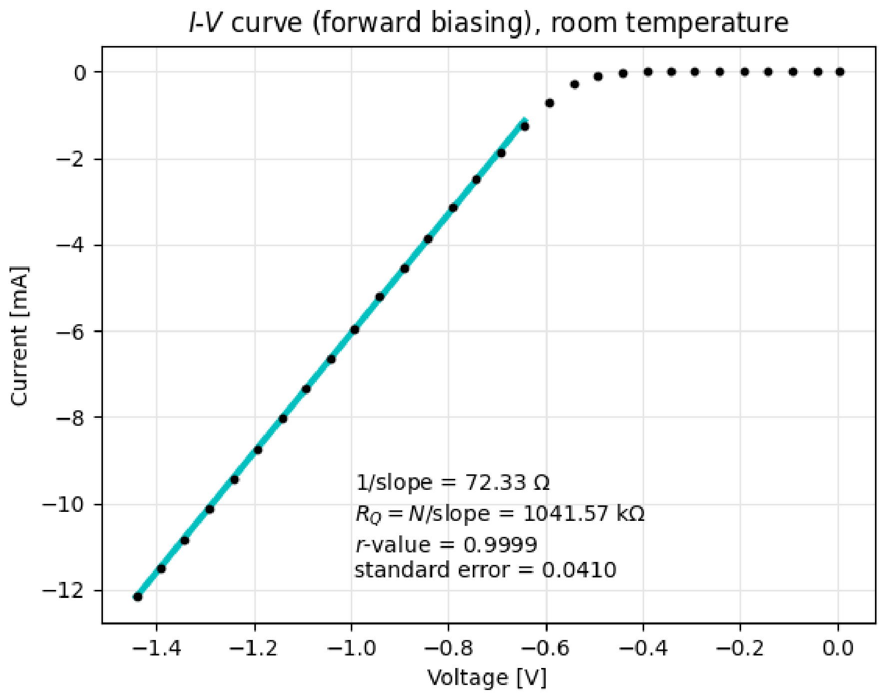


Figure 5.2: Forward I–V characteristic of SiPM. Since the polarity is reversed when measuring, the forward biasing regime is $V < 0$.

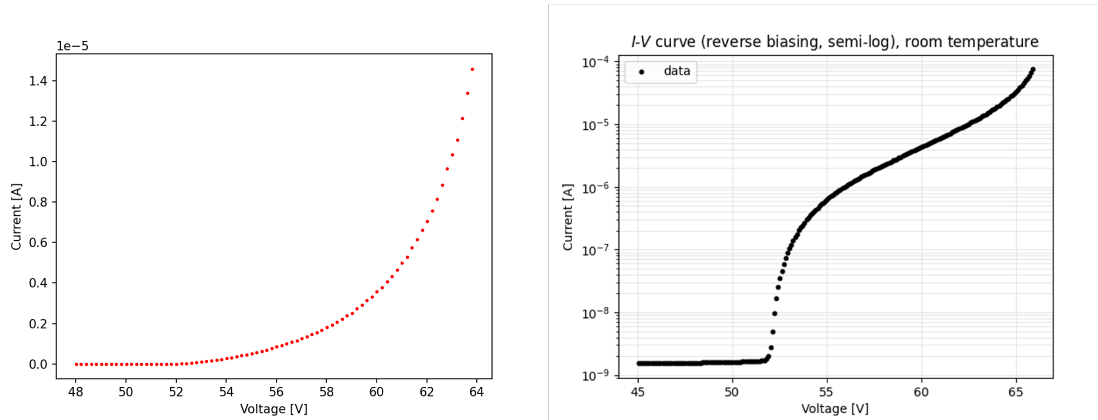
the interval with a sloped straight line in the measured forward bias I–V data points:

$$R_{\mathrm{Q}} = \frac{N_{\mathrm{cell}}}{\mathrm{d}I/\mathrm{d}V}. \tag{5.2}$$

An example of our measurement results is sketched in Fig. 5.2.

5.3 Reverse I–V characteristic

Figs. 5.3a and 5.3b shows the examples of the reverse I–V characteristics, with one plotted in the linear scale, the other plotted in the semi-log scale. Both plots show the rapid increase in current with voltage. But one can hardly identify where the breakdown voltage begins in the linear plot (Fig. 5.3a) . Therefore, the extraction of breakdown voltage usually relies on features of the I–V curve in its semi-log scale (Fig. 5.3b), where a "turning point" clear



(a) I vs. V (b) $\ln(I)$ vs. V

Figure 5.3: Examples of SiPM's reverse I–V curves.

exists.

5.3.1 Proposed methods

Many methods have been proposed to extract the breakdown voltage from the $\ln(I)$ vs. V curve. They can be roughly divided into three categories—the *fitting methods*, the *derivative methods*, and the *model-based methods*.

Fitting methods

The fitting methods split the curve into two segments: the *pre-breakdown region*, fitted using a linear regression, and the *post-breakdown region* with significant curvature, fitted using a polynomial regression. The breakdown voltage is then determined as the intersection point of these two fitted curves.

If the polynomial degree is 1, the method is known as the *tangent method*; if the degree is 2, it's also referred to as the *parabolic method*.[2] For examples, [45] took the data points very near the breakdown point and used the tangent method. [148] considered a larger range of data points and used the parabolic fitting.

[2] The nomenclature of the methods mentioned here are mostly based on the one used in [126] or [108].

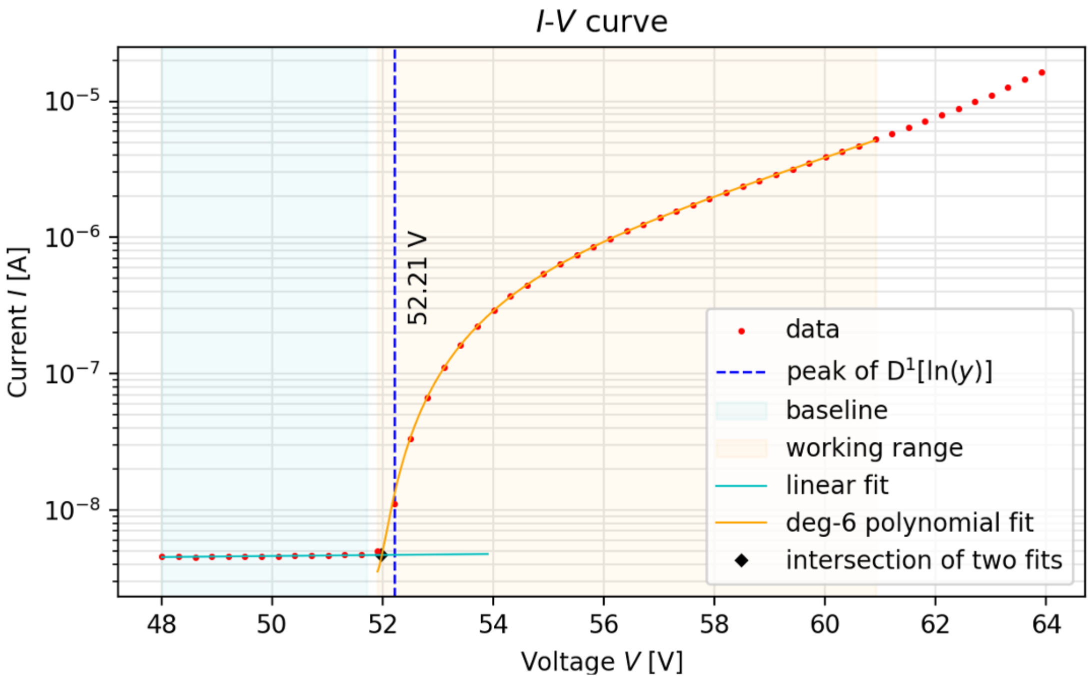


Figure 5.4: An early attempt with a sixth-degree polynomial fit to the post-breakdown region of the an reverse I–V curve (labeled *working range*).

Derivative methods

First deivative method The second category of methods differentiate the function $\ln(V)$ to some order with respect to V and then find the extremes in these derivatives. For exmaple, the *first derivative method*, or the *relative derivative method*, calculates

$$f_{1\mathrm{d}}(V) = \frac{\mathrm{d}}{\mathrm{d}V}(\ln I) = \frac{1}{I}\frac{\mathrm{d}I}{\mathrm{d}V}, \tag{5.3}$$

and then the local maximum of $f_{1\mathrm{d}}(V)$ is treated as the breakdown voltage. This is method is used in [149], and the Hamamatsu's technical note[132], as well. [108] suggest using a peaked and skewed function to fit the region around V_{BD}, e.g., a Landau function.

Second deivative method The *second deivative method* calculates[150]

$$f_{2\mathrm{d}}(V) = \frac{\mathrm{d}^2}{\mathrm{d}V^2}(\ln I), \tag{5.4}$$

and then also finds the local maximum of $f_{2\mathrm{d}}(V)$.

Inverse deivative method The *inverse deivative method*[147, 151] calculates

$$f_{\mathrm{inv}}(V) = \left(\frac{1}{I}\frac{\mathrm{d}I}{\mathrm{d}V}\right)^{-1} = \left(\frac{\mathrm{d}}{\mathrm{d}V}\ln I\right)^{-1}, \tag{5.5}$$

and find the local minimum of $f_{\mathrm{inv}}(V)$ "in some sense". Because of the skewness of $f_{1\mathrm{d}}(V)$ and the finiteness of the maximum of $f_{1\mathrm{d}}(V)$, the true minimum of $f_{\mathrm{inv}}(V)$ may not be the best place to be defined as the breakdown point. [147] used more than one method to define a more suitable minimum value, including using parabola fitting to find the lowest point, linear fitting of the areas on both sides of the lowest point and then taking their intercepts with the x-axis, etc. [147] found consistency in the results of these three methods (errors $\simeq 0.4\,\mathrm{V}$), and this is also true for our results (Fig. 5.5).

Model-based methods

Although based on finding derivative of current, third derivative method suggested by [126] is actually slightly different from the methods mentioned above, where the breakdown voltage is defined as one of the extremes of the derivatives. It is pointed out in [126] that avalanche breakdown in a microcell exhibits a hysteresis of breakdown voltage with two distinct voltages: the *turn-on voltage* V_{01} is where the breakdown is initiated, while the *turn-off voltage* V_{10} is where the avalanche breakdown is quenched. The difference between these two voltages is called the voltage hysteresis h:

$$h = V_{01} - V_{10}. \tag{5.6}$$

Upon deriving their fitting fucntion, [126] introduced several assumptions:

1. the charge released by one microcell during the avalanche breakdown is proportional to $(V - V_{10})$;
2. the avalancge triggering probability P_{trig} is proportional to $(V - V_{01})$;
3. the hysteresis of all microcells are the same;
4. the distribution of V_{01} among all microcells is a Gaussian distribution.

The first two assumptions are based on theory and the last two were a posteriori justified by experiments. Finally, they were able to find the fitting function for the 3rd derivative curves,

$$y = A \cdot \left[2 - \frac{h}{\sigma^2}(V - V_{01})\right] \cdot \exp\left[-\frac{(V - V_{01})^2}{2\sigma^2}\right], \tag{5.7}$$

where A and σ are the amplitude and the standard deviation of the Gaussian distribution.

The other model-based I–V method takes into account the different phyical processes along the entire I–V curve and develops the so-called *inverse I–V curve model*, in which the post-breakdown region is divided into four sub-regions, each of which has its own mathematics based on different processes. Unlike the previous methods that limited to process data in particular region, the inverse I–V curve model method aims to be applicable to a board range, even reaching the so-called *second breakdown*. Our device also exhibited second breakdown at $\sim 65\,\mathrm{V}$ at room temperature (see Fig. 5.3b).

This model is full described in [152] and was realized in [108].

5.3.2 Attempts and discussion

An attempt at various derivative methods

To compare these derivative methods, the analyses based on acquired data were carried out with them, and a set of examples is shown in Fig. 5.5. The dashed lines represent the breakdown voltages derived from those methods. The results appear to be quite consistent.

Unfortunately, it turns out that the algorithm was only stable when treating the relative derivative method and the inverse derivative method. It can sometimes produces an unstable result shown in Fig. 5.6.

Here are some reasons that may have undermined the stability of the algorithm:

1. The settings of voltage sweep of the SMU might play a role. The effect of voltage step size and measurement delay on experimental errors has not been studied in depth. However, it has been observed that the I–V curve swings up and down in adjacent points for certain voltage steps (like the case in Fig. 5.6). This problem was more serious for the pre-breakdown region. Increasing the number of repeated measurements did not improve the situation. This problem together with the immature algorithm is suspected to contribute largely to the erroneous results of differentiation.

2. Some authors applied smoothing on the original data, such as the local regression [150], the cublic spline interpolation[147], and Savitzky–Golay filter[126].

3. For the first, second, and third derivative methods, no fitting was performed after differentiation, thereby deviating from the correct usage of the methods.

Despite the potential problems, the algorithm still produces relatively efficient results through the inverse derivative method, in which the (relative) derivative method is adopted. There-

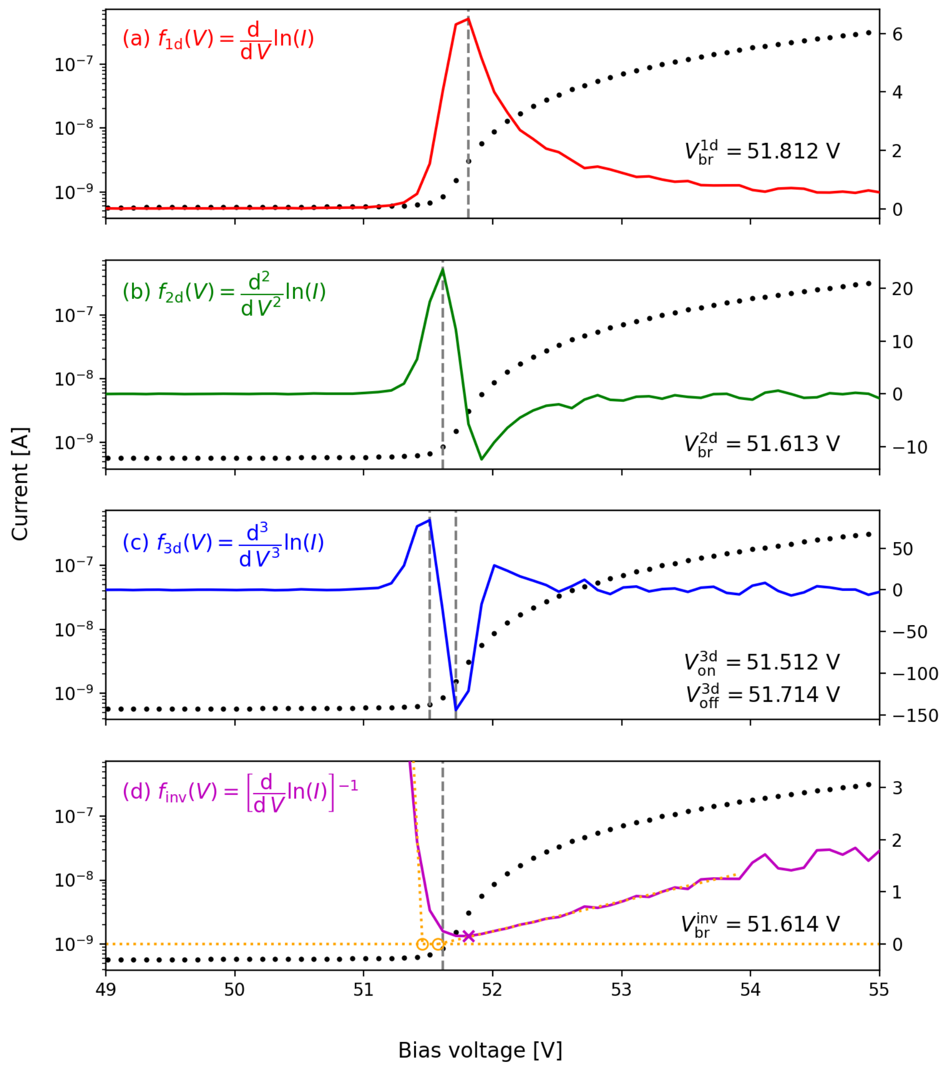


Figure 5.5: Analyses of (a) the *relative derivative method*, (b) the *second derivative method*, (c) the *third derivative method* and (d) *inverse derivative method* of reverse *I–V* characteristics, based on the measurements. The solid data points represent the measured current readings, with values corresponding to the y-scale on the left. The solid curves are the functions used for analysis, with values corresponding to the y-scale on the right. The positions of the breakdown voltages V_{BD} obtained by four methods are marked by vertical dashed lines, and their value is shown in the lower right corner of each subfigure. In Subfigure (d), the dotted lines include two linear fits and the zero line; the two circle markers are the intersection points of the two linear fits and the zero line, and the cross marker is the local minimum of the function $f_{inv}(V)$. The temperature was not measured.

fore, in subsequent analyses, including IV curve analysis measured at different temperatures (Section 5.4), the inverse derivative method is relied on to derive the breakdown voltage.

All in all, although the datasets that failed to be processed can be temporarily filtered out, a stable and efficient breakdown voltage analysis method should still be developed for future large-scale characterization for IACT cameras. The inverse I–V curve model method, which takes into account both physical principles and curve characteristics, may be a worthwhile avenue to explore.

5.4 Temperature-dependent static characterization

In the preliminary stage of this research, when the polynomial fitting method was still used, the I–V curve measurement experiments 0 °C to 70 °C was done with the climate chamber. The results are shown in Figs. 5.7 and 5.8. It can be roughly seen from the I–V curves in Fig. 5.8 that the breakdown voltage V_{BD} has a positive temperature coefficient, which is consistent with statements in Section 4.3.1. Fig. 5.7 also reveals that the quenching resistance R_{Q} has a negative temperature coefficient, which is consistent with the ordinary characteristics of semiconductors.

Later the climate chamber experiment was re-conducted. This time, more data sets were are collected only between 0 °C to 35 °C. And the polynomial fitting method was replaced by the derivative methods. The results are shown in Fig. 5.9. Data sets that failed to analyze were temporarily removed. This graph more clearly shows the linear relationship between breakdown voltage V_{BD} and temperature T, as described by Eq. (4.8). According to linear regression with regression coefficients very close to 1, the temperature coefficient b obtained by the first derivative method and the inverse derivative method can be estimated as 53.1 mV °C^{-1} and 51.5 mV °C^{-1}, respectively, which are very close to the reference value 54 mV °C^{-1} given in [124].

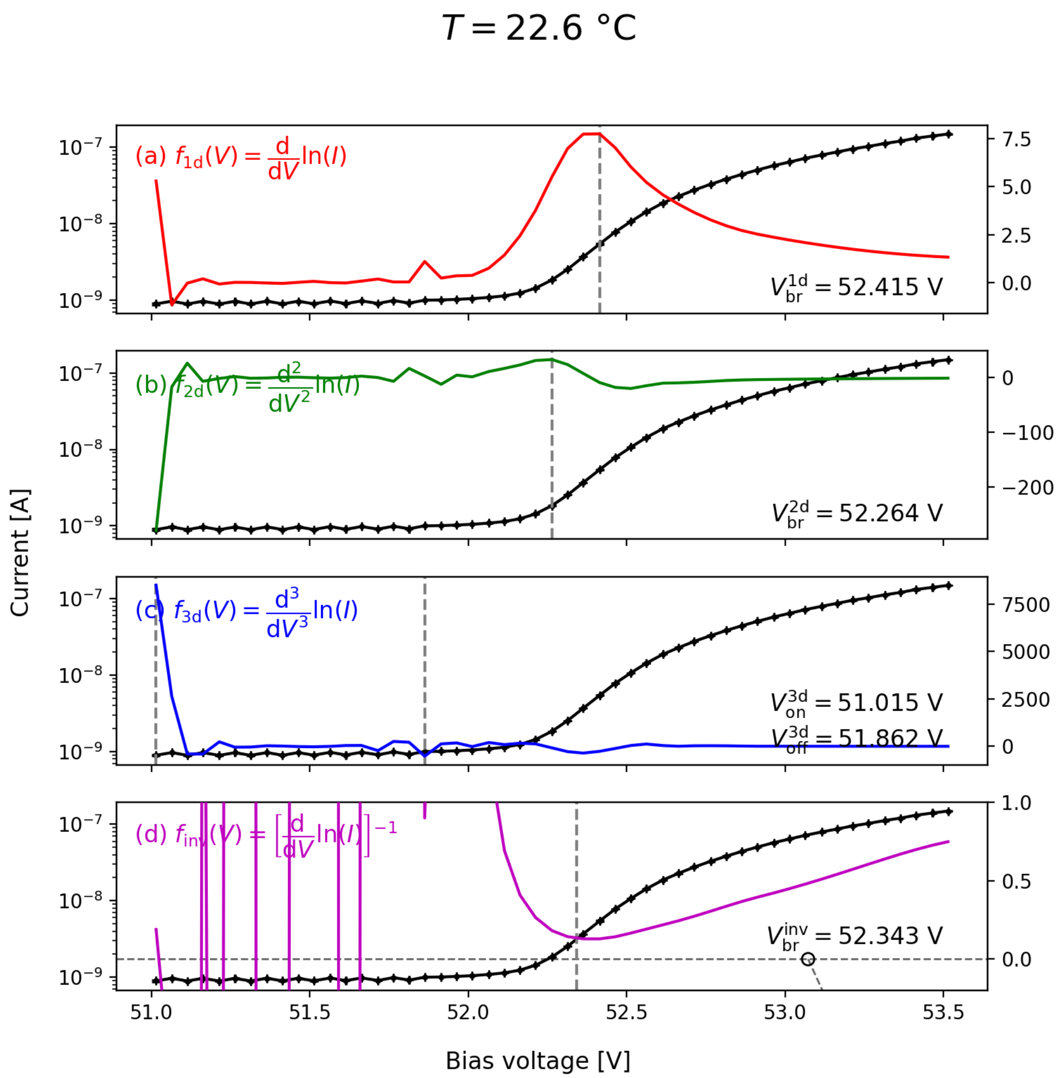


Figure 5.6: A flawed example using the methods as Fig. 5.5. The descriptions about the figure elements are alike.

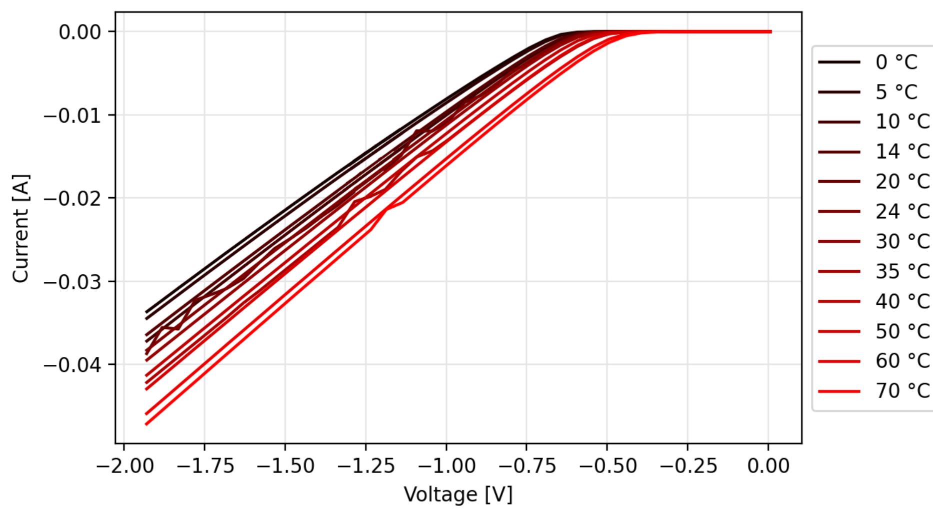


Figure 5.7: Forward I–V curves measured from 0 °C to 70 °C.

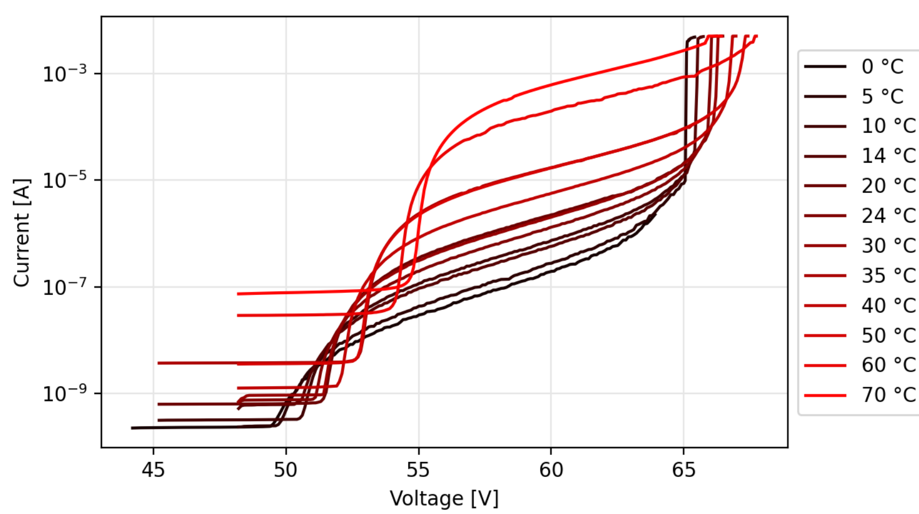


Figure 5.8: Reversed I–V curves measured from 0 °C to 70 °C using the high-degree polynomial method.

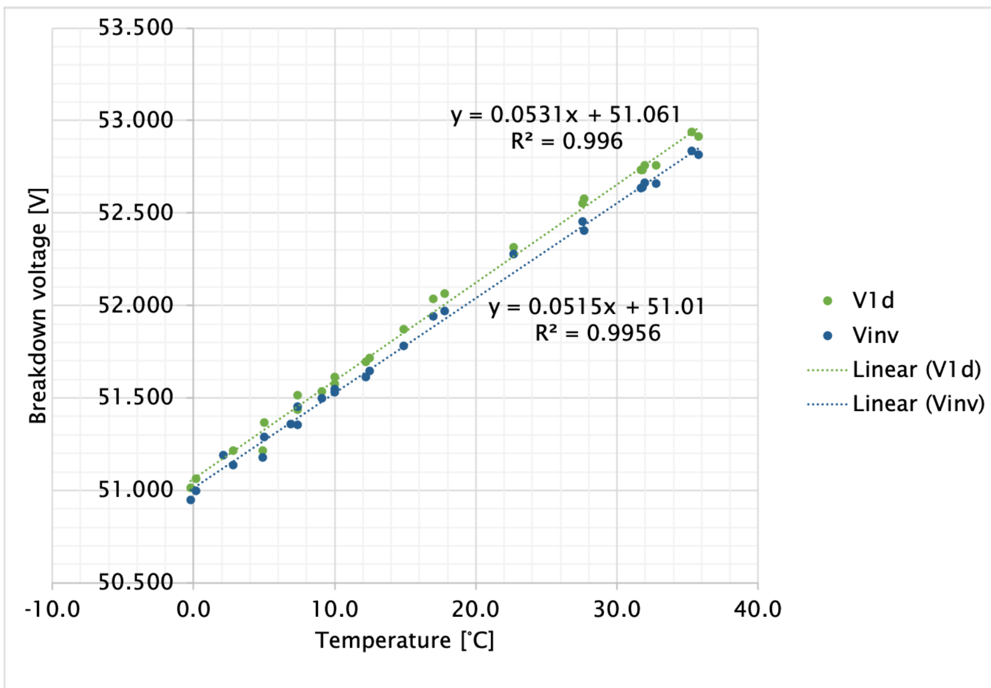


Figure 5.9: I–V curves measured from 0 °C to 35 °C using the first derivative method (labeled V1d) and inverse derivative method (labeled Vinv). Linear regressions are also shown in the plot.

Chapter 6

Dynamic characterization

The dynamic characterization of SiPMs involve the analyze of their temporal behaviors, including the pulses and the parameters extracted from the information carried by pulses. To this end, a proper readout circuit is required for acquiring a large number of pulse samples. Available parameters include gain, DCR, optical crosstalk probability, afterpulsing probability, charge resolution, and so on. In addition, some parameters related to the pulse shape can also be obtained, such as pulse width, fall time constant, etc., although these only partially reflect the nature of SiPM.

6.1 Readout circuit

The readout circuit should be tailored to the application or characterization purpose. As mentioned before, the SiPM signal contains a fast component mediated solely by C_{Q} and a slow component mediated by $C_{\mathrm{Q}}+C_{\mathrm{J}}$. When photon timing is the primary issue, the features of the fast component must be preserved as much as possible, so that the arrival time of a photon can be obtained through the timestamp of the rising edge, allowing the timing resolution to reach tens or hundreds of picoseconds. In application of IACT we lay more stress on the photon counting ability. But because of the existence of NSB, the signal width is also a very important consideration. In short, different users have to design a readout circuit that produces the optimal pulse shape according to their own needs. Here we only consider a relatively simple readout circuit for SiPM dynamic characterization.

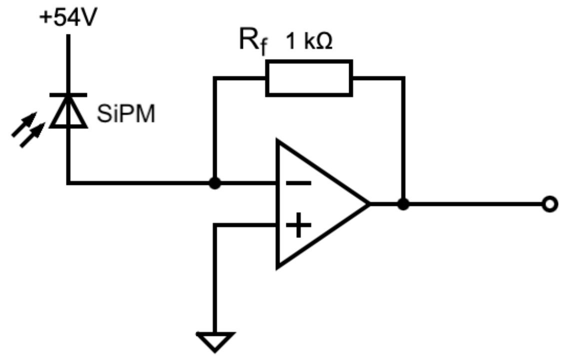


Figure 6.1: SiPM's readout with a TIA.

In general, SiPM signals can be read out in the voltage mode or current mode. In the voltage mode, by sensing the voltage drop between anode and cathode, one will directly observe the fast component[115].

On the other hand, the so-called *current mode* refers to the inclusion of a *transimpedance amplifier* (TIA). It is an operational amplifier (op-amp) with an inverting input and a negative feedback resistor (Fig. 6.1). It converts the current output of the SiPM (I_{SiPM}) to voltage signals (V_{out}) with a gain determined by the feedback resistance R_{f}, as in

$$V_{\mathrm{out}} = -R_{\mathrm{f}} \cdot I_{\mathrm{SiPM}}. \tag{6.1}$$

R_{f} will become the amplifier gain in the later calculation of the gain with Eq. (6.5). TIA is not only a common configuration for acquiring SiPM waveforms, but also adopted by the preamplifier chains of the LST camera[9] and the SST-1M camera[108, 153]. In particular, Hamamatsu's driver circuit module C12332, as shown in Fig. 6.2, contains an op-amp along with several jumpers to modify the preamplifier stage[154, 155]. To read out the current signal, our C12332 module was modified using TIA topology with a feedback resistor $R_{\mathrm{f}} = 1\,\mathrm{k}\Omega$.

Although we did not use it, it is worth mentioning that *pole-zero cancellation* (PZC) networks is often used as the pulse shaper for SiPMs[113, 136, 156] and many other detectors. In circuit analysis, poles and zeros refer to the locations that the transfer function goes to infinity and zero, respectively. A PZC network adjusts the frequency response of a circuit by strategically

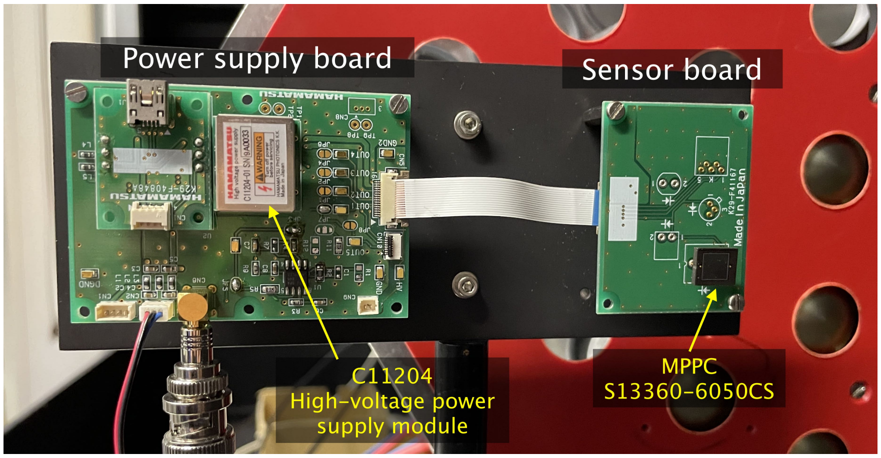


Figure 6.2: Hamamatsu C12332 driver circuit module can implement a TIA in the SiPM's readout. The module comprises a power supply board (*left*) and a sensor board (*right*). The op-amp and the power supply module are mounted on the power supply board, and SiPM is mounted on the sensor board.

placing combinations of resistors and capacitors to cancel out the undesirable poles and zeros. Details about the PZC network may be found in [157], or in [158], the latter being a brief discussion in the context of FlashCam. This approach has been tested or adopted in the front-end electronics of cameras in CTA-LST[159] and CTA-SST[91]. Recently produced SiPMs have also shown the cability to achieve short pulse widths without the usage of the PZC circuits[9].

There may be a digitization stage between the amplifier stage and the data acquisition system. This can be a combination of different types of ADCs that convert analog signals, including timestamp and pulse height, to digital signals. Commonly used are time-to-digital converters (TDCs) and charge-to-digital converters (QDCs); the latter is an integrator followed by an ADC. It expresses charge released by the SiPM in terms of channels, which can be converted back to charge if needed.[1] When using a digitizer to analyze the pulse height, one must find the optimal integrating gate width, in order to cover the entire pulse response and obtain the correct charge through integration. Therefore, it is a routine procedure to study the charge spectrum by varying the gate width. Furthermore, when one seeks to instantly obtain the pulse height spectrum without converting it to the charge spectrum, he or she may consider using a multichannel analyzer (MCA). The digitization devices mentioned here, such as QDC and MCA, can indeed bring some convenience in data acquisition and

[1]Eq. (7) of Ref. [135] is an example of conversion factor between QDC channel and charge.

have been modularized in some educational kits.[2] These devices come with additional costs, of course.

As a matter of fact, the analog waveforms after amplifier stage can be digitally acquired by a digital oscilloscope, and data can be transferred to a personal computer for waveform analysis, achieving the same functionality as hardware-based digitization. The software approach is proven fast and reliable[145]. In our case, python scripts and a library have been developed for the SiPM dynamic characterization. The current version aims to handle the following task:

- Baseline correction
- Modeling the single-photoelectron response (SPR)
- Template subtraction
- Plot the SPE spectrum
- Extraction of parameters, e.g. DCR, crosstalk probability

The explanation and results of these analysis tasks are mainly presented in Section 6.2.

6.2 Dark conditions

6.2.1 Experimental setup

The experimental setup is shown in Fig. 6.3. The setup is placed in dark conditions, like the one in static characterization. Temperature is not monitored during the measurements. However, it is recommended that the entire setup be integrated into the temperature control system in future upgrades.

Hamamatsu's driver circuit module is equipped with a high-voltage power-supply module, C11204, which contains a temperature compensation function and can be configured by a GUI application installed on a computer[160]. Despite these merits, we did not use it to

[2] For example, CAEN SP5600 (`https://www.caen.it/products/sp5600an/`) are used by [135, 144].

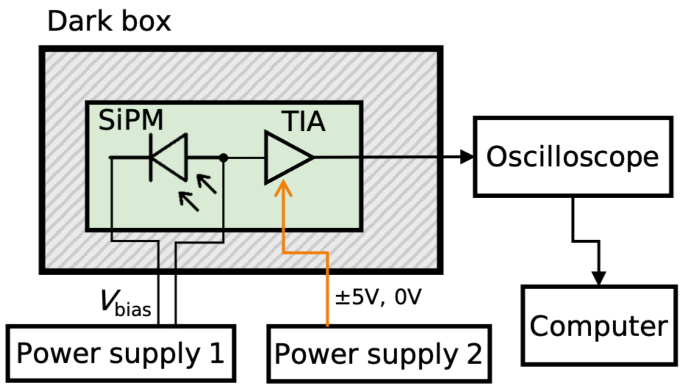
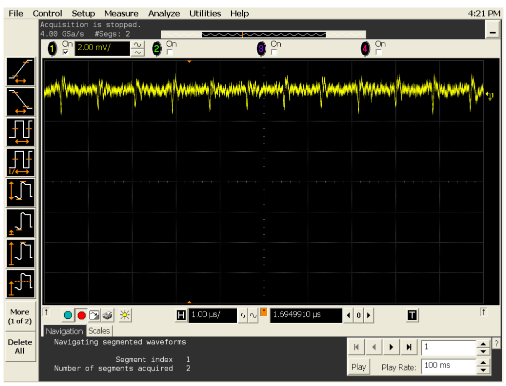
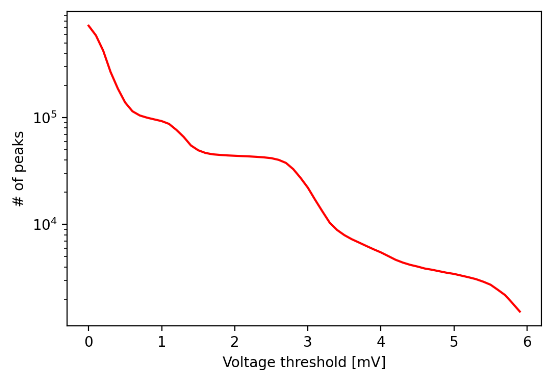


Figure 6.3: Experimental setup for dynamic characterization in dark conditions. The SiPM is biased by Power Supply 1. The TIA is powered by Power Supply 2.



(a) Periodic electrical noise displayed on the oscilloscope screen.



(b) Staircase plot obtained under the influence of periodic electrical noise. An "extra little step" appears near 1 mV.

Figure 6.4: Periodic electrical noise

supply bias voltage to the SiPM because we found that it might cause periodic electrical noise (Fig. 6.4a). The amplitude of this noise is about half the amplitude of the MPPC (Fig. 6.4b); that is to say, there are redundant counts on the pulse height spectrum around 0.5 p.e. This makes it impractical to take 0.5 p.e. as the threshold for the DCR statistics. The reasons for this electrical noise have not been investigated in detail. Our expedient was to dismount this power supply module temporarily and replace it with an external voltage source. In addition to this, the amplifier requires a voltage supply as well. Thus, two voltage supplies were adopted in the setup—one supplies the reverse biasing of the SiPM, and the other supplies 0 V and ±5 V to the driver circuit.

6.2.2 Data acquisition

As already mentioned, the data to be analyzed is acquired by a digital oscilloscope, Agilent Infiniium DSO8104A, stored on an external hard drive and can be accessed by the computer. The maximal sampling rate of the oscilloscope is 4 GSa/s, but a sampling rate of 500 MSa/s is chosen. The signal fed directly from the amplifier is DC-coupled and filtered by a 20 MHz bandwidth-limiting filter. *Segments* of waveform samples are acquired when the oscilloscope is triggered by a falling edge of 1 mV trigger level.[3] In addition, one segment is set to contain 10 000 sampling points, or 20 µs. Up to 200 segments can be saved in a text file with ASCII encoding within the limits of the oscilloscope's memory depth. Hence, each text file (called a *batch*) contains $20\,\mu\text{s} \times 200 = 4\,\text{ms}$. For overvoltages from 1.0 V to 7.0 V in steps of 0.5 V, 10 batches form a complete dataset, which amounts to a total time duration of 40 ms. For the DCR that scales on the order of hundreds of kHz to MHz at room temperature, such a sampling duration would produce tens of thousands of dark pulses, which is reasonably sufficient.

6.2.3 Baseline correction

A baseline drift of order 10^{-4} V to 10^{-5} V can be observed on each waveform segment. In terms of the purpose of application in an IACT camera, if the baseline drift is not too fast, it can be modified in digital electronics in the camera and corrected before triggering. However, in terms of dynamic characterization, an expedient approach was adopted as a preprocess of the software to obtain the single-photon response model quickly.

The idea is, for every segment, a filtered average amplitude of the samples A_i $(i = 1, 2, 3, \ldots)$ in the segment is computed. The filtering is done in following steps:

1. Compute the average $\langle A \rangle$ and the standard deviation σ_A of the all samples in the segment.
2. Define the standard score as $z_i = (A_i - \langle A \rangle) / \sigma_A$ for each sample A_i.
3. Select samples with the empirically assigned mask condition $-1.2 < z_i < 0.8$, then compute the average of these samples as the baseline of the this segment. The mask

[3] In principle, the triggering mode does not affect the waveform segments or subsequent pulse statistics. The only effect is that a triggered pulse is consistently aligned at the center of the waveform segment.

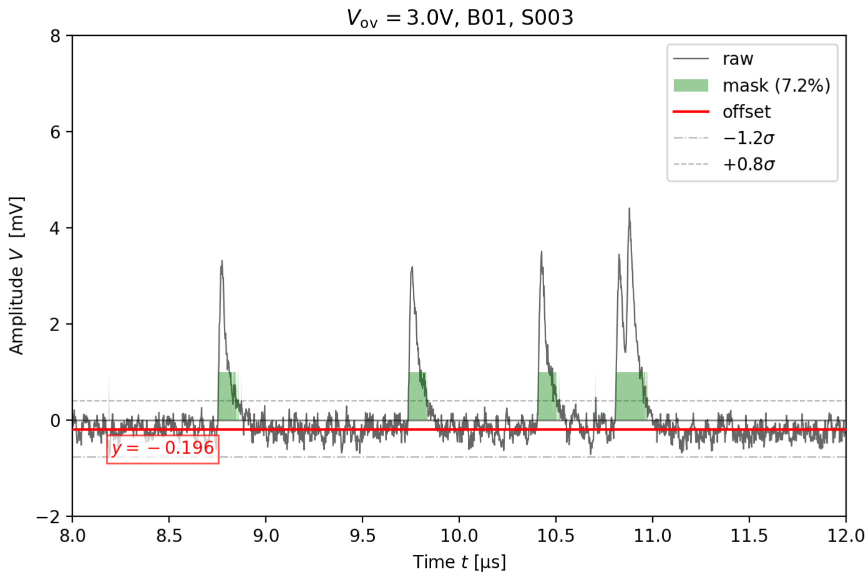


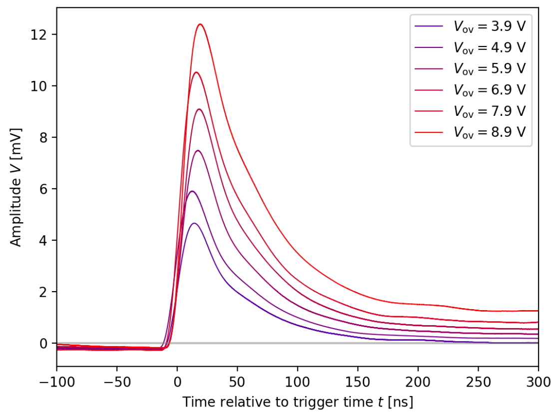
Figure 6.5: The elements of the baseline correction algorithm. The shadowed rectangles represent the masks that are excluded. The red line (labeled as *offset*) is the calculated average $y = -0.196$ (in mV). The dashed lines visualize the conditions of masking. The percentage in the legend is the mask rate, defined by the ratio of mask length to the total segment length.

condition is not symmetric about zero because $\langle A \rangle$ is positive biased from the real baseline level by the signal peaks.

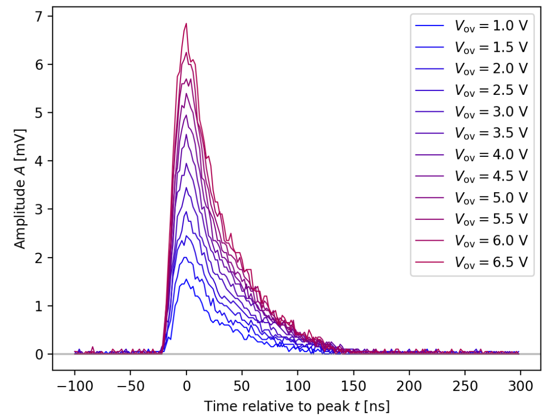
This condition will effectively exclude the pulses that drive up the overall average. The calculated average is then treated as the baseline drift for the entire segment and is deducted from the raw data. The quality of baseline correction is evaluated by calculating the chi-square per for each segment, which is defined as sum of squares of deviation from baseline divided by 10 000. Fig. 6.5 illustrates the elements in this baseline correction approach and compares the waveforms with and without the correction. It was found that this approach reduces up to 40% of the chi-square.

6.2.4 Single-photoelectron response

The next procedure is to build a model for the single-photoelectron response (SPR) of SiPM, which serves as the template for pulse reconstruction and pile-up elimination. Although the



(a) Average waveforms calculated by oscilloscope.



(b) Statistically-modeled SPRs.

Figure 6.6: Average waveforms and SPRs of different overvoltages V_{OV}.

SPR can be seen on the oscilloscope using the waveform averaging function, this is only suitable for observing pulses and roughly estimating pulse heights, and it is not suitable for analysis. This is because the superposition of adjacent pulses makes the voltage levels before and after the triggered pulse deviate from zero, and the rear of the pulse is significantly higher than the front of the pulse. Fig. 6.6a shows such a phenomenon. The time response $A(t)$ decays to a non-zero level, which signifies that the gain cannot be calculated by integrating the response over time.

Pulse finding

Instead of taking the average waveform as in Fig. 6.6a, the SPR is numerically calculated in our software. The results can be previewed in Fig. 6.6b. First, 100 segments are randomly selected from the dataset of 10 000 baseline-corrected segments. Then a peak-finding algorithm based on the scipy function `scipy.signal.find_peaks`[4] is performed on these samples. The peak-finding criteria are the height $\geq 1\,\mathrm{mV}$ and the prominence $\geq 1\,\mathrm{mV}$. Prominence measures how much a peak stands out from its reference baseline;[5] the latter may be the falling edge of some pulse. So combining these two conditions can:

[4] https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html.

[5] https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.peak_prominences.html.

- identify not only single pulses but also piled-up pulses with typical 1 p.e. amplitude ($A_{1\,\mathrm{pe}} \simeq 3.5\,\mathrm{mV}$ for $V_{\mathrm{OV}} = 3\,\mathrm{V}$ in our readout circuit);
- filter out electrical noise no matter if it appears on zero level or the pulse edges;
- preserve the opportunity to find afterpulses with pulse heights lower than $1A_{1\,\mathrm{pe}}$.

Another strategy is to estimate $A_{1\,\mathrm{pe}}$ and use $0.5A_{1\,\mathrm{pe}}$ for the pulse-finding trigger, where $A_{1\,\mathrm{pe}}$ can be given by the average waveforms calculated by the oscilloscope.

Superimposed pulses

All identified pulses are extracted from the waveform segment in an observation window of 200 indices wide (400 ns long) to encompass the complete signal edge. Fig. 6.7 displays all the identified pulses superimposed in the same graph such that the peaks are aligned. Each curve has a certain degree of transparency, so the darker the area, the greater the probability that the pulse has a specific amplitude at a specific point in time. It can be clearly seen that the vast majority of pulses have the same shape, including the same pulse height $A_{1\,\mathrm{pe}}$, demonstrating the fact that SiPM's microcell is a Geiger-mode detector.

Pulses with height $2A_{1\,\mathrm{pe}}$ or $3A_{1\,\mathrm{pe}}$ can be seen in Fig. 6.7. These pulses can, in principle, be attributed to either two or more simultaneous dark events or an event with prompt crosstalk. The probabilities of both cases are evaluated as follows. Suppose the overall dark count rate is 1000 kHz per 14400 cells, then each cell has $(1000\,\mathrm{kHz}/14400) \times (2\,\mathrm{ns}) \simeq 1.4 \times 10^{-7}$ dark counts in a 2 ns time window.[6] The probability of any two cells each independently generating a dark event within 2 ns is

$$\binom{14400}{2} \times \left[\left(1.4 \times 10^{-7}\right) \times e^{1.4\times10^{-7}}\right]^2 \simeq 2 \times 10^{-6}.$$

Suppose the crosstalk probability is 6%. For the case in which two cells are triggered due to prompt crosstalk, the probability is

$$(1000\,\mathrm{kHz}) \times (2\,\mathrm{ns}) \times 6\% = 1.2 \times 10^{-4},$$

[6] Due to the limitation of sampling rate, two pulses with time delay than the 2 ns time window can technically be considered "prompt."

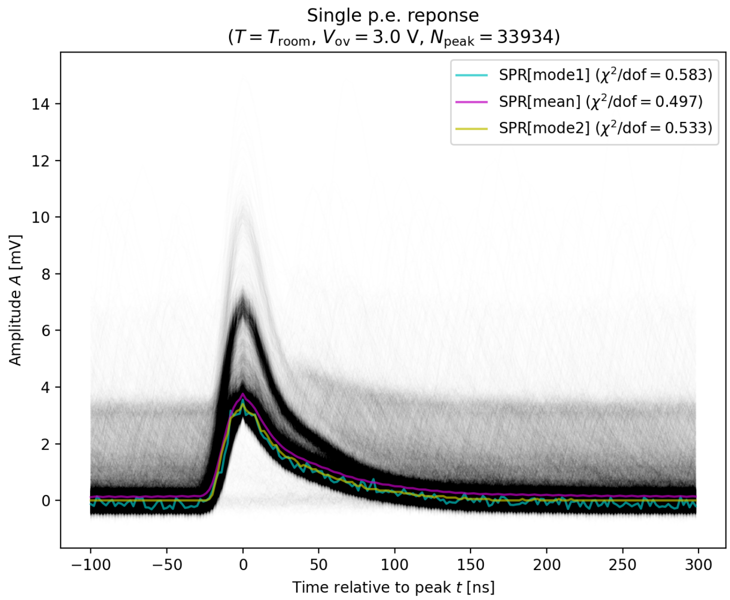


Figure 6.7: Three statistical models of SPR. $V_{\mathrm{OV}} = 3.0\,\mathrm{V}$. To enhance readers' impression, the pulses found in the entire dataset (2000 waveform segments) at this overvoltage are superimposed. However, only 100 samples will be used for subsequent analysis.

which is about 60 times the former case. As a consequence, pulses with height $2A_{1\,\mathrm{pe}}$ or $3A_{1\,\mathrm{pe}}$ can primarily be attributed to prompt crosstalk.

Furthermore, a minority of pulses appear before and after the most apparent pulse shape, forming a large shadow. They cannot all be attributed to afterpulses correlated to primary pulses because causation cannot be inferred simply from temporal proximity.

Attempts was made to fit an analytical function to the data of superimposed pulses. Eqs. (6.2) and (6.3) are two of the exemplary functions that were used.

$$A(t) = \begin{cases} 0 & \text{for } t < t_0, \\ A_1 \cdot e^{-(t-t_0/\tau)} & \text{for } t \geq t_0. \end{cases} \tag{6.2}$$

$$A(t) = \begin{cases} 0 & \text{for } t < t_0, \\ A_1 \cdot e^{(t-t_0)/\tau_1} + A_2 \cdot e^{-(t-t_0)/\tau_2} & \text{for } t \geq t_0. \end{cases} \tag{6.3}$$

Particularly, a result of Eq. (6.3) is shown in Fig. 6.8. Note that only pulses with pulse height

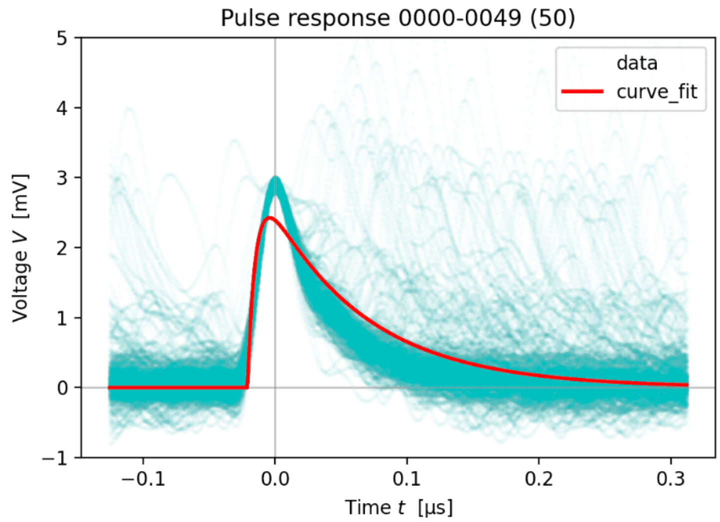


Figure 6.8: Fitting the analytic function described in Eq. (6.3) to an earlier obtained dataset with $V_{\mathrm{OV}} \simeq 3\,\mathrm{V}$. Only pulses with pulse height within ± 1 standard deviation around $A_{1\,\mathrm{pe}}$ are superimposed.

within ± 1 standard deviation around $A_{1\,\mathrm{pe}}$ are superimposed and used for fitting. Although the fitting curve does capture the features of fast rise and slow fall, there is noticeable distortion in pulse height, and the fall time decay constant is also expected to be biased. The form of Eq. (6.3) was inspired by Fig. 7 of [161], which describes the rising and falling edge of the pulse signal using 2 time constants, respectively. With the benefit of hindsight, it is noticed that the readout circuit used in [161] contains a load resistance (i.e., it was the "voltage mode"), which differs from our case. Therefore, it is naturally impossible to achieve a better quality of the fit. In addition, earlier literature also recognized that the analytic formulation of SiPM's pulse shape could not be perfectly found[162].

In summary, finding analytic functions to fit superimposed pulses is not mandatory. The more critical issue is that if we insist on doing this, we must first consider building an electrical model of the readout circuit, which is redundant for our goals. After all, we are interested in the SiPM pulse shape because we want to perform subsequent template subtraction, not the pulse shape parameters themselves.

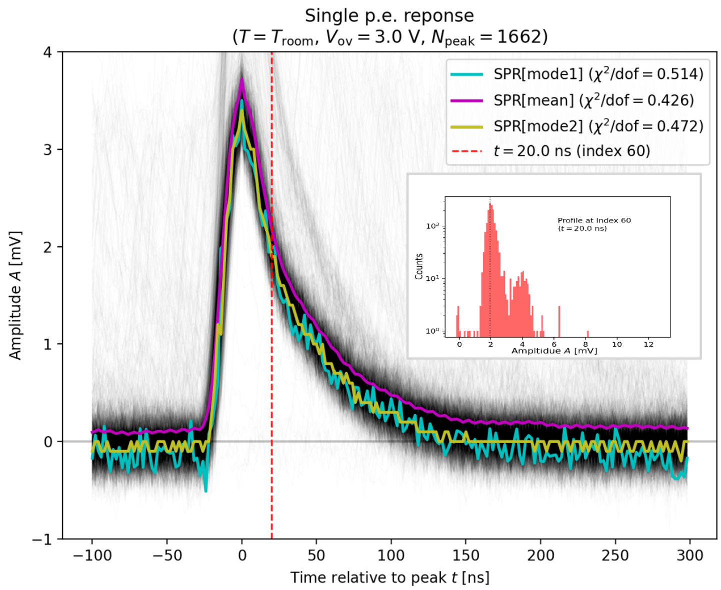


Figure 6.9: Three statistical models of SPR based on 1662 pulses were found in 100 randomly selected waveform segments. The inset is the histogram (with the y-axis in the log scale) of the amplitude profile at $t = 20$ ns, corresponding to the red dashed line in the main panel.

Statistical models of SPR

As a result, we turn our attention to statistical methods that can obtain SPR more efficiently and effectively.

First, we organize all the waveform amplitudes superimposed within the specified window into a two-dimensional array. The typical shape of the array is $N_{\text{peak}} \times 200$. Then, statistics are performed on the amplitude data along axis 1 of the array, representing the time t relative to the peak. It can be imagined that an amplitude profile is made at each time value, resulting in an amplitude distribution at that time, as shown the inset of Fig. 6.9.

Because of the presence of pulse pile-up, the distribution is positively skewed, as expected. So, if we are looking for the most probable value in this distribution, we should calculate the *mode* of this distribution rather than the *mean*, which will be elevated due to the positively skewed tail and will not restore to zero level when decaying.

The most direct way to calculate the mode in an array is to use the function `scipy.stats.mode` (labeled `mode1`), but the SPR obtained in this way suffers from statistical fluctuations and thus appears unstable over time, especially when only a relatively small portion of the segments is sampled.

To avoid the unstable behavior of `mode1`, we can resort to another approach. First we group the amplitude distribution for each t into a histogram with a bin size of 0.1 mV, and then we take the group's maximum value as the mode. The results (Fig. 6.9) show that the "binned mode method" (labeled `mode2`) does not have the problem of fluctuation; compared with the "mean method" (labeled `mean`), it does not have the problem of baseline level elevation. So, qualitatively speaking, the "binned mode method" is the best choice for obtaining the SPR.

To quantitatively evaluate the different statistics-based models, chi-square per degree of freedom

$$\left(\frac{\chi^2}{\text{dof}}\right)_{\text{SPR}} := \frac{\sum_{i=1}^{N_{\text{peak}}} \left(A_i - \hat{A}_i\right)^2}{N_{\text{peak}}} \tag{6.4}$$

are also calculated for three models, where $\hat{A}_i$ denotes the entries of any model. Examples of calculated values are displayed in the legends of Figs. 6.7 and 6.9.

It is reasonable to see that χ^2/dof of `mode2` only ranks second highest because of the positively skewed distribution. But this does not change the fact that `mode2` is representative of the most probable amplitude for any given t. Besides, it turns out that increasing the number of segment samples N_{sample} yields little improvement to χ^2, but will significantly prolong the processing time. As a rule of thumb, $N_{\text{sample}} = 100$ would be adequate.

Results and applications

Now that the SPR is found, we can derive some parameters from the pulse shape:

1. Calculate the gain G by integrating the area under the SPR and consider (cf. Eq. (4.12))

$$G = \frac{\int A_{\text{SPR}}(t)\,\mathrm{d}t}{q_{\text{e}} \cdot R_{\text{f}}}, \tag{6.5}$$

   where $q_{\text{e}} = 1.602 \times 10^{-19}$ C is the elementary charge and $R_{\text{f}} = 1\,\text{k}\Omega$ is the feedback resistance of the TIA in the readout circuit, which connects $A_{\text{SPR}}(t)$ (a voltage function)

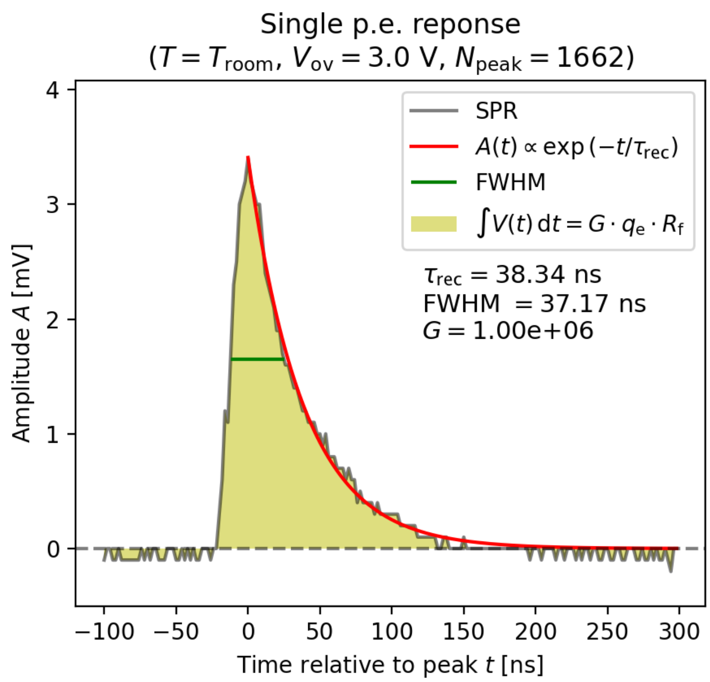


Figure 6.10: The final result of SPR using the `mode2` model. The recovery time constant τ_{rec}, the gain G, and the FWHM are also calculated and annotated.

and $I_{\text{SiPM}}(t)$ through Eq. (6.1). There are many other ways to calculate the gain of SiPM; this is just one of them.

2. Calculate the 1 p.e. amplitude $A_{1\,\text{pe}}$ by find the maximum of the SPR function.

3. Calculate the Full Width at Half Maximum (FWHM) as a representative of pulse width.

4. Calculate the recovery time constant τ_{rec} by fitting to the exponential decay function $A(t) \propto e^{-(t/\tau_{\text{rec}})}$ to the falling edge. τ_{rec} may be helpful information for the designers of the IACT camera.

These results and their corresponding graphical elements are shown in Fig. 6.10. The linear dependence of obtained gain versus overvoltage is proven by Fig. 6.11.

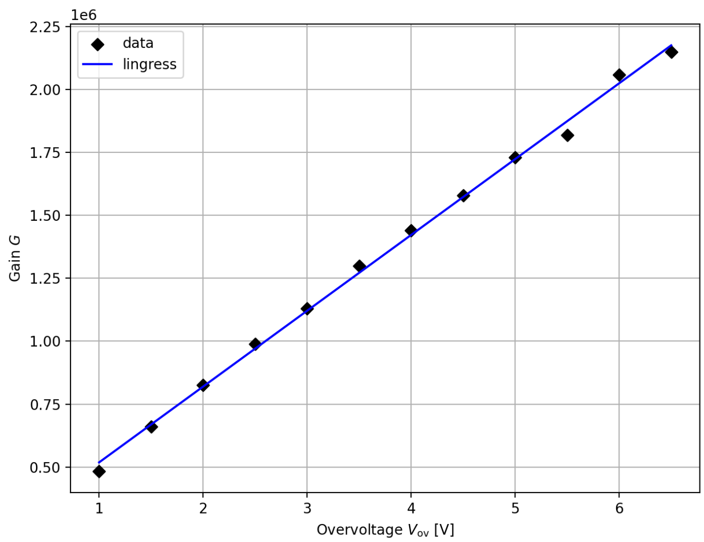


Figure 6.11: Gain vs. overvoltage. The measurement was conducted at room temperature.

Discussion

Deviation from datasheet Comparing Fig. 6.11 with Figure 3-1 of the datasheet[132], it can be found that although the G vs. V_{OV} curves of both are linear, there are some deviations. For example, when $V_{\mathrm{OV}} = 3\,\mathrm{V}$, Fig. 6.11 gives a gain of 1.1×10^6, but the datasheet gives a gain of about 1.8×10^6. The detailed reasons still need to be explored. However, the problem may lie in the transimpedance gain R_{f} of the TIA in the readout circuit, which is assumed to be the nominal value $1\,\mathrm{k\Omega}$ without examination.

Timing resolution Fig. 6.6a reveals that it is rising (leading) edges that aligned perfectly, but not the peaks. This is because the starting time of the edge trigger is taken as $t = 0$ for the waveform segment. The slight difference (of a few ns) in the rising edges under different overvoltages may manifest the timing resolution in the readout circuit. Our readout circuit might be unable to resolve the intrinsic time jitter of SiPM ($< 100\,\mathrm{ps}$, [163]). One reason is that, the maximum sampling rate of our oscilloscope is $4\,\mathrm{GSa/s}$, so the shortest data point spacing is $250\,\mathrm{ps}$. Apart from the limitation of hardware, our template subtraction algorithm (see Section 6.2.5 and the comments therein) introduces another timing uncertainty, which is limited by the rise time of the SPR ($\sim 30\,\mathrm{ns}$).

Possible signs of the fast component? There is a faint tip at the peak of the pulse, which can be seen both in superimposed pulses and in SPRs with large overvoltages. Chances are that this is the fast component inherent in the SiPM pulse. Theoretically, we could increase the bandwidth of the readout circuit to make the fast component more visible, or choose a SiPM with smaller area to reduce the capacitance. However, how to use algorithms to extract information about this fast component constitutes another interesting question.

Necessity of smoothing the SPR? Although the SPR function has a jagged appearance, there is no need to smooth it with spline interpolation[162]. One reason is that the tip of the fast component may be obliterated, reducing the pulse height. Another reason is that the entire waveform at the pulses or elsewhere is contaminated by electrical noise. If the SPR is smoothed, the noise will still come out when the template subtraction algorithm is implemented later, so smoothing does not provide substantial help in this situation.

6.2.5 Template subtraction

The template subtraction algorithm aims to minimize the pile-up effect and to correct the pulse height for each pulse. In addition to template subtraction, there are various algorithms to correct the pile-up effect[164], one of which is the deconvolution method (e.g., `scipy.signal.deconvolve`)[143]. Nevertheless, this usually places some limitations on the functions, e.g., smoothness and non-zero values. In our case, SPR is not smoothed, nor does it rule out zeros. So, the attempts to deconvolve the waveform by the SPR were unsuccessful.

Pulse trains

In order to find the pulses that may be affected by the pile-up, for all the pulses found the pulse finding (before the SPR modeling, see Section 6.2.4), we take the time window[7] 100 ns before and 300 ns after the peak ($-100\,\mathrm{ns} < t < 300\,\mathrm{ns}$). Next we find the unions of these time windows that may contain one or more pulses. These are called the *pulse trains* here.

[7] A proper time window length should take into account the time required to recover 99% of the bias voltage, approximately equal to $5\tau_{\mathrm{rec}}$[121]. τ_{rec} has been defined in Section 4.1.6 and been obtained in Section 6.2.4.

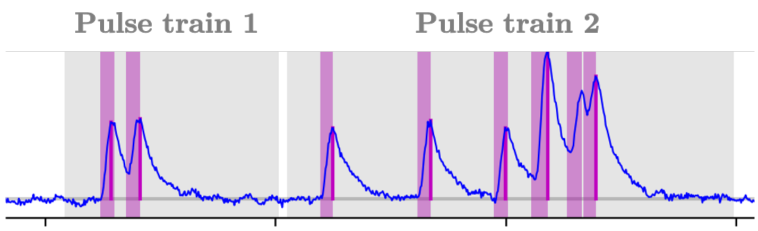


Figure 6.12: Two pulse trains with the pile-up effect. The magenta line segments indicate the pulses found in Section 6.2.4. The magenta strips show the rising edges of the pulses.

Rising edges

Then, a *rising-edge finding algorithm* is performed for each pulse train. The conditions for the rising edges are relatively simple: we pick out the intervals where every amplitude is larger than some threshold (e.g., $> 1\,\text{mV}$) and the amplitude is continuously increasing for at least some time (e.g., $> 6\,\text{ns}$ or 3 data points was chosen in the analysis). Thanks to the fast rise of the SiPM signal, it turns out that the rising-edge finding algorithm is more sensitive than the ordinary pulse finding algorithm within a pulse train.

Fig. 6.12 shows an example of two pulse trains, where Pulse Train 1 contains two rising edges and Pulse Train 2 contains six rising edges. Note that one of the rising edges was not found during the pulse finding because it is not "prominent" enough. However, the rising-edge finding algorithm can remedy the loss counts of peaks.

Template subtraction

After the rising-edge finding, the loop of template subtraction begins. The so-called *template* is precisely the SPR found in Section 6.2.4. Let $W(t)$ be the original waveform of a pulse train, and let $\text{SPR}(t-t_1)$ be the SPR that is horizontally shifted such that it peaks at $t = t_1$, where t_1 is the temporal position of the peak associated with the first (i.e., leftmost) rising edge in the pulse train. Then, $W(t)$ can be expressed by

$$W(t) = A_1 \cdot \text{SPR}(t - t_1) + W_1(t), \tag{6.6}$$

where A_1 is the corrected pulse height and is defined as the ratio between the original height and the height of the SPR,[8] and $W_1(t)$ is the remainder. Although being called a "subtraction," this operation is more like a division, where the $W(t)$ is the dividend, A_1 is the quotient, and SPR(t) is the divisor.

Afterward, $W_1(t)$ becomes the dividend for the next run of template subtraction. The rising-edge finding algorithm is performed again to count the remaining number of rising edges again. Now, $W_1(t)$ is expressed by

$$W_1(t) = A_2 \cdot \mathrm{SPR}(t - t_2) + W_2(t), \tag{6.7}$$

where the definitions of A_2, t_2, and W_2 are the same as above, except that the object to be described is turned into the second peak from the left.

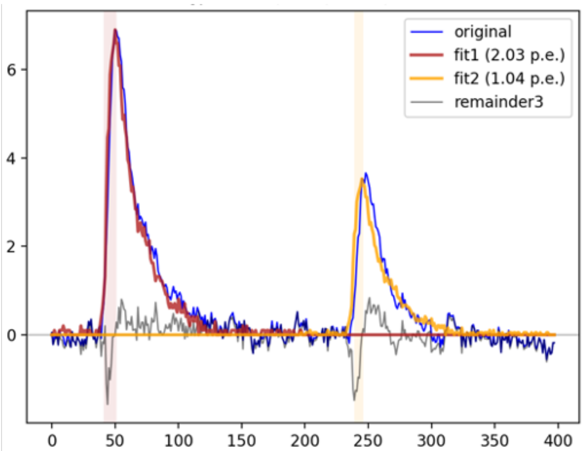
In general, if there are k rising edges in a pulse train, then the original waveform is equivalently deconvolved by the SPR:

$$W(t) = R(t) + \sum_{i=1}^{k} A_i \cdot \mathrm{SPR}(t - t_i), \tag{6.8}$$

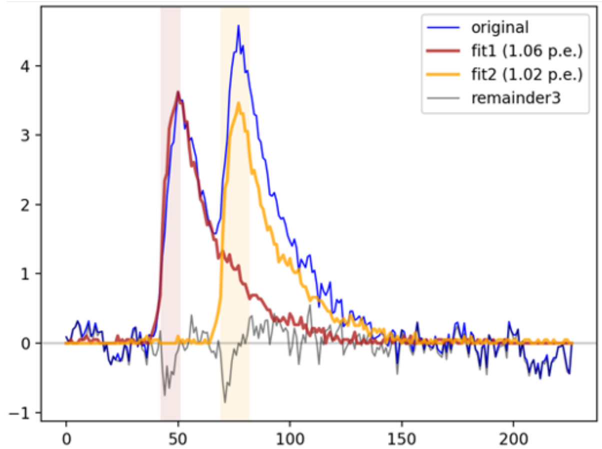
where $R(t) = W_k(t)$ is the last remainder, where no more rising edges will be identified. Ideally, $R(t)$ will contain only noise with amplitude less than the threshold, and the reasonable range for corrected pulse heights A_i is (1.0 ± 0.1) p.e.

Fig. 6.13 shows the potency of the template subtraction algorithm. Fig. 6.13a is a pulse train where the pulses almost do not pile up, and the first pulse is also a crosstalk event. Figs. 6.13b to 6.13d are pulse trains where the degree of the pile-up is increasing. Especially, Fig. 6.13c is a pair of piled-up crosstalk events. Figs. 6.13e and 6.13f are the pulse trains with three piled-up events. If there were no template subtraction algorithms, the peak heights of the (blue) original curves would be mistaken for the pulse heights, further leading to the distortion of the pulse height spectrum.

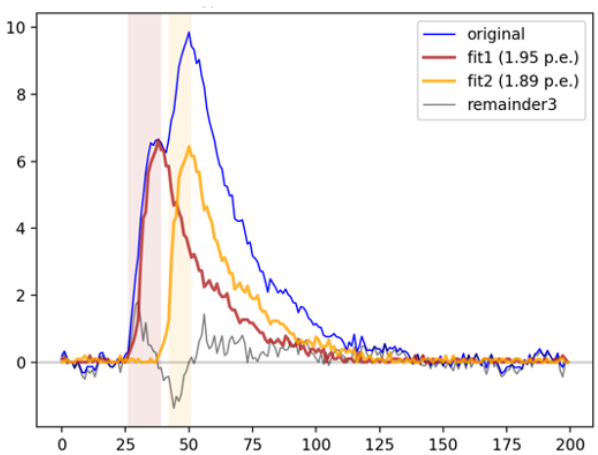
[8] For the leftmost peak in the pulse train, the reconstructed pulse height will be the same as the original pulse height. However, for the other peaks, the reconstructed and original ones will differ significantly as long as the original ones are affected by the pile-up effect.



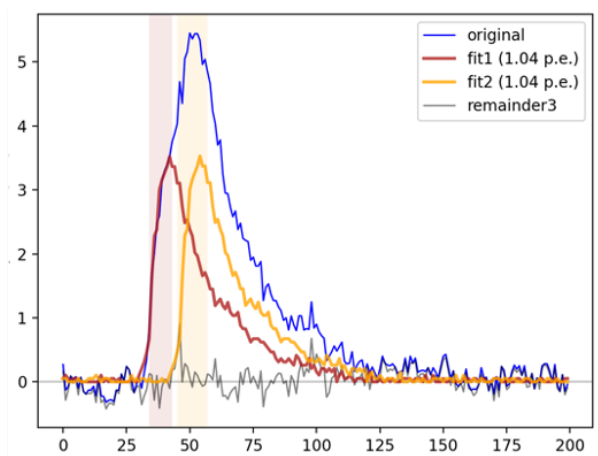
(a) Pulse train with 2 almost separated pulses



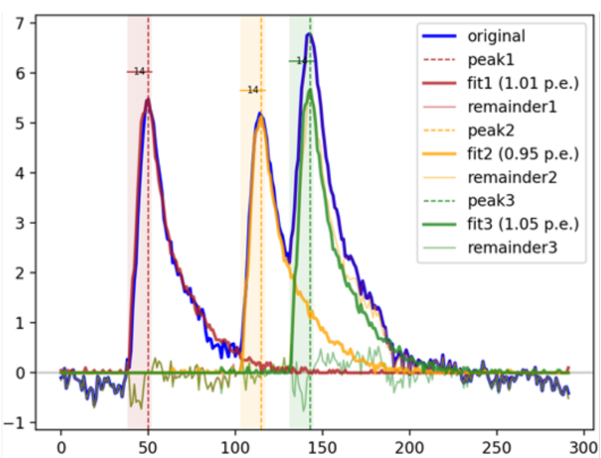
(b) Pulse train with 2 piled-up pulses



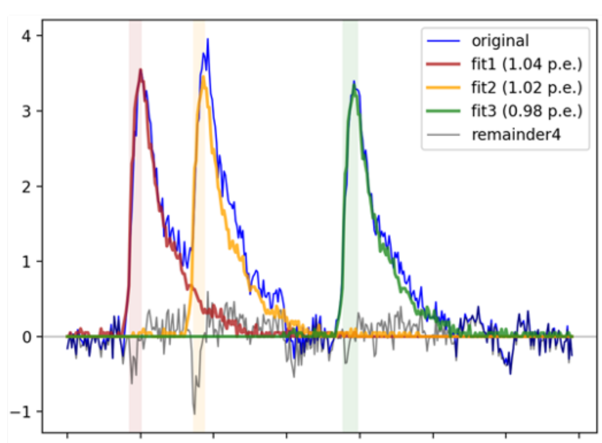
(c) Pulse train with 2 piled-up pulses



(d) Pulse train with 2 piled-up pulses



(e) Pulse train with 3 piled-up pulses



(f) Pulse train with 3 piled-up pulses

Figure 6.13: Results of template subtraction. Each x-axis is data index in the pulse train (1 index spacing = 2 ns). Each y-axis is the amplitude (voltage) in mV. "Original" in legend means $W(t)$, "fit i" means $(A_i \cdot \mathrm{SPR}(t - t_i))$, and "remainder k" means $R(t)$ in Eq. (6.8).

Discussion

The template subtraction algorithm still has some issues worth investigating. The horizontal translation of the template is done in the algorithm such that the maximum of the SPR is aligned with the peak to the maximum of the rising edge; the latter is found by extending a small interval to the right on the rising edge and finding the maximum value. Due to electrical noise, the amplitude and time coordinates calculated in this way have uncertainties.

Amplitude uncertainty In order to reduce the impact of the amplitude uncertainty, each time the template is subtracted on a pulse train, the rising edges will be searched again and compared with the originally found ones; if the numbers of rising edges are inconsistent, a flag variable in the program will be marked as a warning. The current version of the program does not particularly process those pulse trains with warnings, but instead continues the loop of template subtraction until no more rising edges are found. It remains unclear which component dominates in the standard deviation of the subsequent Gaussian fitting: the amplitude uncertainty caused by the electrical noise or the inherent gain fluctuation of the SiPM. The solution may lie in canceling the threshold for electrical noise, allowing the pedestal to appear, and calculating its width σ_{ped} by an independent Gaussian fitting.

Time uncertainty From Fig. 6.10, one can obtain the rise time of SPR is $\sim 20\,\text{ns}$, which sets a limit to the discernibility of piled-up pulses. To break through this limitation, an additional measure is to split the rising edge into two sections on average if the length of the rising edge is $> 32\,\text{ns}$ and the height is < 1.8 p.e. Fig. 6.13d is such an example. The segmentation of excessively long rising edges technically enhances the resolution to 16 ns. Compared to this value, the time uncertainty due to to electrical noise appears to be less problematic.

However, we do not know whether the timing resolution of our current hardware and software settings is sufficient to analyze the delayed correlated noise of our SiPM. The Differential Leading Edge Discriminator (DLED) algorithm, reportedly efficient in acquiring accurate timestamps[145, 165], might be a powerful substitute for the current template subtraction algorithm if the timing resolution is strictly required.

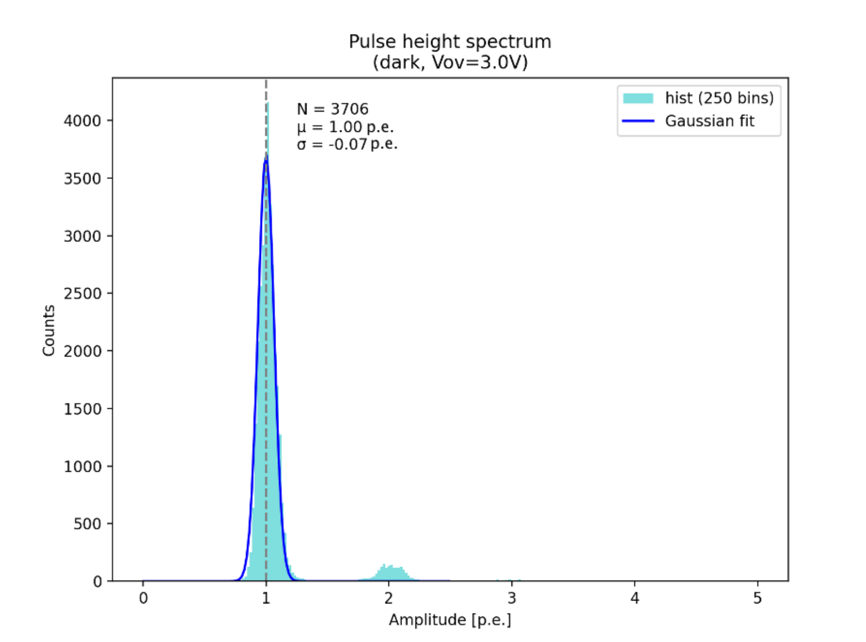


Figure 6.14: Pulse height spectrum (linear scale) in dark conditions ($V_{\mathrm{OV}} = 3.00\,\mathrm{V}$). The template subtraction is applied.

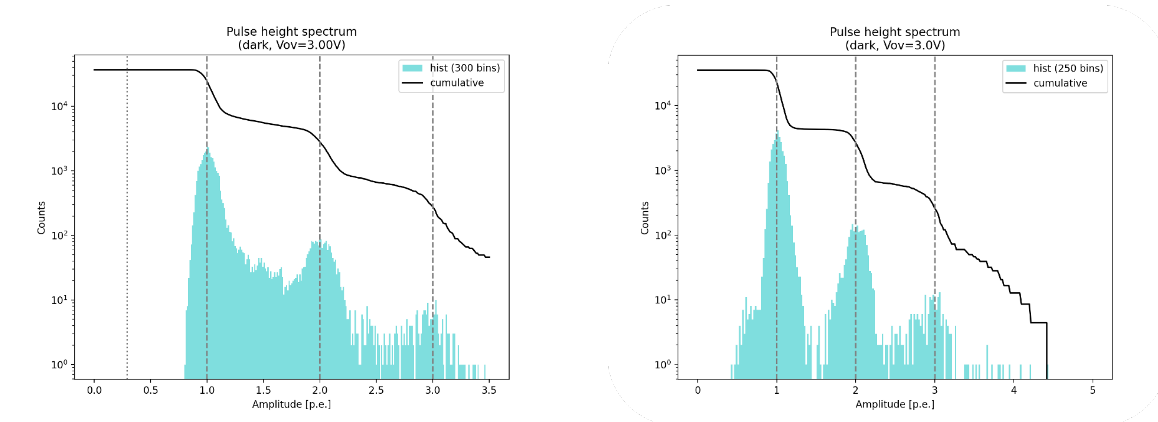
6.2.6 Single photoelectron spectrum

Now that the pulse heights have been corrected, the pulse height spectrum (PHS) can be plotted with more accuracy.

Fig. 6.14 is a PHS obtained in dark conditions and with the overvoltage $V_{\mathrm{OV}} = 3.00\,\mathrm{V}$. For the analysis, one can fit the Gaussian function on the leftmost peak, i.e., the 1 p.e. peak,[9] to examine if the single photoelectron amplitude is equal to 1 p.e. The results justify the applicability of the Gaussian distribution and the consistency of the single photoelectron amplitude. A tiny peak can also be seen at 2 p.e., which represents the optical crosstalk that triggers one adjacent microcell. The 2 p.e. peak is much smaller than the 1 p.e. peak, which shows how excellent this SiPM is in suppressing optical crosstalk.

The two subfigures in Fig. 6.15 are also PHS obtained with the same data and the same conditions as Fig. 6.14. Fig. 6.15a and Fig. 6.15b shows the PHS without and with template subtraction, respectively. Compared with Fig. 6.15a, the counts between the photon peaks in

[9] The rising-edge finding algorithm has already set a threshold on these PHS, so there are no pedestals (0 p.e. peak).



(a) PHS *without* the template subtraction (b) PHS *with* the template subtraction

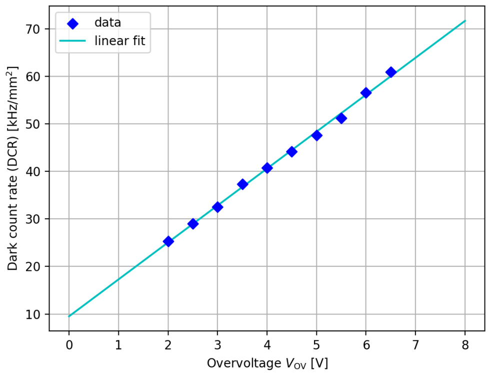
Figure 6.15: Pulse height spectrum (log scale) in dark conditions ($V_{\text{OV}} = 3.00\,\text{V}$). The black curves shows the accumulative accounts of the histogram and can be nicknamed "staircase plots."

Fig. 6.15b are significantly reduced, which can be regarded as template subtraction's ability to eliminate pile-up effects.

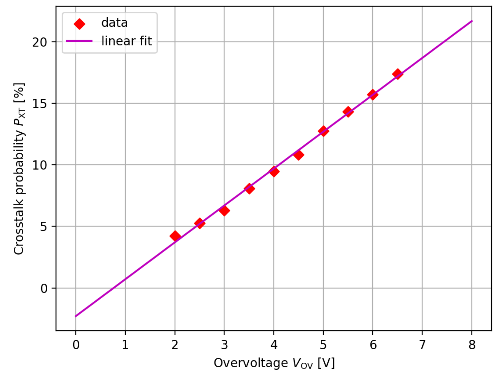
Lastly, the dark count rate (DCR) and the crosstalk probability can be extracted from the PHS such as Fig. 6.15. DCR is the total count in PHS divided by the total length of dataset (40 ms). Measuring DCR directly by counting pulses faces a problem—it cannot isolate the afterpulses. Therefore, further analysis, especially temporal statistical analysis, is required to decompose the afterpulse components and the primary dark counts.

The crosstalk probability P_{XT} is determined by Eq. (4.25). According to analysis, template subtraction reduces the overestimated crosstalk probability by 25%. However, the measured crosstalk probability is approximately 1.8 to 2.4 times the data-sheet reference values[124, 132]. The reasons require more in-depth research.

Fig. 6.16 shows DCR per area and P_{XT} plotted against the overvoltage; both of the parameters exhibit good linearity.



(a) Dark count rate per area



(b) Crosstalk probability

Figure 6.16: Dark count rate per area and crosstalk probability vs. overvoltage.

6.3 Low-light-level measurements

The purpose of this experiment is to get acquainted with the operation of the laser–SiPM measurement system and to obtain PHS at low light levels.

6.3.1 Experimental setup

The main components and wiring of the experimental setup are shown in Fig. 6.17. The experiment is still carried out in the dark box. The SiPM (Hamamtsu MPPC S13360-6050CS) is illuminated by a picosecond diode laser (PDL), which comprises a laser driver module and a laser head.

Pulsed diode laser

The laser head (PicoQuant LDH-D-C-405) is mounted on an optical table in the dark box. The wavelength of laser is (405 ± 10) nm[166], which can match the spectral response of the SiPM.

To generate a pulsed laser beam, the laser driver (PicoQuant PDL 828 "Sepia II"[167]) receives input from a pulse generator (Agilent 33250A), which generates a square wave of a pulse width of 5 ns at a frequency of 1 kHz. The sync output of the pulse generator

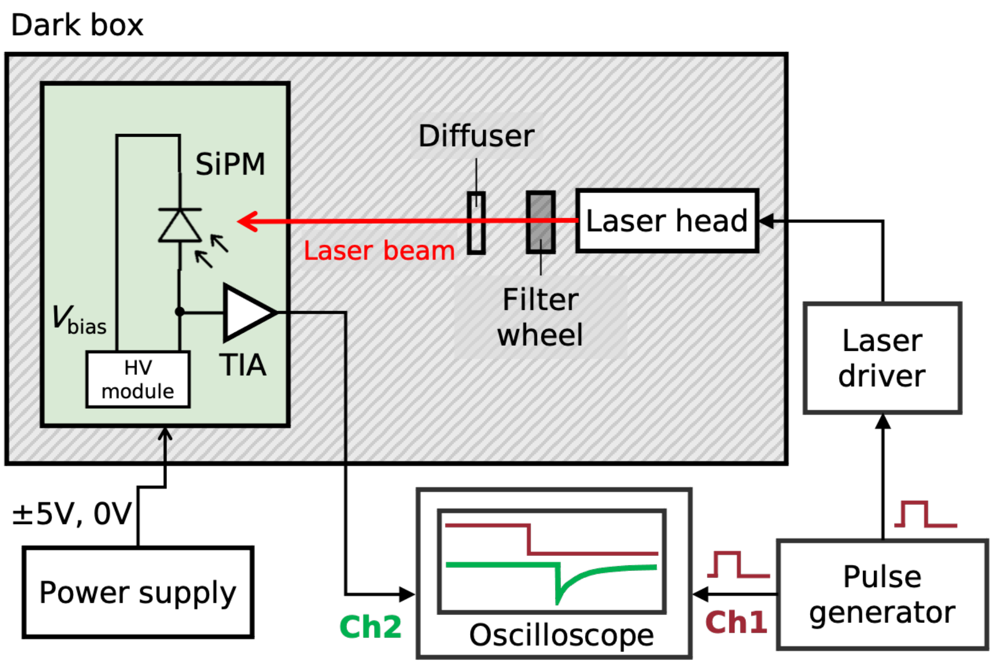


Figure 6.17: Experimental setup for low-light-level measurements. The SiPM is biased by the HV power supply module on the driver circuit module.

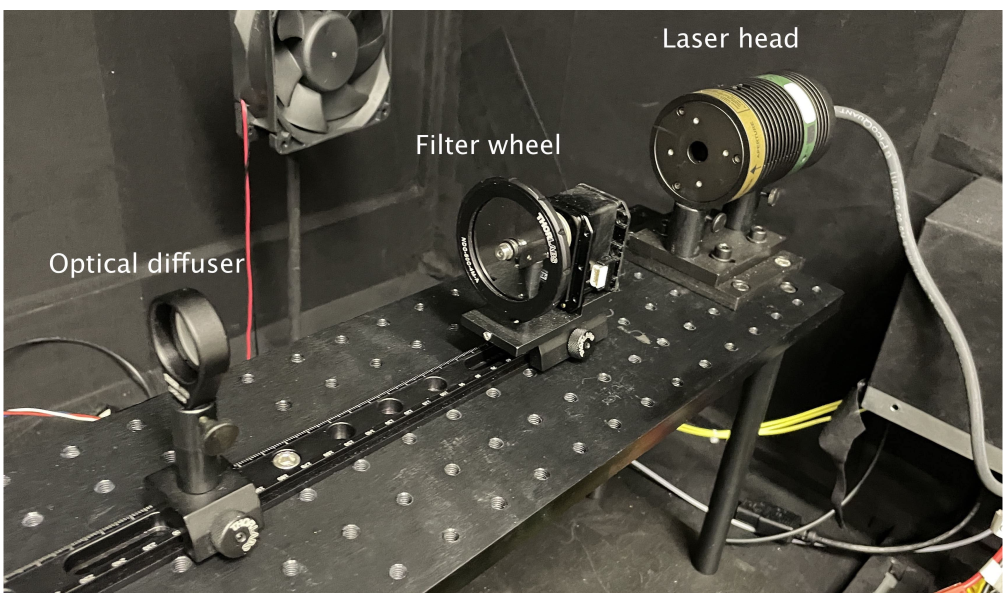


Figure 6.18: Components on the optical table.

is connected to one of the oscilloscope channels (e.g., Channel 1) and serves as the trigger source for the oscilloscope. Synchronizing laser emission and SiPM waveform acquisition with the pulse generator facilitates the differentiation between genuine photon-induced pulses and dark pulses, thereby increasing the reliability of collected data for plotting the PHS.

Filter wheel and optical diffuser

A filter wheel and an optical diffuser are installed on the optical table, as shown in Fig. 6.18. The intermittently emitted laser pulses are attenuated by the filter wheel and diffused by the diffuser, and there are very few photons that can reach the SiPM. The filter wheel (Thorlabs NDC-50C-4M-A) is placed in front of the exit of the laser head to control the light intensity. Its optical density (OD) varies continuously from 0.04 to 4.0 as the angle varies from 0° to 270°[168]. Therefore, when the laser beam is irradiated in a fixed direction and the filter wheel is rotated, the intensity of the emitted laser light can be adjusted. The mechanical device that rotates the filter wheel can be controlled by a computer. Even so, it remains the same settings throughout our measurents.

Additionally, with the optical diffuser placed in front of the filter wheel, the laser intensity can further be reduced and the laser beam does not need to pointing to SiPM with high accuracy.

SiPM and its driver/readout circuit

The SiPM is mounted on the Hamamatsu's driver circuit module, situated on the opposite side within the dark box. In contrast to the setup detailed in Section 6.2.1, the high-voltage power supply module C11204 has been reinstalled onto power supply board. This decision was driven by two key factors: firstly, the necessity of temperature compensation firmware to regulate the overvoltage at a consistent 3.0 V; and secondly, the improved noise immunity achieved by the synchronous triggering and sampling method.[10] The remaining aspects of the driver circuit and readout circuit are consistent with the descriptions provided in Sections 6.1 and 6.2.1.

[10]The sampling method now confines to a short interval (mentioned shortly later) defined by the triggering. As a consequence, the waveform is less susceptible to the periodic electrical noise mentioned in Section 6.2.1

Data acquisition

The SiPM signal is recorded by a different channel of the oscilloscope (e.g., Channel 2). With each trigger initiated by the rising edge of the square wave from the pulse generator, the oscilloscope captures the waveforms from the SiPM. In the acquisition setups, the waveforms are sampled in Segmented Mode with a sampling rate of 2.00 GSa/s. Each wavefrom segment comprises 1000 points, spanning a time interval of 500 ns. A comprehensive dataset comprising 20 000 segments was collected and subsequently subjected to analysis.

6.3.2 Pulse height spectrum at low light levels

The process of obtaining the PHS at low light levels can be divided into four steps. The first three steps are dedicated to the process of *background removal*. Since the baseline drift is seen in the waveforms (recall Section 6.2.3), it is recommended to perform the background removal to correct the pulse height before plotting the PHS. The term "background" here refers to the mean of the amplitudes within a properly chosen interval preceding the rising edge of the triggered pulse (i.e., *pre-trigger interval*).

Step 1: Calculate the average waveform. The 20 000 segments are averaged element-wise, resulting in a pulse-shaped waveform as in Fig. 6.19. A suitable segment length of 500 ns has already been chosen, so that the entire falling tail of SPR ($\sim$ 200 ns) is encompassed in the right half of a segment. While the pulse shape in Fig. 6.19 resembles the SPR (e.g., Fig. 6.10), they are not identical. Since the SiPM can receive more than one photon from the laser beam, the pulse height in a single waveform segment is expected to be an integer multiple of $A_{1\,\mathrm{pe}}$. The smoothness of the pulse shape reflects the precise alignment of the maximum points ($t = 250$ ns) in all pulses. This alignment occurs because the oscilloscope acquires the waveform triggered by the synchronous outputs of the pulse generator.

Step 2: Define two intervals of interest. We need to define two time intervals (time windows) of equal length from the average waveform. With the exception of specific conditions outlined below, the positions and lengths of these intervals are not strictly defined.

1. **Pre-trigger interval.** The first interval, known as the *pre-trigger interval*, is

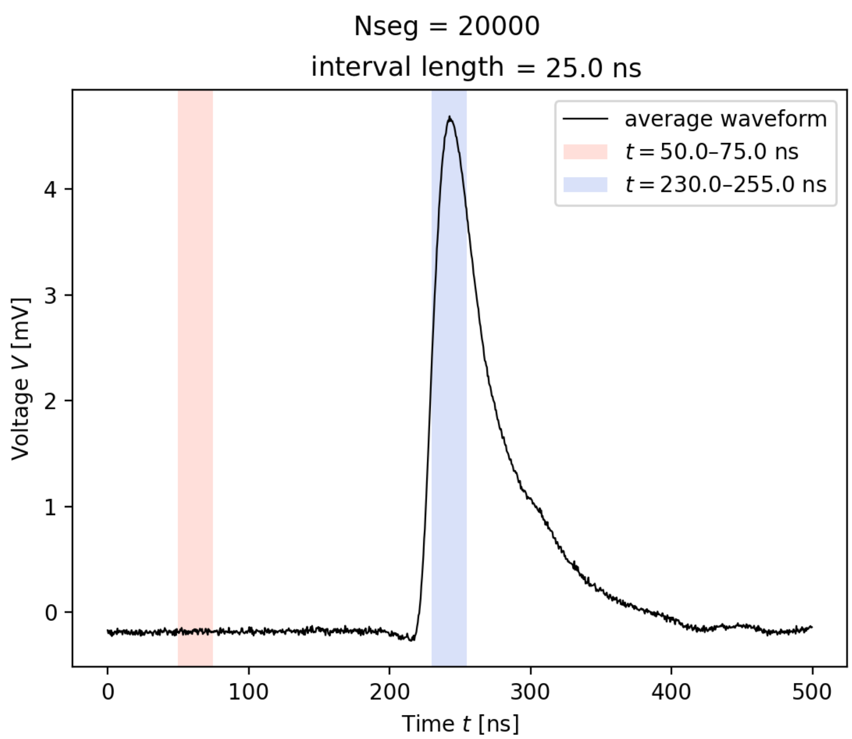


Figure 6.19: Average waveform of 20 000 waveform segments low-light-level measurements. The two vertical spans mark the pre-trigger interval (*red*) and the maximum interval (*blue*), respectively.

delineated by the red vertical span in Fig. 6.19. To prevent including afterpulses or interference from falling tails in the right half of the segment, the pre-trigger interval should be confined to the left half of the segment ($t < 250$ ns). Furthermore, it should exclude the rising edge of the pulse. In our experiment, the pre-trigger interval spans from $t = 50$ ns to $t = 75$ ns.

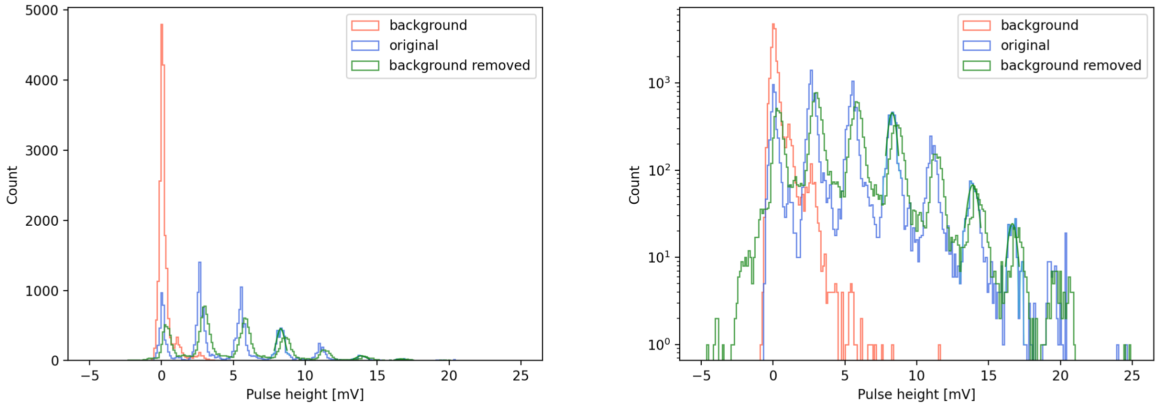
2. **Maximum interval.** The second interval, referred to as the *maximum interval*, is indicated by the blue vertical span in Fig. 6.19. In this interval, the sampling of pulse height occurs, necessitating its coverage of the waveform's maximum point. In our experiment, the pre-trigger interval spans from $t = 230$ ns to $t = 255$ ns. To avoid capturing the afterpulses, it is advisable to restrict the right boundary of the maximum interval to be less than $t = 270$ ns.

Step 3: Deduct the background. As a correction for the pulse height, the background is calculated and then deducted from the amplitudes in the maximum interval.

Step 4: Plot the PHS. After background removal, the PHS can be obtained using the same method as previously described. Fig. 6.20 depicts the resulting PHS , which includes a histogram of the pre-trigger background. A comparison between the PHS with and without background removal is shown to highlight the impact of this process.

It is intriguing to observe that both the background-removed PHS and the original PHS in Fig. 6.20 exhibit a characteristic "finger plot" shape. Notably, the photoelectron peaks in the two PHS can align, suggesting consistency in photon counting. Conversely, the histogram of the pre-trigger background reveals a very significant pedestal at the 0 mV amplitude, meeting the requirements for the pre-trigger interval. The peak-to-trough ratio of the background-removed PHS decreases compared to the original PHS, likely due to error propagation resulting from subtracting the background value from the original amplitude.

However, we made an unexpected discovery while examining the background-removed PHS. It appeared to be slightly shifted to the *right* in comparison to the original counterpart, which was contrary to what we had anticipated. Upon closer scrutiny of Fig. 6.19, it becomes apparent that the amplitude of the pre-trigger interval dips into negative values. While the pre-trigger interval fluctuates across different waveform segments, ideally, it should average to zero. The discrepancy from zero may be attributed to the DC coupling of the oscilloscope. Subtracting negative background values causes the PHS to shift to the right. Subtracting background values that are mostly negative leads to an average rightward shift in the PHS.



(a) in the linear scale

(b) in the logarithmic scale

Figure 6.20: Pulse height spectrum (PHS) with and without background removal, and pre-trigger background histogram.

Consequently, it appears that our background-removal algorithm may benefit from further refinement.

A strategy to assessing the performance of the algorithm involves shifting the focus from the displacement of the PHS to the consistency of the distance between the photoelectron peaks. If consistency is observed, pulse heights can be standardized in unit of p.e., which should not cause adverse effects on subsequent fitting processes. In case of inconsistencies, rigorous evaluation of the algorithm's effectiveness in eliminating baseline shift becomes necessary.

6.3.3 Fitting of pulse height spectrum

While the pulse height data remain uncorrected, in the final phase of the dynamic characterization, we proceed with a fitting procedure on the background-removed PHS. The conventional method involves the fitting of multiplie Gaussian functions to individual photoelectron peaks and the fitting of Poisson[11] distribution to the peak apses. For instance, previous studies have suggested that the following function can provide a general fitting model for PHS to first approximation[93]:

[11] Assuming that prompt crosstalk obeys the Borel distribution, the photoelectron peak height distribution should obey the generalized Poisson distribution.[138]

$$f(A) = C \cdot \sum_{k=0}^{N} \text{Poisson}(k;\mu) \left[\frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{ - \left(\frac{A - k \cdot A_{1\,\text{pe}}}{\sqrt{2}\sigma_k} \right)^2 \right\} \right], \tag{6.9}$$

where C is the normalizing factor, σ_k is the peak width defined in Eq. (4.22), and the explicit formula of Poisson($k;\mu$) is given in Eq. (4.21).

However, as illustrated more distinctly in Fig. 6.21, there is a noticeable presence of background occurrences between the photoelectron peaks that defy accurate modeling by a Gaussian distribution. These inter-peak occurrences stem from various sources of noise, including electrical interference in the readout circuit, sporadic primary dark pulses, and correlated noise. Therefore, it is crucial to carefully select bin intervals for Gaussian fitting.

One feasible method for bin selection involves identifying peaks in the PHS and then fitting the Gaussian function to the bin interval centered at each peak position. In our analysis, pulse criteria and fitting intervals were determined based on rules of thumb. While performing a Fourier transform on the PHS to determine peak-to-peak distances could aid in peak identification, this step was not executed in our analysis.

If we do not rely on excluding inter-peak background events for fitting, but insist on using a single function for fitting, we must accurately understand the origin and nature of the background events, see [138]. Recently, a Python program called PeakOTron based on [138]'s work was developed specifically for fitting the charge spectrum of SiPMs[169]. It may be prudent to consider attempting this program in the future studies, especially in the large-scale characterization for SiPM.

After careful consideration, we have decided to temporarily employ the following steps for performing PHS fitting.

1. Find the peaks in the histogram of pulse heights with `scipy.find_peaks` and manually tuned criteria. Six photoelectron peaks were foud in our case.

2. Fit the Gaussian distribution to the surroundings of each photoelectron peak.

3. Fit the Poisson distribution to the maximal points of Gaussian fits.

The results of these steps are depicted in Fig. 6.21. The mean μ in from the Poisson distribution signifies the expected number of detected photons, with our obtained value

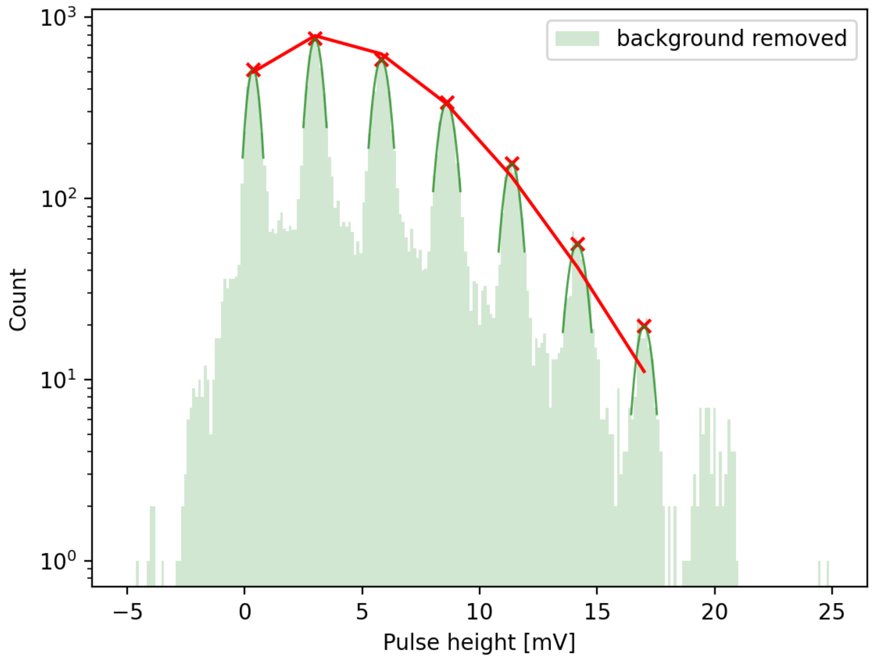


Figure 6.21: Pulse height spectrum (PHS) and its Poissonian fitting. The green histogram represents the background-removed PHS, while the green curves denote Gaussian fits to the photoelectron peak. The red crosses indicate peak apices for the Poissonian fitting, and the red broken line represents the Poissonian fit.

being 1.58. We will not delve into the deviation between the fitting and the data in Fig. 6.21, because we knew in the previous section that the pulse height may not be well corrected.

Chapter 7

Conclusion

In this study, the concepts of very high-energy gamma ray (VHEGR), imaging atmospheric Cherenkov technique (IACT), and silicon photomultiplier (SiPM) are reviewed, and the characteristics of Hamamatsu's MPPC S13360-6050CS were measured. Static characterizations reveal a breakdown voltage V_{BD} of approximately 52 V at room temperature, exhibiting a linear relationship with temperature, with a temperature correction coefficient of $52\,\mathrm{mV\,^{\circ}C^{-1}}$, aligning closely with supplier specifications. Dynamic characterizations unveil performance parameters such as gain, dark count rate (DCR), and crosstalk probability. Despite nice linearity with respect to overvoltage, the measured gain and crosstalk probability deviates from the reference values on the data sheet. The gain is about 40% lower, while the crosstalk probability is about 100% higher than the reference values. Algorithm optimization or an evaluation of the readout circuit characteristics may be required to explore the reasons behind these inconsistencies in SiPM performance parameters. Dynamic characterization at low light levels was also performed to obtain pulse height spectrum (PHS) and gain experience in experimental operation.

Bibliography

[1] Z. Cao et al. "Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray Galactic sources". In: *Nature* 594 (2021), pp. 33–36. DOI: 10.1038/s41586-021-03498-z.

[2] F. A. Aharonian, L. Bergström, and C. Dermer. *Astrophysics at Very High Energies*. Vol. 40. Saas-Fee Advanced Course. Springer, 2013. ISBN: 978-3-642-36134-0. DOI: 10.1007/978-3-642-36134-0. URL: https://link.springer.com/book/10.1007/978-3-642-36134-0.

[3] D. Bose et al. "Ground-based gamma-ray astronomy: history and development of techniques". In: *The European Physical Journal Special Topics* 231.1 (2022), pp. 3–26. DOI: 10.1140/epjs/s11734-021-00396-3. arXiv: 2201.04719 [astro-ph.IM].

[4] P. K. Grieder. *Extensive Air Showers. High Energy Phenomena and Astrophysical Aspects — A Tutorial, Reference Manual and Data Book*. Vol. 1. Berlin Heidelberg: Springer, 2010. ISBN: 978-3-540-76940-8. DOI: 10.1007/978-3-540-76941-5.

[5] G. Pühlhofer, F. Leuschner, and H. Salzmann. "H.E.S.S.: The High Energy Stereoscopic System". In: *Handbook of X-ray and Gamma-ray Astrophysics*. Ed. by C. Bambi and A. Santangelo. Singapore: Springer, 2023. ISBN: 978-981-16-4544-0. DOI: 10.1007/978-981-16-4544-0_69-1.

[6] C. Baixeras. "The MAGIC telescope". In: *Nuclear Physics B - Proceedings Supplements* 114 (6 2003). Proceedings of the XXXth International Meeting of Fundamentals Physics, pp. 247–252. DOI: 10.1016/S0920-5632(02)01910-2.

[7] R. Mukherjee. "Observing the energetic universe at very high energies with the VERITAS gamma ray observatory". In: *Advances in Space Research* 62.10 (2018), pp. 2828–2844. DOI: 10.1016/j.asr.2018.04.002.

[8] G. Hermann. “Towards the future Cherenkov telescope array CTA”. In: *Nucl. Instrum. Methods Phys. Res. A* 623 (1 2010), pp. 408–409. DOI: 10.1016/j.nima.2010.03.018.

[9] H. Iwasaki. “Fundamental development of SiPM photodetector camera for performance improvement of atmospheric Cherenkov telescopes”. Master’s thesis. Kyoto University, 2023. URL: http://www-cr.scphys.kyoto-u.ac.jp/m_ron/iwasaki_M.pdf.

[10] D. Guberman and R. Paoletti. “Silicon photomultipliers in Very High Energy gamma-ray astrophysics”. In: *Journal of Instrumentation* 15, C05039 (2020). International Summer School on INtelligent signal processing for FrontIEr Research and Industry (INFIERI 2019). DOI: 10.1088/1748-0221/15/05/C05039.

[11] The CTA Consortium. *Science with the Cherenkov Telescope Array*. Singapore: World Scientific, 2019. ISBN: 978-981-327-008-4. DOI: 10.1142/10986.

[12] D. Bose et al. “Galactic and extragalactic sources of very high energy gamma rays”. In: *The European Physical Journal Special Topics* 231.1 (2022), pp. 27–66. DOI: 10.1140/epjs/s11734-022-00434-8. arXiv: 2201.06789 [astro-ph.HE].

[13] T. Kifune. “Ground-based gamma-ray astronomy: General remarks”. In: *Il Nuovo Cimento C* 19.6 (1996), pp. 953–957. DOI: 10.1007/BF02508136.

[14] Y. Fukui et al. “Pursuing the Origin of the Gamma Rays in RX J1713.7-3946 Quantifying the Hadronic and Leptonic Components”. In: *The Astrophysical Journal* 915.2 (2021), p. 84. DOI: 10.3847/1538-4357/abff4a.

[15] N. J. Corso, R. Diesing, and D. Caprioli. “Hadronic versus Leptonic Origin of Gamma-Ray Emission from Supernova Remnants”. In: *The Astrophysical Journal* 954.1 (2023), p. 1. DOI: 10.3847/1538-4357/ace699.

[16] A. Albert et al. “HAWC Search for High-mass Microquasars”. In: *The Astrophysical Journal Letters* 912.1, L4 (2021). DOI: 10.3847/2041-8213/abf35a.

[17] H. Netzer. “Revisiting the Unified Model of Active Galactic Nuclei”. In: *Annu. Rev. Astron. Astrophys.* 53.1 (2015), pp. 365–408. DOI: 10.1146/annurev-astro-082214-122302.

[18] M. A. Roth et al. “The diffuse γ-ray background is dominated by star-forming galaxies”. In: *Nature* 597 (2021), pp. 341–344. DOI: 10.1038/s41586-021-03802-x.

[19] M. S. Longair. *High Energy Astrophysics*. 3rd ed. Cambridge: Cambridge University Press, 2011. ISBN: 978-0-521-75618-1. DOI: 10.1017/CBO9780511778346.

[20] F. A. Aharonian. *Very High Energy Cosmic Gamma Radiation*. A Crucial Window on the Extreme Universe. Singapore: World Scientific, 2004. ISBN: 978-981-02-4573-3. DOI: 10.1142/4657. URL: https://www.worldscientific.com/doi/abs/10.1142/4657.

[21] G. Ghisellini. "Synchrotron Self-Compton". In: *Radiative Processes in High Energy Astrophysics*. Heidelberg: Springer, 2013, pp. 89–93. ISBN: 978-3-319-00611-6. DOI: 10.1007/978-3-319-00612-3_6.

[22] S. D. Bloom and A. P. Marscher. "An Analysis of the Synchrotron Self-Compton Model for the Multi–Wave Band Spectra of Blazars". In: *The Astrophysical Journal* 461 (1996), p. 657. DOI: 10.1086/177092.

[23] D. J. Thompson. "Gamma ray astrophysics: the EGRET results". In: *Reports on Progress in Physics* 71, 116901 (2008). DOI: 10.1088/0034-4885/71/11/116901.

[24] W. Atwood et al. "Design and initial tests of the Tracker-converter of the Gamma-ray Large Area Space Telescope". In: *Astroparticle Physics* 28.4 (2007), pp. 422–434. DOI: 10.1016/j.astropartphys.2007.08.010.

[25] R. C. Hartman et al. "The Third EGRET Catalog of High-Energy Gamma-Ray Sources". In: *The Astrophysical Journal Supplement Series* 123.1 (1999), pp. 79–202. DOI: 10.1086/313231.

[26] J.-M. Casandjian and I. A. Grenier. "A revised catalogue of EGRET γ-ray sources". In: *Astronomy & Astrophysics* 489.2 (2008), pp. 849–883. DOI: 10.1051/0004-6361:200809685. arXiv: 0806.0113 [astro-ph].

[27] P. Auger et al. "Extensive Cosmic-Ray Showers". In: *Rev. Mod. Phys.* 11 (3-4 1939), pp. 288–291. DOI: 10.1103/RevModPhys.11.288.

[28] W. Galbraith and J. V. Jelley. "Light Pulses from the Night Sky associated with Cosmic Rays". In: *Nature* 171.4347 (1953), pp. 349–350. DOI: 10.1038/171349a0.

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

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

[31] T. K. Gaisser, R. Engel, and E. Resconi. *Cosmic Rays and Particle Physics*. 2nd ed. Cambridge: Cambridge University Press, 2016. ISBN: 9781139192194. DOI: 10.1017/CBO9781139192194.

[32] H. J. Völk and K. Bernlöhr. "Imaging very high energy gamma-ray telescopes". In: *Experimental Astronomy* 25.1 (2009), pp. 173–191. DOI: 10.1007/s10686-009-9151-z.

[33] M. Amenomori, X. J. Bi, and D. Chen, The Tibet ASγ Collaboration. "Tibet air shower array: results and future plan". In: *Journal of Physics: Conference Series* 120.6, 062024 (2008). DOI: 10.1088/1742-6596/120/6/062024.

[34] H. He, The LHAASO Collaboration. "Design of the LHAASO detectors". In: *Radiation Detection Technology and Methods* 2 (2018), p. 7. DOI: 10.1007/s41605-018-0037-3.

[35] J. Hinton, SWGO Collaboration. "The Southern Wide-field Gamma-ray Observatory: Status and Prospects". In: 23. DOI: 10.22323/1.395.0023. arXiv: 2111.13158 [astro-ph.IM].

[36] J. Goodman and P. Huentemeyer. "The High-Altitude Water Cherenkov Detector Array: HAWC". In: *Handbook of X-ray and Gamma-ray Astrophysics*. Ed. by C. Bambi and A. Santangelo. Singapore: Springer, 2024, pp. 2607–2632. ISBN: 978-981-16-4544-0. DOI: 10.1007/978-981-19-6960-7_65.

[37] D. Morcuende et al. "Relevance of the fluorescence radiation in VHE gamma-ray observations with the Cherenkov technique". In: *Astroparticle Physics* 107 (2019), pp. 26–34. DOI: 10.1016/j.astropartphys.2018.11.003.

[38] M. Tluczykont et al. "HiSCORE: A new detector for astroparticle and particle physics beyond 10 TeV". In: *Nucl. Instrum. Methods Phys. Res. A* 692 (2012). 3rd Roma International Conference on Astroparticle Physics, pp. 246–249. DOI: 10.1016/j.nima.2011.12.075.

[39] P. A. Cherenkov. "Visible luminescence of pure liquids under the influence of γ-radiation". In: *Doklady Akademii Nauk SSSR* 2.8 (1934), pp. 451–454. DOI: 10.3367/UFNr.0093.196710n.0385.

[40] E. P. Cherenkova. "The discovery of the Cherenkov radiation". In: *Nucl. Instrum. Methods Phys. Res. A* 595.1 (2008). Part of Proceedings, 6th International Workshop on Ring Imaging Cherenkov Detectors (RICH 2007), pp. 8–11. DOI: 10.1016/j.nima.2008.07.006.

[41] I. E. Frank and I. M. Tamm. "Coherent visible radiation of fast electrons passing through matter". In: *Doklady Akademii Nauk SSSR* 14.3 (1937), pp. 109–114. DOI: 10.3367/UFNr.0093.196710o.0388.

[42] M. Khurana et al. "Effects of the moonlight on the operating parameters of the MACE γ-ray telescope: a feasibility study". In: *Journal of Astrophysics and Astronomy* 43.12 (2022). DOI: 10.1007/s12036-021-09792-6.

[43] M. Errando and T. Saito. "How to Detect Gamma Rays from Ground: An Introduction to the Detection Concepts". In: *Handbook of X-ray and Gamma-ray Astrophysics*. Ed. by C. Bambi and A. Santangelo. Singapore: Springer, 2023. ISBN: 978-981-16-4544-0. DOI: 10.1007/978-981-16-4544-0_61-1. arXiv: 2401.04460 [astro-ph.HE].

[44] K. Bernlöhr. "Impact of atmospheric parameters on the atmospheric Cherenkov technique". In: *Astroparticle Physics* 12.4 (2000), pp. 255–268. DOI: 10.1016/S0927-6505(99)00093-6.

[45] B. Bi. "Research of the Application of Silicon Photomultiplier on the Wide Field of View Cherenkov Telescope Array of LHAASO". PhD thesis. Chinese Academy of Sciences, 2018.

[46] S. Dorota et al. "An analysis method for data taken by Imaging Air Cherenkov Telescopes at very high energies under the presence of clouds". In: *Astroparticle Physics* 120, 102450 (2020). DOI: 10.1016/j.astropartphys.2020.102450.

[47] S. Sajjad. "The effect of night sky background on source reconstruction in atmospheric Cherenkov telescope arrays and implications for image cleaning". In: *New Journal of Physics* 25, 083038 (2023). DOI: 10.1088/1367-2630/acecd9.

[48] M. L. Knoetig et al. "FACT — Long-term stability and observations during strong Moon light". In: *Proceedings of 33rd International Cosmic Ray Conference (ICRC2013)* (Rio de Janeiro, Brazil, July 2–9, 2013). 2013, 695. arXiv: 1307.6116 [astro-ph.IM].

[49] S. Masuda et al., for the CTA Consortium. "Development of the photomultiplier tube readout system for the first Large-Sized Telescope of the Cherenkov Telescope Array". In: *Proceedings of the 34th International Cosmic Ray Conference (ICRC2015)* (The Hague, The Netherlands, July 30–6, 2015). Ed. by A. van den Berg. 2016, PoS(ICRC2015)1003. DOI: 10.22323/1.236.1003. arXiv: 1509.00548 [astro-ph.IM].

[50] T. Saito, CTA LST project. “Characterization of SiPM and development of test bench modules for the next-generation cameras for Large-Sized Telescopes for Cherenkov Telescope Array”. In: *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)* (Nagoya, Japan, July 26–August 3, 2023). Ed. by T. Saito and K. Okumura. 2023, PoS(ICRC2023)635. DOI: 10.22323/1.444.0635.

[51] A. Bouvier et al. “Photosensor characterization for the Cherenkov Telescope Array: silicon photomultiplier versus multi-anode photomultiplier tube”. In: *Hard X-Ray, Gamma-Ray, and Neutron Detector Physics XV*. Ed. by M. Fiederle et al. Vol. 8852. International Society for Optics and Photonics (SPIE), 2013, 88520K. DOI: 10.1117/12.2023778. arXiv: 1308.1390 [astro-ph.IM].

[52] M. Doering et al. “Measurement of the Cherenkov light spectrum and of the polarization with the HEGRA-IACT-system”. In: *Proceedings of of the 27th International Cosmic Ray Conference (ICRC 2001)* (Hamburg, Germany, August 7–15, 2001). 2985. Copernicus Gesellschaft, 2001. arXiv: astro-ph/0107149 [astro-ph].

[53] C. Benn and S. Ellison. “Brightness of the night sky over La Palma”. In: *New Astronomy Reviews* 42 (6 1998), pp. 503–507. DOI: 10.1016/S1387-6473(98)00062-1.

[54] G. Pareschi et al. “Status of the technologies for the production of the Cherenkov Telescope Array (CTA) mirrors”. In: *Optics for EUV, X-Ray, and Gamma-Ray Astronomy VI* (San Diego, California, United States, September 26, 2013). Ed. by S. L. O’Dell and G. Pareschi. Vol. 8861. International Society for Optics and Photonics (SPIE), 2013, 886103. DOI: 10.1117/12.2025476. arXiv: 1310.1713 [astro-ph.IM].

[55] S. Diebold et al., CTA Consortium. “Actuator development at IAAT for the Cherenkov Telescope Array Medium Size Telescopes”. In: *AIP Conference Proceedings* 1792.1, 080014 (2017). The 6th International Symposium on High Energy Gamma-Ray Astronomy. DOI: 10.1063/1.4969035. arXiv: 1610.01363 [astro-ph.IM].

[56] J. M. Davies and E. S. Cotton. “Design of the quartermaster solar furnace”. In: *Solar Energy* 1 (2 1957), pp. 16–22. DOI: 10.1016/0038-092X(57)90116-0.

[57] K. Tayabaly et al. “Roughness tolerances for Cherenkov telescope mirrors”. In: *Optics for EUV, X-Ray, and Gamma-Ray Astronomy VII*. Ed. by S. L. O’Dell and G. Pareschi. Vol. 9603. International Society for Optics and Photonics (SPIE). SPIE Press, 2015, 960307. DOI: 10.1117/12.2187025. arXiv: 1610.00723 [astro-ph.IM].

[58] V. Vassiliev, S. Fegan, and P. Brousseau. “Wide field aplanatic two-mirror telescopes for ground-based γ-ray astronomy”. In: *Astroparticle Physics* 28 (1 2007), pp. 10–27. DOI: https://doi.org/10.1016/j.astropartphys.2007.04.002. arXiv: 0612718 [astro-ph]. URL: https://www.sciencedirect.com/science/article/pii/S0927650507000527.

[59] F. R. Pantaleo, CTA SCT project. “The prototype Schwarzschild Couder Telescope: a Medium-Sized Telescope for the Cherenkov Telescope Array”. In: *Journal of Physics: Conference Series* 2429, 012021 (2023). 12th Cosmic Ray International Seminar (CRIS 2022). DOI: 10.1088/1742-6596/2429/1/012021.

[60] F. Di Pierro, CTA-LST Project. “Status of the Large-Sized Telescope of the Cherenkov Telescope Array”. In: *Journal of Physics: Conference Series* 2429.1, 012020 (2023). DOI: 10.1088/1742-6596/2429/1/012020.

[61] Hamamatsu Photonics K.K. *Photomultiplier tubes. Basics and Applications*. 4th ed. 2017. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/PMT_handbook_v4E.pdf.

[62] S. Sailor. “The FlashCam camera for CTA: trigger verification and fluorescence light detection capabilities”. Heidelberg University, 2020. URL: https://archiv.ub.uni-heidelberg.de/volltextserver/29105/.

[63] A. Gadola. “Towards the first imaging atmospheric cherenkov telescope camera with continuous signal digitization”. PhD thesis. University of Zurich, 2013. DOI: 10.5167/uzh-85743. URL: https://www.zora.uzh.ch/id/eprint/85743/.

[64] K. Bernlöhr et al. “The optical system of the H.E.S.S. imaging atmospheric Cherenkov telescopes. Part I: layout and components of the system”. In: *Astroparticle Physics* 20 (2 2003), pp. 111–128. DOI: 10.1016/S0927-6505(03)00171-3.

[65] A. Okumura. “Optimization of the collection efficiency of a hexagonal light collector using quadratic and cubic Bézier curves”. In: *Astroparticle Physics* 38 (2012), pp. 18–24. DOI: 10.1016/j.astropartphys.2012.08.008.

[66] C. Alispach et al. “Large scale characterization and calibration strategy of a SiPM-based camera for gamma-ray astronomy”. In: *Journal of Instrumentation* 15, P11010 (2020). DOI: 10.1088/1748-0221/15/11/p11010. arXiv: 2008.04716 [astro-ph.IM].

[67] D. Depaoli, CTA SST project. "Status of the SST Camera for the Cherenkov Telescope Array". In: *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)* (Nagoya, Japan, July 26–August 3, 2023). Ed. by T. Saito and K. Okumura. 2023, PoS(ICRC2023)771. DOI: 10.22323/1.444.0771. arXiv: 2310.01183 [astro-ph.IM].

[68] G. Pühlhofer et al., CTA Consortium. "FlashCam: A fully digital camera for CTA telescopes". In: *AIP Conference Proceedings* 1505.1 (2012), pp. 777–780. DOI: 10.1063/1.4772375. arXiv: 1211.3684 [astro-ph.IM].

[69] G. Hermann et al. *Medium-Sized Telescope FlashCam Technical Design Report*. Version 1.20. FlashCam Team, 2016.

[70] G. Lamanna et al. *Cherenkov Telescope Array Data Management*. 2015. arXiv: 1509.01012 [astro-ph.IM].

[71] D. Malyshev and L. Mohrmann. "Analysis Methods for Gamma-Ray Astronomy". In: *Handbook of X-ray and Gamma-ray Astrophysics*. Ed. by C. Bambi and A. Santangelo. Singapore: Springer, 2023. ISBN: 978-981-16-4544-0. DOI: 10.1007/978-981-16-4544-0_177-1.

[72] M. Gaug et al. "Using Muon Rings for the Calibration of the Cherenkov Telescope Array: A Systematic Review of the Method and Its Potential Accuracy". In: *The Astrophysical Journal Supplement Series* 243.1 (2019), p. 11. DOI: 10.3847/1538-4365/ab2123. arXiv: 1907.04375 [astro-ph.IM].

[73] R. M. Wagner. "Measurement of Very High Energy Gamma--Ray Emission from Four Blazars Using the MAGIC Telescope and a Comparative Blazar Study". PhD thesis. Technical University of Munich and Max Planck Institute for Physics, 2006. URL: https://mediatum.ub.tum.de/doc/604099/.

[74] A. M. Hillas. "Cerenkov light images of EAS produced by primary gamma". In: *Proceedings from the 19th International Cosmic Ray Conference (ICRC19)*. Vol. 3. 1985, pp. 445–448.

[75] W. Hofmann et al. "Comparison of techniques to reconstruct VHE gamma-ray showers from multiple stereoscopic Cherenkov images". In: *Astroparticle Physics* 12.3 (1999), pp. 135–143. DOI: 10.1016/S0927-6505(99)00084-5.

[76] M. T. Rissi. "Detection of pulsed very high energy Gamma-rays from the Crab Pulsar with the MAGIC telescope using an analog sum trigger". Doctoral Thesis. ETH Zurich, 2009. DOI: 10.3929/ethz-a-005901896. URL: https://www.research-collection.ethz.ch/handle/20.500.11850/151444.

[77] M. de Naurois. "Analysis methods for Atmospheric Cerenkov Telescopes". In: *7th Workshop on Towards a Network of Atmospheric Cherenkov Detectors 2005*. 2006, pp. 149–161. DOI: 10.48550/arXiv.astro-ph/0607247.

[78] M. de Naurois and L. Rolland. "A high performance likelihood reconstruction of γ-rays for imaging atmospheric Cherenkov telescopes". In: *Astroparticle Physics* 32.5 (2009), pp. 231–252. DOI: 10.1016/j.astropartphys.2009.09.001.

[79] R. Parsons and J. Hinton. "A Monte Carlo template based analysis for air-Cherenkov arrays". In: *Astroparticle Physics* 56 (2014), pp. 26–34. DOI: https://doi.org/10.1016/j.astropartphys.2014.03.002.

[80] J. Aschersleben et al., the CTA Consortium and the CTA Observatory. "Application of pattern spectra and convolutional neural networks to the analysis of simulated Cherenkov Telescope Array data". In: *Proceedings of the 37th International Cosmic Ray Conference (ICRC2021)* (Berlin, Germany (Online), July 12–23, 2021). Ed. by A. Kappes and B. Keilhauer. Sissa Medialab, 2022, PoS(ICRC2021)697. DOI: 10.22323/1.395.0697. arXiv: 2108.00834 [astro-ph.IM].

[81] A. Donath et al. "Gammapy: A Python package for gamma-ray astronomy". In: *Astronomy & Astrophysics* 678, A157 (2023). DOI: 10.1051/0004-6361/202346488. arXiv: 2308.13584 [astro-ph.IM].

[82] B. Acharya, M. Actis, and T. Aghajani. "Introducing the CTA concept". In: *Astroparticle Physics* 43 (2013). Seeing the High-Energy Universe with the Cherenkov Telescope Array - The Science Explored with the CTA, pp. 3–18. DOI: 10.1016/j.astropartphys.2013.01.007.

[83] Cherenkov Telescope Array Observatory. *CTAO Releases the Layout of its Arrays for the Alpha Configuration*. Press Release. 2022. URL: https://www.ctao.org/news/stefano-stanghellini-joins-ctao-as-telescope-coordinator/ (visited on 05/20/2024).

[84] W. Hofmann and R. Zanin. "The Cherenkov Telescope Array". In: *Handbook of X-ray and Gamma-ray Astrophysics*. Ed. by C. Bambi and A. Santangelo. Singapore: Springer, 2023. ISBN: 978-981-16-4544-0. DOI: 10.1007/978-981-16-4544-0_70-1.

[85] Cherenkov Telescope Array Observatory. *CTAO Releases the Layout of its Arrays for the Alpha Configuration*. Press Release. 2022. URL: https://www.ctao.org/news/ctao-releases-layouts-for-alpha-configuration/ (visited on 05/20/2024).

[86] O. Gueta, the CTA Consortium and the CTA Observatory. "The Cherenkov Telescope Array: layout, design and performance". In: *Proceedings of the 37th International Cosmic Ray Conference (ICRC2021)* (Berlin, Germany (Online), July 12–23, 2021). Ed. by A. Kappes and B. Keilhauer. Sissa Medialab, 2022, PoS(ICRC2021)885. DOI: 10.22323/1.395.0885. arXiv: 2108.04512 [astro-ph.IM].

[87] Cherenkov Telescope Array Observatory. *CTAO Performance*. 2024. URL: https://www.ctao.org/for-scientists/performance/ (visited on 04/19/2024).

[88] V. Vassiliev. "Extragalactic background light absorption signal in the TeV γ-ray spectra of blazars". In: *Astroparticle Physics* 12 (4 2000), pp. 217–238. DOI: 10.1016/S0927-6505(99)00107-3.

[89] S. Vercellone et al. "ASTRI Mini-Array core science at the Observatorio del Teide". In: *Journal of High Energy Astrophysics* 35 (2022), pp. 1–42. DOI: 10.1016/j.jheap.2022.05.005.

[90] J. Zorn, CTA SST Collaboration. "CHEC — A compact high energy camera for the Cherenkov Telescope Array". In: *Nucl. Instrum. Methods Phys. Res. A* 936 (2019). Frontier Detectors for Frontier Physics: 14th Pisa Meeting on Advanced Detectors, pp. 229–230. DOI: 10.1016/j.nima.2018.09.138.

[91] J. Lapington, CTA SST Collaboration. "The silicon photomultiplier-based camera for the Cherenkov Telescope Array small-sized telescopes". In: *Nucl. Instrum. Methods Phys. Res. A* 1055, 168433 (2023). DOI: 10.1016/j.nima.2023.168433.

[92] J. Aguilar et al. "Characterization and commissioning of the SST-1M camera for the Cherenkov Telescope Array". In: *Nucl. Instrum. Methods Phys. Res. A* 845 (2017), pp. 350–354. DOI: 10.1016/j.nima.2016.05.130.

[93] M. Heller et al. "An innovative silicon photomultiplier digitizing camera for gamma-ray astronomy". In: *European Physical Journal C* 77.47 (2017). DOI: 10.1140/epjc/s10052-017-4609-z.

[94] J.-L. Dournaux et al. "Performance of the Mechanical Structure of the SST-2M GCT Proposed for the Cherenkov Telescope Array". In: PoS(ICRC2015)946 (2016). Ed. by A. van den Berg. DOI: 10.22323/1.236.0946. arXiv: 1508.06415 [astro-ph.IM].

[95] T. Tavernier et al. “Analysis of commissioning data from SST-1M : A Prototype of Single-Mirror Small Size Telescope”. In: *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)* (Nagoya, Japan, July 26–August 3, 2023). Ed. by T. Saito and K. Okumura. Vol. 444. 2023, PoS(ICRC2023)741, p. 741. DOI: 10.22323/1.444.0741.

[96] T. Tavernier, J.-F. Glicenstein, and F. Brun. “Status and performance results from NectarCam — a camera for CTA medium sized telescopes”. In: *Proceedings of the 36th International Cosmic Ray Conference (ICRC2019)* (Madison, WI, U.S.A. July 24–August 1, 2019). Ed. by P. Desiati, T. Gaisser, and A. Karle. 2021, PoS(ICRC2019)805. DOI: 10.22323/1.358.0805.

[97] F. Bradascio. “Status of the Medium-Sized Telescopes for the Cherenkov Telescope Array Observatory”. In: *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)* (Nagoya, Japan, July 26–August 3, 2023). Ed. by T. Saito and K. Okumura. 2023, PoS(ICRC2023)859. DOI: 10.22323/1.444.0859. arXiv: 2310.02127 [astro-ph.IM].

[98] G. Pühlhofer et al., H.E.S.S. Collaboration. “Science verification of the new FlashCam-based camera in the 28m telescope of H.E.S.S.” In: *Proceedings of the 37th International Cosmic Ray Conference (ICRC2021)* (Berlin, Germany (Online), July 12–23, 2021). Ed. by A. Kappes and B. Keilhauer. Sissa Medialab, 2022, PoS(ICRC2021)764. DOI: 10.22323/1.395.0764. arXiv: 2108.02596 [astro-ph.IM].

[99] C. B. Adams, G. Ambrosi, and M. Ambrosio. “Design and performance of the prototype Schwarzschild–Couder telescope camera”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 8, 014007 (1 2022). DOI: 10.1117/1.JATIS.8.1.014007. arXiv: 2203.08169 [astro-ph.IM].

[100] C. Adams et al. “Detection of the Crab Nebula with the 9.7 m prototype Schwarzschild–Couder telescope”. In: *Astroparticle Physics* 128, 102562 (2021). DOI: https://doi.org/10.1016/j.astropartphys.2021.102562. arXiv: 2012.08448 [astro-ph.IM].

[101] J. Sitarek et al. “Analysis techniques and performance of the Domino Ring Sampler version 4 based readout for the MAGIC telescopes”. In: *Nucl. Instrum. Methods Phys. Res. A* 723 (2013), pp. 109–120. DOI: 10.1016/j.nima.2013.05.014. arXiv: 1305.1007 [astro-ph.IM].

[102] S. Abe et al., CTA-LST project. “Multiwavelength study of the galactic PeVatron candidate LHAASO J2108+5157”. In: *Astronomy & Astrophysics* 673, A75 (2023). DOI: 10.1051/0004-6361/202245086. arXiv: 2210.00775 [astro-ph.HE].

[103] Cherenkov Telescope Array Observatory. *LST-1 Discovers the Most Distant AGN at Very High Energies*. Press Release. 2023. URL: https://www.cta-observatory.org/lst1-makes-first-scientific-discovery-by-detecting-the-most-distance-agn/ (visited on 01/10/2024).

[104] Cherenkov Telescope Array Observatory. *Technical Specifications*. 2024. URL: https://www.ctao.org/for-scientists/technical-specifications/ (visited on 04/19/2024).

[105] R. White et al., the CTA SST Collaboration. "The Small-Sized Telescopes for the Southern Site of the Cherenkov Telescope Array". In: *Proceedings of the 37th International Cosmic Ray Conference (ICRC2021)* (Berlin, Germany (Online), July 12–23, 2021). Ed. by A. Kappes and B. Keilhauer. Sissa Medialab, 2022, PoS(ICRC2021)728. DOI: 10.22323/1.395.0728. arXiv: 2110.14527 [astro-ph.IM].

[106] W. R. Leo. *Techniques for Nuclear and Particle Physics Experiments. A How to Approach*. 2nd ed. Berlin/Heidelberg: Springer, 1994. ISBN: 978-3-540-57280-0. DOI: 10.1007/978-3-642-57920-2.

[107] K. W. Böer and U. W. Pohl. "Carrier Generation". In: *Semiconductor Physics*. Cham: Springer, 2023, pp. 1205–1239. ISBN: 978-3-031-18286-0. DOI: 10.1007/978-3-031-18286-0_29.

[108] A. Nagai et al. "Characterization of a large area silicon photomultiplier". In: *Nucl. Instrum. Methods Phys. Res. A* 948, 162796 (2019). DOI: 10.1016/j.nima.2019.162796. arXiv: 1810.02275 [physics.ins-det].

[109] S. M. Sze and M. K. Lee. *Semiconductor Devices. Physics and Technology*. 3rd ed. John Wiley & Sons, 2012. ISBN: 978-0470-53794-7.

[110] S. Cova et al. "Avalanche photodiodes and quenching circuits for single-photon detection". In: *Applied Optics* 35.12 (1996), pp. 1956–1976. DOI: 10.1364/AO.35.001956.

[111] S. Tisa et al. "Electronics for single photon avalanche diode arrays". In: *Sensors and Actuators A: Physical* 140 (1 2007), pp. 113–122. DOI: 10.1016/j.sna.2007.06.022.

[112] D. A. Neamen. *Microelectronics. Circuit Analysis and Design*. 4th ed. New York: McGraw-Hill, 2007. ISBN: 978-0-07-338064-3.

[113] F. Acerbi and S. Gundacker. "Understanding and simulating SiPMs". In: *Nucl. Instrum. Methods Phys. Res. A* 926 (2019), pp. 16–35. DOI: 10.1016/j.nima.2018.11.118.

[114] C. L. Anderson and C. R. Crowell. "Threshold Energies for Electron-Hole Pair Production by Impact Ionization in Semiconductors". In: *Physical Review B* 5 (6 1972), pp. 2267–2272. DOI: 10.1103/PhysRevB.5.2267.

[115] S. Gundacker and A. Heering. "The silicon photomultiplier: fundamentals and applications of a modern solid-state photon detector". In: *Physics in Medicine & Biology* 65 (2020), 17TR01. DOI: 10.1088/1361-6560/ab7b2d.

[116] *Avalanche Photodiodes. A User's Guide*. PerkinElmer, Inc. 2003. URL: https://www.chem.uci.edu/~unicorn/243/papers/AvalanchePD.pdf. via the course website of "Chem 243: Instrumental Analysis (Winter Quarter 2021)" of the University of California, Irvine (UCI).

[117] C. Grupen and B. Shwartz. *Particle detectors*. 2nd ed. Cambridge: Cambridge University Press, 2008. ISBN: 978-0-521-84006-4. DOI: 10.1017/CBO9780511534966.

[118] H. T. J. Meier. "Design, Characterization and Simulation of Avalanche Photodiodes". PhD thesis. University of Zurich, 2011. URL: https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/152656/eth-2682-02.pdf.

[119] S. M. Sze, Y. Li, and K. K. Ng. *Physics of Semiconductor Devices*. Chinese. Trans. English by H. S. Koo and M. Chen. 4th ed. Vol. 1. Hsinchu: National Yang Ming Chiao Tung University Press, 2022. ISBN: 978-986-5470-28-9.

[120] A. Otte. "The Silicon Photomultiplier - A new device for High Energy Physics, Astroparticle Physics, Industrial and Medical Applications". In: *Proceedings of International Symposium on Detector Development for Particle, Astroparticle and Synchrotron Radiation Experiments (SNIC 2006)* (April 3–6, 2006). Ed. by V. Luth. Vol. C0604032. 0018. Stanford, California, U.S.A., 2006. URL: https://www.slac.stanford.edu/econf/C0604032/papers/0018.pdf.

[121] N. Dinu. "8 - Silicon photomultipliers (SiPM)". In: *Photodetectors. Materials, Devices and Applications*. Ed. by B. Nabet. Woodhead Publishing, 2016, pp. 255–294. ISBN: 978-1-78242-445-1. DOI: 10.1016/B978-1-78242-445-1.00008-7.

[122] R. H. Haitz. "Model for the Electrical Behavior of a Microplasma". In: *Journal of Applied Physics* 35.5 (1964), pp. 1370–1376. DOI: 10.1063/1.1713636.

[123] S. Piatek. *What is an SiPM and how does it work?* URL: https://hub.hamamatsu.com/us/en/technical-notes/mppc-sipms/what-is-an-SiPM-and-how-does-it-work.html (visited on 08/04/2023).

[124] Hamamatsu Photonics K.K. *MPPC® (Multi-Pixel Photon Counter) S13360 series*. MPPCs for precision measurement. Document No. KAPD1052E05. 2022. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/ssd/s13360_series_kapd1052e.pdf.

[125] P. Léna et al. *Observational Astrophysics*. Trans. French by S. Lyle. 3rd ed. Astronomy and Astrophysics Library. Berlin, Heidelberg: Springer, 2012. ISBN: 978-3-642-21814-9. DOI: 10.1007/978-3-642-21815-6. URL: https://link.springer.com/book/10.1007/978-3-642-21815-6.

[126] F. Nagy et al. "A model based DC analysis of SiPM breakdown voltages". In: *Nucl. Instrum. Methods Phys. Res. A* 849 (2017), pp. 55–59. DOI: 10.1016/j.nima.2017.01.002.

[127] R. Pagano. "Operative Parameters of Si Photomultipliers". Dottorato di ricerca. Università di Catania, 2011.

[128] K. G. McKay. "Avalanche Breakdown in Silicon". In: *Physical Review* 94 (4 1954), pp. 877–884. DOI: 10.1103/PhysRev.94.877.

[129] M. Shayduk et al. "Light sensors selection for the Cherenkov Telescope Array: PMT and SiPM". In: *Nucl. Instrum. Methods Phys. Res. A* 695 (2012). New Developments in Photodetection NDIP11, pp. 109–112. DOI: 10.1016/j.nima.2011.12.010.

[130] S. Vinogradov. "Analytical models of probability distribution and excess noise factor of solid state photomultiplier signals with crosstalk". In: *Nucl. Instrum. Methods Phys. Res. A* 695 (2012). 6th International Conference on New Developments in Photodetection (NDIP11), pp. 247–251. DOI: 10.1016/j.nima.2011.11.086.

[131] P. Eckert et al. "Characterisation studies of silicon photomultipliers". In: *Nucl. Instrum. Methods Phys. Res. A* 620.2 (2010), pp. 217–226. DOI: 10.1016/j.nima.2010.03.169.

[132] Hamamatsu Photonics K.K. *MPPC®*. Technical note. 2023. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/ssd/mppc_kapd9008e.pdf (visited on 10/09/2023).

[133] W. Oldham, R. Samuelson, and P. Antognetti. "Triggering phenomena in avalanche diodes". In: *IEEE Transactions on Electron Devices* 19 (9 1972), pp. 1056–1060. DOI: 10.1109/T-ED.1972.17544.

[134] F. Zappa et al. “Principles and features of single-photon avalanche diode arrays”. In: *Sensors and Actuators A: Physical* 140 (1 2007), pp. 103–112. DOI: 10.1016/j.sna.2007.06.021.

[135] R. Bayerlein and I. Fleck. *Characterization of a Silicon-Photomultiplier Using Ultra-Fast Pulsed LED*. Laboratory manual. Department of Physics, University of Siegen. 2021. URL: https://www.hep.physik.uni-siegen.de/teaching/masterlab/manuals/SiPM_Manual.pdf (visited on 11/18/2023).

[136] G. Ambrosi et al. “Measurements and tests on FBK silicon sensors with an otpmized electronic design for a CTA camera”. In: *Proceedings, Science with the New Generation of High Energy Gamma-ray experiments, 10th Workshop (SciNeGHE2014)* (June 4–6, 2014). Lisbon, Portugal: Sissa Medialab, 2015, PoS(Scineghe2014)004. DOI: 10.22323/1.218.0004. arXiv: 1411.5241 [astro-ph.IM].

[137] Broadcom Inc. *AFBR-S4NxxPyy4M Single-Photon Measurements*. Application Note. 2023. URL: https://docs.broadcom.com/doc/AFBR-S4NxxPyy4M-Single-Photon-Measurements (visited on 12/30/2023).

[138] V. Chmill et al. “On the characterisation of SiPMs from pulse-height spectra”. In: *Nucl. Instrum. Methods Phys. Res. A* 854 (2017), pp. 70–81. DOI: 10.1016/j.nima.2017.02.049.

[139] A. Lacaita et al. “On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices”. In: *IEEE Transactions on Electron Devices* 40.3 (1993), pp. 577–582. DOI: 10.1109/16.199363.

[140] A. Gola et al. “SiPM optical crosstalk amplification due to scintillator crystal: effects on timing performance”. In: *Physics in Medicine & Biology* 59.13, 3615 (2014). DOI: 10.1088/0031-9155/59/13/3615.

[141] F. Nagy et al. “Afterpulse and delayed crosstalk analysis on a STMicroelectronics silicon photomultiplier”. In: *Nucl. Instrum. Methods Phys. Res. A* 759 (2014), pp. 44–49. DOI: 10.1016/j.nima.2014.04.045.

[142] L. Gallego et al. “Modeling crosstalk in silicon photomultipliers”. In: *Journal of Instrumentation* 8, P05010 (2013). DOI: 10.1088/1748-0221/8/05/P05010.

[143] J. Rosado et al. “Modeling crosstalk and afterpulsing in silicon photomultipliers”. In: *Nucl. Instrum. Methods Phys. Res. A* 787 (2015). 7th International Conference on New Developments in Photodetection (NDIP14), pp. 153–156. DOI: 10.1016/j.nima.2014.11.080.

[144] M. Caccia, R. Santoro, and G. Stanizzi. "A simple and robust method to study afterpulses in Silicon Photomultipliers". In: *Journal of Instrumentation* 9, T10004 (2014). DOI: 10.1088/1748-0221/9/10/T10004.

[145] C. Piemonte et al. "Development of an automatic procedure for the characterization of silicon photomultipliers". In: *Proceedings, 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC 2012)* (October 27–November 3, 2012). Ed. by B. Yu. Anaheim, California, U.S.A.: IEEE, 2012, pp. 428–432. ISBN: 978-1-4673-2030-6. DOI: 10.1109/NSSMIC.2012.6551141. URL: https://www2.pv.infn.it/~cattaneo/Automatic_Dark_procedure.pdf.

[146] Tektronix, Inc. *How to Write Scripts for Test Script Processing (TSP®)*. Application note. 2018. URL: https://download.tek.com/document/1KW-61540-0_Script_with_TSP_Application_Note_031121.pdf (visited on 07/15/2023).

[147] V. Chmill et al. "Study of the breakdown voltage of SiPMs". In: *Nucl. Instrum. Methods Phys. Res. A* 845 (2017). Proceedings of the Vienna Conference on Instrumentation 2016, pp. 56–59. DOI: 10.1016/j.nima.2016.04.047.

[148] Z. Li et al. "A gain control and stabilization technique for Silicon Photomultipliers in low-light-level applications around room temperature". In: *Nucl. Instrum. Methods Phys. Res. A* 695 (2012). New Developments in Photodetection NDIP11, pp. 222–225. DOI: 10.1016/j.nima.2011.12.037.

[149] G. Bonanno et al. "Characterization Measurements Methodology and Instrumental Set-Up Optimization for New SiPM Detectors—Part I: Electrical Tests". In: *IEEE Sensors Journal* 14.10 (2014), pp. 3557–3566. DOI: 10.1109/JSEN.2014.2328621. URL: https://www.researchgate.net/publication/265130574_Characterization_Measurements_Methodology_and_Instrumental_Set-Up_Optimization_for_New_SiPM_Detectors-Part_I_Electrical_Tests.

[150] M. Simonetta et al. "Test and characterisation of SiPMs for the MEGII high resolution Timing Counter". In: *Nucl. Instrum. Methods Phys. Res. A* 824 (2016). Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors, pp. 145–147. DOI: 10.1016/j.nima.2015.11.023.

[151] E. Garutti et al. *Characterization and X-Ray Damage of Silicon Photomultipliers*. 2015. DOI: 10.22323/1.213.0070.

[152] N. Dinu, A. Nagai, and A. Para. “Breakdown voltage and triggering probability of SiPM from IV curves at different temperatures”. In: *Nucl. Instrum. Methods Phys. Res. A* 845 (2017). Proceedings of the Vienna Conference on Instrumentation 2016, pp. 64–68. DOI: 10.1016/j.nima.2016.05.110.

[153] J. Aguilar et al. “The front-end electronics and slow control of large area SiPM for the SST-1M camera developed for the CTA experiment”. In: *Nucl. Instrum. Methods Phys. Res. A* 830 (2016), pp. 219–232. DOI: 10.1016/j.nima.2016.05.086.

[154] Hamamatsu Photonics K.K. *Driver circuit for MPPC® C12332-01 (Instruction manual)*. Document No. K29-B61647. 2015.

[155] Hamamatsu Photonics K.K. *Driver circuit for MPPC® C12332-02 (Datasheet)*. Cat. No. KACC1287E02. 2022. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/ssd/c12332-02_kacc1287e.pdf (visited on 03/07/2024).

[156] F. Zhu et al. “Design and Test of SiPM Readout Circuit”. In: *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)* (Nagoya, Japan, July 26–August 3, 2023). Ed. by T. Saito and K. Okumura. 2023, PoS(ICRC2023)489. DOI: 10.22323/1.444.0489.

[157] C. H. Nowlin and J. L. Blankenship. “Elimination of Undesirable Undershoot in the Operation and Testing of Nuclear Pulse Amplifiers”. In: *Review of Scientific Instruments* 36.12 (1965), pp. 1830–1839. DOI: 10.1063/1.1719475.

[158] A. Manalaysay. *FlashCam Signal Reconstruction*. 2012. URL: https://cta.physik.uzh.ch/public/presentations/files/flashTiming-manalaysay-amsterdam-2012.pdf (visited on 03/13/2024). CTA Consortium Meeting, Amsterdam: FlashCam Signal Reconstruction.

[159] D. Depaoli et al. “Development of a SiPM Pixel prototype for the Large-Sized Telescope of the Cherenkov Telescope Array”. In: *Nucl. Instrum. Methods Phys. Res. A* 1055, 168521 (2023). DOI: 10.1016/j.nima.2023.168521.

[160] Hamamatsu Photonics K.K. *Bias power supply with built-in high precision temperature compensation for MPPCs®*. Cat. No. KACC1203E10. 2023. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/ssd/c11204-01_kacc1203e.pdf (visited on 03/07/2024).

[161] R. M. Gutierrez et al. "The silicon photomultiplier as a metasystem with designed electronics as metadevice for a new receiver-emitter in visible light communications". In: *Metamaterials, Metadevices, and Metasystems 2015*. Proceedings of SPIE. Ed. by N. Engheta, M. A. Noginov, and N. I. Zheludev. Vol. 9544. International Society for Optics and Photonics (SPIE). SPIE Press, 2015, 95441P. DOI: 10.1117/12.2190161. arXiv: 1508.01924 [physics.ins-det].

[162] Y. Du and F. Retière. "After-pulsing and cross-talk in multi-pixel photon counters". In: *Nucl. Instrum. Methods Phys. Res. A* 596 (3 2008), pp. 396–401. DOI: 10.1016/j.nima.2008.08.130.

[163] M. Nemallapudi et al. "Single photon time resolution of state of the art SiPMs". In: *Journal of Instrumentation* 11, P10016 (2016). DOI: 10.1088/1748-0221/11/10/P10016.

[164] M.-R. Mohammadian-Behbahani and S. Saramad. "A comparison study of the pile-up correction algorithms". In: *Nucl. Instrum. Methods Phys. Res. A* 951, 163013 (2020). DOI: 10.1016/j.nima.2019.163013.

[165] A. Gola, C. Piemonte, and A. Tarolli. "The DLED Algorithm for Timing Measurements on Large Area SiPMs Coupled to Scintillators". In: *IEEE Transactions on Nuclear Science* 59.2 (2012), pp. 358–365. DOI: 10.1109/TNS.2012.2187927.

[166] PicoQuant GmbH. *LDH Series. Picosecond Laser Diode Heads for PDL 800-D / PDL 828*. 2019. URL: https://www.picoquant.com/images/uploads/downloads/ldh_series.pdf.

[167] PicoQuant GmbH. *PDL 828 "Sepia II". Computer Controlled Multi-channel Diode Laser Driver*. 2021. URL: https://www.picoquant.com/images/uploads/downloads/pdl828_sepiaii.pdf.

[168] Thorlabs, Inc. *NDC-50C-4M-A. — Mounted Continuously Variable Neutral Density Filter*. 2012. URL: https://www.thorlabs.us/thorproduct.cfm?partnumber=NDC-50C-4M-A.

[169] J. Rolph et al. "PeakOTron: A Python module for fitting charge spectra of Silicon Photomultipliers". In: *Nucl. Instrum. Methods Phys. Res. A* 1056, 168544 (2023). DOI: https://doi.org/10.1016/j.nima.2023.168544. arXiv: 2301.11833 [astro-ins-det].

[170] C. Bambi and A. Santangelo, eds. *Handbook of X-ray and Gamma-ray Astrophysics*. Singapore: Springer, 2022. ISBN: 978-981-16-4544-0. DOI: 10.1007/978-981-16-4544-0.

[171] A. van den Berg, ed. *Proceedings of the 34th International Cosmic Ray Conference (ICRC2015)* (The Hague, The Netherlands, July 30–6, 2015). 2016.

[172] A. Kappes and B. Keilhauer, eds. *Proceedings of the 37th International Cosmic Ray Conference (ICRC2021)* (Berlin, Germany (Online), July 12–23, 2021). Sissa Medialab, 2022.

[173] T. Saito and K. Okumura, eds. *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)* (Nagoya, Japan, July 26–August 3, 2023). 2023.