Design, Simulation and Development of a Miniature Detector System for Radiation Monitoring on Satellites

Daohua (Hubert) Hu

A dissertation submitted in partial fulfilment of

the requirements for the degree of

Doctor of Philosophy

of

University College London.

Mullard Space Science Laboratory

Department of Space and Climate Physics

University College London

June 23, 2021

I, **Daohua Hu**, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Space environment is particularly harsh for human activity and human assets. As human activity expands rapidly into space, it becomes ever more urgent to monitor the space environment in situ and in real time. This thesis aims to study a low resource, miniature detector system to be employed on the satellites to monitor the in-situ space environment particularly for the hazard high energy particles. The vision is that every satellite even at small scale can afford to carry one such detector system. The core technology used in the detector system is scintillator + SiPM. The scintillator is a mature and traditional radiation detector and the SiPM is the latest generation of photon detector, which rapidly replaces the role of traditional PMT. Development of such detector system requires knowledge over space environment, detector material and geometry, electronics, data processing and more. This thesis tries to have a as complete coverage as possible to demonstrate the flow and the structure of such development. The thesis closely couples the theoretical knowledge to engineering realization. A detector system prototype for pioneering study is fully built and tested from end to end. This serves as the first stepping stone for the path to the full-scale space instrument.

Impact Statement

In 1492, Christopher Columbus first crossed the Atlantic and discovered America. Behind this huge success was the mature of a range of technologies for human sea sailing activity. One of the technologies was the weather observation and prediction. Sailors carry instruments to monitor the wind direction, humidity and pressure. This information gave prediction to the incoming weather condition, upon which, the captain could make knowledge-based decision to optimize the route and avoid the hazards ahead as much as possible. History always amazingly repeats itself. In the dawn of the great space exploration, we are rapidly developing the technology to monitor the space weather. This serves the similar purpose as the weather to the sea sailors. Our satellite operator or future spacecraft captain need the information of space weather to make knowledge-based decision.

The space weather is part of a wider Space Situational Awareness (SAA) operation. ESA has established a large operation in this area. Industry body such as satellite providers continuously demand a reliable and affordable radiation monitoring instrument. There is a wide range of scientific instruments developed to study the space environment from different aspects but none of them are purposely developed for the need of routinely monitoring of hazard environment. None is manufactured in great quantity to form a monitoring network around the Earth orbit. None is compact and light enough while maintaining a good reliability and capability for the very small satellite to carry. This thesis aims to fill in this gap with a new type of radiation monitor instrument. The thesis has completed the pioneering concept study and established the framework for the development of such instrument. Following this thesis, a standardized, highly modulated, low resources and highly reliable radiation monitor instrument will be developed. The vision is to equip it with nearly every Earth orbiting satellite to form a huge network. This network will provide dynamic, real time and ever abundant radiation environment data. With such large scale of data, new space environ-

ment model can be created, or the current model can be greatly extended and improved. Our understanding of the near-Earth space environment can be extended. Combining the latest data mining technology and artificial intelligence technology, a second development of the data can yield a flourish of new applications.

Such development will be one of the many technology evolvements that pave the way for the great human exploration of space. We stand at the door step of a very exciting era, when our civilization is spreading outside of our beloved blue planet. It is very exciting that the work done by this thesis and idea behind it can contribute to such great historical movement.

Our step will not stop after this study. We will continue to optimize our design, solve the problems discovered in this study and ultimately develop the instrument. To achieve this goal, more efforts and more expertise are required. We welcome funding agency, industry body or any individual who share the same vision to join the force.

Acknowledgements

In 2011, I joined MSSL as an electronic engineer under the “push” of my wife Maggie to pursue a higher research degree. In 2012, I formally registered as a part-time research student and started to study a detector system using SiPM sensor under the supervision of Dhiren Kataria and Alan Smith. The scope of the work and the exact goal was not defined very clearly at the beginning, probably because this was a part-time project. This gave me much freedom to explore anything that the project encountered. As the full-time job kept my life and my family supported, I didn’t have the same living pressure as a full-time student had. The only big hurdle is the limited amount of time that I have got besides the busy work and family life. The research area I take has very little overlapping with my day job. This puts great pressure on my time and knowledge. If I had a chance to do this again, I would probably study something closely linked to the work content. However, as every coin has two sides, nothing is completely bad. I gradually built up the knowledge I needed for this project. I gained insight to the area that I probably would never encounter in my normal career path as an electronic engineer. This is a wonderful new world to me. I found that I quite enjoyed reading the academic paper, enjoyed studying the physics about the space environment, the particle interaction with matter, the detector, the radioactive source and more, enjoyed learning the new simulation tools like Geant4, the magical mathematic algorithm that can recover the contaminated signal. I enjoyed almost every area that I explored for this project. The reward is indescribable. The research transformed my knowledge, my view of life and my understanding of the world. I greatly thanked my wife, who “kicked” me onto this path.

Over the journey, we welcomed two lovely children, Ryan and Ivy. They brought great joy to my life together with many sleepless nights. I enjoyed every moment with them. Incredibly, I found myself even more motivated when my life is ever busier. I

felt a deep sense of responsibility that I should care them, should create a good environment for them to grow and flourish, should show them how to be good and how to pursue a meaningful life. I took the elder boy to the lab on Sunday, when I carried on my PhD work and he watched. Sometimes he helped me to write down some test data. So, he is well qualified to my thank to list. After family dinner, we often study together as my son does his homework, my daughter does her craft work or little drawing and I do my PhD research. We enjoy the company of each other and the momentary peace. I sincerely thank my children, who are the ultimately source of power that drives me to strive forward.

I would like to thank MSSL, a lovely place to work. The department gives me ultimate support on the PhD work. The department grant allows me to buy all the necessary materials to build the prototype detector system. For this, I particularly thank the two department directors: Alan Smith and Andrew Fazakerley. Alan is also my principle supervisor. He never missed a PhD meeting with me and always gave me valuable suggestions. I benefited a lot from every discussion with him. His great insight has led to many developments in this thesis work, for example, the study of the scintillation timing.

I should not forget those who support my study behind the scene: David Linder, Kawata Daisuke, Dimitra Stamogiannou, Philippa Elwell and many more. Their help (some I know, some I even not know) ensure I can focus my energy in the study.

I would like to particularly thank my first supervisor Dhiren Kataria. I worked the first instrument ChaPs for TDS-1 mission with him closely together. This project inspired me to embrace on the study of an in-situ instrument. We continued working closely together on the successor projects (QB50, HOPE, LGR) in parallel to the PhD project. We sometimes argue, sometime disagree, sometimes laugh, but we always work closely together and support each other. No doubt, without his lead, the thesis won’t be presented.

I would like to thank Mark Hailey, Gary Davison, Phil Guttridge and Richard Cole who are my line managers and project manager. They gave me generous support to pursue the PhD research along the daily job. Their support and sponsorship of the purchase of some of the research materials and equipments are vital to the success of this project.

I would like to thank Andy Malpuss, Vitor Botelho, who helped me on the two tests with radioactive source. Thank our lovely office ladies: Julia Wehrle, Suzanne Winter, Rosalind Medland, who helped me on the purchase, travel booking etc.

The list can continue. Limited by the space and my memory, I just would like to say thank you to all those who have helped me either directly or indirectly for this PhD study. No doubt, I will share all the outcomes to MSSL and to the colleagues to pay back the uncountable help.

A special thanks to my mum, a woman who not only gave me life but also shaped my personability. She devoted all she had to my growth and bore lots of hardship to give me a stable environment to flourish. This is a thesis for her.

Table of Contents

List of Figures

List of Tables

Abbreviation

BF	Branching Fraction
BR	Branching Ratio
BJT	Bipolar Junction Transistor
CAD	Computer Aid Design
CEM	Channel Electron Multipliers
COTS	Commercial of the Shelf
CSDA	Continuous Slowing Down Approximation
CSV	Comma-separated Values
CTE	Coefficient of Expansion
DC	Dark Count
DCR	Dark Count Rate
DFT	Discrete Fourier Transform
DRS	Domino Ring Sampler
EDA	Electronic Design Automation
ESP	Emission of Solar Protons
FEE	Front End Electronics
FT	Fourier Transform
FWHM	Full Width at Half Maximum
GCR	Galactic Cosmic Ray
GDML	Geometry Description Markup Language
Geant4	GEometry ANd Traking
GPS	General Particle Source
GWP	Gain Bandwidth Product
GUI	Graphic User Interface
IFT	Inverse Fourier Transform
JFET	Junction Gate Field-effect Transistor
KDE	Kernel Density Estimation
MCP	Micro-channel Plate
PCB	Printed Circuit Board
p.e.	photon electron
PDE	Photon Detection Efficiency
PDF	Probability Density Function
PET	Positron Emission Tomography
PHA	Pulse Height Analysis
PMT	Photon Multiply Tube
RF	Radio Frequency
ROOT	Open-source Data Analysis Framework
SEP	Solar Energetic Particle
SiPM	Silicon Photon Multiplier
S/N	Signal to Noise
SPAD	Single Photon Avalanche Diode
SPENVIS	Space Environment Information System

SPICE	Simulation Program with Integrated Circuit Emphasis
SSD	Solid State Detector
STEP	Standard for the Exchange of Product Data
TB	Total Branching Ratio
TDS	TechDemoSat
TIA	Trans-Impedance Amplifier
TOF	Time of Flight
XML	Extensible Markup Language

Preface

Where there's a will, there's a way.

----- by a wiseman

March to the universe is our ultimate destiny.

----- by Hubert Hu

Chapter 1 Introduction

1.1 Thesis Overview

Since the first satellite launched into space, humanity have built up extensive knowledge about the hazard of the space environment. The intensive, highly penetrable and energetic particles impose great threat to the human safety in the near-Earth space and is the prime reason for the failure of electronics on the satellites. So, it is of great interest to develop the space situation awareness capability to monitor the dynamic and hostile near-Earth Space environment. The in-situ particle measurement instruments are a key part of such capability. Such instruments have a history as long as that of human space exploration. From Faraday Cups, Langmuir Probes to Plasma Analysers, Mass Spectrometers and Solid-State Detector Instruments, in-situ instruments become more complex and capable. This thesis describes the development of a miniature, low resources, solid-state detector instrument. The long-term vision is to make it an affordable (in terms of both resources and cost) payload to the emerging miniature satellite constellations, such as the OneWeb constellation. The benefit is both ways. One way is that the Satellites made mainly of commercial of the shelf (COTS) components can receive in-situ real time warning of the surrounding space environment and probably some autonomous action can be pre-programmed to protect the satellite electronic system upon sever radiation interaction. The other way is that the vast number of instruments with standard design can form a network to provide a complete coverage of the near-Earth space where the satellite constellation is. The data from the network can then feed into a global space weather monitor system.

Developing a space instrument is a complex system project which requires efforts beyond a single person’s expertise. Therefore, this thesis focuses on the core detector system of the instrument. The chosen detector system is the scintillator + SiPM platform. A scintillator made of high z material can be very radiation hard and be able to stop high energy particles within a small volume. A SiPM is the most sensitive photon

sensor and does not require a high bias voltage, unlike the traditional PMT. This brings the benefits of both high photon detection efficiency (hence improving the overall detector system efficiency) and low demand on power resource.

The remaining of the *Introduction* chapter will introduce the near-Earth space environment, a brief review of the common detector technology and a brief survey of similar in-situ instruments. Following the introduction, there are six major chapters, which thoroughly discuss the research, development and test of the Scintillator+SiPM detector system prototype. These chapters follow the nature order of an instrument development cycle as:

Chapter *Scintillator Design* studies the principles of scintillators. The understanding helps us to guide the selection of scintillator material. 18 popular scintillators are surveyed and compared for all the relevant properties. After three iterations of design, we finalized our scintillator module for manufacturing. The manufactured scintillator module is presented. The last part of the chapter discussed the concept of the full instrument based on our scintillator module design.

Chapter *Simulation* presents the Geant4 simulation outcome of the first scintillator module design. The simulation covers the energy resolution, angular resolution and energy range to different particle species. Geant4 is a very powerful simulation tool. The accuracy of the simulation output depends on the good understanding of the code function and the proper configuration of the physics process and detail setup of detector geometry. ROOT toolkit is used to analyse the Geant4 simulation output data.

Chapter *Electronics* discusses about the SiPM preamplifier circuit design. This is the first and very important stage of the readout electronics chain. Two common preamplifier architectures (current mode and voltage mode) are thoroughly studied and compared. The study shows the current mode preamplifier has better bandwidth while maintaining the system stability. A theoretical noise model of the current mode preamplifier is studied. This paves the way to deliver a low noise circuit design. The last section presents the development, manufacturing and testing of a compact preamplifier board.

Chapter *Signal Processing* is about data processing. The idea is to use some smart software filters to tackle the signal pile-up issue. The adopted filters are Deconvolution filter and Wiener filter. Deconvolution filter can sharpen the signal, i.e. narrow the pulse,

at the cost of increasing noise. Wiener filter can improve the signal to noise ratio, so it serves to compensate the increasing noise introduced by Deconvolution filter. The two filters can be constructed as a single combi filter. The second half of the chapter presents the other general data processing software developed in Python language. This software helps to decode the raw data and run statistical analysis. This customized software provides powerful tool for the analysis conducted in the next chapters.

Chapter *System Characterization* explores the data collected from the system test with a LED source to characterize the SiPM and its readout system performance. We have found the major time response limit comes from the readout electronics and software after the preamplifier stage. The SiPM system is characterized for after-pulse, optical crosstalk, photon detection efficiency, linearity and dynamic range. The results are compared to those published in the manufacturer datasheet. We are pleased to see the close match between them with the relatively simple test setup. The SiPM dark count is a very important parameter that affects the overall system performance. We develop a statistic model to describe the occurrence of dark count. The model proved that without after-pulse and cross-talk, the chance of a signal >2 photo electrons (p.e.) is negligible. In another word, the dark count shall be a 1p.e. signal. Those 2p.e. or above signal that we captured at dark condition are, therefore, all from the effect of after-pulse and cross-talk. Based on the statistic model, we develop a simulation program to simulate the dark count event with a random generator. The result matches the model prediction well. The real dark count rate is also measured.

Chapter *Radiation Test* discusses the final, end-to-end system test with a radioactive source. The test is carried out in air with Am-241 alpha source. This limits the signal strength as alpha particle can only penetrate a short distance in air. However, we indeed recorded the system response as we alter the distance between the source and our detector system. The Am-241 source also emits gamma ray as a by-product of the alpha decay at a ratio of 1:0.4. We use timing to separate the output signals from the two different scintillators. The ratio of the number of signals collected from the two scintillators match the same 1:0.4. We run a theoretical calculation to predict the signal height from each scintillator. The prediction matches our observation statistically well. Overall, partially limited by the test in air and partially because of the low photon collection efficiency, the signal strength requires further improvement.

The last *Conclusion* chapter summarizes the project and points out the next step to the development of an operational instrument.

1.2 Space Environment

Up to now, human's activity in space has concentrated on the near-Earth region. Naturally, this near-Earth environment is the most important and of the most interest to our study. In this environment, there are three major sources of high energy particles: (a) trapped charged particles in the Earth's radiation belts (ERBs), (b) solar energetic particles (SEP) from our sun and (c) galactic cosmic ray (GCR) from the external of our solar system [1]. All three sources are impacted by the Earth's magnetic field. Figure 1-1 presents the three sources and their interaction with Earth's magnetic field.

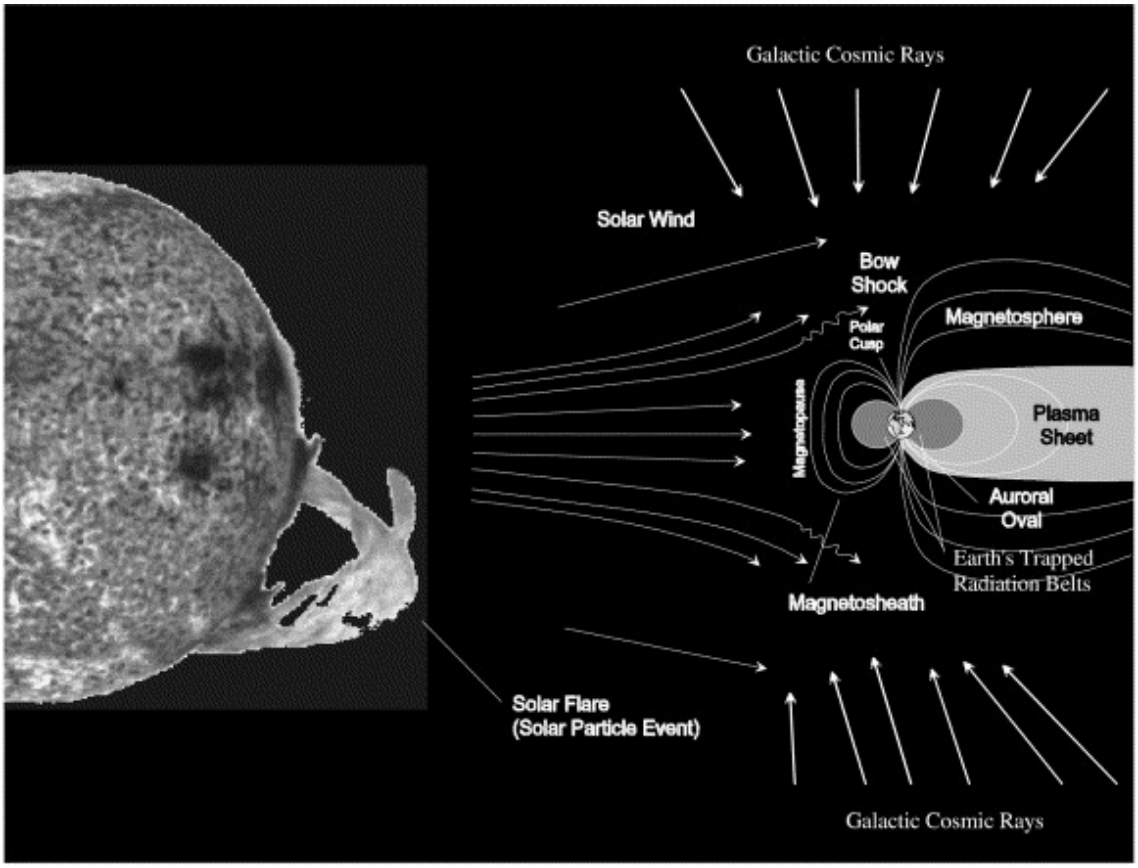


Figure 1-1 The illustration of the interaction of Earth Magnetosphere to all three high energy particle sources. [2]

1.2.1 Trapped charged particles

The Earth radiation belts were discovered in 1958 from the Explorer 3 mission by Van Allen. The detail description of the radiation belts and the related physical process can be found in [3][4][5]. Here we will only briefly describe its main characteristics.

The radiation belts are formed by the Earth magnetic field capturing the charged particles (mainly electron and proton) from the solar wind. There is mechanism that these particles are accelerated in the magnetosphere to reach even higher energy level. According to the Maxwell's electromagnetic theory, the acceleration should be done by some electric field not by the magnetic field as magnetic field only changes the charged particle direction of motion and it doesn't apply work to the particles.

There are mainly two radiation belts around the Earth. The outer belt locates between 3 Earth radii (R_E) and as far as 10 Earth Radii. It is mainly made up of high energy electrons (0.1 -10 MeV). The inner belt is located between 0.2 R_E and 2 R_E. It contains both high energy electrons (hundreds of keV) and energetic protons with energy exceeding 100MeV. Due to the non-concentricity of Earth and its magnetic dipole, the inner belt comes closest to the Earth's surface at South Atlantic area. The existence of excessive high energetic particle flux in this area is named South Atlantic Anomaly (SAA).

Decades of data has been collected through various missions to study the radiation belts. Based on the data, a few models for the radiation belt charged particle population, flux and energy distribution are created. The most widely used models are NASA's AP-8[6] and AE-8[7] because only these two models have full coverage over spatial and spectral range of the radiation belts. The models present the omnidirectional, integral electron and proton fluxes in the energy range 0.04MeV to 7Mev for electrons and 0.1MeV to 400MeV for protons in the form of maps. Figure 1-2 shows some example.

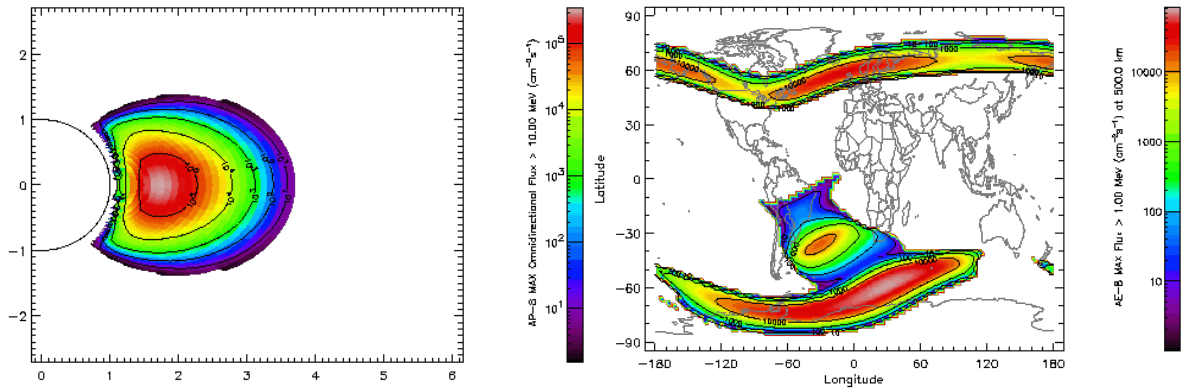


Figure 1-2 The left is the section view of integral proton flux > 10MeV. The coordinate is the number of Earth radii. The right is the world map of the AE-8 integral electron flux >1MeV at 500km altitude. Source: SPENVIS help page[8]

The 11-year solar cycle has a periodic impact on the low altitude trapped particle flux as the Earth's neutral atmosphere expands during solar maximum, which erodes into the lower edge of the inner radiation belt. This will change the proton and electron flux level periodically.

To compensate the inadequacy and error in the AP-8 and AE-8 models, other models specializing in, e.g. low altitude or a certain energy range have been developed. They are also available in the SPENVIS tool [9] and a brief introduction of them can be found in the SPENVIS help page.

In the magnetosphere, besides the radiation belts, there is also a drift current, called Ring Current[10][11], circulating the Earth longitudinally at altitudes of ~10,000 – 60,000km. The current is made up of medium-energy particles, with proton drifting westward and electron drifting eastward. Protons/ions ranging from ~10keV to a few hundreds of keV comprises most to the current. The electron contributes little because of their negligible energy density. The ring current is closely linked to magnetic storms in the near-Earth space, which is known as a cause of the disturbance or even permanent damage of telecommunication and navigation satellites [12].

The structure of the Earth Magnetosphere is shown in Figure 1-3.

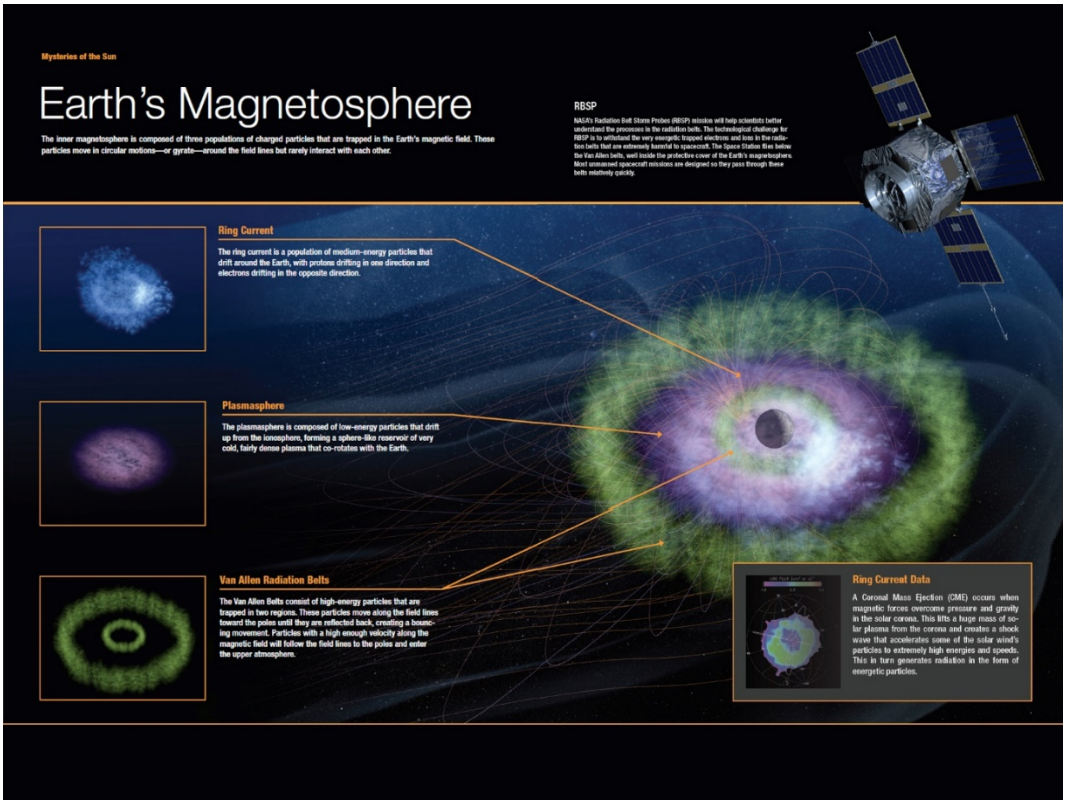


Figure 1-3 The illustration of Earth Magnetosphere and its major components. [13]

1.2.2 Solar Energetic Particles

Solar Energetic Particles (mainly protons and some ions) originate from our Sun as its name suggests. Their energy ranges from a few tens of keV to many GeV. The study [14] so far shows two mechanisms that create these energetic particles: *impulsive* and *gradual*. The impulsive SEP is associated with Solar Flares and Jets and is rich in electrons. The duration is only at the order of hours. The frequency is estimated to be ~1000/year at the sun, and we can observe ~100/year at the Earth. The gradual SEP is accelerated in shock waves driven by CME [15] and is rich in protons. The duration can reach several days. It is less frequent and is estimated to be ~10/year. However, the gradual SEP carries much higher energy and imposes a much larger threat. Figure 1-4 describes the mechanism of both SEP.

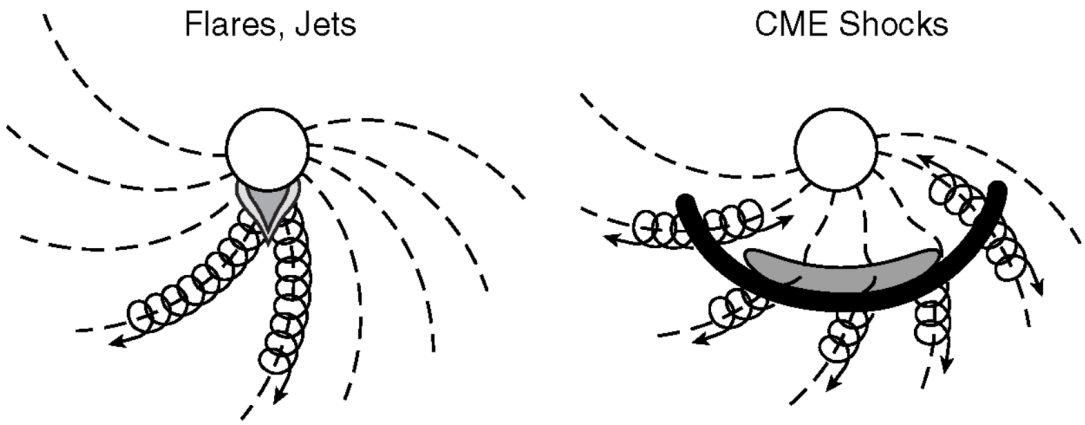


Figure 1-4 The left is Impulse SEP and the right is gradual SEP. The solid black on the right is the front of the shock waves driven by the CME (grey). Re-produced from [14].

Both classes of the SEP travel along the interplanetary magnetic field line, also called the Parker Spiral [16]. Figure 1-4 illustrates the gyration trajectory of the SEP along the magnetic field line. This gyro-motion is governed by the Lorentz force.

It is not possible to predict the onset of an SEP event. However, it is possible to provide an early warning. The reason is that it takes some time for the shock wave carrying hazardous SEPs to arrive while the occurrence of a CME-driven shock wave on the solar surface can be observed by radio burst with a delay as little as about 8 mins. The time window is about 12 hours [17].

We should differentiate SEPs from the Solar Wind [18]. The Solar Wind is a stream of energized, charged particles, primarily electrons and protons, originating from the Corona (the upper atmosphere of the Sun). Differentiated by the position of their origin on the Sun, two fundamental states of solar wind are observed, which is termed the

slow solar wind and the fast solar wind. The slow solar wind has a typical velocity of 300-500km/s and the fast solar wind has a typical velocity of 750km/s. Extreme case of 1850km/s associated with CME driven shock are also reported [19]. Unlike SEPs, the Solar Wind is a continuous process although the energy level of the charged particles it carries is relatively low (between 0.5 and 10keV) and, therefore, imposes little risk to the human space assets or activity.

Based on the collective mission data, a few SEP models were developed. The King model [20] was created using data from the active years of solar cycle 20 (1966 – 1972). The model mainly covers protons in the energy range 10-100MeV. The JPL model was developed by Feynman [21] . Its latest version JPL-91 model covers data from the IMP and OGO spacecraft between 1963 and 1991. The model has energy thresholds of 1, 4, 10, 30 and 60MeV. The ESP model [22] is based on more complete data covering the 3 full cycles (20, 21 and 22). In addition, the model was developed on a newer and better theoretical basis called maximum entropy theory [23]. All three models are for long term total Solar proton fluence use with the ESP model the default choice in SPENVIS. Only active years (7 out of the 11 year cycle) are used and only high energy protons are counted.

For single event upset rate prediction, CREME96 model [24] is widely used. It covers flux data for the worst week, worst day or worst 5 min for a large SEP event observed during 19-27 October 1989.

1.2.3 Galactic Cosmic Ray

Galactic Cosmic Rays (GCRs) [25][26] are very high-energy charged particles that enter the solar system from the outside. GCRs are composed of mainly protons at 90%, helium nuclei at ~9% and the remaining heavier elements at ~1%. GCRs are known for their extreme high energy (starting at 1MeV and continuing to around 10^{21}eV [26]) compared to the trapped charged particles and solar energetic particles. GCR's flux is relatively much lower than them, and this is partly attributed to the protection of the solar interplanetary magnetic field, carried by the solar wind. In plain word, the solar wind can "push" back the GCR flux from reaching the inner solar system. As solar wind strength varies through the solar cycle, this modulates the flux intensity of GCRs. Therefore, GCR flux is at a peak when solar activity is at minimum [27]. Figure 1-5 shows the flux of all nuclear components present in GCR.

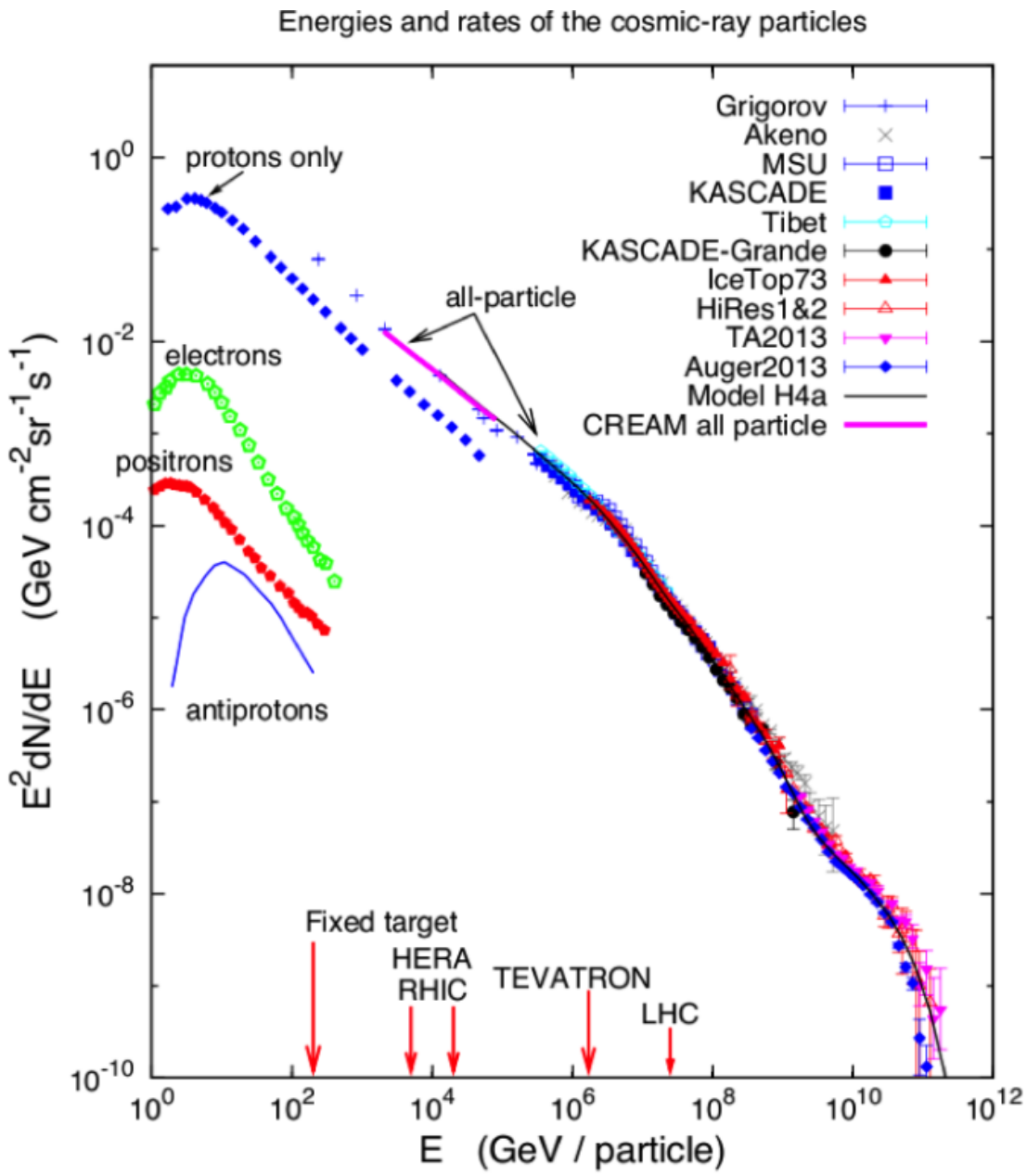


Figure 1-5 Spectrum of cosmic rays observed at Earth [25]

SPENVIS provides four models for estimating the GCR flux intensity: the ISO-15390 model (an ISO standard), the CREME96 model [28], the CREME86 model (part of CREME86 software package) and the Nymmik model[29]. Figure 1-6 compares the GCR flux intensity predicted by different models.

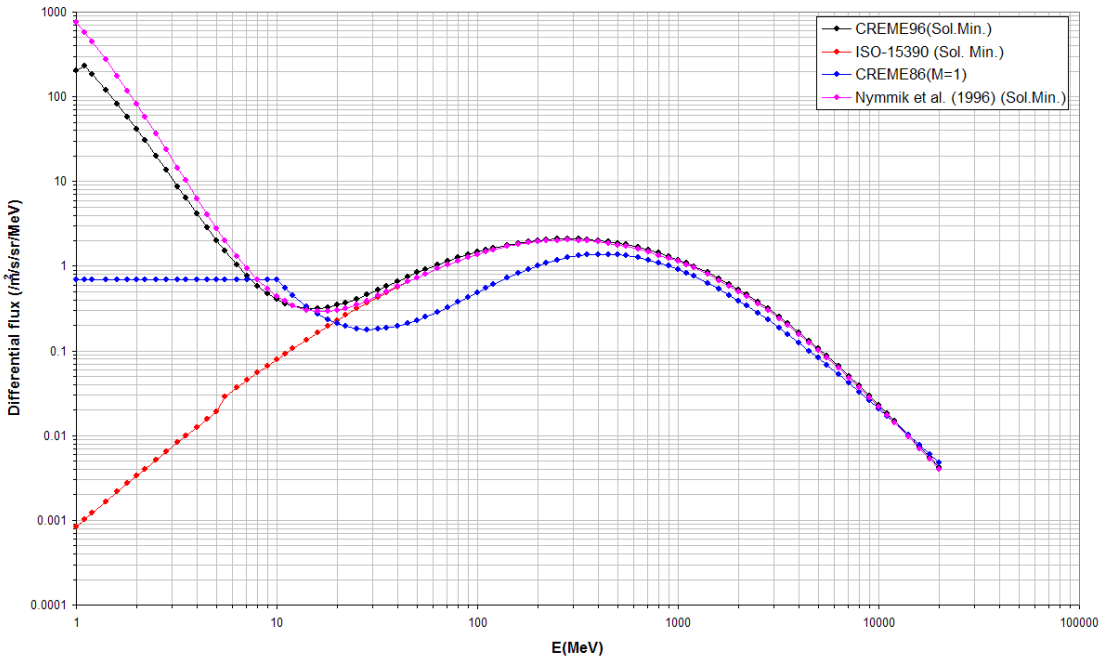


Figure 1-6 The Hydrogen GCR flux near Earth predicted by the different GCR particle flux models at solar minimum [30]

1.2.4 Environment for the target mission orbit

We need to establish a set of detector requirements to narrow down the detector design choice and shape the detector geometry. We select a near-Earth satellite mission that we previous participated in. The purpose is to demonstrate the process that the detector requirements such as the energy range and flux rate, are determined from the target mission orbit environment. The selected mission is TDS-1 [31] launched by SSTL [32] in 2014. TDS-1 satellite operates in a sun-synchronous circular orbit, altitude of ~650km with a high inclination angle of ~98.3°. Its orbit two-line code (two lines of orbit parameters that unambiguously determine the satellite orbit) is:

1 40076U 14037H 16252.22481556 .00000205 00000-0 33619-4 0 9994

2 40076 98.3252 346.4000 0004790 258.8803 101.1897 14.81248253117283

We use SPENVIS web tool to generate all the mission related space environment data.

The first step is to generate the mission profile and satellite trajectories. It is straightforward with just entering the two-line codes and the duration of mission. Figure 1-7 shows the TDS-1 trajectory on the world map for a complete 24 hours.

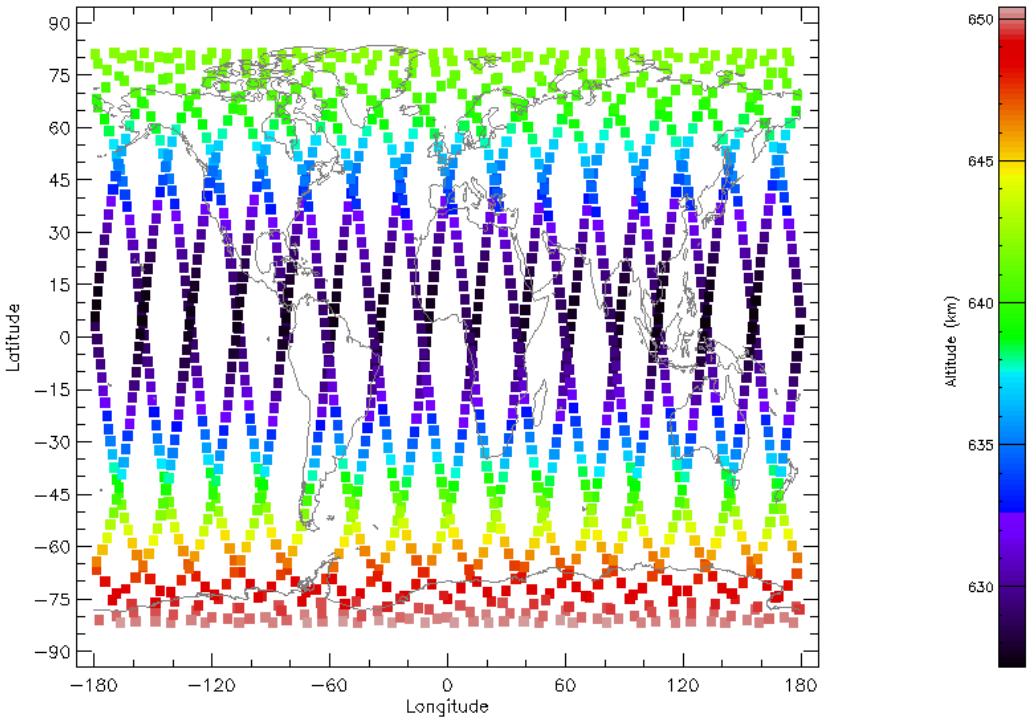


Figure 1-7 TDS-1 orbit trajectory on the world map for a full day. The colour bar shows the altitude. Together with Longitude and Latitude, this plot completes the 3D position of the satellite on a 2D format.

The second step is to calculate the trapped proton and electron fluxes using the AP-8 MIN and AE-8 MAX models. The result is shown in Figure 1-8 and Figure 1-9.

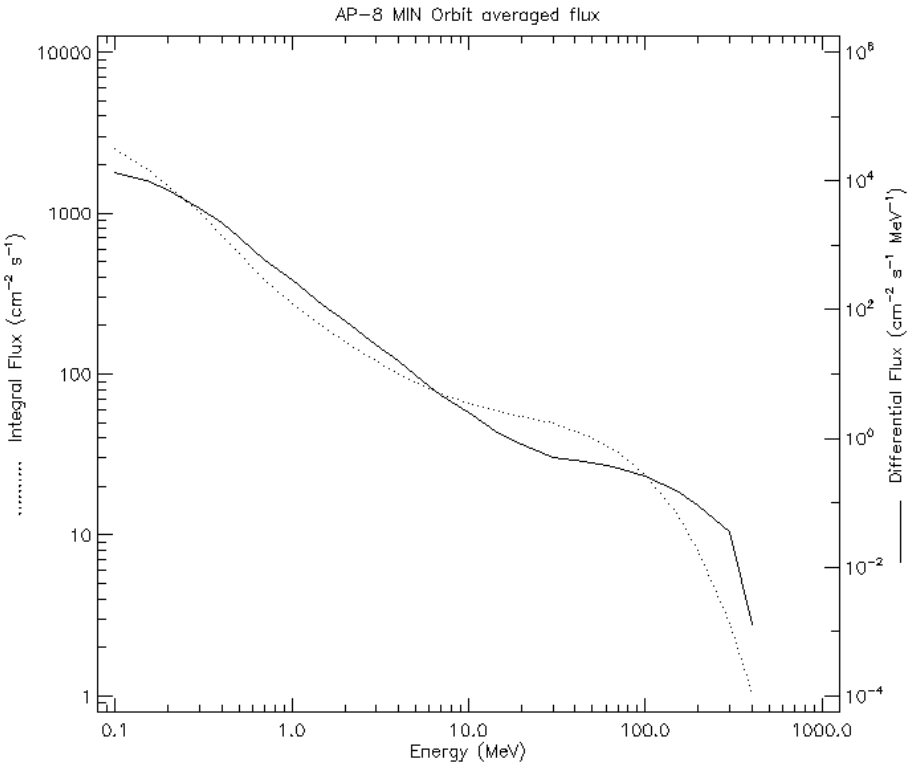


Figure 1-8 The proton flux for the TDS-1 satellite orbit calculated by the AP-8 model

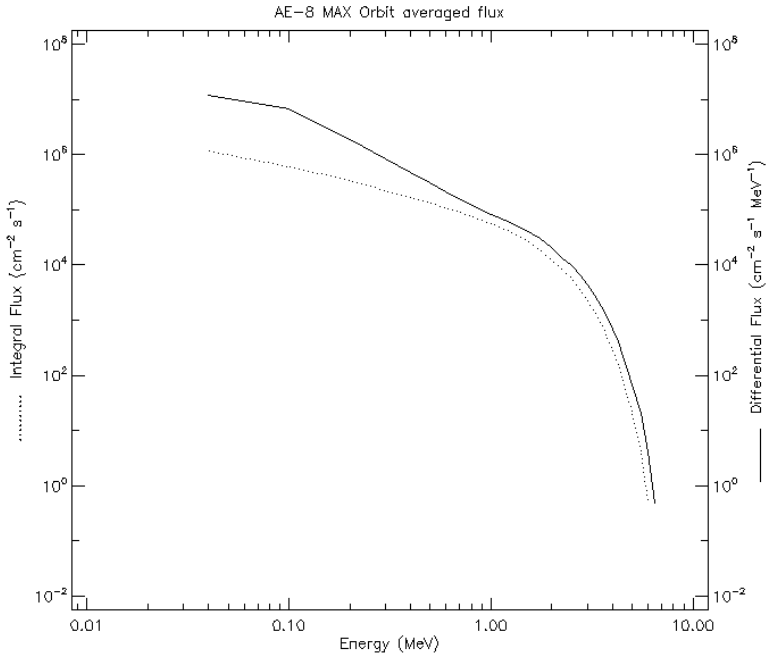


Figure 1-9 The electron flux for the TDS-1 satellite orbit calculated by the AP-8 model

SPENVIS can also provide world map view to intuitively show the high flux density area, which is linked to the intrusion of the inner radiation belt due to the irregular of the Earth magnetic field. The highlight of most famous SSA area and polar region is obvious as shown in Figure 1-10.

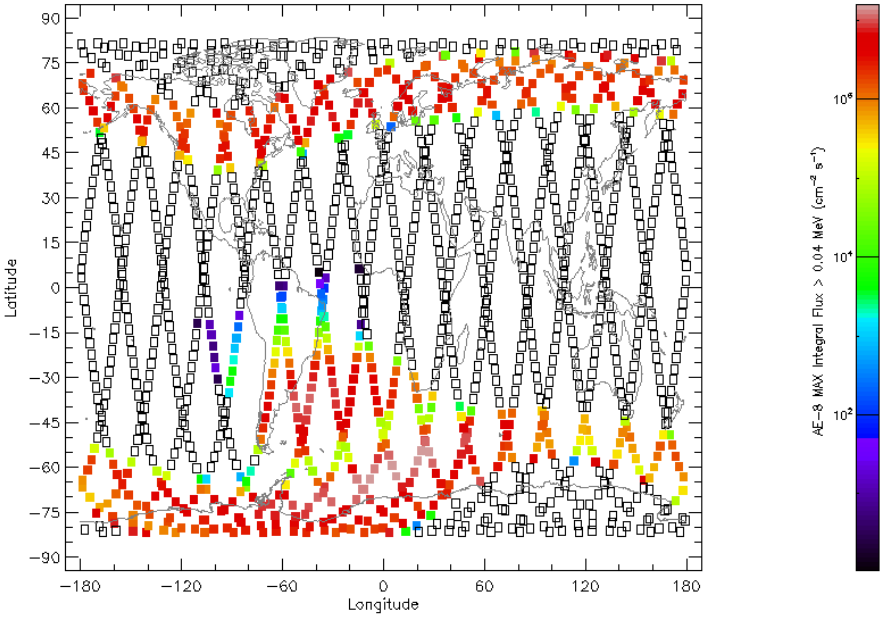


Figure 1-10 The world map view to show the electron flux density distribution along the orbit trajectory. The SSA and polar region with outstanding high intensity is clearly shown.

The third step is to calculate the solar energetic particle flux for both long term (total influence) and short term (worst week). The long-term data from ESP-PSYCHIC model is used for estimating the mission dose and the short-term data from CRÈME-96 is

used for evaluating the single event upset rate. Figure 1-11 and Figure 1-12 show the model calculation results.

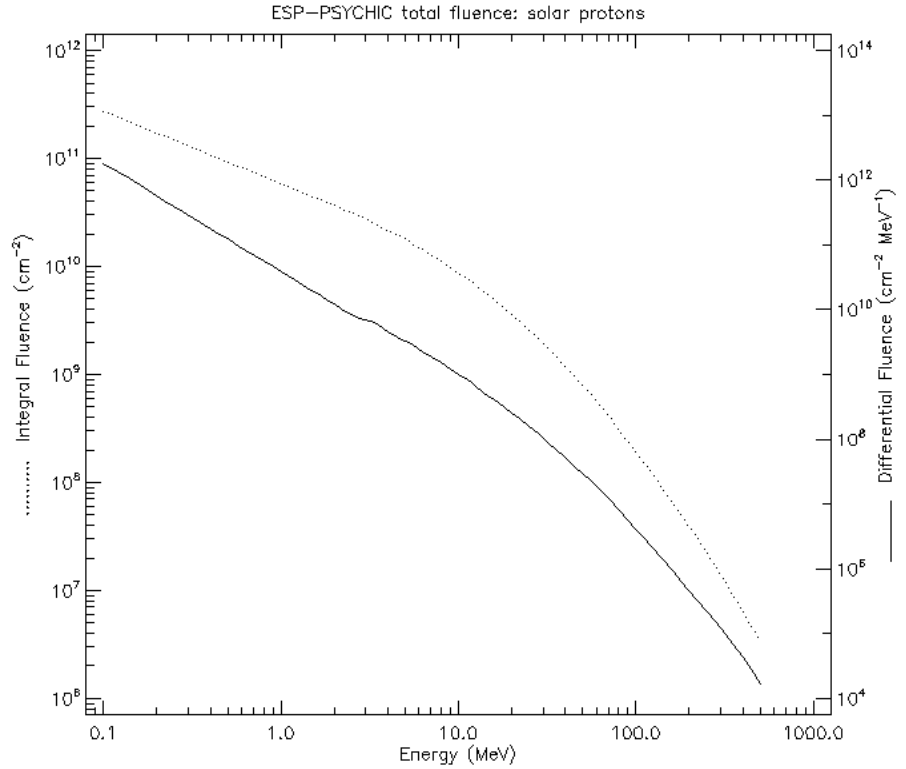


Figure 1-11 The total proton fluence for the TDS-1 satellite mission life calculated by ESP-PSYCHIC model

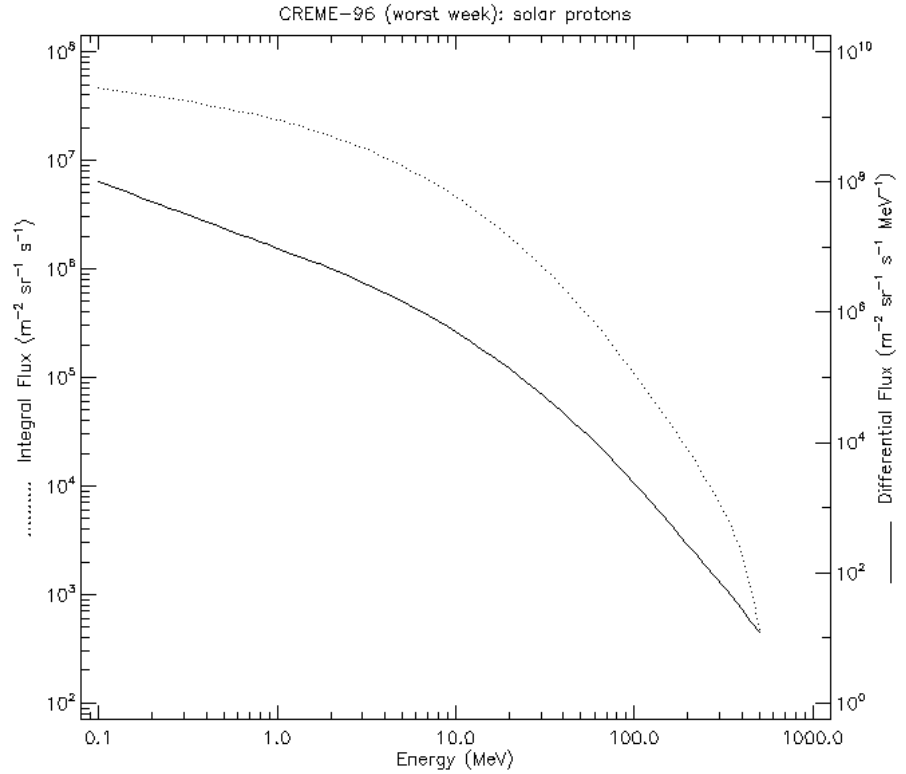


Figure 1-12 The worst week proton flux intensity for the TDS-1 satellite calculated by the CRÈME-96 model

The fourth step is to calculate the galactic cosmic ray flux using ISO-15390 standard model with Solar Minimum (May 1996) data. The model generates flux spectrum covering ions from atomic number z=1 to z=92. We know hydrogen (z=1, also called proton) and helium (z=2, also called alpha) made 90% and 9% of the entire GCR population as the dominant species. Therefore, we only plot the flux of these two particles in Figure 1-13 and Figure 1-14.

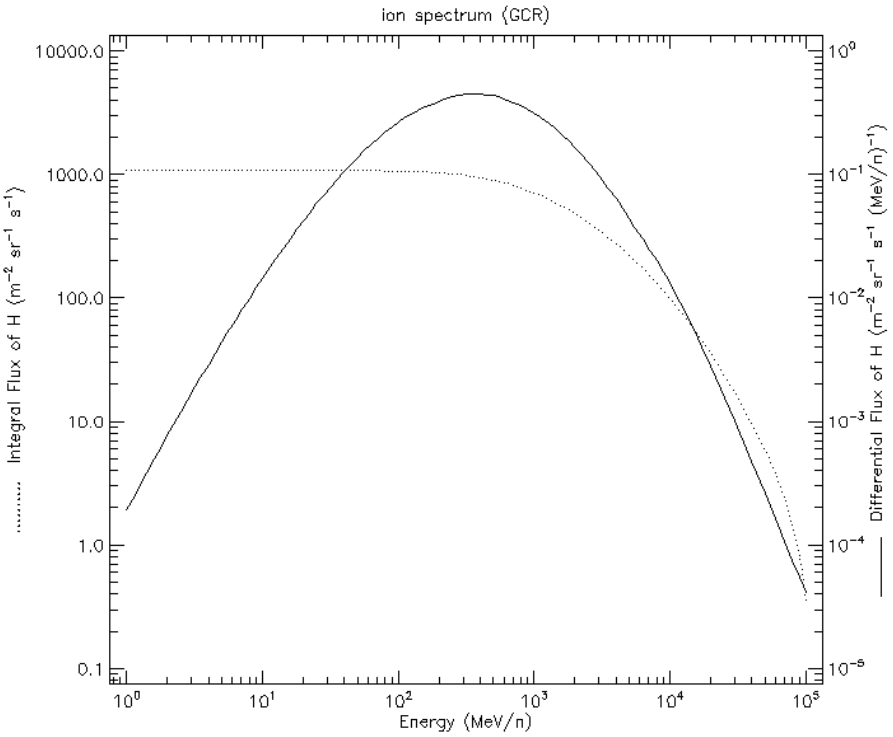


Figure 1-13 The GCR high energy particle H (Z=1) flux intensity spectrum

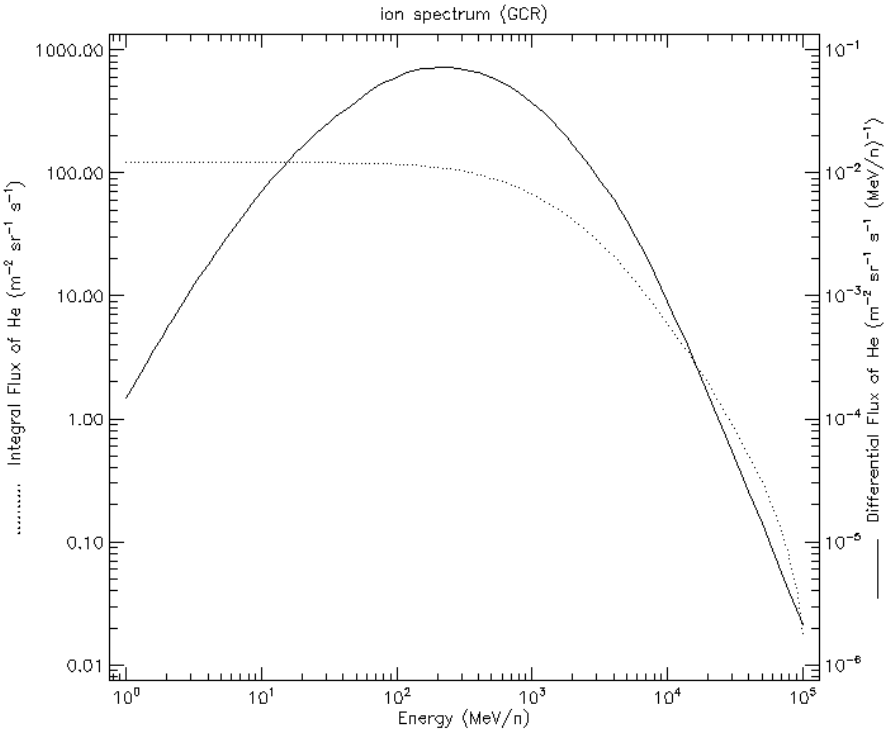


Figure 1-14 The GCR high energy particle He (Z=2) flux intensity spectrum

The summary of data from the SPENVIS calculation is presented in Table 1-1. It shall be noted that trapped particles data and SEP data quote the peak differential flux to represent the most abundant energy level. The GCR data quotes the integral flux for the dominant species proton and helium nuclei. Iron nuclei is also listed as it is exceptionally abundant at such high atomic number. We inherited the original unit that SPENVIS quoted. Be aware the subtle difference in the units used in different rows.

Table 1-1 Summary of TDS-1 satellite mission flux (peak level) for the charged particles from three main sources

Source	Species	Energy range	Peak flux	Unit
Trapped particles in belts	Protons	100keV-400MeV	1.3675E+4[1]	$cm^{-2}\ s^{-1}\ MeV^{-1}$
	Electrons	40keV – 6MeV	1.1939E+7[2]	$cm^{-2}\ s^{-1}\ MeV^{-1}$
Solar Energetic Protons	Total fluence	100KeV – 500MeV	1.777E+12[3]	$cm^{-2}\ MeV^{-1}$
	Worst week	100KeV – 500MeV	1.019E+8[4]	$m^{-2}\ sr^{-1}\ s^{-1}\ MeV^{-1}$
Galactic Cosmic Ray	Proton	1MeV – 100GeV	1.076E+03[5]	$m^{-2}\ sr^{-1}\ s^{-1}$
	Helium	1MeV – 100GeV	1.209E+02[6]	$m^{-2}\ sr^{-1}\ s^{-1}$
	Iron	1MeV – 100GeV	3.6357E-1[7]	$m^{-2}\ sr^{-1}\ s^{-1}$

Note:

1. Differential proton average flux over total mission at 100KeV.
2. Differential electron average flux over total mission at 40KeV.
3. Differential solar proton total mission fluence at 100keV.
4. Differential solar proton worst week flux at 100keV.
5. Integral GCR proton average flux from 1MeV.
6. Integral GCR helium nuclei average flux from 1MeV.
7. Integral GCR iron nuclei average flux from 1MeV.

These energy ranges are limited by the data model capability. In the real space environment, there is no definite range.

The SPENVIS tool provides a simple and quick method to estimate the mission environment. This gives a good indication about what will be seen for the orbit that our instrument will be launched into over the mission life.

It is worth pointing out that all the models used in the calculation are only statistic model based on historical data. These data are limited in energy range, resolution and flux range. To compensate, scientists developed the mathematical model to fit the limited observation data and then they used the model to extend the energy coverage and resolution. This inevitably is imperfect as the physical process is so complex and dynamic that it is not possible to be represented by just a few mathematic formulas. However, on a macro scale and over the average of mission life, the model predicted data make statistical sense.

1.2.5 Other orbits

The TDS-1 orbit is a typical Low Earth orbit or Polar orbit. To complete our survey but not at the same level of environment analysis as for the TDS-1 orbit, we list the definition and feature of other orbits in Table 1-2.

Table 1-2 Classification of orbits and description of orbit characteristics [33]

Name	Altitude (km)	Inclination (deg)	Orbit description	Satellite applications	Operational satellites **[34]**
Low earth orbit	100-1000	<65	Cold, dense, ionospheric plasma; dense, supersonic neutral atmosphere; solar ultraviolet(UV); orbital debris; South Atlantic Anomaly(SAA)	CubeSat, Surveillance, ISS, Communication, Earth observation	471(47%)

Name	Altitude (km)	Inclination (deg)	Orbit description	Satellite applications	Operational satellites [34]
Medium Earth orbit	1000-36000	<65	Solar UV; Trapped radiation belts; Plasmasphere;	GPS, Satellite broadband,	69(7%)
Polar orbit	>100	>65	Solar UV; Cold, dense ionosphere; Supersonic neutral atmosphere; Orbital debris; Auroral particles; Solar flares; Cosmic rays; SAA; Horns of radiation belt;	Imaging, space weather	35(4%)
Geosynchronous orbit	~36000	0	High energy plasmasheet; Substorm plasma; UV radiation; Outer radiation belts; Solar flares; cosmic rays	Communication, Weather forecast, TV broadcast, Signal relay	419(42%)
Interplanetary orbit	Outside magnetosphere	N/A	Solar-wind plasma; Solar flares; cosmic rays	Large science exploration mission	

1.3 General measurement requirements

The environment analysis gives us the knowledge of the target particles that our instrument will encounter. It is ideal to be able to detect all these particles simultaneously with good energy resolution and time resolution. The reality is that current technology of particle instrument and satellite downlink rate can't support this goal. Therefore, a trade-off must be made to focus on the most valuable particle species and their energy range for a single instrument design. Alternatively, a couple of instruments forming a suite can possibly cover the full spectrum.

We list a few general design considerations for a particle instrument in space environment.

Flux, Energy Ranges and Time Resolution

These are the most basic requirements that a particle instrument shall consider. Depending on the application, the instrument shall have the non-saturation max detection rate above the peak flux level. If it is not possible to cope with the very rare high flux event, the instrument shall be designed to be safe and be able to resume normal operation when the flux level drops. If an instrument can vary its geometry factor, algorithm can be developed to adjust the geometry factor to cope with different flux range.

Energy range is another challenge to the detector design. From Table 1-1, we can see GCR particles have an energy range of five orders of magnitude and the trapped particles have nearly four orders of magnitude. Such a wide range imposes difficulty not only on the detector design but also on the readout electronics. The wider the range that the instrument needs cover, in general, the poorer the energy resolution will be.

Time resolution highly depends on the application. If dynamic process of the particles to be measured are very important, the time resolution is critical. For example, to measure the substructures in the solar wind, the instrument needs a time resolution less than the time that the Solar Wind passes the spacecraft. For other application like dosimetry the time resolution is not critical and, generally, an instrument with one second resolution can meet most of the need.

Angular Coverage

A 4π angular coverage is always a good feature to have. However, this often require a spinning spacecraft or multiple instruments to achieve. Therefore, the instrument design or the number of instruments can be simplified if the angular coverage requirement can be tailored down.

A general approach to assess how much angular coverage is required is to consider three characteristic velocities: the thermal velocity of the particle group $\mathbf{V_{th}}$, their bulk velocity, $\mathbf{V_b}$ and the spacecraft velocity, $\mathbf{V_{sc}}$.

If $\mathbf{V_b}$ is much larger than $\mathbf{V_{th}}$, for example the ions in the solar wind, the particles will be confined in a small cone of directions and a small range of velocities (kinetic energy). In this case, the instrument can have a small angular and energy coverage and be optimized for a finer angular and energy resolution. The pointing accuracy and stability become more critical in this type of application.

If $\mathbf{V_b}$ is much smaller than $\mathbf{V_{th}}$, for example the solar wind electron or most of the plasma in Earth's magnetosphere, the particles will encounter the instrument from all directions. In this case, a 4π angular coverage becomes necessary. As the thermal energy becomes more significant, the total particle energy (thermal energy + kinetic energy) will be more diverse. Therefore, the instrument shall have a wider energy range to cover the disperse.

If $\mathbf{V_{sc}}$ is much larger than $\mathbf{V_{th}}$ and $\mathbf{V_b}$, for example the plasma in the low-Earth orbit, the particles appear to be still to the spacecraft. In this case, the dominant direction is the spacecraft "ram direction" (the moving direction of the spacecraft) and the particle energy is determined by $\mathbf{V_{sc}}$ and the particle mass.

Particle identification

Some applications require the capability of identifying the ion species. This knowledge is useful to trace the source of the particles. For example, a high portion of Helium nuclei indicates the origin of the Solar Wind while a high portion of oxygen ion indicates the origin of ionosphere. The current standard instrument for an accurate ion species identification is the mass spectrometer. This kind of instrument incorporates time of flight (ToF) technique and is usually very bulky and complex.

Background particle rejection

In some sense similar to the particle identification, background particle rejection allows the detection to focus on the target particles only. Often it is difficult to distinguish the incident particle merely from the detector output signal. Therefore, we design the instrument in such a way that only the wanted particles can enter and trigger the detector response. Many ways can be deployed to achieve this from the simple passive method such as shielding, to complex active method such as anti-coincidence.

1.4 Survey of particle instruments in space

The ISSI scientific report by M.Wuest et al. provides a very good review [35] of the particle instruments in space. Wuest points out that a typical particle instrument consists of these elements shown in Figure 1-15. However, not all elements are necessarily present in a single instrument. The choice depends on the application needs and the instrument technique.

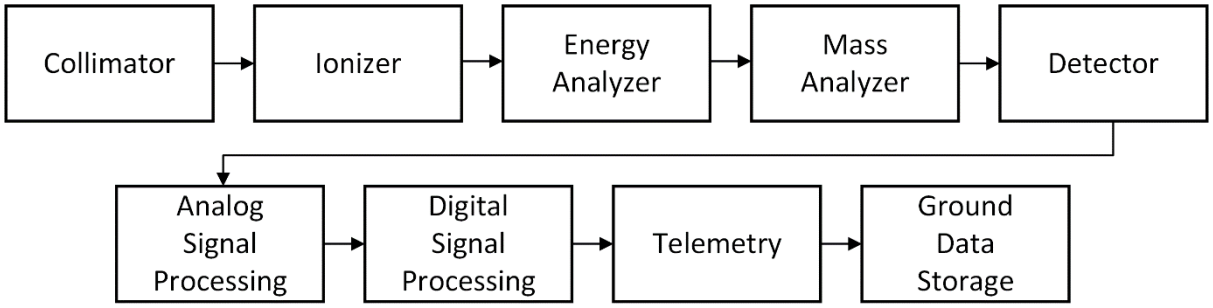


Figure 1-15 The major elements of a typical particle instrument

The collimator is the first mechanical structure that shapes the instrument inlet. It largely defines the instrument field of view (FoV). The exception is instrument with input deflector plates which can achieve FoV larger than the limit of the instrument opening/aperture as the electric field inserted by the deflector plates can bend the incident charged particles' flow/trajectory.

The ionizer is only present in the neutral particle instrument. It converts the neutral particle into ions, which are influenced by an electromagnetic field. The reason to amend the trajectory of the particles is to allow the detector to be not in the line of sight to the instrument inlet. This can reduce the signal contamination by unwanted UV or other sources that can trigger the detector.

The energy analyser can be a solid detector like scintillator or an electrostatic analyser. This stage can convert the particle energy into secondary signals or filter the particles by energy. The purpose is to complete the measurement of particle energy.

The mass analyser is a second analyser section that serves for ion mass discrimination. This can be the total E detection while the previous analyser serves as ΔE detection. Together, the two form a ΔE/E detector system. Also, this can be the stop section of the time-of-flight instrument while the previous analyser serves as the start section.

The detector is the stage that converts the secondary signals or the prime particles into electrical signal, which can be further processed by the readout electronics.

The initial electric signal is always in an analogue form. So, the function of the analogue signal processing is mainly to amplify, shape and condition the signal. The unavoidable electronic noise in this stage shall be kept as low as possible. The bandwidth shall be as large as possible to preserve the raw signal information. The processed analogue signal is then ready to be converted to the digital signal.

The digital signal then can be further processed. In digital form, the signal is much more robust against electric noise interference. Digital processing is very flexible and various filtering algorithm with adjustable parameters can be implemented. The processed signal is then ready for passing to the spacecraft and being relayed to the ground via telemetry.

On the ground, the data can be further processed, e.g. applied complex calibration to correct the data. Lastly, the data is stored and achieved into a database.

For more detail instrument technique discussion, books like Pfaff et al. [36][37] provide a very good coverage.

1.4.1 Langmuir Probes

Langmuir probes (LP) [38][39][40] is one of the first instruments that have been used in the early day's rocket and satellite experiments. It is probably the simplest instrument for plasma measurement. The sensor can be a wire or a cylindrical or of any other geometry metal probe. During measurement, the sensor/probe is inserted into the target plasma and electrically biased with a constant or varying voltage in respect to the plasma potential. The current is measured against different bias voltage to plot an I-V curve. The presence of the plasma will affect the ideal I-V curve in a predicable way, therefore, the plasma parameters like electron density, electron temperature, ion mass and ion density can be derived from the I-V curve characteristics. Figure 1-16 shows the principle block diagram of a typical Langmuir probe and its theoretical I-V curve. Interestingly, this curve is very similar to a semiconductor device, Diode's I-V curve, though their principle is completely different.

The Langmuir probes are suitable for low energy (i.e. a few eV for electrons) plasma measurement. When targeting the low-density space plasma like the magnetosphere, solar wind etc., the probe is usually placed on a boom to reach outside of the Spacecraft sheath, which is characterized by the Debye length. This is to avoid the impact of spacecraft potential on the probe measurement accuracy.

The mathematic description for the interaction between the probe and the plasma is very complex [41][42] and we still have some problem of interpreting the I-V curves observed in real experiments [43][44]. This might be due to the complex real situation or our theoretical understanding is not yet complete. This makes interpreting the I-V curves in-situ and in real time very difficult. The data must be processed on the ground

with some human intervention. This makes this instrument not the ideal candidate for real time warning system in space.

Another limiting factor is the Langmuir has no intrinsic gain. The current is only tens of nA or even down to pA. This means the readout electronics must be very sensitive and extremely low noise (as the noise shall be kept 10 times below the signal to have a good signal to noise ratio.) This makes the instrument very sensitive to the surrounding interference, such as electromagnetic interference.

However, this is a proven instrument with long history. The simple structure means it is very robust and it can survive very high radiation environment. It is still widely used in science mission. Beyond that, on the ground, it can also be found in use in the semiconductor manufacturing process.

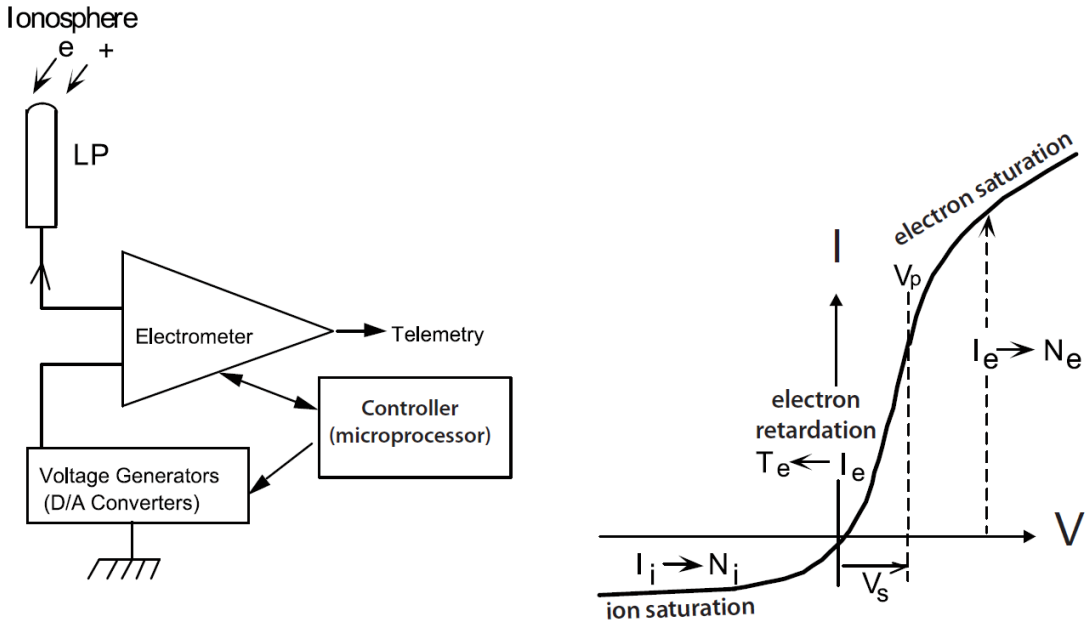


Figure 1-16 The block diagram of the Langmuir probe instrument and the theoretical I-V curve. [38]

1.4.2 Faraday Cups

Faraday cup (FC) is another plasma instrument with a long history. The principle is very simple: an ion hitting the metal collector will deposit a single charge. Therefore, a continuous flow of ions can deposit a continuous charge onto the collector. By measuring the discharge current rate of the collector, the rate of ions can be calculated. The real instrument is more complicated with extra elements to improve the detection efficiency and rejection of unwanted particles. For example, an entry “modulator section”

can filter the ions with right energy. A “suppressor” grid with negative bias to prevent the emission of electrons due to ions striking the collector plate. A split design of collector plate will help to determine the ions flow angle by comparing the current deposited on the equal half. Figure 1-17shows a Faraday cup instrument [45] flown on Wind spacecraft.

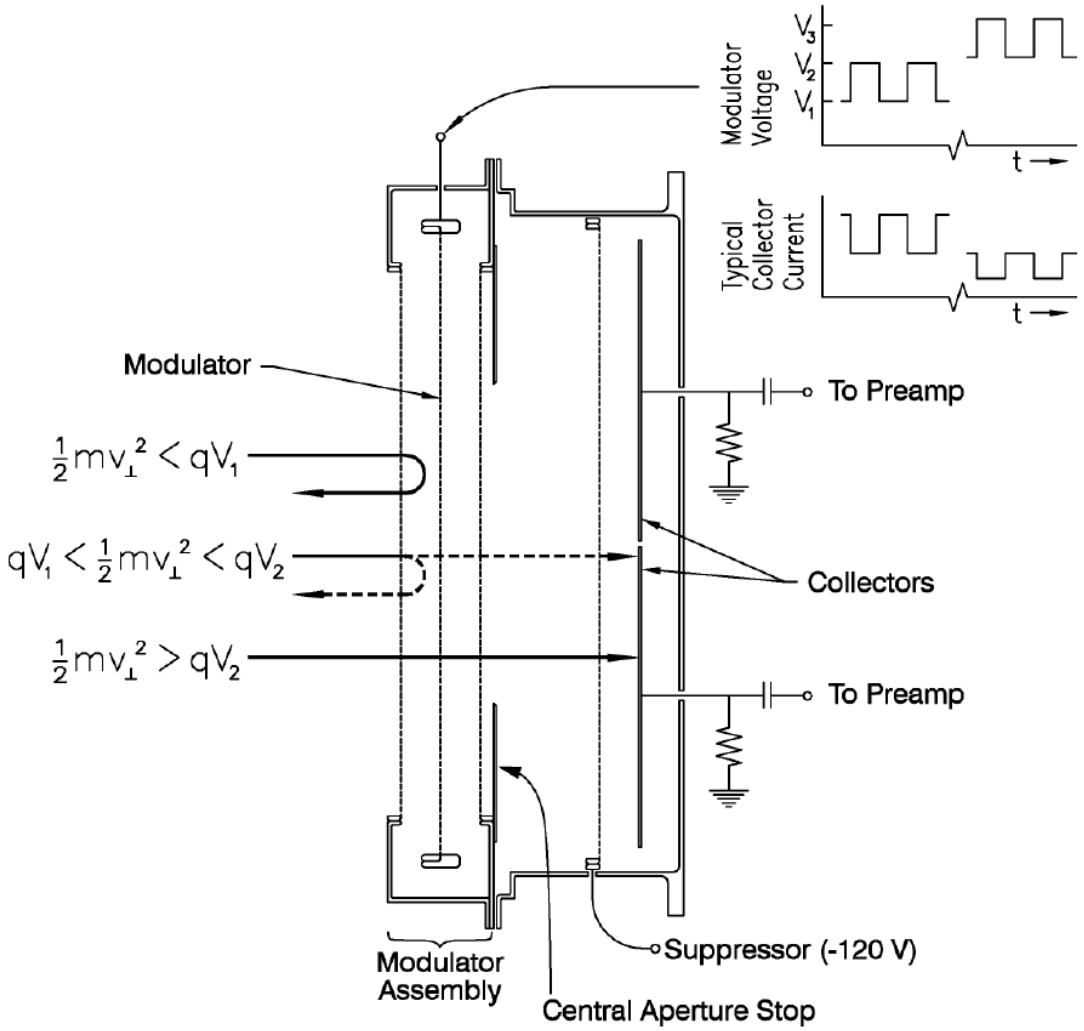


Figure 1-17 The schematic cross section of the Wind FC instrument.

The Faraday cup is commonly used for measuring the Solar Wind. As a direct measurement instrument, there is also no intrinsic gain in the detector system. This means the instrument readout electronics must be very sensitive and low noise like that of Langmuir probe. The difference is that a Faraday cup can provide real time measurement. The output data is simple to interpret, so it can be used to provide in-situ, real time warning for space awareness monitoring.

The Faraday cup is faster than the Langmuir probe as it doesn't need to complete the full I-V scan. However, it still requires time to integrate the ion induced electric current. This limits the instrument response time. This is the side effect of direct measurement as the signal is too weak. However, the benefit is the accuracy. Therefore, a Faraday cup is often used as a calibration instrument in an ion beam facility.

1.4.3 Electrostatic Analyser

The Electrostatic Analyser (ESA) is also used for low energy plasma measurement like Langmuir probe and Faraday cup, while it is much more sensitive and complex. Young[36] provides a good survey of various types of electrostatic analyser. The ESA can be designed in different geometry such as cylindrical curved plate[46] , spherical sector analyser[47] and Top-Hat analysers[48].

The electrostatic analyser uses electric field to steer and select the species and energy of incident charged particles. The choice of electric field rather than magnetic field gives multiple benefits: easy to control, lower mass and lower power. The right particles will be collected by an active detector with intrinsic gain as high as 10^6, e.g. a micro-channel plate (MCP). This allows the detection of single particle. The detector output signal is then processed by the readout electronics. Two modes of operation are commonly used: count mode and PHA (pulse height analysis) mode. In count mode, the signal pulse height is not recorded and only the number of pulses above a threshold is counted. This can keep the electronics simple and further data processing simple. In PHA mode, the pulse height is measured. Assuming the gain of the detector is constant at a given bias and ignoring the gain fluctuation, from the pulse height, the captured particle's fine energy can be extracted. However, PHA mode requires more complex electronics and more stringent noise control.

Top-Hat analyser allows a 360° field of view with fine angle resolution in the azimuth plane. To increase the FoV in the elevation direction, a pair of electrostatic deflection plate can be deployed. Since the deflectors were exposed to sunlight, only positive bias voltages can eliminate photoelectron emission. In addition, two aperture grids are recommended to shield the leakage fields from the deflectors. Figure 1-18 shows the section view of a Top-Hat instrument from MSSL[49] .

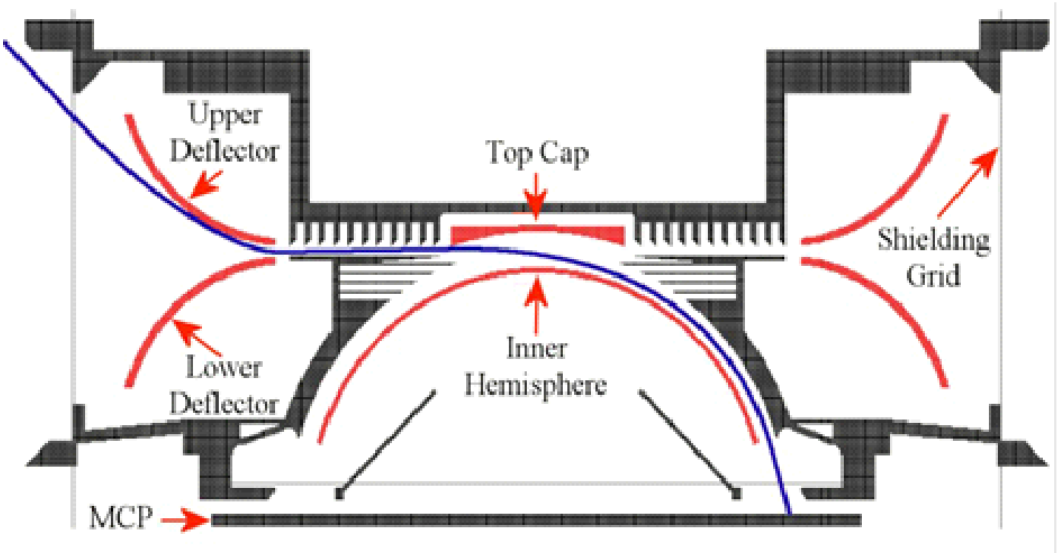


Figure 1-18 MSSL's improved plasma analyser

With extra Time-of-Flight (ToF) elements, the enhanced ESA can measure the ion species. A detail example of such instrument can be found by D.T. Young et. al. [50].

ESA has a capability of detecting single particle and has a dead time limited by the detector (MCP) and readout electronics recovery time. Generally, a detection rate of 1MHz is achievable. The energy range can cover from sub eV to a few 10s keV/nuclei for both electron and ions.

1.4.4 Solid-state detector instruments

Solid-state instruments are used above 20keV energy (as technology evolves quickly, the threshold is being lowered continuously). This threshold is usually defined by the thermal noise of the solid-state detector which shall be below the minimum signal that the deposited energy can generate. Solid-state instruments include both semiconductor (like silicon) detectors and scintillator detectors. The configuration of two layers of solid-state detectors in a stack creates a solid-state detector telescope. Usually, the upper detector is very thin to only absorb a portion of energy, defined as **ΔE**. The lower detector is thick enough to absorb all the residual energy, defined as **E**. The combination of **ΔE** and **E** signal can be used to determine the incident particle energy and even particle species.

The basic structure of two-layer detector can be extended to form a more advanced telescope like RAD[51] on Mars Science Laboratory (MSL) mission. Figure 1-19 shows

the cross section of the RAD instrument, which deploys six different solid-state detectors marked as A to F, which include both silicon detectors and scintillator detectors.

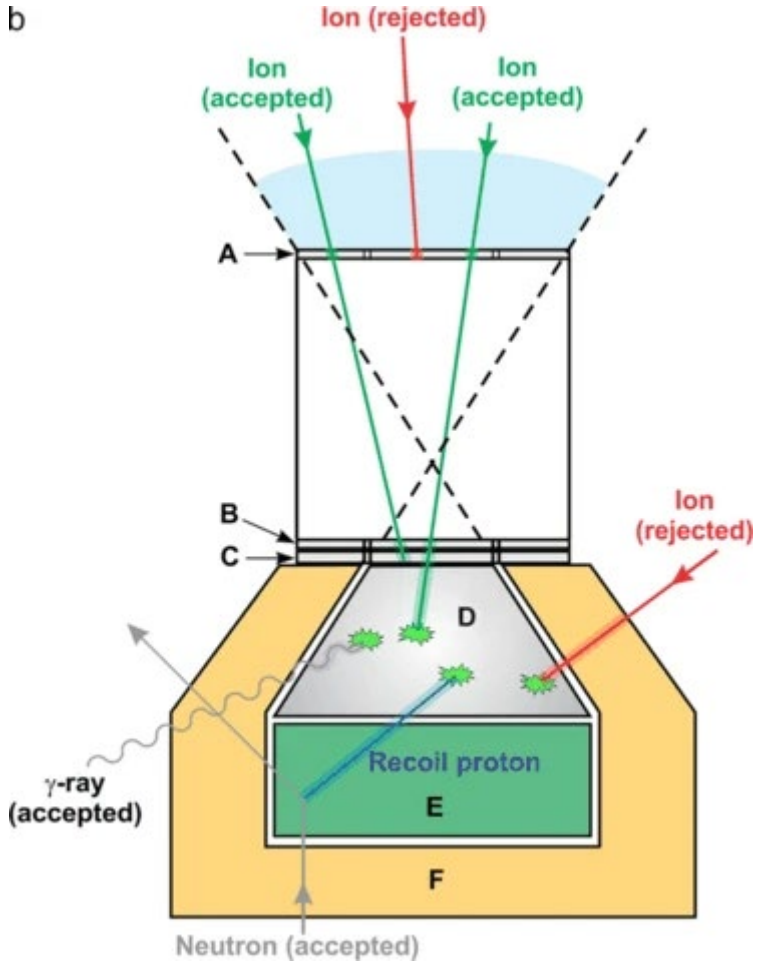


Figure 1-19 the cross-section of RAD instrument [51].

The solid-state instrument can cover a very wide energy range, but as a result, the energy resolution is relatively poor. The instrument response time can be extremely fast as the solid-state detector dead time is in the order of 10s ns or even shorter. The electronics can operate at equivalent bandwidth to preserve the fast detector signal. However, wide bandwidth electronics generally mean higher power consumption. To confine the instrument data volume, the detector signal needs to be heavily processed in the instrument onboard computer. The data product usually includes pulse height analysis (PHA) and histograms.

Table 1-3 lists the survey of some recent miniature solid-state instruments.

Table 1-3 Miniature radiation monitor instruments for space application

Instrument	Principle	Target Species	Energy Range	Application	Features
SREM[52]	Silicon detector	Protons Electrons	10-600MeV 0.5-10MeV	Radiation data collection, radiation warning	3 solid state detectors cover wide range of energy and species
HMRM[53]	APS sensor	Electrons Protons	0.06 – 6MeV 1-500MeV	Radiation monitor, early alert	"Chip sized " ASIC solution
RAD[51]	Silicon detector, scintillator detector	Electrons Protons Heavy ions Neutrons Gamma	~0.2 –12MeV ~10 – 100MeV/nuc ~5 – 100MeV ~5 – 100MeV	Measure particle spectra, dose on Mars and transmission to it	3 silicon PIN diodes 1 CsI(TI) + photon diode 2 Plastic scintillator
LUCID[54]	Hybrid silicon pixel detector	Protons Electrons Neutrons Heavy ions Photons	All energy range *	Cosmic Rays particle species identification	Timepix hybrid sensor APS
RADFET[55]	MOS structure sensitive to ionising radiation	Protons Heavy ions		Local real time dosimetry	Specially designed P-MOSFET [56]
EPD[57]	Solid state Si detectors	Electrons Ions	50keV–4MeV 50keV-500keV	Radiation monitor	Two detectors to differentiate electron and ion population
RADEM[58]	Si-microstrip	Electrons Proton and Ions	100keV-20(40)MeV 2MeV-200MeV	Radiation monitor	Radiation hard design

* Quote from http://www.thelangtonstarcentre.org/index.php/capability

1.5 Detector technology

The core of the particle instrument is the detector. Here follows a brief survey of the detector technology to understand their principle, characteristic and application.

1.5.1 CEM

Channel electron multipliers (CEM) [59] is the first device developed under the concept of continuous-dynode electron multiplier. This contrasts with discrete dynode electron

multiplier like the PMT in that the dynode is formed on a continuous high resistance and high secondary electron yield surface. It is usually constructed as a tube using special glass, such as lead silicate glass, which has a stable property on exposure to air. When in operation, the tube is biased with a high voltage (2 to 4kV) This produces a uniform and strong electric field along the tube. When a primary particle strikes onto the tube entry, a few secondary electrons are emitted from the dynode surface. Driven by the electric field, these secondary electrons continue to strike the opposite surface and more secondary electrons are emitted. The process repeats itself and the number of electrons multiplies after every strike. When the cloud of electrons reaches the exit of the tube, the number can reach 10^6 to $>10^8$. The shape of the tube is often curved to avoid ion feedback. Figure 1-20 shows a typical CEM structure and operation principle.

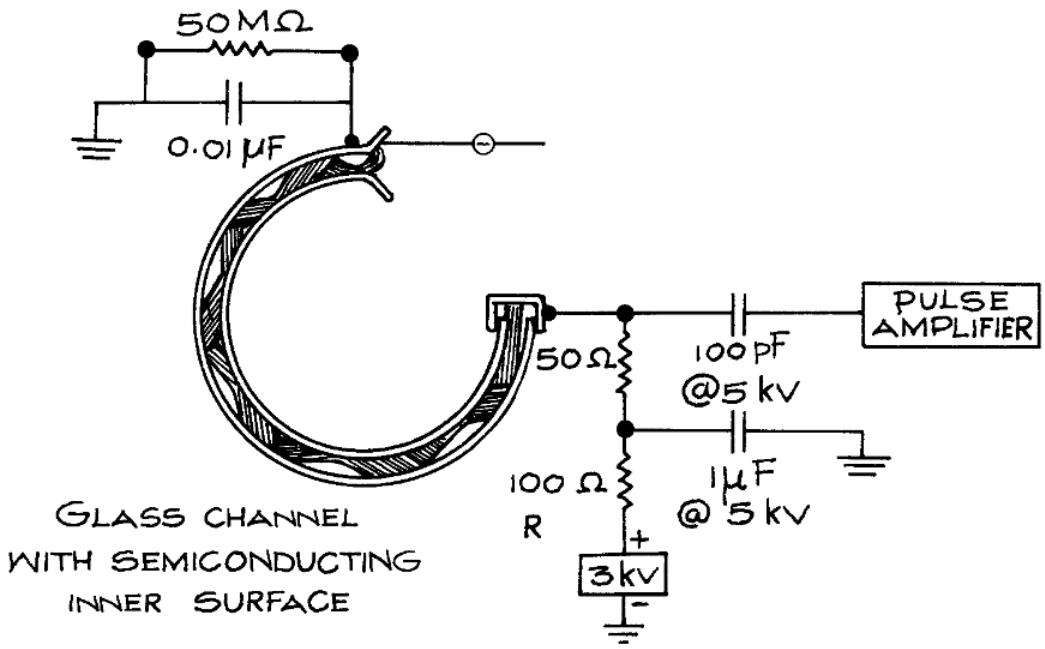


Figure 1-20 A typical configuration of a CEM detection system, from Moore et al. [60]

A CEM is generally operated in count mode with the gain saturated. Uniform gain can be maintained if the pulse current is kept less than 10% of the standing current. This is one of the factors that limit the count rate. Generally, a count rate >1MHz is achievable with background rates < 0.5Hz.

A CEM can also be operated in analogue mode, in which the CEM current is proportional to the particle flux. As the fluctuation of the high gain introduces a very large variation, the CEM output pulse height is not used to measure the particle energy.

A CEM has a different detection efficiency (also called quantum efficiency) depending upon particle species and energy level. Generally, the efficiency is higher for higher energy impact, higher for ions than electrons and very low for photon radiation like UV or gamma ray. Figure 1-21 shows one study [61] of different ion detection efficiency. More can be found on the CEM manufacturer Dr. Sjuts [62] website. The published results show large discrepancy which suggests that the measurement result is very sensitive to the test setup and any subtle difference in configuration parameters.

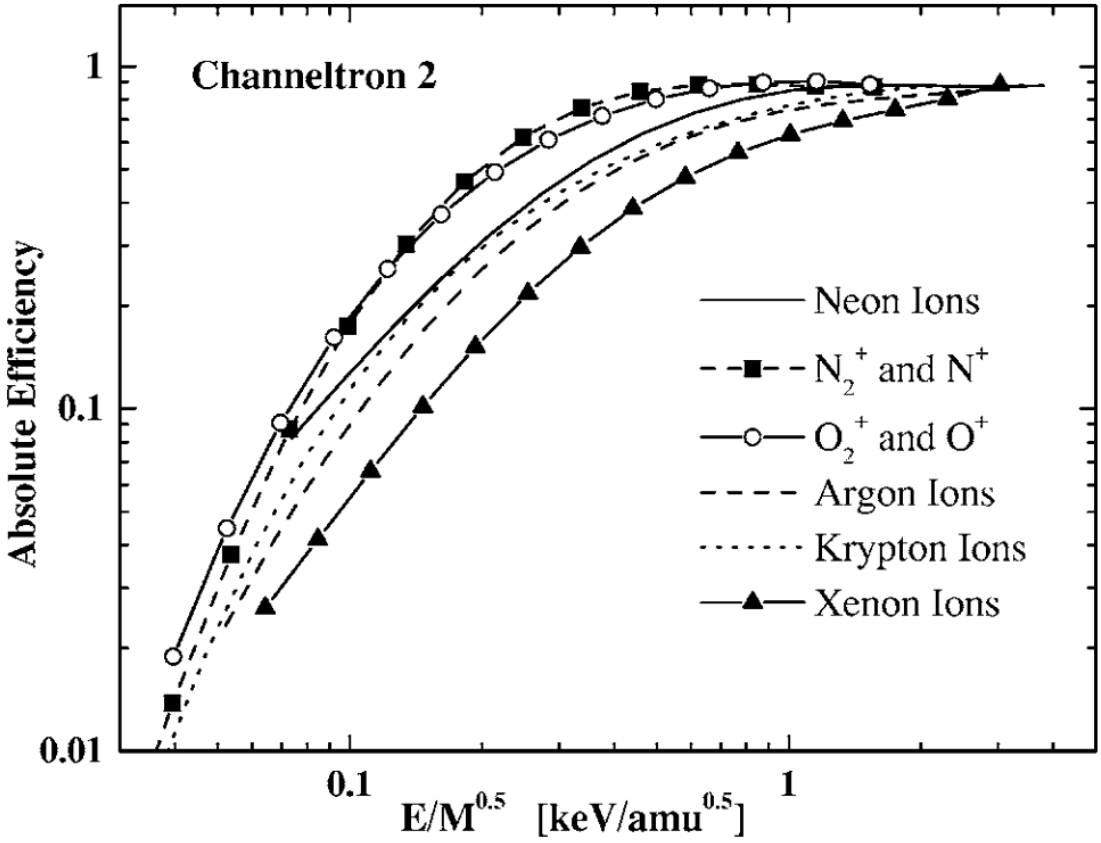


Figure 1-21 Ion detection efficiency measured by a CEM. The x-axis is scaled to energy per square root of the ion mass [61].

1.5.2 MCP

A micro-channel plate (MCP) [63] is an array of micro electron multipliers channels (made of a glass tube) bundled together. The micro channel operates in similar principle as the CEM, but in much smaller size. The typical channel diameters are in the range 10-100μm and the length is less than 1mm. Usually the length to diameter ratio (usually defined as α) is between 40 and 100. This parameter is very important for the performance and operation condition of the micro channel.

All the channels are physically and electrically connected in parallel. The total resistance of the plate is at the order of 10^7-10^9 Ω. Given a plate is made up of tens of

thousands of channels, each channel resistance is as high as 10^{14} Ω. With the channel capacitance, this RC constant determines the channel recovery time, which can be as long as 10s of millisecond. However, as usually only one or a few channels are triggered at a time, the remaining channels are still active. This effectively reduces the average dead time of the total plate.

A single MCP typically has a gain of 10^3 - 10^4 at a bias of ~1000V. This is not high enough. The solution is to stack two identical plates with the channel bias angle in opposite direction. This configuration is named Chevron. The Chevron can reach a total gain as high as 10^7 without suffering from ion feedback. The Z-stack configuration with three plates stacked can yield even higher gain.

Figure 1-22 illustrates the micro-structure and the basic operation principle. Figure 1-23 presents a typical MCP configuration in an electrostatic analyser instrument.

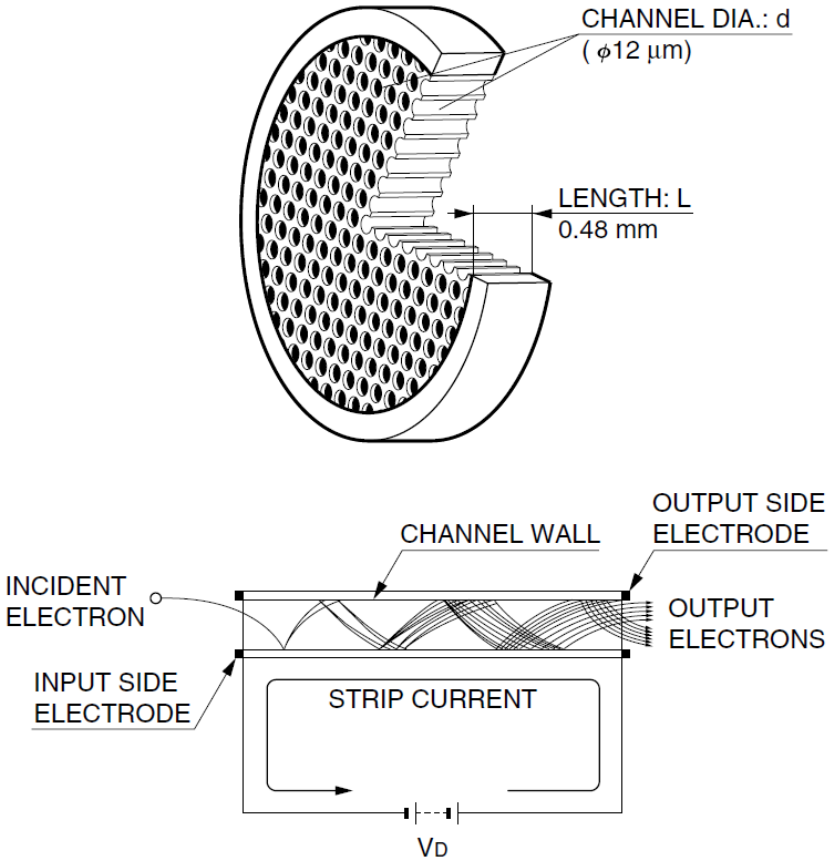


Figure 1-22 MCP structure and operation principle [64]

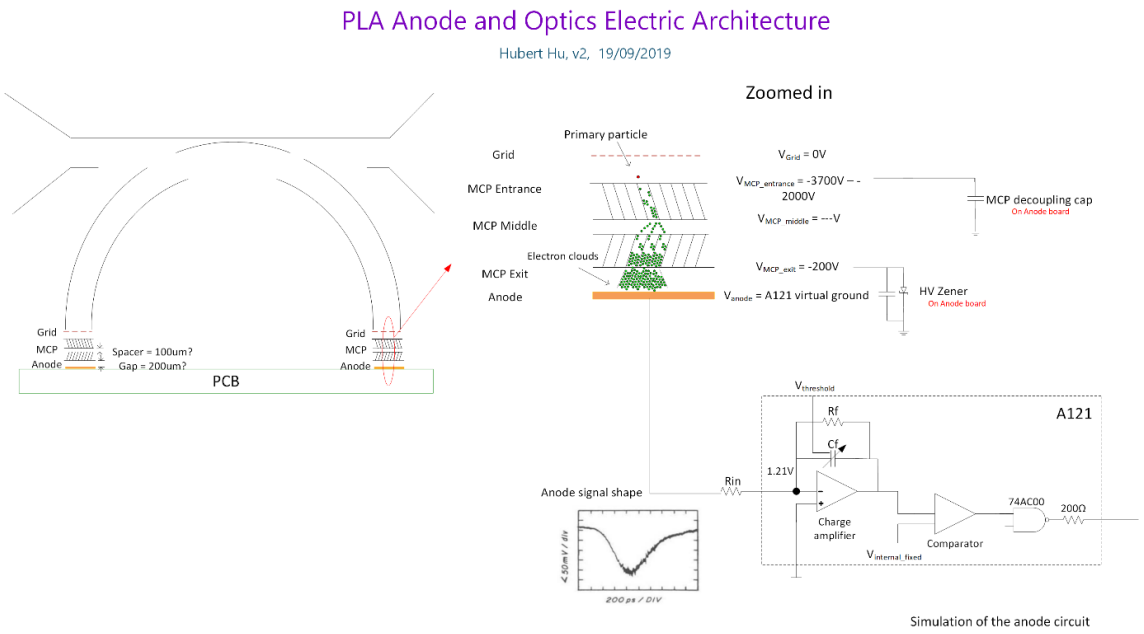


Figure 1-23 The MCP and its readout configuration from a space particle instrument.

The MCP can detect a wide range of particles, including ions, electrons, UV, X-ray and even neutron. The detection efficiency is highly dependent on energy or wavelength. Table 1-4 shows the typical detection efficiency of MCP manufactured by Hamamatsu.

Table 1-4 Typical detection range and efficiency of a MCP [63]

Types of Radiation	Energy or Wavelength	Detection Efficiency (%)
Electron	0.2keV to 2keV	50 to 85
	2keV to 50keV	10 to 60
Ion(H^+, He^+,Ar^+)	0.5keV to 2keV	5 to 58
	2keV to 50keV	60 to 85
	50keV to 200keV	4 to 60
UV	300 Å to 1100Å	5 to 15
	1100 Å to 1500 Å	1 to 5
Soft X-ray	2 Å to 50 Å	5 to 15
Hard X-ray	0.12 Å to 0.2 Å	Up to 1
High energy particle(π,ρ)	1GeV to 10GeV	Up to 95
Neutron	2.5MeV to 14MeV	0.14 to 0.64

From the table, we can clearly see that the best operation range for electron detection is 0.2keV to 30keV and for ion is 2keV to 50keV. Compared to CEM, MCP is less efficient in ion detection.

1.5.3 Silicon Solid-State Detectors

Silicon solid-state detectors (SSD) [65] are built using pure silicon crystals. As a particle traverses the crystal, it ionizes along the path and creates many electron-hole pairs. The electron is promoted from the valence band to conduction band, where it can freely move. The hole created by the missing electron still locates in the valance band, but it can also “move” under the influence of the electric field in the opposite direction to the electron in the conduction band. The combination of movements forms an electric signal (a pulse) that can be detected.

The magnitude of the signal is proportional to the number of pairs that is created. Further, the number of pairs is proportional to the energy of the incident particle. This creates a link between the signal magnitude and the particle energy. Over a good energy range, the relationship is linear. In silicon, it requires 3.6eV to create an electron-hole pair for its bandgap of 1.12eV. This determines the principle energy resolution of SSD at 3.6eV. However, such a signal is well below the thermal noise and therefore it is not detectable. An important fact that limits the minimum energy that SSD can detect is a layer of thin material on the surface of the SSD. The layer doesn’t contribute to the signal generation, so it is named as dead layer. A particle must penetrate this layer to enter the active volume of the SSD. With the current technology, an SSD with typical dead layer can reach about 20keV. This is much higher than MCP or CEM’s lowest detectable energy.

An SSD is very compact in its form. The thickness is under 1mm. This means for higher energy particle, it doesn’t have enough depth to stop the particle and absorb the full energy. However, partial energy can still be absorbed and detected. Therefore, in a high energy detector telescope, an SSD is often used as a ΔE detector.

The electric field over the detection volume is moderate and it takes 10-20ns for the electrons and holes to move towards electrodes for a distance 300μm. This means the response time of an SSD is slower than those detectors with intrinsic gain who has a very high electric field to accelerate the ionized electrons and holes.

Damage by ionizing and non-ionizing radiation is the major limitation for the use and life of an SSD in space. Part of the surface damage caused by ionization from charged particles and X-ray photons are recoverable by annealing at room or elevated temperature. Damage by excessive current (over heating) can be irreversible. Non-ionizing

damage like the displacement damage is an accumulated effect and is hardly reversible.

1.5.4 Scintillator

A scintillator is a traditional material, which can convert a photon or a charged particle to a large number of secondary photons in visible or UV range. The initial interaction process with the particle in some form is similar to that of an SSD, but because of different material conductivity, the later process goes in different direction. The silicon is a semi-conductive material, in which the ionized electron-hole pairs can be collected by an external electric field, while the scintillator is an insulator material, in which the electron-hole pairs can't move. Eventually the promoted electrons lose energy in the form of photon emission and return to the valence band. The emitted photon needs to be collected by a photon sensor and converted to an electrical signal. This two-stage detection means much poorer energy resolution and overall detection efficiency when compared to SSD.

More about the scintillator detector is discussed in depth in chapter 2.

1.5.5 PMT

A photomultiplier tube (PMT)[66] is a traditional single photon detector with intrinsic gain at the order of 10^6 – 10^7. It belongs to the type of discrete dynode electron multiplier. Figure 1-24 shows the typical structure of a PMT. The first stage of the PMT is called photocathode, which converts the photon to a few photoelectrons. The material choice of the cathode determines the spectral range and quantum efficiency. This crucial stage dominants the overall PMT performance like spectrum response and gain. Follow that is the multiple stage of dynodes. The dynodes are biased with a uniform high voltage and are positioned in the way that the secondary electrons can only emit towards the next dynode. The dynodes stage basically serves as a multiplier. The last stage is anode, which collects the electron clouds from the last dynode. The output signal from anode is a negative pulse with Gaussian shape and area equal to the total charge.

A PMT is sensitive to magnetic field as the trajectory of the secondary electron can be directed away from the target path under the field. A dark current exists in a PMT even when it is operated in a completely dark condition. The total dark current increases with an increasing supply voltage and the change is exponential. Different sources dominate the level of dark current at different bias voltage region. But overall, the level of dark current is very low ranging from 10^{-10} to 10^{-6}A.

A PMT has a superb time response with typical pulse width (FWHM) around a few ns. The response time is determined primarily by the transit time that takes the initial photoelectrons liberated in the cathode to reach the anode. Hence, different PMT structure design particularly the dynode type and position have a major impact on timing.

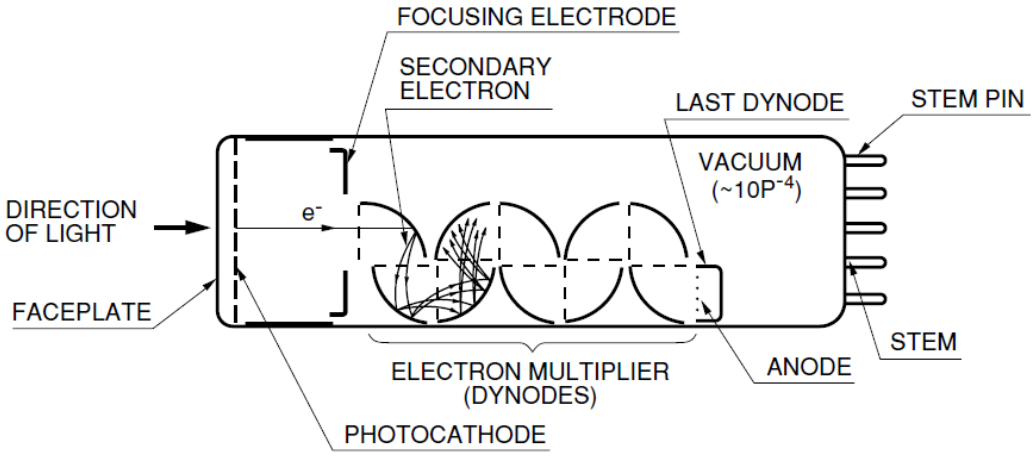


Figure 1-24 Typical PMT structure [66]

1.5.6 SiPM

A silicon photon multiplier (SiPM) is an array of avalanche photon diodes operating in Geiger-Mode. Each pixel operates independently, like the microchannel of an MCP. The electrodes of all the pixels are connected in parallel and to a common supply and output. Hence, there is only one signal output from a SiPM, which is the sum of all the pixel outputs. This is different to an MCP, which maintains individual outputs for each microchannel. The common connection arrangement of a SiPM makes it easy to interface with but the down side is the loss of spatial/imaging function as the information of which pixel is triggered is lost.

The SiPM development originates from a Russian group [67] from early 2000s. After more than just a decade of efforts, the process and design of the device have progressed significantly. Almost all the performance characteristics of the device, from bias voltage, gain, dark count, efficiency etc. have been improved.

Many paper and technique notes have studied the principle and characteristics of SiPM in detail [68][69] [70][71] [72]. Here we discuss them briefly to provide a quick insight.

The architecture of SiPM pixel varies from different design by different manufacturers. Figure 1-25 presents a typical architecture. This architecture is called n^+-p- π-p^+. From the top, the Si resistor (red) is the quenching resistor. The Al-conductor (small yellow block on top) is an electrode for bias voltage. The two form the optical dead region. The SiO_2 layer is an optical window, which is transparent to the incident photon. The guard ring is for reducing the optical cross-talk. The n^+ (blue) region is a heavily doped n-type layer. It is made very thin (about 0.1-1.5μm) to minimize Auger and SHR (Shockley-Read-Hall) recombination. The p (green) region is a thin p-type layer next to the n+ layer. The junction of the two regions forms a depletion region of about 1μm thick. The π layer (white) is a thick and very lightly doped p-type layer (almost intrinsic). The p^+ substrate (big yellow block at the bottom) is another heavily doped p-type layer. The p, π and p^+ layer help to modify the electric field distribution across the structure. The field strength is at maximum at the n^+p junction and decreases slowly through the π layer until it vanishes in the p^+ layer. Figure 1-26 shows the configuration of the electric field along the x-axis.

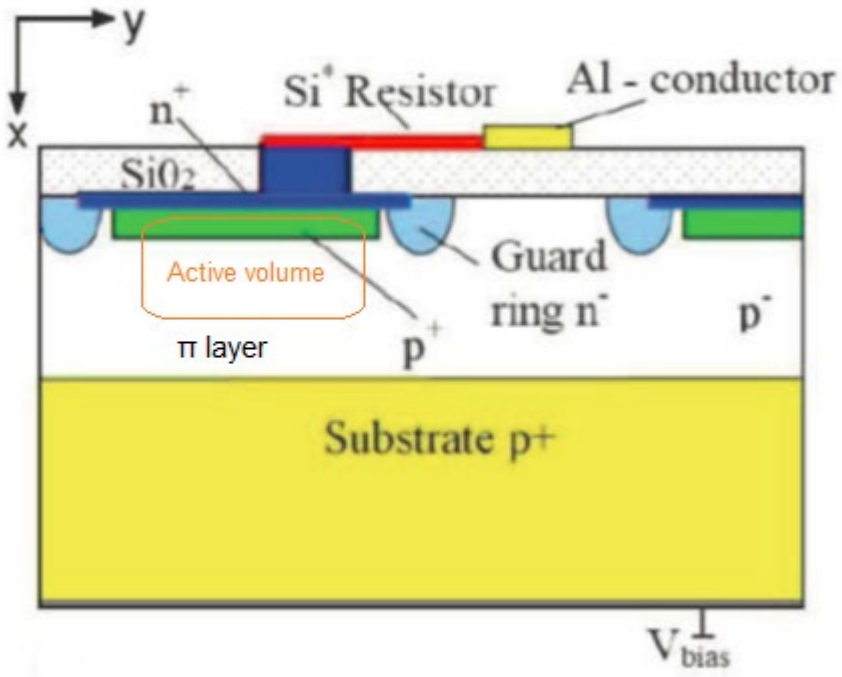


Figure 1-25 Schematic structure of a typical avalanche microcell of SiPM [70]

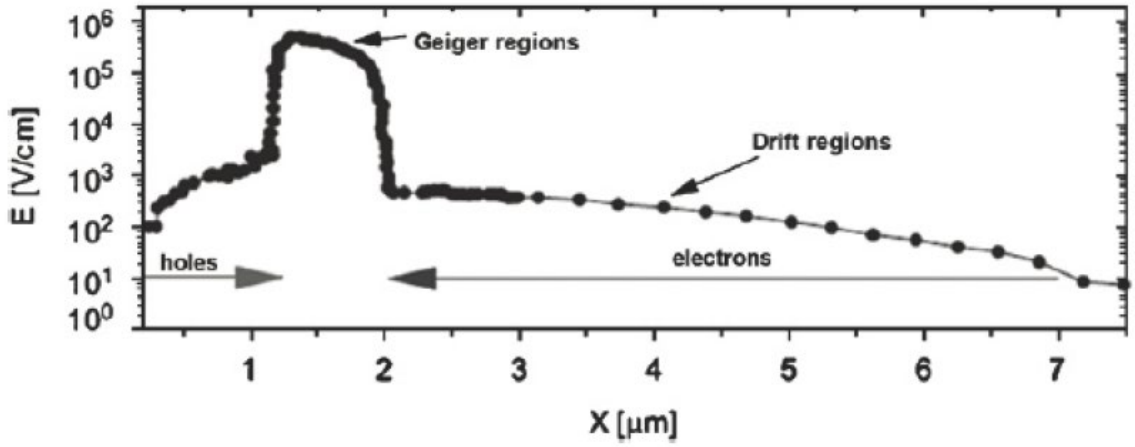


Figure 1-26 Electric field distribution across the structure [70]

The absorption of photons at wavelength of around 400nm takes place in the π-layer. Under the nearly uniform electric field in π-layer, the electron-hole pairs are separated. Electrons drift towards the positive electrode and holes drift towards the negative/ground electrode. The electrons pass through the p region where field becomes stronger. The electrons are accelerated under the stronger field. When electrons reach the n^+p junction, where the field becomes abruptly high, the avalanche process is triggered. Millions of new electron-hole pairs are created. Thus, a large current flow is formed. The avalanche process will be terminated by the voltage drop across the quenching resistor as current surges. The termination point can be set by the value of

the quenching resistor and the bias voltage. Usually, the amplification of the initial photoelectron can reach the order of 10^6 before termination.

A SiPM has a higher noise level than a PMT and it has many different noise sources. In count mode, the thermally triggered counts are the dominant noise source. This type of counts is often defined as dark count. The typical temperature coefficient for dark count is -5%/°C. So, every drop of 10°C can see the reduction of dark count rate by a factor of 2. After-pulse and cross-talk are the other two sources of noise. After-pulse is the delayed event by the trapped electron in a crystal defect, impurities or generation-recombination centre. Hence, better process and quality of crystal can reduce the after-pulse rate. Cross-talk is because some of the few emitted photons in the avalanche process may successfully travel to another pixel and initiate an avalanche process there. Better optical isolation between the pixels such as an optical trench, can reduce this effect.

An SiPM has even better time response than a PMT. This is because all the multiplication processes occur within a very small region (a few μm), and therefore, it takes less time for the ionized electron-hole pairs to travel to the electrode.

1.5.7 Summary

We can see all the detectors can be categorized into two general classes. One has a very high gain and the other has no intrinsic gain. The detectors with high gain are capable to detect a single particle or photon, but due to the fluctuation of the gain or even saturation, the detectors mostly operate in count mode. They don’t have a linear or interpretable energy response. CEM, MCP, PMT and SiPM belong to such class. The detectors with no intrinsic gain have a very good linear energy response, but without any gain, the detector signal is very small so it is much more susceptible to background noise, like the thermal noise. The response time is also slower due to the lack of strong electric field to accelerate the secondary electrons, which forms the output signal. SSD and scintillator are in this class.

Another categorization method is to divide the detectors by the output and input signal type. CEM, MCP and SSD are one class that can accept a wide range of charged particles or energetic photon and output electric signal directly. Scintillators is another

class which output in the form of photon. As photon cannot be processed by the electronic system directly, another photon to electron conversion sensor is required. Of the range of photon sensors, PMT and SiPM have the highest sensitivity due to their huge gain. So, it is common to find a scintillator coupling to PMT or SiPM to form a complete detector system. The second class of the detector system has much poorer detection efficiency compared to the first class.

Detectors with high gain generally operate with a high bias voltage. However, due to the downsize of the detector, the required voltage scales down significantly. CEM, MCP and PMTs have an acceleration distance in the order of mm. They require bias voltage of the order of 1000V. A SiPM's depletion region is in the order of μm. It requires bias voltage at only a few 10s volts. This is because to reach the same level of electric field strength, the shorter distance requires less voltage difference across it. To maintain a sustainable multiplication, the electric field strength needs to reach 10^5 – 10^6 V/cm. This seems to be a common level for all the detector technologies.

The general trend for detector development is smaller in size, better efficiency and faster in response. However, some of these features usually come at the cost of increasing noise or higher power. There is yet no detector that can outperform the others in every aspect, which is probably intrinsically impossible. Therefore, the choice of the detector remains highly dependent on the application.

Table 1-5 provides a general overview of the surveyed detectors.

Table 1-5 Overview of the detector performance

Detector	Intrinsic gain	Bias Volt (V)	Re-sponse time (ns)	Energy res-olution	Operation mode	Background noise	Detect species	Energy range	Efficiency	Operation Environ-ment
CEM	10^6-10^8	2000-4000	1-10	No	Saturated Counting	<0.5s^{-1}	e-, i+ uv	>100eV e- >10keV i+	25% - 85% e- 50% - 90% i+ 0.1% -10% uv	Vacuum <10^{-5}mBar
MCP	10^6	2000-4000	0.2	No	Saturated Counting	<1 $cm^{-2}s^{-1}$	e- i+ uv	0.2 to 30keV e- 2 to 50keV i+ 30-150nm	10-60% 5-85%	Vacuum <10^{-5}mBar
SSD	No	100 – 300	10-100	<200eV (e-) ~30keV (i+)	Linear	Low	e- i+ uv	>10keV (e-) <14MeV (p)		Ambient
Scintilla-tor	No	No	10-1000	moderate	Linear	Almost zero	e- i+ x/γ n			Ambient
PMT	10^6	1000	1	No	Saturated Counting	10^{-7} - 10^{-8}A	Photon	Up to 650nm	<30%	Ambient
SiPM	10^6	<100	<0.1	No	Saturated Counting	<1MHz	Photon	300nm – 800nm	<40%	Ambient

Note:

1. the numbers quoted in the table are all typical values collected from various documents, some of which might be out of date. Also, individual product will differ a lot from the typical values. So, these numbers are here for comparison purpose only.
2. e- means electron, i+ means ions, p means proton. Ambient means detector can be operated in both ambient and vacuum.

1.6 System requirement and the proposed solution

The survey of space environment near Earth helps to identify the target charged particles or energetic photon's energy range, flux and abundance. It is beyond a single instrument's capability to cover the full range. On the other hand, different applications have different target ranges. For example, only high energy particles like GCRs, SEPs or the high energy charged particles from the radiation belts impose threat to the safety of satellite and astronauts. An application to monitor such particles only need to focus on the medium to high energy range. This means that the scintillators and SSDs are the best detector choice. The average flux of high energy particles is low, but when a solar event occurs, the burst of high energy particles can reach very high intensity in a short period. This means the instrument shall have fast time response and wide dynamic range so that it doesn't saturate under intense flux. The instrument can operate in a reduced capability mode in a quiet time to preserve power and it automatically switches to a full capability mode as soon as detecting the onset of a large event. This arrangement can lower the average instrument power. The instrument shall be able to withstand extreme space weather event and remain functional during such event even if the host spacecraft is in partial shut-down.

By studying the data from SPENVIS toolkit, we further simplify Table 1-1 as

Table 1-6 Refined Near-Earth Environment Specification

Particle	Source	Min E (MeV)	Max E (MeV)	Flux coverage	Avg Flux ($cm^{-2}sr^{-1}s^{-1}$)	Model
Proton	Trapped	0.1	10	99.99%	197	AP8
Electron	Trapped	0.04	1	99.74%	14218	AE8
Proton	Solar	0.1	10	99.96%	169	ESP-PSYCHIC
Proton	Solar	0.1	10	99.78%	4634	CREME96 (wrt. week)
Ion(He)	GCR	1	100	30.56%	0.1076	ISO15390
Gamma*	Solar, GCR					

* No gamma data is found in SPENVIS database.

It shall be possible to design an instrument to cover most of the specification in Table 1-6. In addition, the instrument shall have:

- Good energy resolution, i.e. <10%

- Wide solid angle, i.e. 2π
- A degree of particle identification/background rejection capability
- Moderate spatial resolution
- Good temporal resolution

Both SSD and scintillator detectors can meet the target energy range as specified in Table 1-6 with sufficient response capability for the peak flux. We opt for scintillator as it is more robust to radiation damage and it can be made thicker to detect higher energies. Our choice for the scintillation photon collection is SiPM. With SiPM, the readout electronics and bias voltage power supply can be much simplified. This saves the instrument resources in both power and mass. The combined Scintillator+SiPM detector platform can be realized in a very compact format. Many experiments using the same platform for different applications have been published and the results are very encouraging. Table 1-7 shows the survey of a few such experiments with key information listed.

Table 1-7 some similar study on the Scintillator+SiPM solution

Application	Scintillator	SiPM	Energy Range	Energy Resolution	Timing res.	Ref.
Ground high energy experiment	LYSO	Hamamatsu-SiPM	300MeV	6MeV	350ps	[73]
Gamma spectrometer	CsI(Tl)	SensL	~7.1MeV	7.88%@662keV		[74]
PET	LYSO NaI(Tl) CsI(Tl)	Hamamatsu-SiPM		~14.4keV@511keV ~27.5%@122keV ~32.7%@122leV		[75]
Gamma spectrometer	CsI(Tl)	SensL	30-3000keV	~6%@662keV		[76]
Nuclear physics experiment	Plastic	Hamamatsu-SiPM	Very high		~1ns	[77], [78]

Chapter 2 Scintillator

2.1 Introduction

This chapter will discuss the design of the stacked scintillator detector module. First, the basic category and theory of the scintillator is introduced. Detail underlying physics can be found in [79]. For particle identification, in-depth knowledge of the scintillation mechanism in relation to different particle species are helpful. Also, knowing the physical process of scintillation, the cause of non-linearity can be understood. This knowledge can help to improve signal processing and overall instrument design.

The second part focuses on the discussion of the scintillator module we designed and manufactured for this project. The design takes a sandwiched approach (commonly also known as phoswich), of which two different scintillators are glued together. The configuration can provide good particle identification of different species within an energy range. This concept is proven by simulation and we can also find support from the literature[80]. We have gone through three major iterations of the scintillator design. The driving reason for the change of design is the manufacturing capability. Each iteration has been validated with a range of simulation tools to confirm the performance. The penetration depth/energy range can be calculated with data from NIST web tool [81][82]. SRIM[83] is also a very easy to use tool for simulating the charged particle interaction with matter. It provides visual illustration of some particle trajectories and the output also shows the lateral dispersion. Geant4[84] is the universal tool that can cover all the particle species in any material and geometry, but it requires dedicated programming. The summary of the simulated detector performance is presented in this chapter. More detail about the simulation setup and results, particularly for the first scintillator module design, are discussed in Chapter 3. By simulation, we compared the geometry impact on the photon collection efficiency and we also tested the idea of particle identification.

The last section proposes the concept of a new instrument, which is a hemisphere structure holding a group of cylinder geometry scintillator modules to cover 2π field of view. This design has a good angular resolution and a good redundancy benefiting from the modular design approach. This concept design is the first step towards a flight instrument prototype. The design is flexible and expandable, allowing it to be tailored to different mission needs.

2.2 Scintillator

The property of a material that can emit photon under a mechanism of energization is called luminescence. If the mechanism is via chemical reaction, it is called chemiluminescence. If the mechanism is via an incident energetic photon, it is called photoluminescence. Photoluminescence is further divided into fluorescence and phosphorescence. The main difference is the time that it takes to release photon. Fluorescence releases the photon immediately while phosphorescence can store the energy for some time and release the photon later. This is because the phosphorescence can go through a middle triplet stage before returning to the ground state. The principle is illustrated by the Jablonski diagram in Figure 2-1.

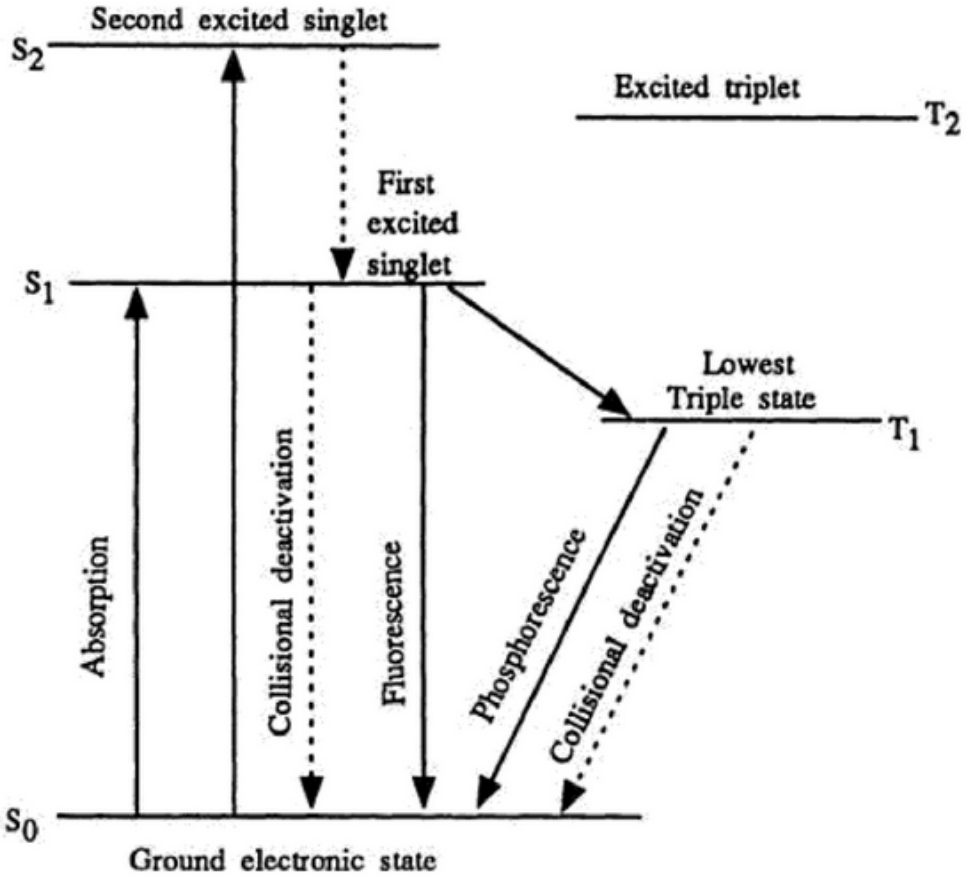


Figure 2-1 The Jablonski diagram for the energy transfer in a diatomic molecule [85]

Scintillation is a type of fluorescence. A scintillator material can absorb energy from either a charged particle or energetic photon. The emitted photon has a wide range of wavelengths. The rule is the absorbed energy must be higher than the emitted photon energy to compensate for the energy loss in the scintillation process.

Figure 2-2 shows the hierarchy of the terms we have discussed.

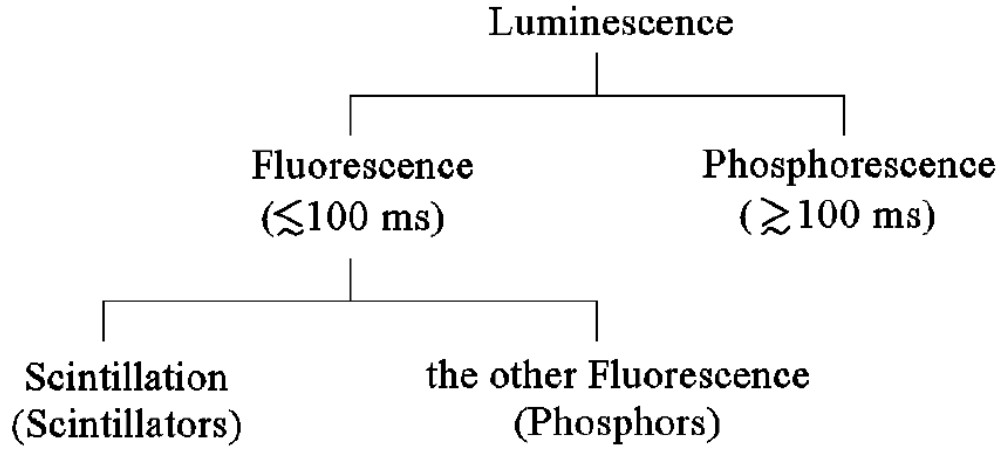


Figure 2-2 the diagram shows the relationship of the terms for different types of luminescence.

The most important characteristics of scintillation materials are listed below:

- Scintillation efficiency
- Photon yield
- Linearity of the light response
- Stopping power
- Response time
- Emission spectrum, (aim to match the photon detector)
- Chemical and mechanical stability
- Radiation hardness
- Cost (a concern for large scale)

There are abundant types of scintillator material available, but none is perfect in all aspects. So, the choice of scintillator shall be tailored to the application needs. For example, in this thesis study, the scintillator aims to be used in space to monitor the radiation environment. This target means the chosen scintillator shall be radiation hard and suitable for use in a vacuum. The emission spectrum of the scintillator shall match the absorption spectrum of the chosen SiPM. Short decay time is desired to match the fast timing of SiPM. The space environment contains full spectrum of X/γ-ray and full range of high energy particles (i.e. solar energetic particles, cosmic high energy particles, high energy electron/protons in the trapped Van-Allen Belts etc.). It is impossible to cover the full spectrum and range, so the chosen scintillator shall be designed to focus on the key part that matters to the target science study.

To understand the characteristics of the scintillator and help to choose the right scintillator, we first provide a brief review of the fundamentals of the scintillator. We will keep the discussion to an essential minimum. For more detail and thorough discussion, these are very useful reviews [86][87][88][89].

2.2.1 Interaction with energetic photon and charged particles

Scintillator can interact with a wide range of particles/photons, from energetic photon (including X- and γ-rays), charged particle to neutron. All the interactions lead to photon emission, but the intermediate process is different. The process follows the principle of how the photon or particle interacts with matter.

In the case of energetic photons, the interaction is covered by three major processes: *photoelectric absorption*, *Compton scattering*, and *pair production*. Each dominates in

a different photon energy range with the boundary overlapping. The *photoelectric absorption* dominates low energy section. *Compton scattering* dominates medium energy section. And *pair production* dominates high energy section. For the energy range that our application focuses on, the *photoelectric absorption* dominates. In the photoelectric absorption process, the photon energy is absorbed by the host atom completely. In return, an energetic *photoelectron* is ejected by the atom. This “hot” electron often comes from the most tightly bound K shell of the atom. Generally, one *photoelectron* is generated per event. This “hot” *photoelectron* can continue to excite many secondary electrons via Coulomb scattering until its energy is exhausted. This Coulomb scattering process is very similar to the way a charged electron directly interacts with the crystal. The similarity is observed in the Geant4 simulation, which shows the scattering effect in the crystal response to both γ-ray and charged electron.

In the case of charged particles, particularly the charged, heavy particles, such as proton, alpha and other heavier ions, the interaction with scintillator is a continuous loss of its kinetic energy via mostly Coulomb scattering. The direct impact of incident particles to the nucleus makes up a very small portion due to the low possibility. The lost/deposited energy will create a string of secondary electrons. These secondary electrons will eventually recombine with holes and release the energy in the form of scintillation photon. Figure 2-3 from [88] compares the processes of the charged particle and energetic photon in an illustration. Note, the energy loss of the heavy charged particle along its track is continuous but not linear. Most of the energy will be dumped near the end of the track. The differential energy loss along the track is described by the *Bragg curve* [90].

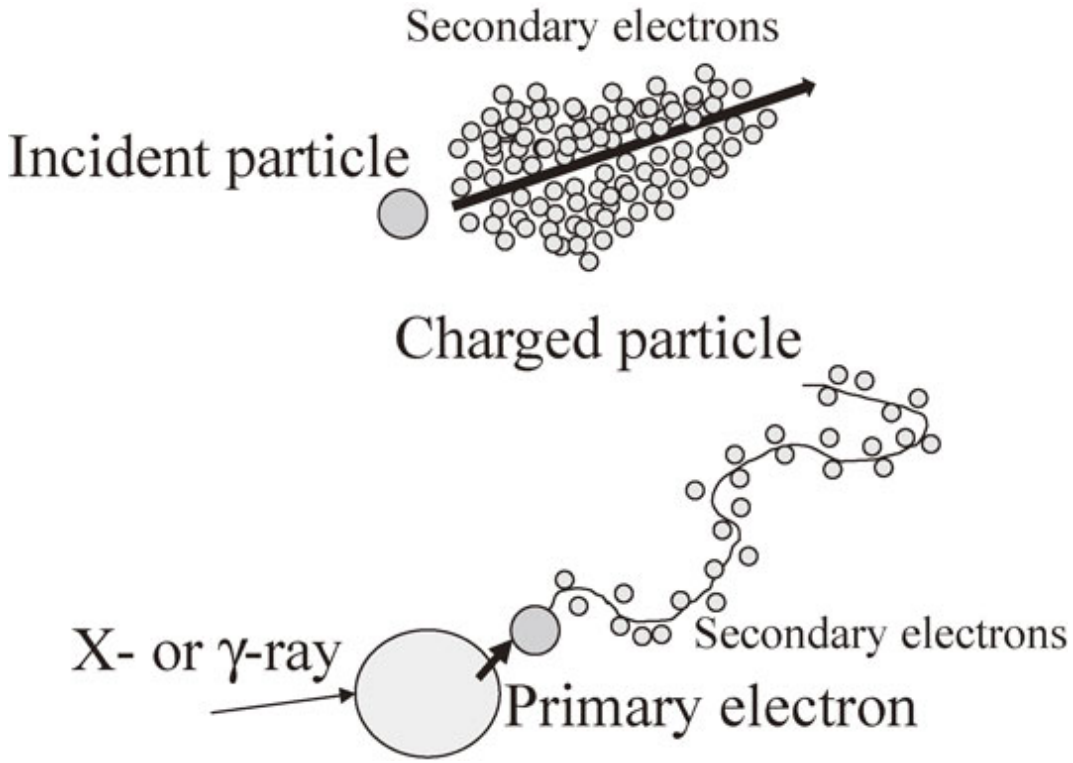


Figure 2-3 the illustration of the interaction between the scintillator and different incident species: charged particle and energetic photon (reproduced)

Since the interaction process is different, the way to describe the interaction in mathematic form is also different. For the energetic photon, it is governed by the probability that any of the three processes (**photoelectric absorption, Compton scattering,** and **pair production**) occur within the volume of the absorber. If one occurs, the primary photon deposits its energy and disappear or is scattered (depends on the process). If not, the photon leaves the absorber unchanged. The probability is affected by the absorber property, thickness and the primary photon energy. For an initial flux of I_0, the number of photons that pass through the crystal (with thickness t), I can be described as:

$$I = I_0 e^{-\mu t} \tag{2.1}$$

where μ is called the **linear attenuation coefficient**.

The **linear attenuation coefficient** is limited in use as it varies with the absorber density. So, a modified term called **mass attenuation coefficient** is used more widely and is defined as:

$$mass\_attenuation\_coefficient = \frac{\mu}{\rho} \tag{2.2}$$

Where ρ is the density.

Then, eq.(2.1) can be expressed as:

$$I = I_0 e^{-(\mu/\rho)\rho t} \tag{2.3}$$

In summary, for energetic photon, once we know the **mass attenuation coefficient** and the absorber thickness, we can calculate the detection efficiency in the form of the percentage of flux that will be absorbed.

For the heavy charged particle, the continuous energy loss along the track in the absorber in unit length is defined as *linear stopping power* ***S*** in the form:

$$S = -\frac{dE}{dx} \tag{2.4}$$

The negative sign indicates the energy is diminishing. The calculation of the stopping power for heavy ion in a defined absorber can be precisely described by the famous **Bethe-Bloch formula** [91]**:**

$$S = -\frac{dE}{dx} = Kz^2 \frac{Z}{A}\frac{1}{\beta^2}[\frac{1}{2}\ln\frac{2m_e c^2\beta^2\gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}] \tag{2.5}$$

Where,

A: atomic mass of absorber; $K/A = 4\pi N_A r_e^2 m_e c^2 / A = 0.307075 MeV g^{-1} cm^2$ for A = 1g mol^{-1}

z: atomic number of incident particle

Z: atomic number of absorber

W_{max}: max energy transfer in a single collision with $W_{\max} = \frac{2m_e c^2\beta^2\gamma^2}{1 + 2\gamma m_e / M + (m_e / M)^2}$

δ(βγ): density effect correction to ionisation loss.

r_e: classical electron radius = 2.817x10^{-13}cm

m_e: electron mass

I: mean excitation energy

β: v/c of the incident particle

γ = $1/\sqrt{1-\beta^2}$

x = ρs, surface density or mass thickness, with unit g/cm^2, where **s** is the length

dE/dx has the unit MeVcm^2/g.

This formula is valid in the region $0.1 \le \beta\gamma \le 1000$ with an accuracy of a few percent.

For light charged particle as electron, the energy loss due to ionization and excitation (the “collisional loss”) can be described with a modified eq. (2.5):

$$\left\langle -\frac{dE}{dx} \right\rangle = \frac{1}{2} K \frac{Z}{A} \frac{1}{\beta^2} [\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 \{m_e c^2 (\gamma - 1)/2\}}{I^2} + (1 - \beta^2) - \frac{2\gamma - 1}{\gamma^2} \ln 2 + \frac{1}{8} (\frac{\gamma - 1}{\gamma})^2 - \delta] \quad (2.6)$$

With these proximity formulas, we can work out the analytic solution given the knowledge of absorber type and geometry and the incident particle species and energy. However, modern Monte-Carlo based simulation software can generate results very similar to the analytic solution and that can save us from the complicated and sometimes intractable calculation.

Table 2-1 summarizes the principle process that is used to describe the stopping power, range and the secondary distribution for the interaction with three major primaries.

Table 2-1 Summary of stopping power, range and the secondary distribution for proton/ion, electron and gamma

	Proton/low Z ion	Electron	Gamma
Stopping power **S(E)**	**Median energy**: Basic Bethe-Bloch **Low energy**: add Shell + Barkas + Bloch corrections. **High energy**: add Density corrections. (MeV g^{-1} cm^2)	**Below critical energy**: Collision loss(Møllor cross section) (MeV g^{-1} cm^2) **Above critical energy**: Bremsstrahlung loss (Y.S.Tsai[92])	**Low energy**: photoelectron effect **Middle energy**: Compton scattering **High energy**: pair production **Ultra-high energy**: pair production reduced
Range	Mean range (cm) $\int_0^{E0} \frac{1}{S(E)} dE$ Bragg curve	Radiation length (Y.S.Tsai[92])	Absorption length λ (g/cm^2) $\frac{I}{I_0} = e^{-(\mu/\rho)x\rho t}$
Secondary distribution	Landau distribution		

2.2.2 Inorganic and organic scintillator

The scintillator materials can be categorized as solid, liquid and gases by the physical state. It can also be categorized as inorganic and organic by the scintillation mechanism. The second category method is more useful because it distinguishes the scintillator material by two major characteristics: decay time and photon yield. For the organic scintillators, the decay time is short, and the photon yield is low. For the inorganic

scintillators, the decay time is long, and the photon yield is high. The two characteristics are mutually exclusive.

2.2.2.1 Inorganic scintillator

Inorganic scintillator material is generally in the form of solid crystal. The discovery and application of the inorganic scintillator can be traced back to more than 100 years. The review by M.Weber[86] presents the timeline of the major inorganic scintillator discovery in Figure 2-4.

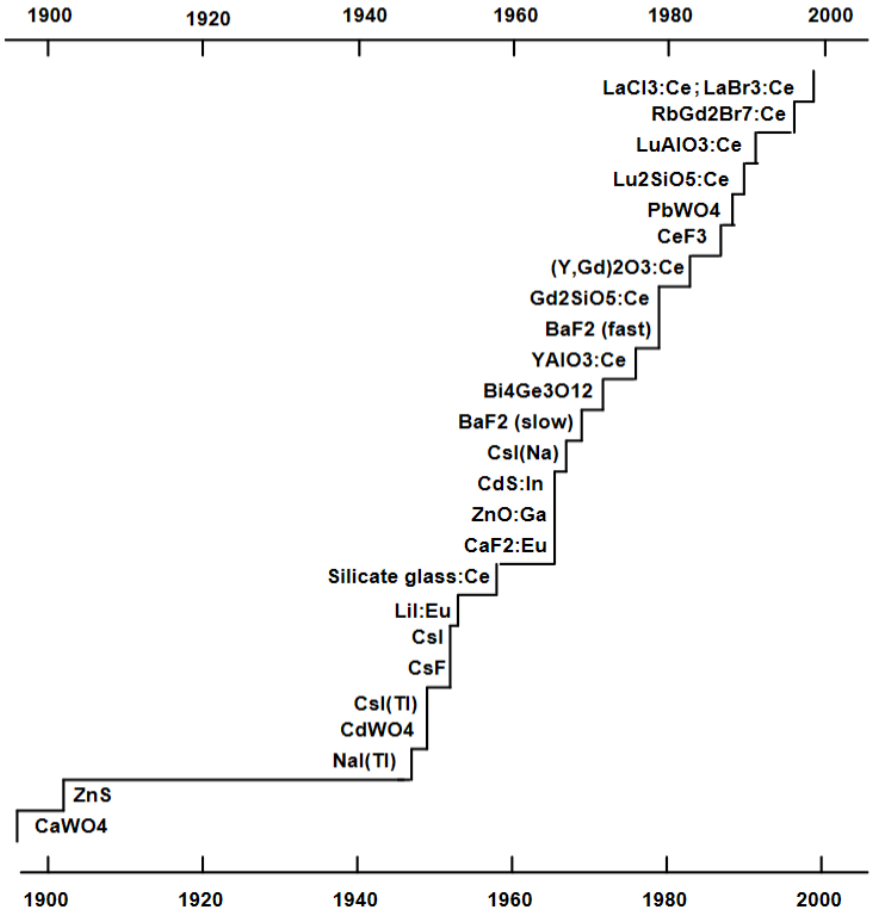


Figure 2-4 the history of the major inorganic scintillator material discovery (reproduced)

We can see the widely known and used alkali halide [93] based crystal like sodium iodide (NaI), caesium iodide (CsI) are discovered in the 1950s. With thallium activator, NaI(Tl) and CsI(Tl) possess close to intrinsic photon yield. However, both scintillators exhibit strong afterglow, which limits their use in medical imaging application. Since 1980s, Ce-doping becomes a standard approach for new scintillator development. Lutetium oxyorthosilicate with Ce-doping (LSO:Ce) and its Y-admixture LYSO becomes standard materials. LYSO now is the first choice for the positron emission tomography

(PET) scanner (a medical imaging equipment) for its balanced excellence in both photon yield and decay time. The lead tungstate ($PbWO_4$) scintillator [94] is the choice for the large hadron collider (LHC) experiment at CERN. It is chosen for its high density (8.28g/cm3), short decay time and good radiation hardness, which are all tailored for the need of the high energy physics experiment. Bismuth germanate (BGO) is another high density and high Z material scintillator before the introduction of LYSO:Ce. BGO is widely used in gamma-ray detector. However, it is intrinsically radioactive, which increases the background noise.

Apart from these insulator-based scintillators, there are also semiconductor-based scintillators under development because of their intrinsically high photon yield and short decay time (the two properties hardly found co-exist in another insulator-based scintillator). Among them, Ga_2O_3 [95] shows promising performance for practical application.

Emission mechanisms

In the inorganic scintillator, the luminescence can occur without activators (also called intrinsic luminescence) or with activators (also called extrinsic luminescence). Various subgroups of mechanism under each catalogue are found. Takayuki Yanagida [88] summarizes the classification clearly as shown in Figure 2-5.

Luminescence without activators (Intrinsic luminescence)	
I	Free exciton luminescence
II	Self-trapped exciton (STE) luminescence
III	Auger free luminescence
IV	Self-activation
Luminescence with activators (Extrinsic luminescence)	
V	$1s \leftrightarrow 2p$ transition (e.g., F-center)
VI	$ns^2 \leftrightarrow nsnp$ transition (Ga^+, In^+, Tl^+, Ge^{2+}, Sn^{2+}, Pb^{2+}, Sb^{3+}, Bi^{3+}, …)
VII	$3d \leftrightarrow 3d$, $4d \leftrightarrow 4f$ transition (transitional metals)
VIII	$4f \leftrightarrow 4f$, $5f \leftrightarrow 5f$ transition (rare earth, Actinoid)
IX	$4f \leftrightarrow 5d$ transition (Ce^{3+}, Eu^{2+}, …)
X	charge transfer luminescence (Yb^{3+}, VO_4^{3-}, …)

Figure 2-5 Classification of the photon emission mechanism under both Intrinsic luminescence and Extrinsic luminescence (reproduced)

Understanding the details of these subgroup is out of the scope of this thesis and generally it has little contribution to the selection of scintillator for our experiment. For those who are interested to know the fundamental details, G.Blasses and B.C. Grabmaier have made a thorough discussion in the book Luminescent Materials [79]. The book expresses two observations: a) generally scintillator with activators will see improvement of photon yield. For example, CsI pure crystal has a photon yield of 2000 photons/MeV, while CsI with Tl activator sees the photon yield jump to 65,000 photons/MeV. b) The scintillators incorporating 4f-5d transition often shows good photon yield and short decay time simultaneously. This typically exists in Ce-doped scintillator.

2.2.2.2 Organic scintillator

Organic scintillator is featured by short decay time (<10ns), and low effective atom number and density. This makes them a good candidate for time of flight application, but not so good for the high penetrating photon like γ-ray. The photon yield is generally very low, which will limit the energy resolution. One advantage of organic scintillator is it can be made to relatively large size at economical cost.

The fluorescence mechanism in organic scintillator occurs at the molecular level. In comparison, the mechanism in inorganic scintillator occurs at the atom level. Upon the deposit of the incident energy, the molecular state will go through excitation, transfer and settlement. Imprecisely speaking, this process is similar to that of the inorganic crystal. But, the energy excitation occurring at molecular level requires more energy than that is required at atomic level. This can explain why the organic scintillator has much lower photon yield, because more energy is required to produce one-unit photon.

As the fluorescence can be observed independently of the physical state, the organic scintillator can be found in gas, liquid or solid form. The common material is stilbene, anthracene and plastic scintillator. They can be made into different shapes and sizes. Cast film sheet is a popular form that has good photon yield and good internal light transmission.

Review by F.D.Brooks[96] explains some interesting features of the organic scintillator. First, there exists prompt and delayed emission. The prompt component is exponential and independent of the species of the incident radiation or particles. The delayed component is non-exponential with mean life typically above 300ns. Figure 2-6 shows the resolution of an organic scintillation into prompt and delayed component.

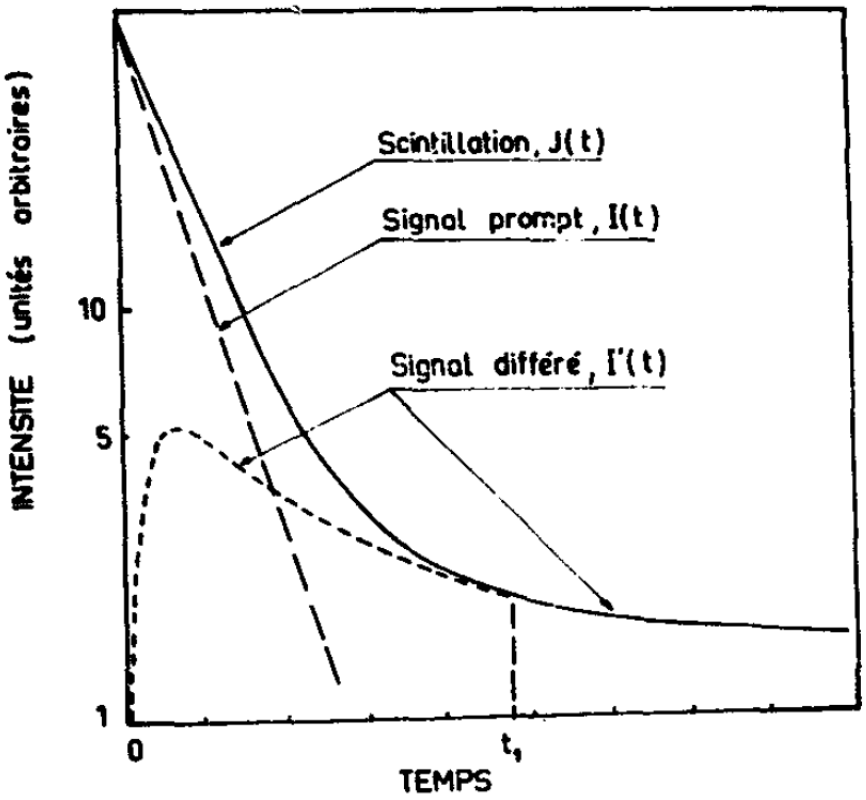


Figure 2-6 The illustration of the two components of an organic scintillation emission. I(t) is the prompt component, I'(t) is the delayed component and J(t) is the sum. This figure is requoted from [97].

The delayed component is less sensitive to ionization quenching (a saturation effect) than the corresponding prompt excitations. Hence, the relative integrated intensities of the two components can be linked to differentiate incident particle species. This provides a route for particle identification by pulse shape discrimination.

Second, the response function L(E) of organic scintillators to heavily ionized particles is a non-linear function of the particle energy. And when comparing different particle species, the heavier ones at the same energy show lower response function output. This non-linear effect is caused by ion quenching. Birks [98] described this effect with an empirical formula:

$$\frac{dL}{dx} = \frac{S\frac{dE}{dx}}{1 + kB\frac{dE}{dx}} \tag{2.7}$$

Where, **dL/dx** represents the fluorescence energy emitted per unit path length and **dE/dx** represents the energy deposit per unit length. **S** is the nominal scintillation efficiency. **kB** represents the fraction of the density of quenching centres that will not produce scintillation photons.

2.2.3 Scintillator property

A good knowledge of the definition and interlink of the scintillator property can help us to choose the right scintillator for the application. The key properties i.e. photon yield efficiency, timing characteristics and emission spectrum will be discussed in this section. Beyond these, other properties might be of interest to other specific applications. The survey in Appendix B: Survey of popular scintillators provides a comprehensive coverage of property for many popular scintillators.

2.2.3.1 Photon yield efficiency

Generally, a higher photon yield efficiency means higher overall detection efficiency, better energy resolution and better detector sensitivity.

The photon yield efficiency is closely related to the physical process of scintillation. Now, we understand the process can be divided into three stages [99]:

a) The conversion stage, in which the energy of the incident particle or radiation is converted to a large number of electron-hole pairs.
b) The transfer stage, in which the energy of an electron-hole pair is transferred to the luminescence centre.
c) The emission process, in which the luminescent ion returns from the excited state to the ground state and emits photon.

The overall photon conversion efficiency is then described as the product of the efficiency of the three stages as [99]:

$$\eta = \gamma Sq \tag{2.8}$$

Where **γ** denotes the conversion efficiency, **S** is for the transfer efficiency and **q** denotes the quantum efficiency of the luminescent centre.

It is discovered that the energy required to create the electron-hole pair is much larger than the bandgap energy of the material. Here we can introduce a coefficient β to represent the ratio of the actual required energy **E** to the bandgap energy $\mathbf{E_g}$, i.e. $E = \beta E_g$. The expression to calculate γ is defined as [99]:

$$\gamma = \frac{E_m}{E_g}\frac{E_g}{E} = \frac{E_m}{E_g}\frac{1}{\beta} \tag{2.9}$$

Where $\mathbf{E_m}$ is the maximum energy of the emitted photon.

The transfer efficiency is the most unpredictable particularly in solid material, because the electron-hole pair might recombine before reaching the luminescent centre, for example, it can be captured by the defects and impurities. The quantum efficiency **q** can be determined by experiment and many materials are known with **q** close to 1.

It is noted not all the radiation energy are absorbed by the scintillator. The loss will be either reflected or transmitted. So a more general expression of eq. (2.8) is:

$$\eta = (1-r)\gamma Sq \tag{2.10}$$

Where **r** denotes the amount of radiation that is lost.

Thus, the full expression to calculate the photon yield is:

$$N_{ph} = E_i(1-r)\gamma Sq \tag{2.11}$$

Where $\mathbf{E_i}$ is the incident energy, and $\mathbf{N_{ph}}$ represents the number of photons emitted.

As most of the parameters in Eq. (2.11) is material dependant and constant, for convenience, manufacturer often groups them together and quote with a single parameter called photon yield coefficient in a unit of photons/MeV.

2.2.3.2 Timing characteristics and time resolution

The scintillation photon decay time can be described with a simple exponential formula. If there are more than one decay components, they can be joined by linear superposition. A two components decay model can be written as:

$$N(t) = Ae^{\left(-\frac{t}{\tau_f}\right)} + Be^{\left(-\frac{t}{\tau_s}\right)} \tag{2.12}$$

Where **N** is the number of emitted photons, τ_f and τ_s is the fast(prompt) and slow(delayed) decay constants. A and B are the scale factor.

Eq.(2.12) is quite basic and it ignores the rise time. A more comprehensive equation [100] including the rise time for a single decay component is:

$$N(t) = \frac{N_0}{\tau_d - \tau_r}[e^{-t/\tau_d} - e^{-t/\tau_r}] \tag{2.13}$$

Where $\mathbf{N_0}$ is the total number of photons generated by the energy deposit. τ_r and τ_d are the intrinsic rise time and decay time of the scintillator.

The rise time and decay time is mostly determined by the scintillator type but the photon output also has a small impact as discussed by P.Lecoq [101]. The high photon output will improve the rise time and decay time.

For time critical application, the detector size will have an impact on the timing in two aspects. One is the added delay by the photon travelling time within the scintillator. For those photons born with initial direction not aiming at the photon detector, the potential reflection path will add more delay. The other is that the time dispersion of the scintillation photon is proportional to the detector size (mainly the dimension in the direction of travelling). The dispersion will worsen the time resolution as it effectively prolongs the decay time.

2.2.3.3 Emission spectrum

The scintillator emission spectrum must match the absorption spectrum of the coupling PMT cathode or SiPM or any chosen photon sensor as close as possible to maximize the photon detection efficiency. A single response factor **η** can be obtained by integrating over the spectrum:

$$\eta = \frac{\int \phi(\lambda) F(\lambda) d\lambda}{\int F(\lambda) d\lambda} \tag{2.14}$$

Where **ø(λ)** is the absorption spectrum and **F(λ)** is the emission spectrum.

2.2.4 Scintillator application

2.2.4.1 Scintillator detector vs other semiconductor detectors

In a scintillator detector system, usually we will find two stage detection process before the signal can be processed by the readout electronics. The first stage is the conversion of incident particle/photon to photons by the scintillator and the second stage is the conversion of photon to electric signal by a photon sensor. Each stage will introduce signal loss, delay and error. The overall impact is cascaded. In contrast, a semiconductor detector like the silicon solid state detector has only one stage process, which can convert the incident particle/photon directly into electrical signal. In this sense, a semiconductor detector is intrinsically more sensitive, faster and more accurate than the scintillator detector system. Moreover, the bandgap of semiconductor is lower than scintillator material. This enables the semiconductor detector to generate

more secondaries than a scintillator for the absorption of the same energy. This is reflected as a better energy resolution.

However, a scintillator still has some unique features that a semiconductor doesn't possess. For example, the lower bandgap of semiconductor also means worse thermal noise as it is easier for the electron to break the bandgap barrier. In comparison, a scintillator is less prone to thermal noise due to its higher bandgap. A scintillator can also be made with a large volume while it is very rare to see large silicon detector, particularly in large depth. For a similar reason, semiconductor detectors suffer more radiation damage than a scintillator, because the defects created by the radiation in the detector causes more thermal noise problem in silicon than in a scintillator. In detection of high energy and high penetrating particles/photons, scintillators with high z materials are the best choice, because a high z material has higher stopping power and provides better energy/flux absorption. These features make scintillator detector system still the first choice for high energy applications, for example, LXSR scintillator is the first choice as gamma detector in the PET instrument.

2.2.4.2 Scintillator in space instruments

There are many past and existing space instruments equipped with a scintillator detector. A short survey is given in Appendix A: Survey of scintillator detector application in space instrument.

The survey focused on the large-scale flagship space instrument missions. We can see the traditional scintillator NaI and CsI remain the first choice for gamma and hard x-ray detection. In recent years, many smaller-scale, low-cost space instruments based on scintillator detection system have flown or are in development. For example, HERMES[102] used a silicon detector "siswich" with a Ce:GAGG scintillator as the main detector for x-ray and gamma-ray. CAMELOT[103] carries CsI detectors coupling with SiPM for Gamma Ray Burst (GRB) detection. Another GRB application BurstCube [104] also carries CsI (Tl) scintillator coupling with SiPM sensor.

Therefore, the scintillator detector remains a popular choice for space instrument application. Particularly, the fast development of the new generation photon detection sensor SiPM greatly improves the overall scintillator detector system performance.

2.3 Scintillator module prototype design

This section will discuss about the scintillator prototype module that we studied, designed, simulated and manufactured for this project. Our goal is to develop a scintillator module that possesses the following features:

- basic energy resolution
- wide solid angle
- medium to high energy range
- sensitive to a range of particles
- particle identification
- direction identification

Also, it is desirable for it to be compact, low resources and radiation hard, as required by space instruments.

2.3.1 Initial scintillator module design

The initial concept of the scintillator module is shown in Figure 2-7.

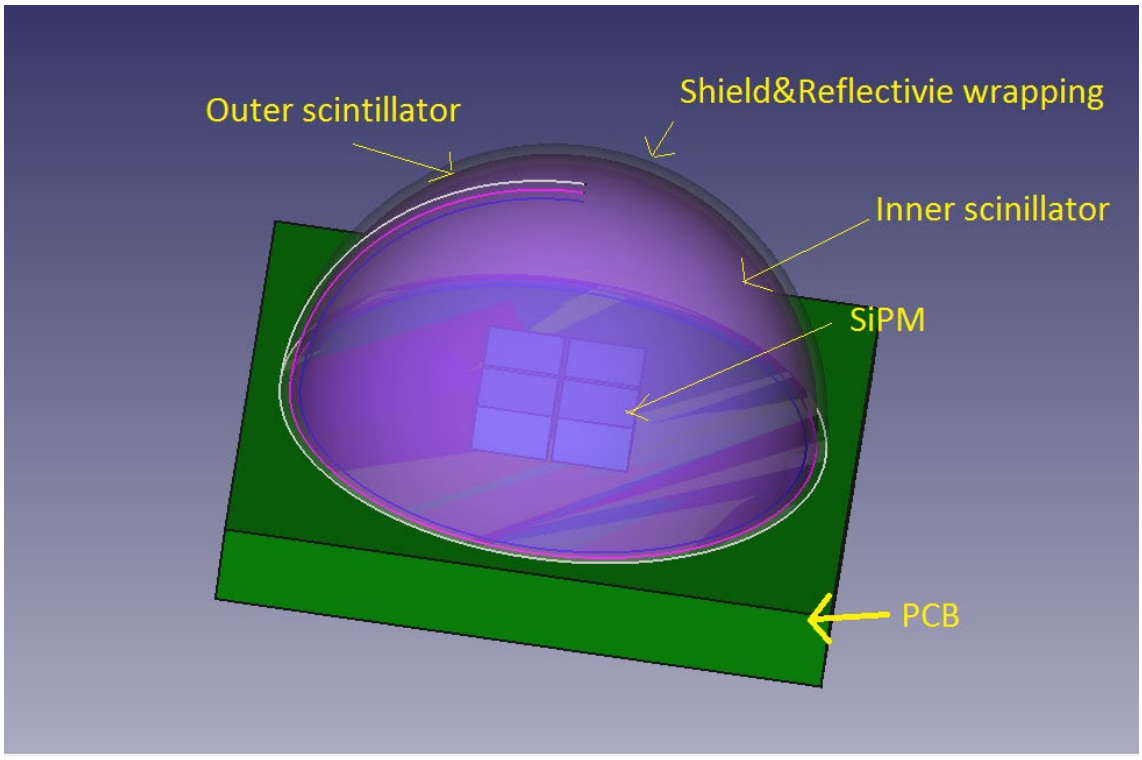


Figure 2-7 Initial concept of the scintillator detector module

The hemisphere scintillator module is composed of two layers. The outer layer is made of a fast response organic scintillator and the inner layer is made of a slow response

inorganic scintillator. The outer layer is wrapped with a thin reflective layer. Four SiPMs are positioned in the centre of the hemisphere.

The hemisphere geometry provides a uniform 2pi solid angle coverage. All the scintillator photons (except those being reabsorbed) shall be focused by the reflection layer onto the centre of the sphere, where the photon sensors are located. This geometry can maximize the photon collection efficiency optically. The principle is illustrated in Figure 2-8.

The four SiPM sensors each shall detect a portion of the total photons depend on the incident direction of the incoming particle. The ratio of the signal strength among the four SiPM provides a coarse angle detection. For example, if the signal of the four sensors are equal, the incident particle must come from the zenith of the scintillator module.

The initial configuration of the module is shown in Table 2-2.

Table 2-2 The basic configuration of the first scintillator module design

Property	Inner Scintillator	Outer Scintillator	Wrap
Material	CsI	BC408	Teflon
Yield (#/MeV)	54,000	500	n/a
Thickness/Radius	1.5mm	0.2mm	0.055 mm
Shape	Semi sphere	Semi sphere	Semi sphere
Refractive Index	Spectrum (avg.1.79)	1.58	1.35

The choice of the two scintillators are based on two fundamental purposes. One is to achieve the ΔE/E discrimination. Second is to have a differentiable decay time. This is because the two scintillators will share the common SiPM photon sensors. The sensor will rely on the timing of the photon to identify their source. Therefore, the outer scintillator shall have a very short decay time and serves as triggering detector. As a ΔE detector, it shall be made very thin to allow the bulk energy of the primary to penetrate through. The inner scintillator shall have a relatively long decay time and high photon yield. As an E detector, it shall have high density and be made to large volume in order to absorb the full energy. Hence, we chose BC408 as the outer scintillator and

CsI as the inner scintillator. Another reason is that these two scintillators are commonly used and widely available.

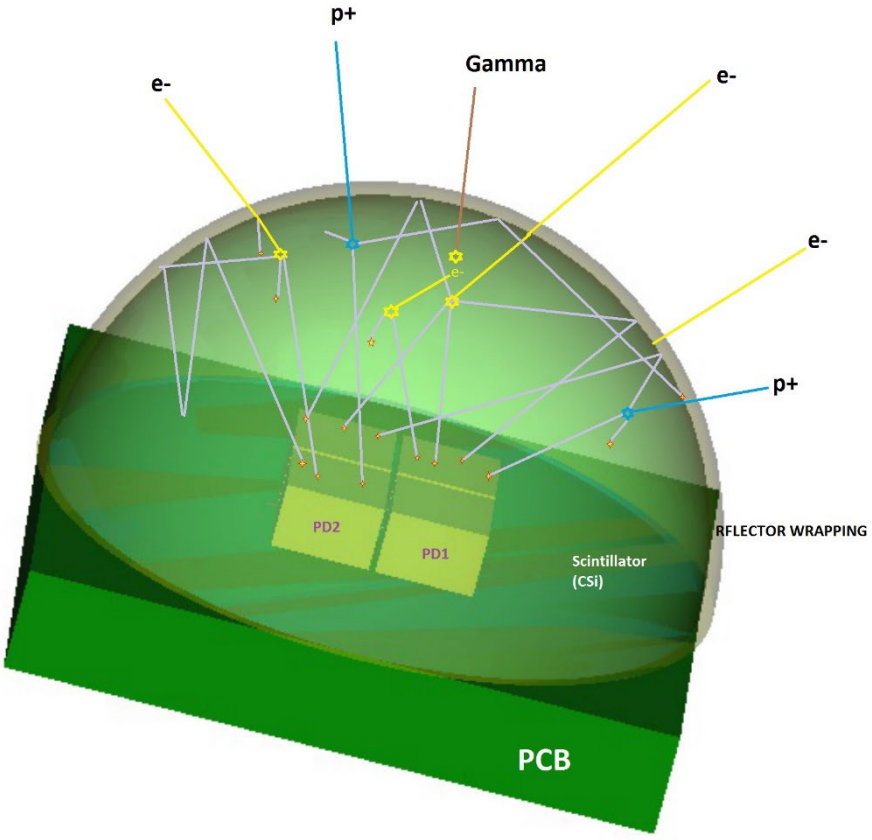


Figure 2-8 The secondary photon reflection in the hemisphere geometry (Created with FreeCAD software)

Various methods and tools are used to check the performance of the detector module. The stopping-power&range tables for electrons, protons and alpha from NIST[81] and the x-ray mass attenuation coefficients also from NIST[82] are used for analytic estimation of the penetration depth of electron, proton, alpha and gamma ray in our detector module. SRIM software[83] is used to simulate and calculate the response for heavy ions. Geant4 [84] is the main simulation tool that is useful to cover all the aspects of simulation we need.

The full detail of the simulation setup and results will be discussed in Chapter 3. Here we only present some brief results. Table 3-11 lists the energy range of the detector module for different incident species. It should be noted that these thresholds are not a clean cut off. They only mean that beyond that energy, the detection efficiency will greatly drop.

2.3.2 2nd scintillator module design

Unfortunately, a hemisphere scintillator in crystal format is not practical to manufacture within the scope of this work, let alone a double layer configuration. Therefore, the detector module design is changed accordingly. The revised geometry is cylinder based as shown in Figure 2-9.

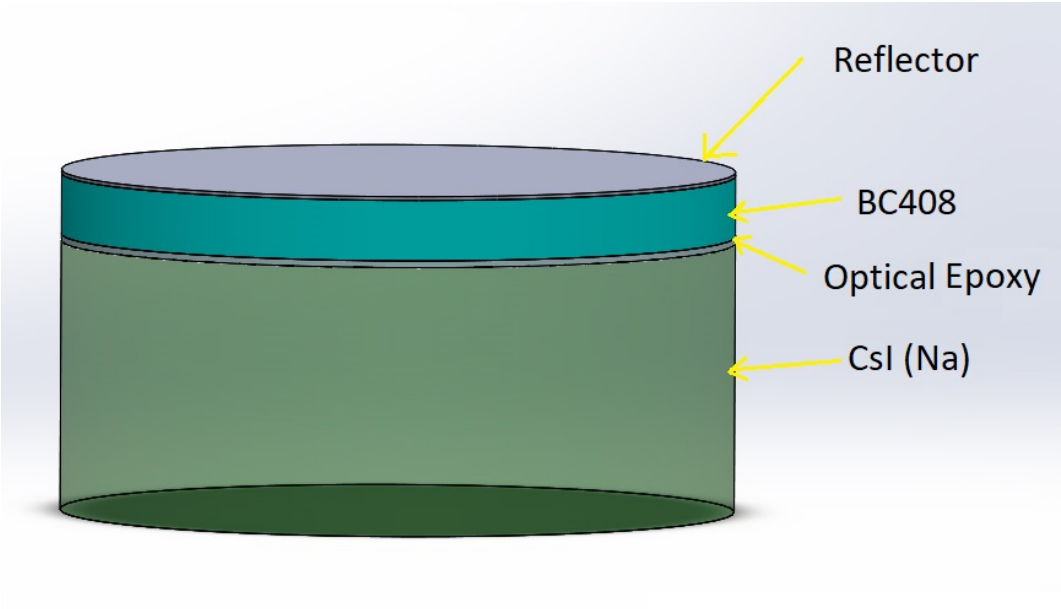


Figure 2-9 The revised 2nd scintillator module design (created with SolidWorks CAD software)

The property and size of the 2nd scintillator module is listed in Table 2-3.

Table 2-3 The configuration of the 2nd scintillator module design

Property	Lower Scintillator	Glue	Upper Scintillator	Wrap
Material	CsI(Na)	Epoxy	BC408	Teflon
Yield	8,500	n/a	500	n/a
Thickness	4mm	0.1mm	0.9mm	0.055 mm
Radius	5mm	5mm	5mm	5mm
Shape	Cylinder	Cylinder	Cylinder	Thin layer
Refractive Index	Spectrum (avg. ~1.79)	1.56	1.58	1.35
Response time	630ns	n/a	2.5ns	n/a
Max emission	420nm	n/a	425nm	n/a
Function	High-z slow detector	Glue two scintillator	Low-z fast detector	Blacking, reflection

The original CsI(Tl) scintillator was replaced by the CsI(Na), which has the output photon peak emission wavelength matching the SiPM max response wavelength better. The downside is the reduced photon yield.

The energy response range is quickly re-evaluated using SRIM and NIST database. Some evaluation results are shown in Table 2-4.

Table 2-4 Some evaluation results for the energy response range

Particle	Energy	Range	Lateral	Stop	Notes	Tool
Proton	10 MeV	1.21mm	24.2um	CsI		SRIM
Alpha	100MeV	8.14mm	213um		Penetrate the full detector	SRIM
Alpha	75MeV	4.8mm	119um	CsI	Max E to stay in CsI	SRIM
Proton	100KeV	1.18um	1270A	Teflon		SRIM
Proton	2.5MeV	71um	1.44um	BC408	Min E to penetrate the Teflon	SRIM
Alpha	10MeV	68um	6300A	BC408	Min E to penetrate the Teflon	SRIM
Proton	18MeV	4.44mm	213um	CsI	Max E to stay in CsI	SRIM
e-	2MeV	3.4mm		CsI	Max E to stay in CsI	NIST
e-	200keV	57.8um		BC408	Min E to penetrate the Teflon	NIST
Gamma	30keV	45.1%			Skipped Teflon	NIST
Gamma	200keV	99.9%		CsI	Absorbed by CsI	NIST

2.3.3 3rd scintillator module design

After sourcing an American scintillator manufacturer SEMicro[105], we further modified our design. The cylinder geometry hasn't changed, but the scintillator material, thickness and wrapping have. The upper scintillator is revised to an organic scintillator disc BC-404 of 3mm radius and 1mm thickness. This is due to the manufacturer has the 1mm thickness disc as one of the standard products. The radius is cut to the customer requirements. The lower scintillator is revised to a single crystal LXSR of 3mm radius and 4mm thickness. This scintillator has similar photon yield as the CsI(Na) but much shorter decay time (~42ns), which is still slow enough compared to the 1.8ns

decay time of the BC-404. The two scintillators are bonded with NOA 61 optical glue. Aluminium is coated around the scintillator module except for the exit face (the SiPM bonding face). The coating thickness is 50nm for the top face (particle entry face) and 100nm for the cylinder body. Figure 2-10 shows the configuration.

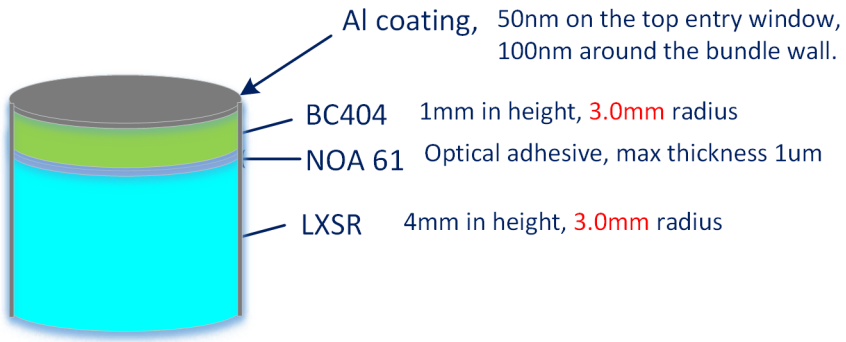


Figure 2-10 The revised 3rd scintillator module design

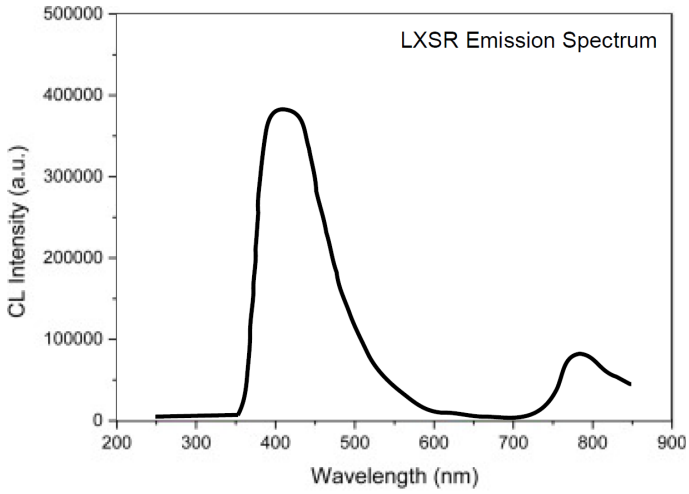
This design replaced the previous wrapping layer with a reflective coating layer. The coating thickness can be controlled to bare minimum. In our case, 50nm is achieved for the input face. This greatly reduced the dead layer and extended the minimum energy that can be detected.

BC-404 and BC-408 come from the same Polyvinyltoluene based organic scintillator family. Their performance is very similar except BC-404 is slightly faster and has higher photon yield. The wavelength of BC-404 is 408nm and that of BC-408 is 425nm, which matches the chosen SiPM's peak sensitivity wavelength at 440nm better. Also, the replacement was a practical choice as the manufacturer only had stock of BC-404.

The CsI(Na) is hygroscopic. The replacement LXSR has no such problem. In addition, LXSR has much faster response comparable to the popular LYSO and high photon yield comparable to CsI (Tl). The chemical composition of LXSR remains undisclosable as manufacturer's intellectual property. Table 2-5 shows its basic property and emission spectrum.

Table 2-5 LXSR property and emission spectrum

Property	LXSR	Emission Spectrum
Light Output(%NaI)	80	
Decay Constant	42 ns	
Peak Emission	420 nm	

Property	LXSR	Emission Spectrum
Index of Refraction	1.80	
Density	4.45	
Hygroscopic	NO	

The optical adhesive used to glue the two scintillator is NOA 61 from Norland [106]. It is a clear, colourless, liquid photopolymer that can cure under ultraviolet light. It has a typical refractive index in cured form of 1.56, which matches the index of the SiPM entry window well, although not that perfect for LXSR's. The photon transmission spectrum of NOA 61 is shown in Figure 2-11. We can see its peak transmission range from 300nm to 2000nm, which is enough to cover both scintillators' emission spectrum.

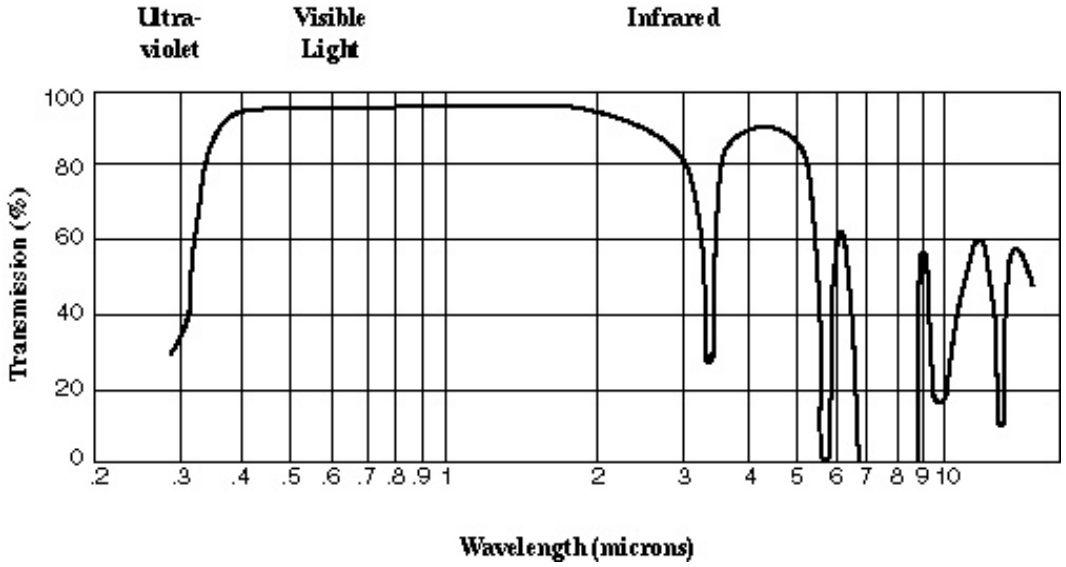


Figure 2-11 NOA 61 transmission spectrum

The energy response threshold is rechecked for charged particles and the results are summarized in Table 2-6.

Table 2-6 The energy response threshold for charged particles for the 3rd scintillator module design

Layer	Depth	Proton	Alpha	Electron	Ion C
		mass 1	mass 4		mass 12
Aluminium	50nm	3.5keV	5.5keV	~1keV	
BC404	1mm	10MeV	40MeV		200MeV
LXSR	4mm	35MeV	140MeV		

The gamma response is calculated using NIST data for LXSR scintillator at 4mm thickness as shown in Figure 2-12. We can see after 200keV, the absorption rate drops below 100%.

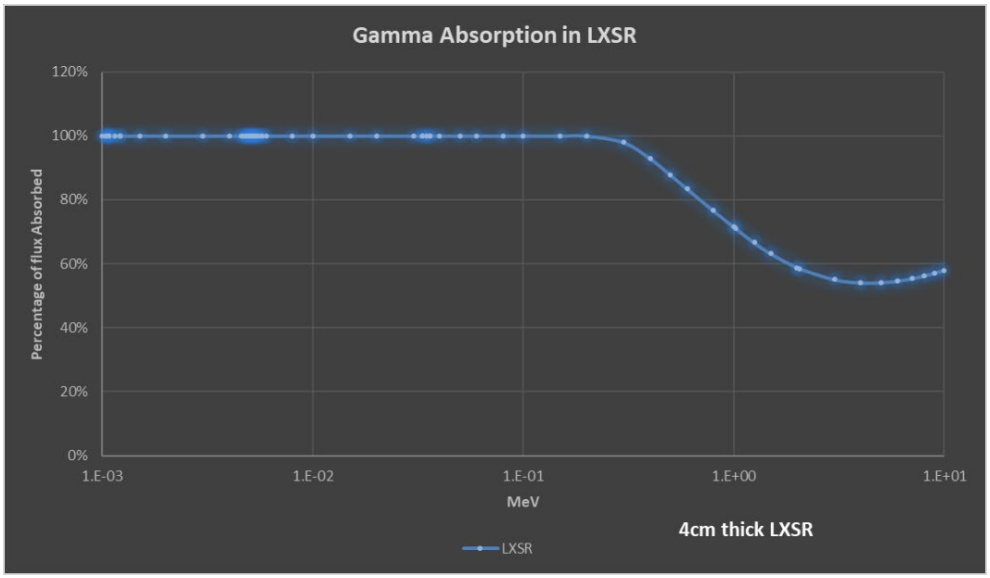
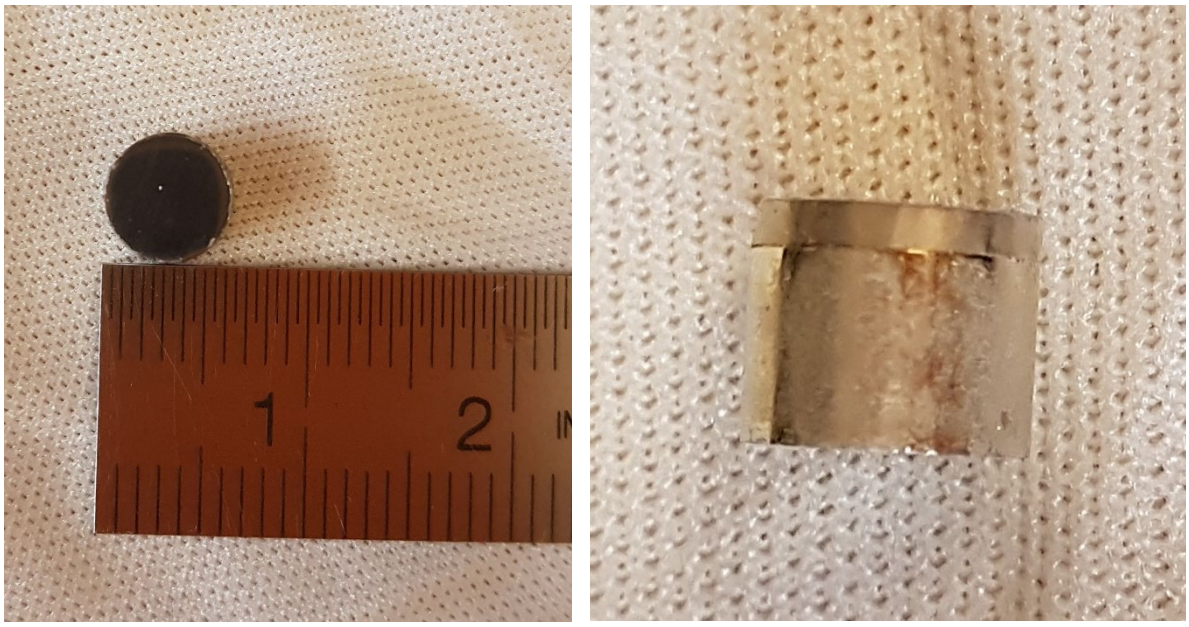
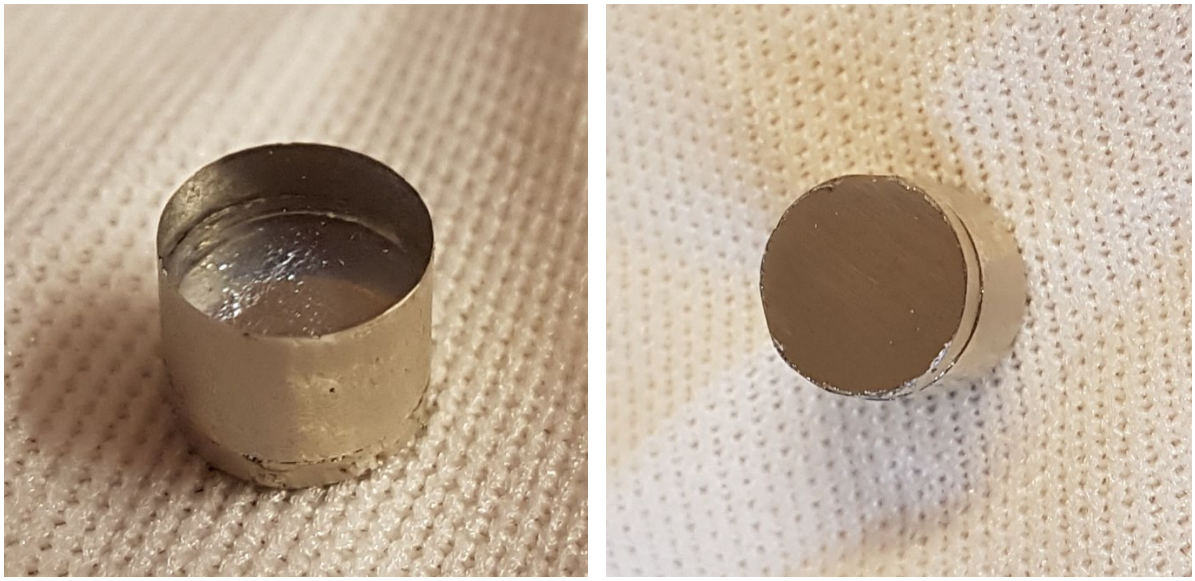


Figure 2-12 Gamma absorption in LXSR

This module was manufactured by SEMicro. Figure 2-13 shows the finished module.



(a) (b)



(c) (d)

Figure 2-13 the manufactured scintillator module. (a) is the dimension, (b) is the side view (the bonding line of the two scintillators can be clearly seen), (c) is the bottom view (clear entrance), (d) is the top view (covered by 50nm Al coating).

2.3.4 Photon collection efficiency comparison between hemisphere and cylinder geometry

We have run Geant4 simulation on both geometries. The results proved our prediction that the hemisphere is more efficient for photon collection than the cylinder, because in the hemisphere geometry, the optical path can be focused to the origin with the shortest path or least number of reflections. This minimized the photon transmission loss and reduced the chance for the photon to be reabsorbed. illustrates the photon trajectory in both geometry with a single primary particle hit. We can clearly see that in the cylinder, the secondary travels through multiple reflections with long path before it reaches the exit face, while in the hemisphere, the secondary reaches the exit face with just one reflection.

To quantify the comparison, we setup simulation (Figure 2-14) to check the photon collection efficiency of the two geometries. We defined the photon collection efficiency as the number of collected photon divided by the theoretical number of emitted photons. Then we run a few tests at different primary energy. The test results are shown in Table 2-7 with key figure highlighted in yellow. We find at the same incident particle species and at the same energy, the efficiency for the hemisphere is almost 100% for

both scintillators, while for the cylinder, the efficiency drops to 39% for the fast scintillator and 66% or 69% for the slow scintillator. The poorer efficiency of the fast scintillator (on the top) is probably because it has longer distance to reach the exit surface, hence more photon loss on the path.

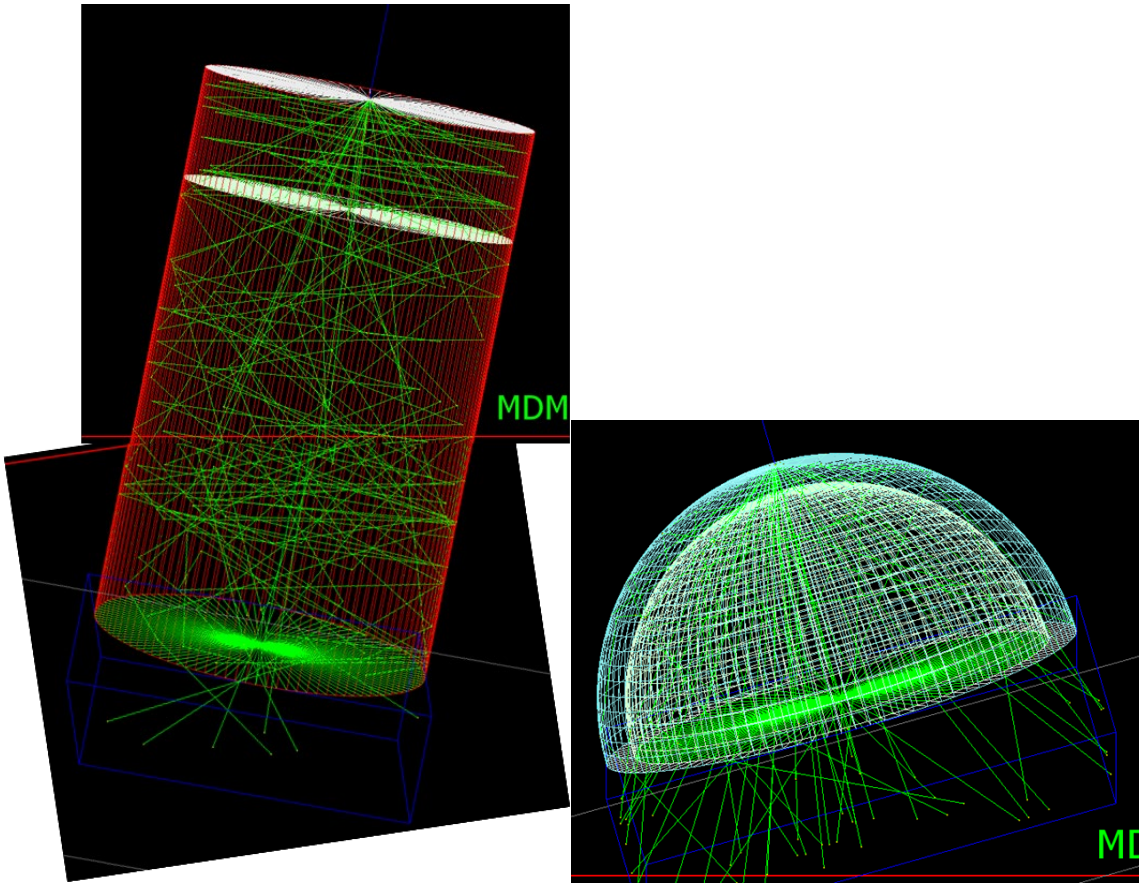


Figure 2-14 Secondary photon trajectory (green) in cylinder geometry (left) and hemisphere geometry (right)

Table 2-7 The comparison of scintillation photon detection efficiency of different geometry shape

Energy (Alpha)	Energy Deposit in Fast Scint. peak	Energy Deposit in Slow Scint. peak	Photon from Fast Scint. peak	Photon from Slow Scint. peak	Total SiPM Photon peak		Fast Scint. Photon Yield	Slow Scint. Photon Yield	Photon Yield in Fast Scint.	Photon Yield in Slow Scint.		Geometry eff. For Fast Scint.	Geometry eff. For Slow Scint.
Hemisphere Geometry: Teflon=55um, BC408 = 0.2mm, CsI = 1.5mm													
keV	keV	keV	#	#			#/KeV	#/keV					
10000	2748	0	1365	0	1369		0.5	54	1374	0		99.34%	
20000	9944	6746	4916	3.61E+05	3.66E+05		0.5	54	4972	364284		98.87%	99.12%
50000	3287	45135	1621	2.43E+06	2.43E+06		0.5	54	1643.5	2437290		**98.63%**	**99.57%**
100000	1800	35484	875	1.90E+06	1.90E+06		0.5	54	900	1916136		**97.22%**	**99.30%**
Cylinder Geometry: Al=50nm, BC408 = 1mm, LXSR=4mm													
keV	keV	keV	#	#			#/KeV	#/keV					
5000	4999	0	941	0	941		0.5	32	2499.5	0		37.65%	
10000	9999	0	1881	0	1881		0.5	32	4999.5	0		37.62%	
20000	20000	0	3833	0.00E+00	3.83E+03		0.5	32	10000	0		38.33%	
50000	18433	31561	3573	6.66E+05	6.72E+05		0.5	32	9216.5	1009952		**38.77%**	**65.96%**
100000	9275	90784	1800	2.00E+06	2.01E+06		0.5	32	4637.5	2905088		**38.81%**	**68.91%**

2.3.5 Particle identification

Particle identification is a very desirable but quite challenging task. There are a few approaches to address this task. However, due to the wide range of energy and species, it is impossible for a single approach to cover all the cases. The common approaches are:

At instrument level:

- Time-of-flight mainly used to identify the ion species
- Anti-coincidence mainly used to reject particles not from the designed FoV
- Active electric or magnetic field, mainly to select the energy and direction
- Shielding, mainly to reject particles not from the designed FoV

At detector level:

- Pulse shape analysis for a single and certain scintillator crystal
- Stack of multiple scintillator/silicon detectors

We will discuss the two approaches at detector level here.

One approach taken by F. Benrachi[107] is to investigate the pulse shape of the two decay components of the CsI(Tl) scintillator. The two decays can be described as the linear sum of two exponential functions:

$$L(t) = \frac{h_f}{\tau_f} \mathrm{e}^{-\frac{t}{\tau_f}} + \frac{h_s}{\tau_s} e^{-\frac{t}{\tau_s}} \tag{2.15}$$

Where τ_f and τ_s are the time constant of the two decays and h_f and h_s is the pulse height of them. After experiment with different charged particles, the following features are discovered:

a) The ratio $R=h_s/(h_s+h_f)$ increases as the ionisation density of the incident particle decreases;
b) τ_f increases as the ionisation density decreases;
c) τ_s is almost constant regardless of the particle species;

Then, by measuring the pulse height of the two decays and working out the ratio, good particle identification can be achieved. See Figure 2-15 for some measurement results.

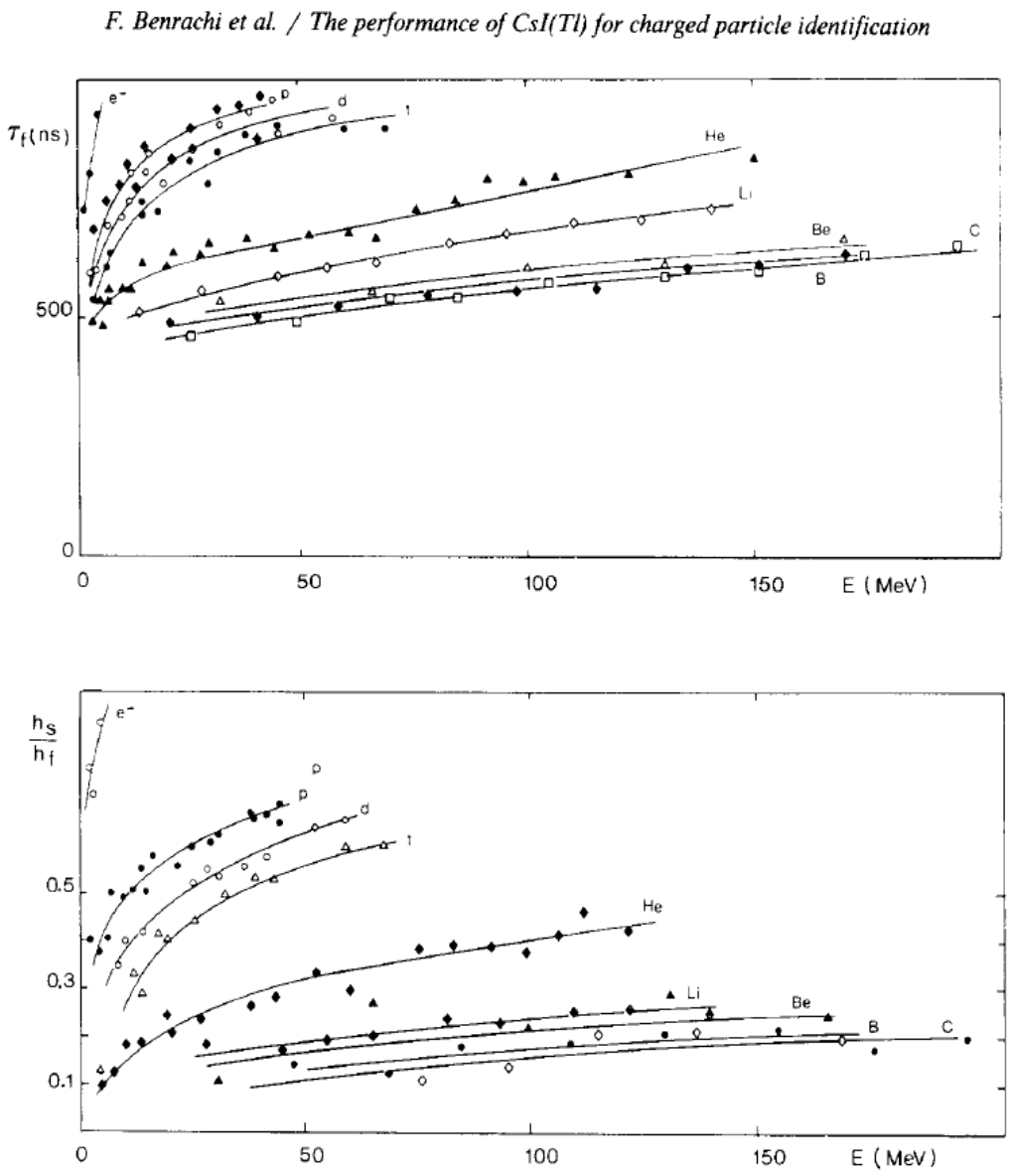


Figure 2-15 The trajectory of fast decay time τ_f (top) and the ratio h_s/h_f (bottom) against different particles across a medium energy range. (reproduced from [107])

The second approach is through a multiple scintillator stack, which is also the approach we take. In this approach, two or more detectors can be staked together. Each detector will capture a portion of the incident particle energy or stop it fully within an energy threshold. The ratio of captured energy between the detectors can help to determine the incident particle species and energy. This method seems to work well for proton and heavy ions, but not quite well for electrons and gamma.

In our case, we opt for two detectors with a thin and fast scintillator on top of a thick and slow scintillator. This sandwich structure shares a common SiPM sensor coupling to the exit face of the slow scintillator. When a prime particle hits, the fast scintillator will be triggered first. The emitted photons will travel through the slow scintillator and

be registered by the SiPM sensor. The pulse height of this leading signal represents the portion of energy captured by the fast scintillator. This signal is defined as **ΔE**. This leading signal quickly diminishes and SiPM is ready for the new signal. Assume the prime particle hasn't lost all its energy in the fast scintillator, it will continue to travel to the slow scintillator. With much higher density and bigger volume, the slow scintillator is likely to absorb all the remaining energy. The emitted photons from it will form the second, wider span (due to slow decay) and higher amplitude signal. This signal is defined as **E**. The ratio of **ΔE/E** provides key information to identify the particle species.

Commonly, an **organic scintillator** featuring fast response and low photon yield is a good candidate for the fast scintillator. An **inorganic scintillator** featuring slow response and high photon yield is a good candidate for the slow scintillator. We changed the specific scintillator through the design iteration, but the choice always follows the same principle.

We tested this strategy in Geant4 simulation with several common particle species across a median energy range. The findings are very encouraging as shown in Figure 2-16. We can see the **ΔE/E** ratio trajectories for different particle species are clearly separated. This effect appears very similar to Figure 2-15 although the underlying mechanism is different.

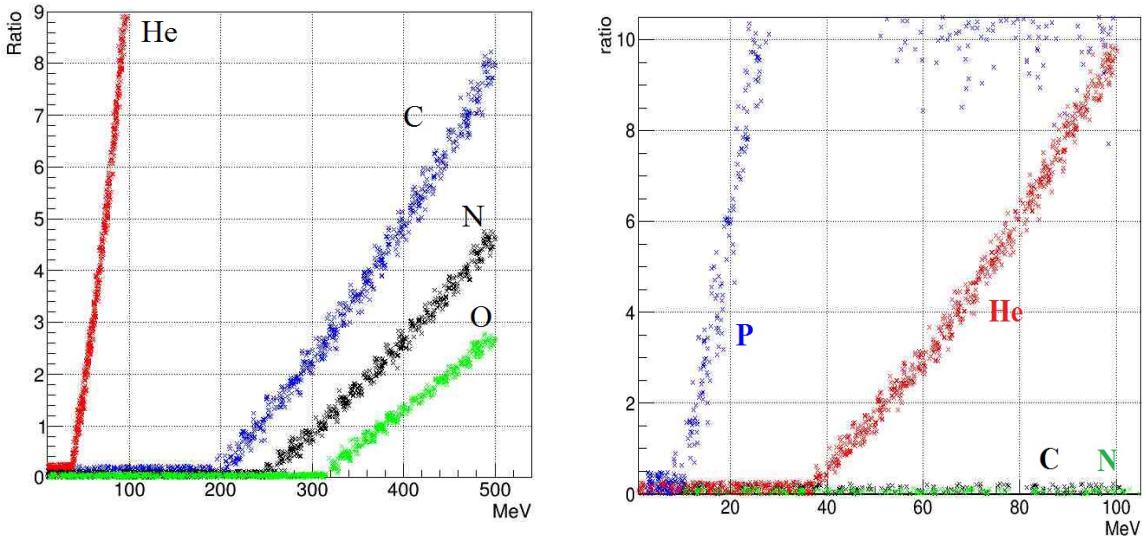


Figure 2-16 Geant4 simulation shows different ΔE/E ratio trajectory for different ion species. The turning corner of each line is the minimum energy that is required for the ion to penetrate the fast scintillator and enter the slow scintillator.

The simulation assumes a perfect photon detection and a perfect signal separation of the two scintillators by the single SiPM detector. In the real experiment, we will find that the signal separation remains a big challenge.

Through literature study, we have found that our approach has a successful predecessor. J.Alarja[108] reported a successful charge separation of up to Z=19 with a phoswich structure (a 0.2mm NE102A scintillator on top of a thick CsI(Tl) scintillator). This further increases our confidence that our design approach shall provide a promising particle identification, particularly for the charged heavy ion.

2.3.6 Scintillator choice and their radiation hardness

As mentioned a few times in the previous discussion, the fundamental criteria for the selection of scintillators for our detector module is the timing. As we want to use a single SiPM sensor, the only way to separate the scintillation photon is by timing. Hence, we require the top scintillator to have a very short decay time and the bottom scintillator to have a comparably long decay time. Naturally, scintillator of short timing tends to have low photon yield and that of long timing has high photon yield. This perfectly matches the demand of the **ΔE/E** configuration.

Other factors are all driven by the engineering practise. For example, to maximize the SiPM detection efficiency, the chosen scintillator emission spectrum peak shall be as close to the SiPM reception spectrum peak as possible. The refractive index of the scintillator shall match that of the SiPM as close as possible to reduce the reflection at the boundary. The scintillator material shall be manufacturable to the geometry we require.

One special consideration for space application is the radiation harness. BC404, CsI and NaI scintillators are commonly used for space mission for their good radiation tolerance. In general, a scintillator is much more radiation hard than silicon detector for it has much higher band gap. As irradiation builds up, the scintillator photon yield will drop for various reasons such as nuclei displacement damage. Yu.M.Protopopov [109] reported BC404 scintillator photon yield drops by 50% post 3.4x104Gy (= 3.4Mrad) γ-irradiation, and by 58% post 1x105Gy(10Mrad) γ-irradiation. It heals a little after 23 days recovery. There is no published data found for LXSR scintillator. We suspect LXSR is a variant of LYSO with very similar performance. A study carried out by Chen [110] show the photon emission spectrum is not affected and photon yield drops at about 12% by γ-irradiation up to 1Mrad.

While, as an insulator, care should be taken to avoid the charge building up in the scintillator. A conductive path shall be provided through the coating or the close contact material.

2.4 Full instrument concept

The cylinder geometry lost the 2π solid angle coverage that the previous hemisphere geometry has. To compensate the missing field of view (FoV), we designed a concept of using multiple cylinder detector modules to form a big hemisphere structure as shown in Figure 2-17. The grey hemisphere structure is filled with a number of evenly positioned scintillator module (in blue). Each module can cover a small FoV. Together the full 2π FoV is covered. This way, the angle of the incident particle can also be clearly detected.

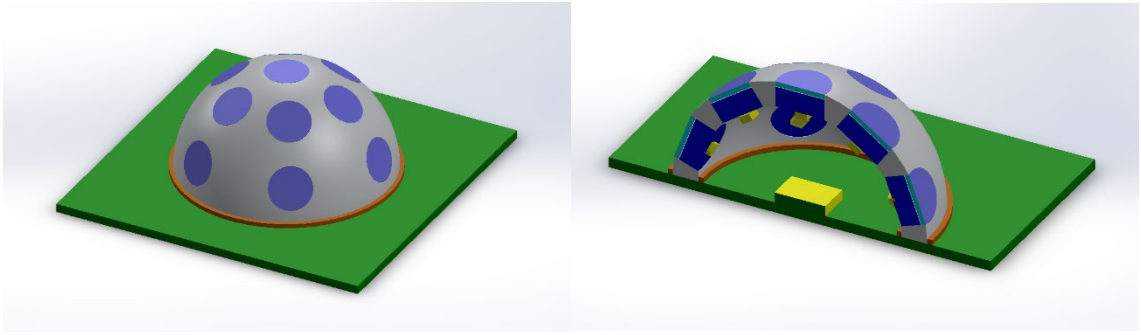


Figure 2-17 The concept design of an integrated structure. Left is the full view and right is the section view. Blue is the scintillator module, grey part is the hemisphere structure. Green is the support PCB.

Based on the stacked scintillator module shown in Figure 2-10, we add a collimator to refine the FoV. It is possible to add extra electric field in the collimator to deflect the unwanted low energy electron and possibly very low energy proton. The SiPM (in yellow) is directly glued to the scintillator module. A small PCB can be used to both support the SiPM and host a compact preamplifier circuit. The dimension of the module can be adjusted to fit the energy range and target particle species. The ratio of dimension to depth/height of the module shall be studied to optimize the optical path. The ratio of radius (R) to depth (D) of the collimator determines the FoV as:

$$FoV_{(half-width-conical)} = \arctan(R/D) \tag{2.16}$$

The required number of modules can also be adjusted to fit the requirement of angular resolution. The max number of modules are limited by the power and processing resources that the instrument host (the spacecraft) can provide. Figure 2-18 and Figure 2-19 shows the concept instrument design in detail.

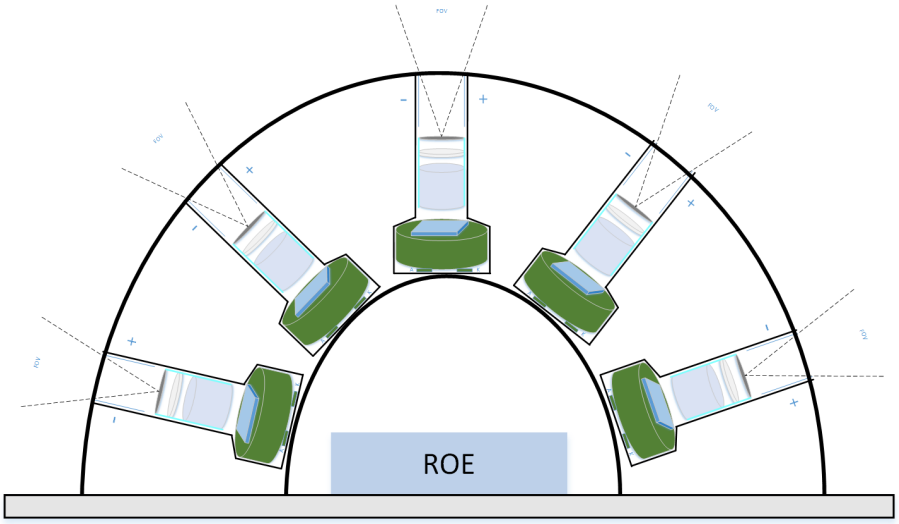


Figure 2-18 the section view of the concept instrument

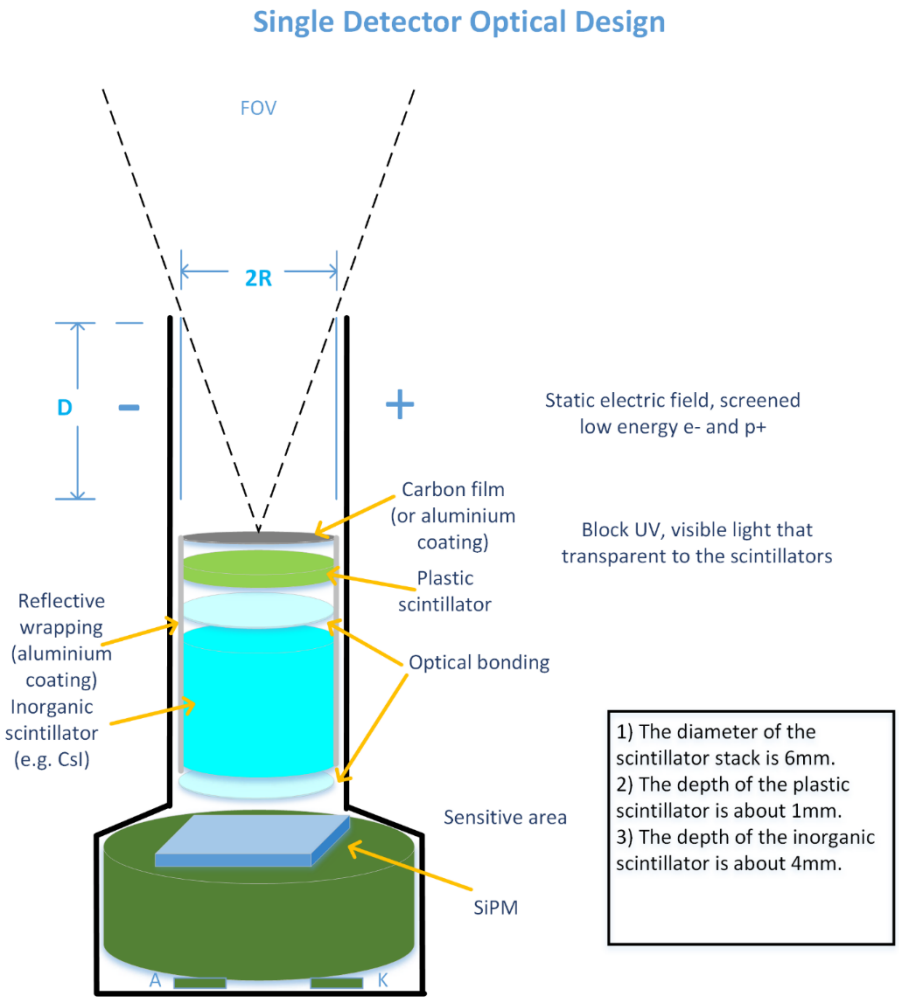


Figure 2-19 The detail of the enhanced single detector module

2.5 Conclusion

We successfully manufactured our first stacked scintillator module using the 3rd design. This design will be the baseline for the development of the remaining thesis except the work in Chapter 3.

The 3rd design introduces a very thin layer of coating, which is the dead layer at the detector entry and determines the minimum energy that the detector can receive. This minimum energy is shown in Table 2-6. The dimension of the module and each scintillator is designed to address the target particles defined in Table 1-6. However, the dimension is also limited by the manufacturing capability and the simulation accuracy. It is not possible to have one dimension fitting all the different particles and their energy range. So, we adopt a convenient and common dimension of 1mm for the BC404 and 4mm for LXSR at 3mm radius. The dimension can be further improved once our understanding of the system is better and the final application needs is more focused.

The key motivation of the stacked scintillator module design is to make particle identification. This is a very important but also very difficult aspect for the scintillator-based particle instrument. We briefly explore the effectiveness of particle identification with our stacked scintillator by simulation. We have found the strategy works well for proton and heavy ions, but less well for electron and gamma. The simulation has assumed some ideal situation for example, the source of scintillation photon is unmistakably known. In the real detector circuit, this will be a great challenge due to the unavoidable overlap in the two scintillators' photon signal.

To fulfil the 2π FoV with the cylinder shape detector module, a concept is proposed to assemble an array of the modules into a hemisphere structure. The number of modules can be expanded, which can improve the angular resolution. However, the increased total power, data and mass must be restricted by the host satellite's capability. The concept will be left as the next step of the study. It is better to have a detail application in place so that the scintillator module dimension and the instrument concept can be further optimized for a clear target.

Chapter 3 Simulation

3.1 Introduction

As mentioned in the last chapter, the design of scintillator detector module has evolved through three iterations. Thorough Geant4 simulation has been done for the first iteration only. For the second and third iteration, only delta simulation has been done to confirm the performance of the revised design. This chapter will present the simulation setup, method and results based on the first design.

Geant4 toolkit[84] is an open source, free license, C++ code assembly for the simulation of the passage of particles through matter using Monte Carlo methods. This code toolkit is developed and maintained by CERN. The original initiative was to assist the design of the nuclear high energy experiment at CERN. Over decades of development, Geant4 evolves into the most sophisticated and accurate simulation tool. Many large nuclear experiments have built their simulation programs using Geant4 toolkit and the simulation results have been checked against the experiment results. This further helps to refine the toolkit.

Geant4 toolkit packs all the underlying physics process and core program into a library with open interface for user to access. Geant4 aims to keep the user task as simple as possible. So, a Geant4 program with minimum user configuration can run itself. To make the simulation meaningful, the minimum task that the user shall define is the detector geometry (including material), the primary particle information (position, direction, energy, species) and the physics list. More advanced application can dynamically interact with the simulation process to acquire the middle step information. Geant4 provides a large collection of example applications, which is also the best place to start building the user application.

The output data from the Geant4 application program is saved in ROOT format. Data in this format can be analysed by another open source toolkit, ROOT[111]. ROOT is

also an open-source, C++ code assembly, which packages all the functions in a standalone library. ROOT is developed and maintained by CERN as well with initiative to process its huge amount of experiment data. ROOT is like a sister code package to Geant4 with seamless connection between them. ROOT toolkit comes with a basic graphic user interface (GUI) program for visual interaction with the data. For batch processing, the user can develop customized code with ROOT functions embedded. In this research, we both used GUI program and developed some simple batch processing scripts to process the simulation data.

Like any simulation program, computation time is a bottleneck. It increases exponentially with the number of simulation runs and the complicity of the geometry. We initially run the simulation on local computer. Later, we immigrate the program to UCL's super computer network: Legion. This network contains a large cluster of CPUs and huge memory. In the cluster, each CPU isn't particularly fast. The advantage is for the parallel processing with multiple CPUs and the access to big memory pool. Hence, if the simulation program can be configured to run in multiple-threads, the simulation time can be greatly reduced. The older version of Geant4 can't support multiple-thread until v10.0 (Dec. 2013). So, we migrate our original program to the newer Geant4 version and compile it with multiple-thread enabled. This greatly speeds up our simulation work.

The great challenge working with Geant4 toolkit is to understand its interface and particularly to understand the definition of the parameters for the interface functions that the toolkit offers for access. The difficulty arises from the lack of detail documentation for each function. The user manual that Geant4 community provides doesn't necessary cover all the functions. In many cases, user has to learn the features from the example code or even directly check out the source code. For example, the available options to configure the material reflectivity property are not documented.

Another challenge is the lack of material data and the lack of understanding of the Geant4 physical processes. To have an accurate simulation, the material properties are very important. These properties, like refractive indies, need to be sourced from the open data, which is rare and is limited to a few common materials. Geant4 has many physical process packages which defines the detail interaction process between particles and matter. For different energy level, the physical process or model is different. For example, for electromagnetic physics, Geant4 provides five options. Each

option represents some different physical process or uses a different mathematical model. It is not straightforward to understand which option the best for our application is.

Geant4 is a very powerful and versatile tool, but it requires a deep learning curve. To simplify the use, many third parties have done second development on top of the Geant4 toolkit to provide much simpler user interface. However, direct development with the Geant4 library provides us the best understanding of the simulation process and the full control and access to every detail.

Another important function that Geant4 lacks, but catching up, is the support to a modern CAD model. A modern detector is so complex that its geometry is usually built in a modern CAD software. Geant4 lacks capability to directly import and convert such CAD models. This function is still under development by Geant4 and it might be available by the time of this thesis writing. However, during the time of the simulation work development, this function is not available. So, our detector geometry is developed using Geant4's built-in geometry definition functions.

3.2 Geant4 at a glance

Geant4 is the short of **GE**ometry **AN**d **T**raking. The name suggests this is a tool about tracking particle that traverses a geometry. This is the core function of the tool. Today, it has grown to a full-scale tool including functions covering all the aspects that are required from the start to the end of the simulation. The simple list for the type of functions is:

1. the geometry of the system,
2. the materials involved,
3. the fundamental particles of interest,
4. the generation of primary events,
5. the tracking of particles through materials and electromagnetic fields,
6. the physics processes governing particle interactions,
7. the response of sensitive detector components,
8. the generation of event data,
9. the storage of events and tracks,

10. the visualization of the detector and particle trajectories, and
11. the capture and analysis of simulation data at different levels of detail and refinement.

This toolkit architecture is described in Figure 3-1. Each block represents a cluster of classes (a term in C++ language). This architecture provides an overview of the functionality coverage of the toolkit. The link between the blocks indicates the logical relationship between them, but the interaction is not restricted to those links shown.

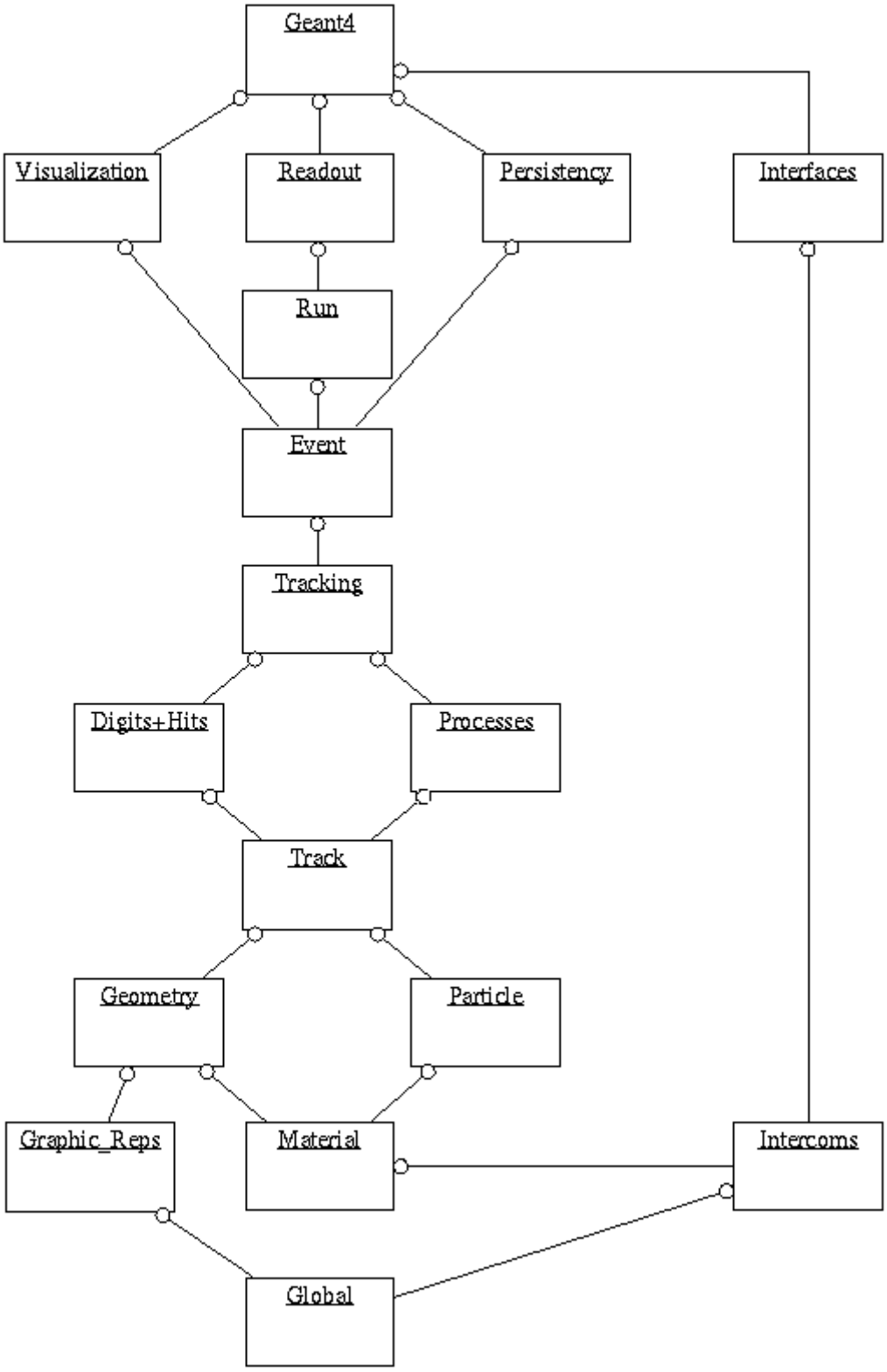


Figure 3-1 Geat4 toolkit architecture[112]

3.2.1 Basic functions

As a minimum, to run the Geant4 simulation, the user needs to understand and define the following three aspects:

1. Detector definition and response;
2. Tracking and physics;
3. User actions.

3.2.1.1 Detector definition and response

The detector definition includes the detector geometry and material, which form the base of the detector to be simulated. As mentioned in the introduction, Geant4 doesn't support importing the CAD model directly from the mainstream CAD software. It provides a set of built-in functions to generate some basic shapes like cube, sphere, cone etc. To build the detector geometry, the user needs to code in each basic shape, and position them precisely in a defined coordination system. For example, our detector module is made up of three semi-spheres and one cube. If the overlap detection option is enabled, Geant4 will report any volume confliction.

This approach is only practicable for simple geometries. Also, it is not convenient to share the hard-coded geometry with other programs like ROOT. To address this, Geant4 provides another approach, called Geometry Description Markup Language (GDML) [113]. GDML is an XML based script language, which allows the geometry to be defined in a simple human readable description language in a structured format. The GDML script can be directly imported by Geant4. In the other direction, Geant4 can also export the hard-coded geometry in GDML format.

The GDML is still not the ultimate solution, as modern complex detector is often developed in CAD software, which doesn't support GDML format. Some third-party software is trying to bridge the gap by providing an intermediate conversion function. For example, GUIMesh[114] makes use of the FreeCAD library to convert the geometry from STEP format (a common format that most CAD software support) to GDML format. CADMesh[115] can load triangular mesh based CAD files into Geant4. FastRad[116] provides a commercial solution.

Each component of the geometry needs a material associated to it. The material property plays an import role in affecting the physics interaction between particle and matter. Geant4 has a good size predefined material database, which covers all the elements in the periodic table and NIST material database[117]. However, not all the

needed properties are included and many newer materials particularly composites are not included. For those, Geant4 provides function to manually enter the data.

Detector response in Geant4 is enabled by the definition of sensitive volume. A component of the geometry that is defined as sensitive volume can collect and output the data of any particles i.e. secondary photons, that enters the volume of the component. For example, in our detector geometry, the cube representing the SiPM sensor is defined as a sensitive volume, which aims to collect the scintillation photons. The Teflon wrapping layer is left as non-sensitive volume as we are not interested in any secondary particles that interact with that layer. The sensitive volume can be further tailored to be only “sensitive” to a particular particle. In our case, we configure the SiPM sensor to be only “sensitive” to photons. This feature helps the user to reduce the amount of simulation computation and memory consumption.

3.2.1.2 Tracking and Physics

Geant4 kernel tracks the life of every primary or secondary particles during the interaction. A new track thread is created when the particle is born. The thread is killed when the energy of the particle is below a pre-set threshold. Along the track, the velocity, position of the particle at each step is recorded. Users can access the step data interactively, but it is not recommended as too frequent access slows down the simulation. It is more efficient to access the data through a sensitive volume mentioned above.

Geant4 maintains a full range of physics modules that define the particle interaction with matter at different energy, different process and different particles. There are following types of physics modules:

1. electromagnetic physics;
2. extra physics processes for gamma and leptons;
3. decay;
4. hadron elastic;
5. hadron inelastic;
6. stopping particles capture processes;
7. ion nuclear interactions;
8. step limiters;

9. others.

Each application shall call at least one physics module or its submodule. Multiple modules or submodules can be grouped to form a PhysicsList. Geant4 predefines a few PhysicsLists for some typical application.

Each physics module is built on some mathematic model, which simulates the particle interaction numerically. Sometimes, there are a few models for the same physics process. Each model is optimized for different aspects. So, a careful selection of the best module for the application requires a good knowledge of the module definition and the application needs.

To maximize the computation efficiency, enable only those physics modules or submodules which are relevant. Even within the module, non-relevant or non-interest particle and process can be disabled. For example, for our application, we only enable electron, proton, gamma and heavy ions. The other particles like positron, muons etc and their associated process are all disabled.

3.2.1.3 User actions

User actions are a set of interface Classes that allow user's code to interact with the Geant4 kernel. There are three mandatory actions that must be realized in any user application, which are DetectorConstruction, PhyscisList and PrimaryGenerationAction. Besides, there are five optional actions:

1. UserRunAction
2. UserEventAction
3. UserStackingAction
4. UserTrackingAction
5. UserSteppingAction

The optional user actions provide user interface to the different level of the simulation run. For example, UserRunAction is called only when a run starts and ends. UserEventAction is called when an Event starts and ends. An event is defined as a tracked particle hits the sensitive volume. In our application, UserEventAction is the main access to collect the event data. UserStackingAction and UserTrackingAction provides access to manage the track. UserSteppingAction is called when a step starts and ends.

A step is defined as one move of the primary or secondary particle. Figure 3-2 is the state machine of a simulation run. The logic relationship of all the five user actions can be seen clearly in the state machine.

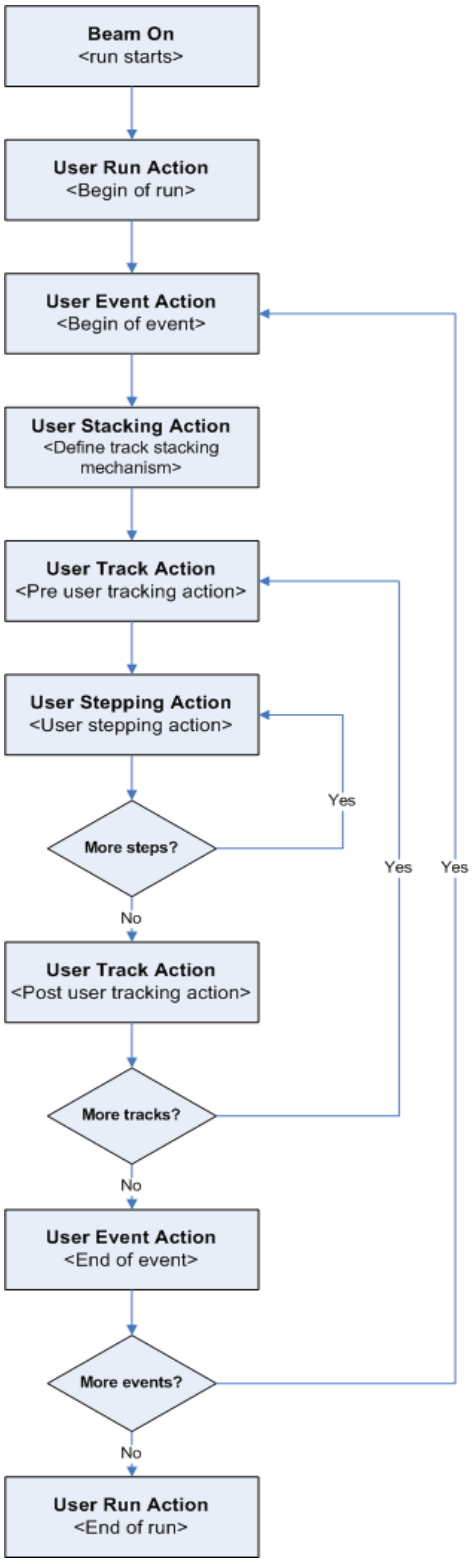


Figure 3-2 State machine of a complete run cycle in Geant4 (reproduced from Geant4 user manual)

3.2.2 Monte Carlo Methods

To solve the particle transport in a geometry, there are two approaches. One is to solve the analytic equation and the other is to use the numeric approximation. For complex 3D geometry, the latter is more efficient in computation.

The numeric method adopted by the Geant4 toolkit is Monte Carlo (MC). Monte Carlo is a very old method. The first example can be dated as early as 1777 by Comte de Buffon. He developed a needle tossing experiment to calculate the π. The method is further developed and applied first on the computer during the Manhattan project (1940s). M.Berger[118] developed the first complete coupled electron-photon transportation code using MC methods.

The basic idea of Monte Carlo method follows the following steps:

1. Define the distribution of the input
2. Generate a random input according to the input distribution
3. Calculate a score function with the input and check the output to a predefined criterion. Accept the input if the criteria are met and reject it if not.
4. Repeat from step 2 until statistically enough samples are collected.
5. Analyse statistically the results.

This is basically how Geant4 calculates each step of the particle traverse. The feature of random sampling with Monte Carlo method also matches the random nature of the particle motion. The accuracy of Monte Carlo method improves as the number of samplings increases.

Certainly, the underlying mathematics process and implementation of Monte Carlo methods are very complicated. More details can be found in publications [119][120][121].

3.2.3 Optical Photons

The heart of our simulation lays in the scintillator photon creation, transmission and boundary interaction. In Geant4, a photon is defined as optics (a wave) when its wavelength is much greater than the typical atomic spacing. While, for high energy optical photon with very short wavelength like gamma ray, the treatment in Geant4 is different.

When a photon is treated as optics, all the physics about wave propagation in and through media apply, like Snell's law.

An optical photon is produced when a charged particle or gamma traverses:

1. a dielectric material with velocity above the Čerenkov threshold;
2. a scintillation material.

Here two physical processes are involved: Čerenkov radiation and scintillation. Čerenkov radiation occurs when a charged particle passes through a dielectric medium at a speed greater than the phase velocity of light in that medium. We know light speed in a vacuum is **c**. The speed of light in a medium will be less than **c**. For example, light in water is only 0.75**c**. So, particle with sufficient energy can possibly exceed the light speed in the matter. As the high energy particles traverse, it disrupts the local electromagnetic field in the medium. If the particle is slow (below speed of light in that media), the medium will return to mechanical equilibrium in a short time. If the particle is fast (above speed of light in that media), the disturbance will be left in the wake of the particle and the energy contained in this disturbance will be released in a form of photon. Usually, the emitted photon is in ultraviolet spectrum. This process is called Čerenkov radiation. In our experiment setup, the involved charged particle or gamma don't have high enough energy to cause Čerenkov radiation, so we disable this process in the simulation.

Scintillation is another physical process that produces photons. It occurs in some special crystal or materials that can fluorescent upon absorbing of incident energy. The detail mechanism is discussed in the Chapter 2. Here, we focus on the Geant4 related setup about scintillation process.

In Geant4, to produce a proper scintillation reaction, we need to set the optical property of the involved volume materials correctly. A few key properties are listed below:

1. Photon yield
2. Emission spectrum
3. Emission angle distribution
4. Absorption length spectrum
5. Refractive index spectrum
6. Emission decay time

To be closer to reality, Geant4 also allows the user to define the statistical yield fluctuation of photon generation through a parameter called RESOLUTIONSCALE.

Geant4 defines three kinds of interactions for optical photons with matter:

1. Elastic (Rayleigh) scattering;
2. Absorption;
3. Medium boundary interactions.

Of the three, only 2 and 3 are closely related to our experimental setup.

3.2.3.1 Absorption

Material will absorb photons. The mean path for a photon at a wavelength/energy is defined as absorption length. Geant4 simply kills the photon when it is deemed “absorbed”. Absorption is a material property that user can input in a spectrum array.

3.2.3.2 Medium boundary interactions

This is a critical and complex part to set in Geant4 for optics transport. The nature of the two materials that form the boundary determines the photon boundary behaviour. Geant4 defines three types of boundaries:

1. dielectric to dielectric
   where photon can be refracted or reflected.
2. dielectric to metal
   where photon can be absorbed by the metal or reflected into the dielectric
3. dielectric to black material
   A black material means an artificial material that user doesn’t define any optical property. In this case, the photon is absorbed undetected immediately by the black material.

The surface property of the dielectric can be further defined into different types. For each type, the photon transport at the boundary is treated differently. The collection of process is called UNIFIED model in Geant4. Figure 3-3 presents the structure of UNIFIED model.

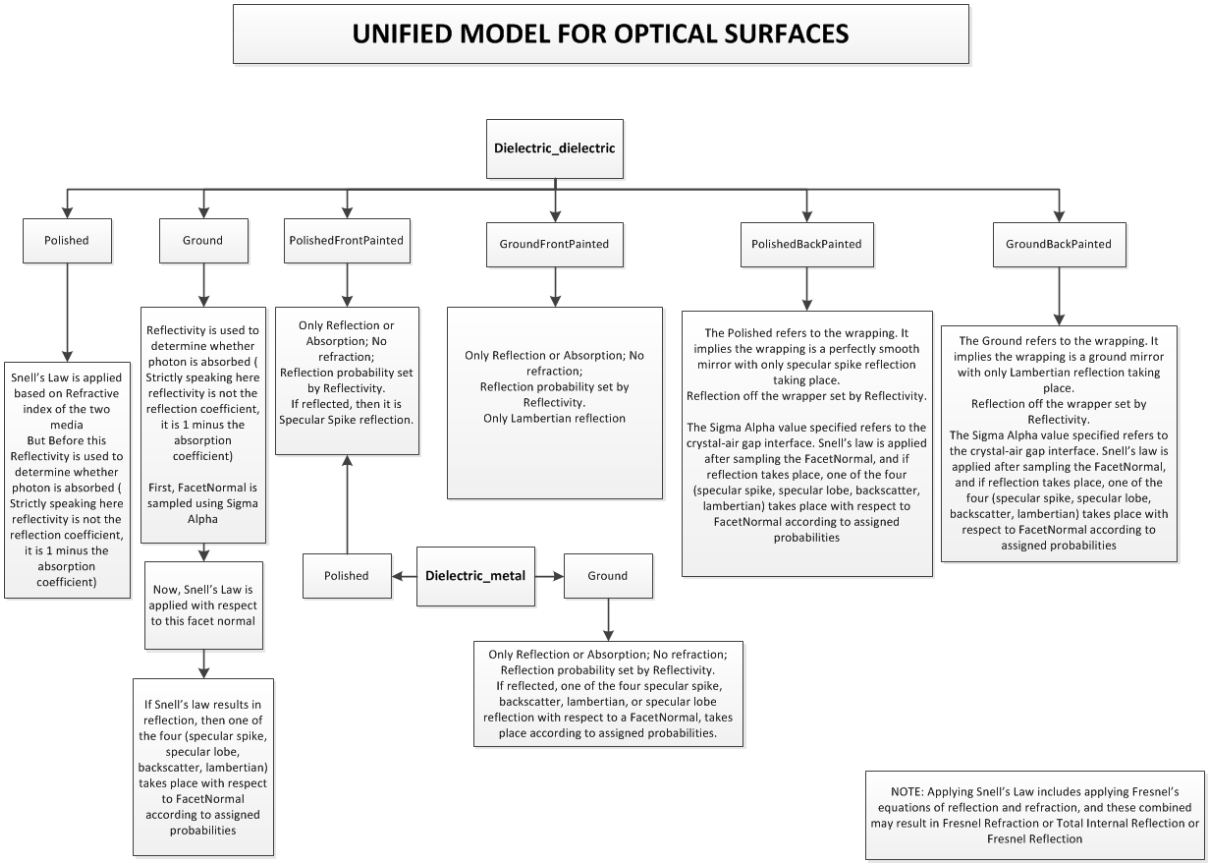


Figure 3-3 UNIFIELD model for optical surfaces [122] (reproduced)

3.3 Simulation Overview

We construct the simulation work in three parts: the first is the Geant4 program, which runs all the simulation and generate all the data. All the detector module information and the interested physics process are included in the code. The second is a small C++ utility which converts the SPENVIS particle flux data into a macro script as the input data to Geant4 program. The aim is to test the detector response in a real environment. The third is the post data analysis program interfaced to ROOT. At this stage, we only plot all the results and we haven't gone very deep into mining the data.

Figure 3-4 shows the work flow of the simulation and highlights the major function of the full system. The upper block is the Geant4 program. Four blue circles represent the user configuration and inputs. Once all set, the Run function operates the Monte Carlo simulation. The lower block is the data storage and post data analysis in ROOT. The further data mining in Python or MATLAB are for future work.

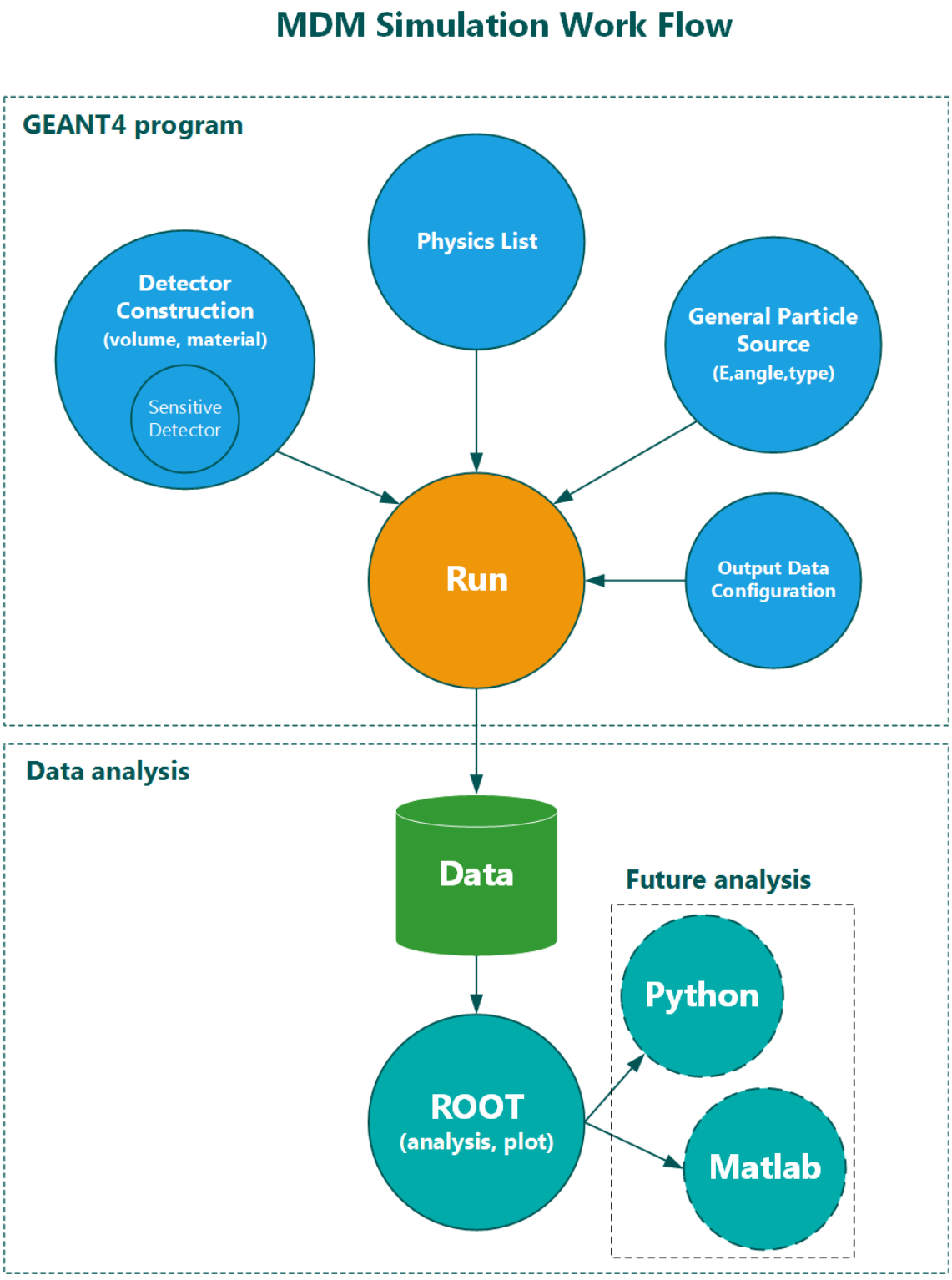


Figure 3-4 Simulation work flow and complete program structure

3.4 Simulation Setup

It is important to setup the detector module with as much detail as possible to achieve a good simulation accuracy, but this comes at the cost of computation time and complexity of the Geant4 program. So, a trade-off needs to be made. In our program, we put as much detail as possible into the scintillator detector module while we keep the SiPM detector as simple as a big silicon block, whose role is only to collect the scintillation photon.

Any Geant4 application must realize three compulsory classes: detector construction, physics list and particle source. Geant4 provides default version but user needs to inherit them and add application specific codes. Next, we will discuss what is included in each class.

3.4.1 Detector Construction

This class includes the definition of geometry and material. As introduced above, there are two ways to construct the detector geometry in Geant4. Since our detector module is made up of a few simple volumes, we simply hard-coded it.

Each volume shall be assigned with a material type. For each material, a list of property shall be defined. Some basic properties can be called from Geant4's built-in material database. Special properties like those relates to the optics behaviour can be manually input with data collected from open publication. For those missing data, simplified or approximate data are used. For example, if the refraction index spectrum is missing, a single constant figure can be used across the spectrum.

With both volume and material defined, the next step is to specify the exact position of each volume in the application coordinate system. Geant4 can check the spatial confliction to avoid miscalculation of any volume position.

The last step is to define the sensitive volume. This is simply to assign a previous defined physical volume as a sensitive volume and reference it to the assigned response function.

Figure 3-5 shows the construction of the detector module and the SiPM sensor. It is made of four basic volumes: fast scintillator (cyan and transparent), slow scintilla-

tor(green), reflection wrap layer(transparent) and the SiPM sensor(blue). The four volumes are placed together with no gap between any of two adjacent volumes. Table 3-1 describes the base material and size of each volume. A large aluminium block (grey) is place behind the SiPM sensor to simulate the shielding effect of the spacecraft body.

Table 3-1 Detector construction in the simulation

Composite	Material	Shape	Max Depth
Fast Scintillator	BC408	Semi-sphere	0.2mm
Slow Scintillator	CsI(Tl)	Semi-sphere	1.5mm
Reflection Wrap Layer	Teflon	Semi-sphere	0.055mm
Photon Detector	Si	Box	1mm

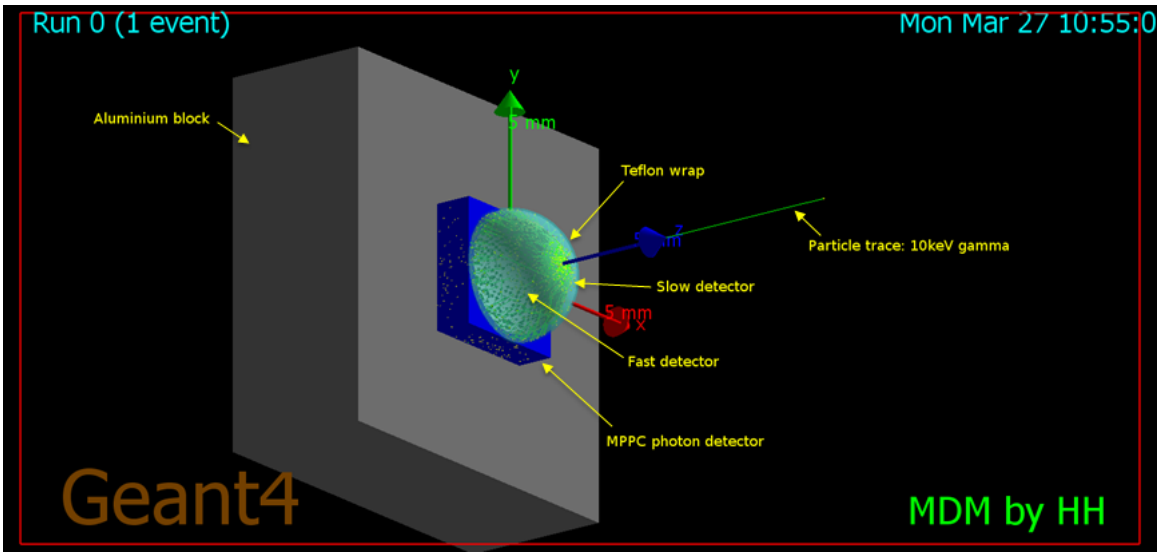


Figure 3-5 Detector module prototype construction in Geant4 program

The optical property of each volume is critical for the photon detection efficiency. Any incident particle with sufficient energy to penetrate the Teflon layer will interact with either the fast scintillator or the slow scintillator and the later subsequently emits a series of scintillation photons. These photons will either be reabsorbed or find a path to reach the SiPM sensor, where they are counted. The basic process as described above is very simple. The complexity is at the boundary of two mediums/volumes, where Snell's law dominates the process. In Snell's law, the refractive index of the medium is important to determine whether the photon will be reflected back or refracted. For some materials like CsI, its refractive index has a wavelength dependence

as shown in Figure 3-6. For other materials like BC408, whose dependence is neglectable or data is just not available, a single constant refractive index across the wavelength has been used. The key optical property of each volume is defined in the following tables (Table 3-2,Table 3-3,Table 3-4 and Table 3-5).

Table 3-2 Optical property of the reflection wrapper - Teflon

Property	Value	Unit
Absorption length	100	cm
Refractive index	1.35	
Reflectivity	100	%

Table 3-3 Optical property of the SiPM sensor – Si based with window cover

Property	Value	Unit
Refractive index (for window)	1.5	

Table 3-4 Optical property of the slow scintillator – CsI (Tl)

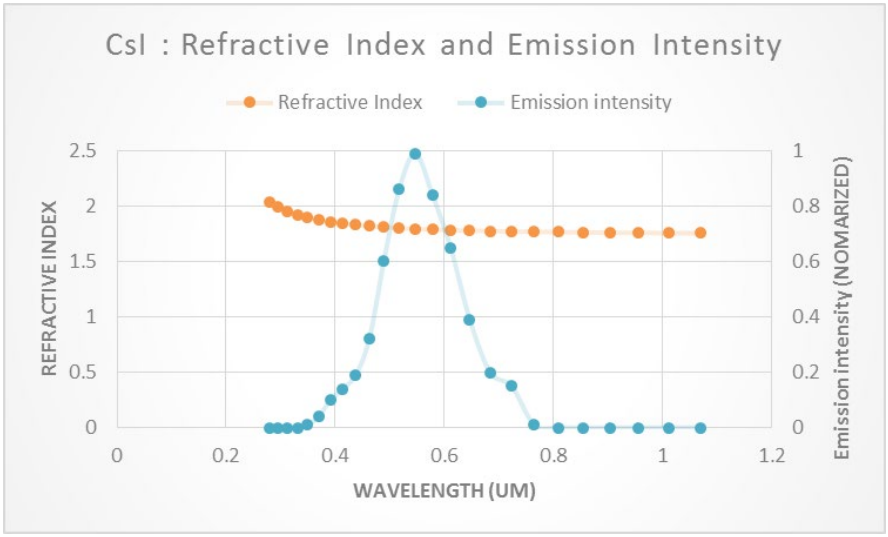
Property	Value	Unit
Photon yield	54000	#/MeV
Density	4.51	g/cm^3
Decay time (mean)	1000	ns
Absorption length	39.3	cm
Refractive index		

Figure 3-6 CsI refractive index and emission intensity against wavelength

Table 3-5 Optical property of the fast scintillator - BC408

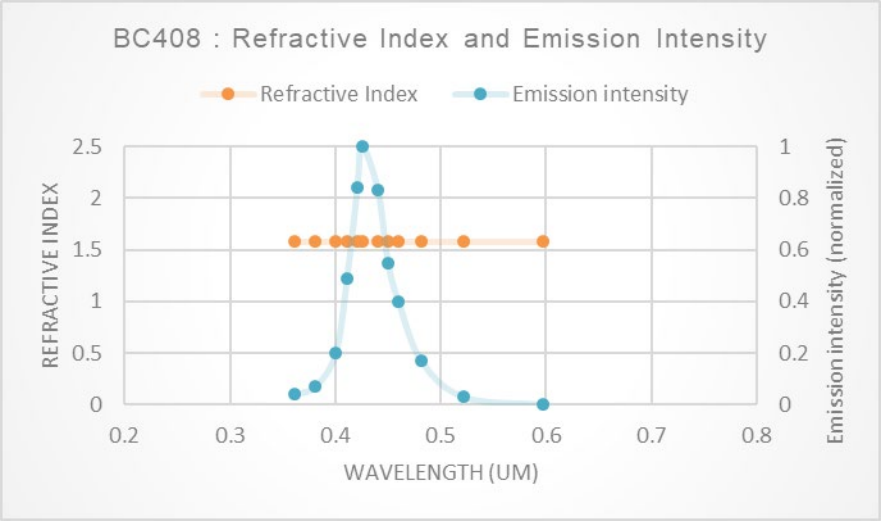
Property	Value	Unit
Photon yield	500	#/MeV
Density	1.032	g/cm^3
Decay time(mean)	1	ns
Absorption length	210	cm
Refractive index		

Figure 3-7 BC408 refractive index and emission intensity against wavelength

3.4.2 Physics List

This class defines the physics processes and the particles used in the simulation. Geant4 has a large library of physics modules derived from the established standards and theoretical models. To save simulation time and make accurate simulation, it is important to tailor the selection of physics processes and particles to the application need.

Our application involves scintillation and basic particle interaction with matter in medium to low energy range. So, our simulation program only registers two modules: G4Opticsphysics and G4EMStandardPhysics.

The electromagnetic physics package (G4EMStandardPhysics) defines the interaction with matter for low to medium energy for γ, $e-$, $e+$, $\mu-$, $\mu+$, $\tau-$, $\tau+$, and all stable charged ions.

The optical physics package (G4Opticsphysics) defines the scintillation and optical behaviour.

The main physics processes involved in the application are described in Table 3-6.

Table 3-6 Main physics process used in the application

Particle	Physics Module	Range
Electron	Ionization, continuous energy loss defined by **Berger-Seltzer** formula	<1GeV
Proton, Alpha	Ionization, continuous energy loss defined by **Bethe-Bloch** formula	<1GeV
Gamma	Photoelectric effect; Compton scattering	<10MeV

3.4.3 Particle Source

This class defines the primary particle type, energy, position and direction. These parameters can be defined in one of the three ways: hard-coded, GUI interface and macro script via the specific class GeneralParticleSource (GPS). We adopted the third way for its flexibility and programmability.

As a wide range of particle type, energy, position and direction needs to be covered, we developed a small utility C++ program to generate the GPS macro script automatically. It was found particularly useful for the angular response simulation, where the particles are injected from a 4π sphere breaking into 648 blocks.

3.5 Two Simulation Tests

Two simulation tests are performed on the detector module as **energy response** and **angular response**.

3.5.1 Energy Response

In energy response test, the primary particle is placed at a fixed position and direction. Then its energy is swept across a predefined range with either a number of discrete steps or a continuous distribution by random sampling. At each step/sampling point, 10000+ runs are committed to provide sufficient statistics. Geant4 allows user to specify the energy spectrum in different distribution, i.e. uniform, Gaussian, logarithm or even arbitrary (user defined shape in a two-column list). In our test, we mainly use a

uniform distribution. Figure 3-8 shows an example primary particle energy distribution used in one of the energy tests. The fluctuation shows the effect of random sampling and the macro profile matches the uniform distribution. In this example, 5000 primary particles are generated.

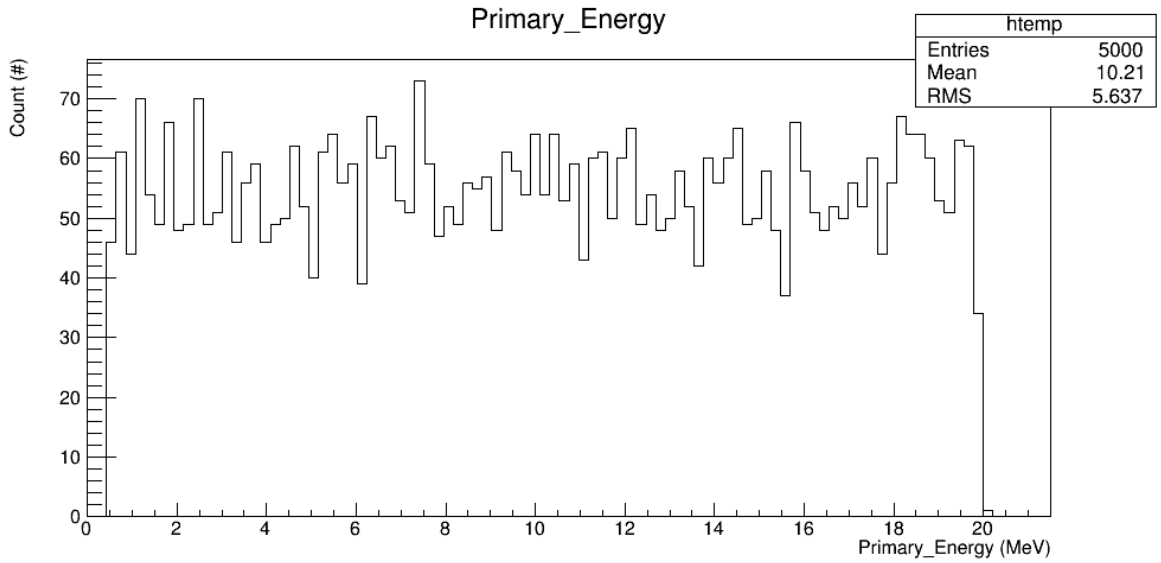


Figure 3-8 An example of primary energy distribution for a proton test

The predefined energy range is tailored for the detector size and the primary particle type. Both has an impact on the required energy to penetrate through. The minimum end of the range is defined by the energy required to penetrate the wrapping layer minus some margin. The maximum end is defined by the energy required to penetrate the entire detector module plus some margin. Table 3-7 summarizes the energy range defined for each type of primary particle.

Table 3-7 The energy range of primaries

Particle	Min	Max	Unit	Beam On*
Electron	0.01	20	MeV	50000
Proton	0.5	20	MeV	5000
Alpha	10	80	MeV	2000
Gamma	0.001	1	MeV	500000

* Beam on means the number of primary particles fired onto the detector.

The number of primary particles shall be high enough for the results to be statistically meaningful. However, the computation time shall also be kept within reasonable limit. It is relatively quick to calculate the electron and gamma interactions, but it is much

slower for protons and alpha particles. Therefore, the chosen particle numbers are different.

The stopping power and penetration range for the charged particles

Both are energy and material dependent. NIST [81] database has the data for electron (Figure 3-9), proton (Figure 3-10) and alpha particles (Figure 3-11) for the materials used in our experiment. In each figure, the penetration range of the particular particle in each material is plotted against the energy. The thickness of each detector volume (the straight lines) is drawn for comparison. The intersections region (marked with red vertical dashed bar) are the effective energy range that the full detector module can capture. The NIST data is used as a preliminary prediction for the effective energy range, which is fed into the Geant4 simulation for fine checking. This is the basis of the data selected for Table 3-7.

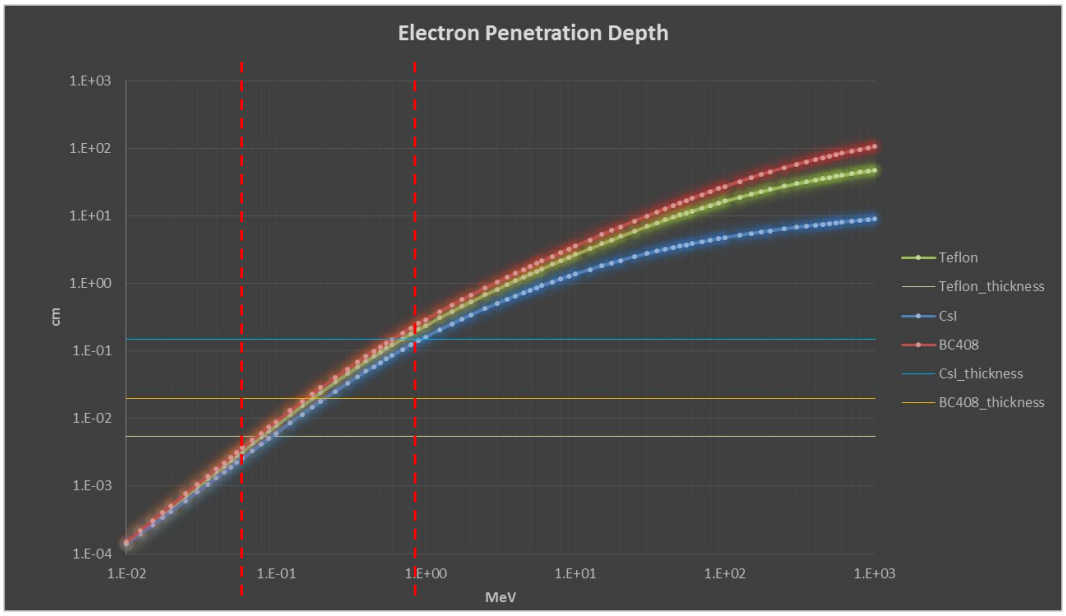


Figure 3-9 Electron penetration depth against energy. Data is sampled from NIST ESTAR database.

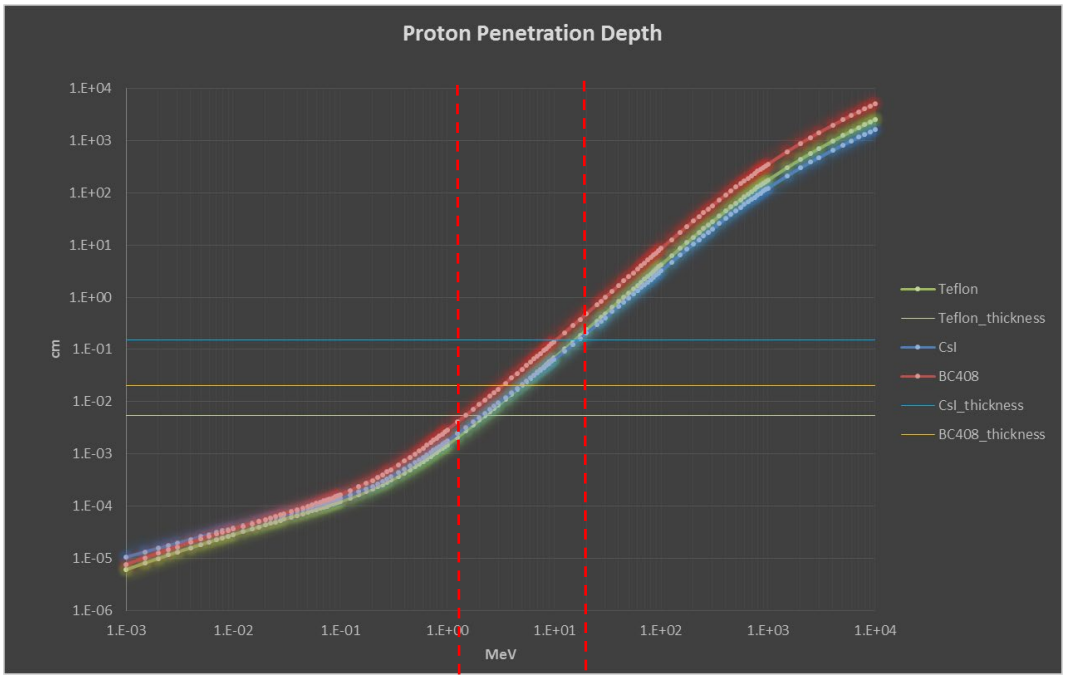


Figure 3-10 Proton penetration depth against energy. Data is sampled from NIST PSTAR database.

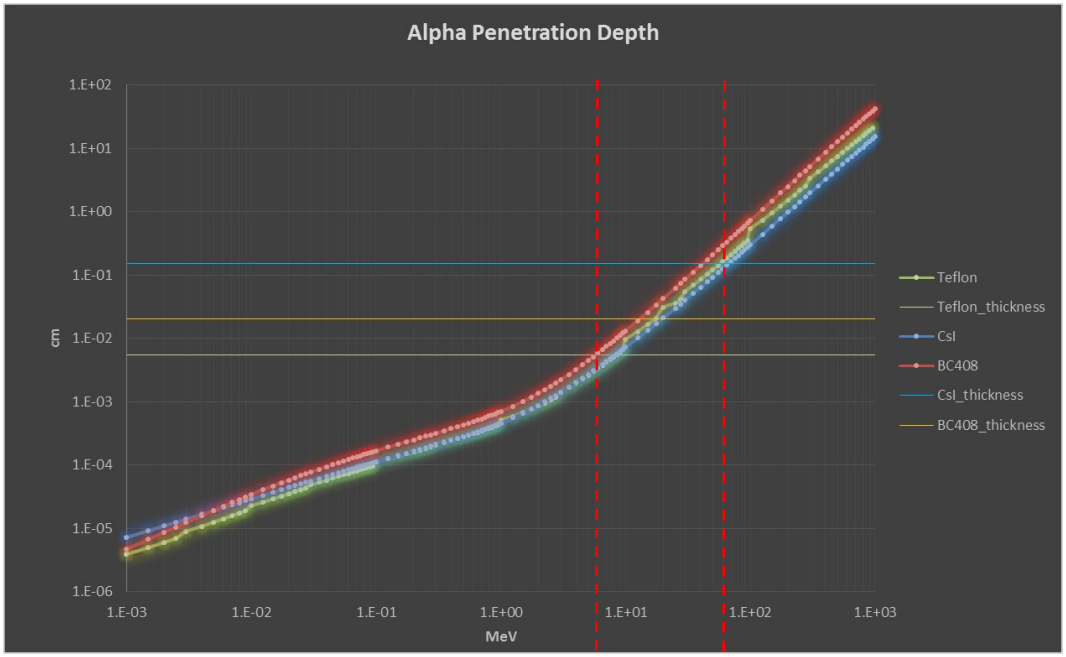


Figure 3-11 Alpha penetration depth against energy. Data is sampled from NIST ASTAR database.

The mass attenuation length for photon

The treatment for high energy photon (gamma in this case) penetration range is different from the charged particles. The process is described as a percentage of the photon flux that will be absorbed by the target medium with thickness **t** over its mass attenuation length **λ**. The percentage depends on the target material density, thickness and mass attenuation coefficient.

$$\lambda = 1/(\mu \times \rho) \tag{3.1}$$

$$I = I_0 \exp(-t/\lambda) \tag{3.2}$$

Where,

ρ is the density (g/cm^3)

μ is the mass attenuation coefficient (cm^2/g)

λ is the attenuation length (cm)

t is the thickness of matter that the photon is traveling through

I_0 is the initial intensity

I is the intensity after traversal of the thickness t.

Using the NIST XCOM database[82], the gamma absorption in CsI is calculated. From Figure 3-12 we can see gamma energy up to 60keV can be 100% absorbed for CsI thickness of 1.5mm (radius of the slow scintillator.)

A very small percentage of photons will be absorbed by the Teflon volume and the BC408 volume. Their contribution to the total loss is so small that it can be safely ignored.

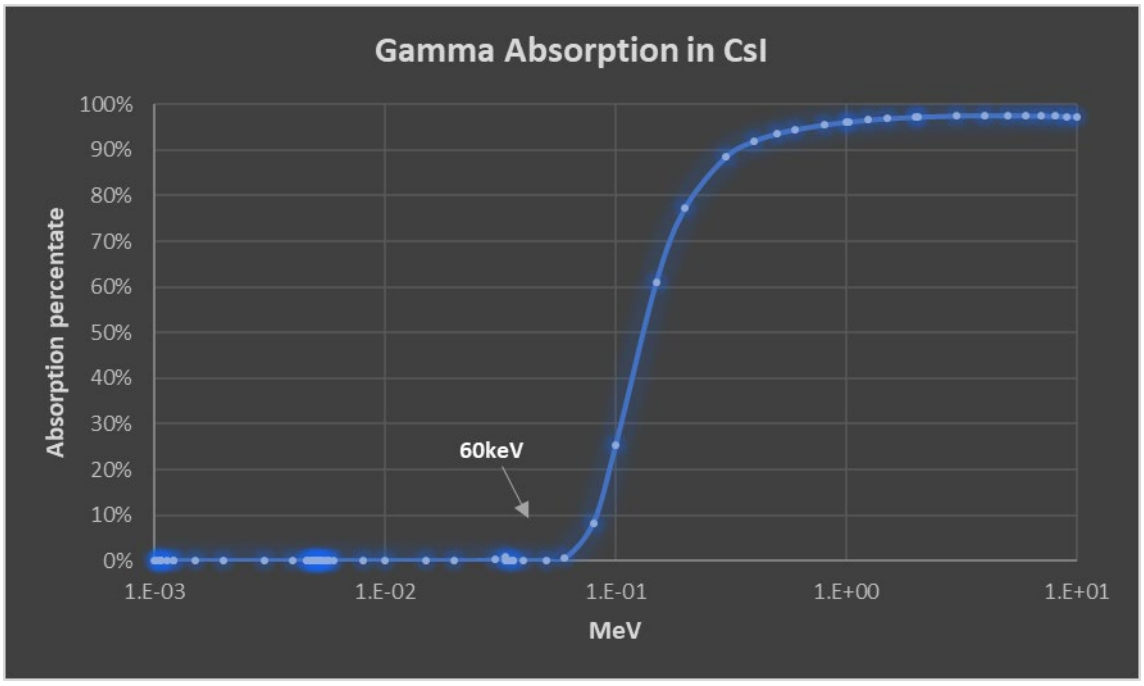


Figure 3-12 Gamma absorption length in CsI against energy. Raw data is collected from NIST database XCOM.

3.5.2 Angular Response

In angle response test, the incidence direction of the primary particle is swept through a 2π or 4π solid angle. To simplify the test setup, the 4π sphere is divided into 648 blocks of size 10º x 10º in altitude θ and azimuth φ. With the same block size, a 2π hemi-sphere is divided into 324 blocks. Note, the surface size of each block in altitude direction is different due to the curvature of the sphere. The nominal direction of each block determines the incident angle of the primary particle. This setup means the angular response will be tested at 10° step in both altitude and azimuth direction.

We use a circle plane source to cover as much area of each block. The primary particles are launched from the circle plane with a uniform spatial distribution. The source plane is placed far enough from the detector module at the origin so that the circle plane source can be as big as the cross-section of the detector module from the view of source. This is to ensure a uniform illumination.

The macro script for the 648 or 324 plane sources are generated by a C++ utility program. Figure 3-13 shows the concept and the steps that the utility program generates the macro script. Appendix B shows how the circular plane source orientation is derived from its position vector and a typical micro script for a circular plane source in a block.

This setup does have two limitations. One is that the circle plane doesn't cover the entire area of the block. The other is the density of blocks in the 4π sphere is not uniform. There are more blocks concentrated in the polar region than in the equator region.

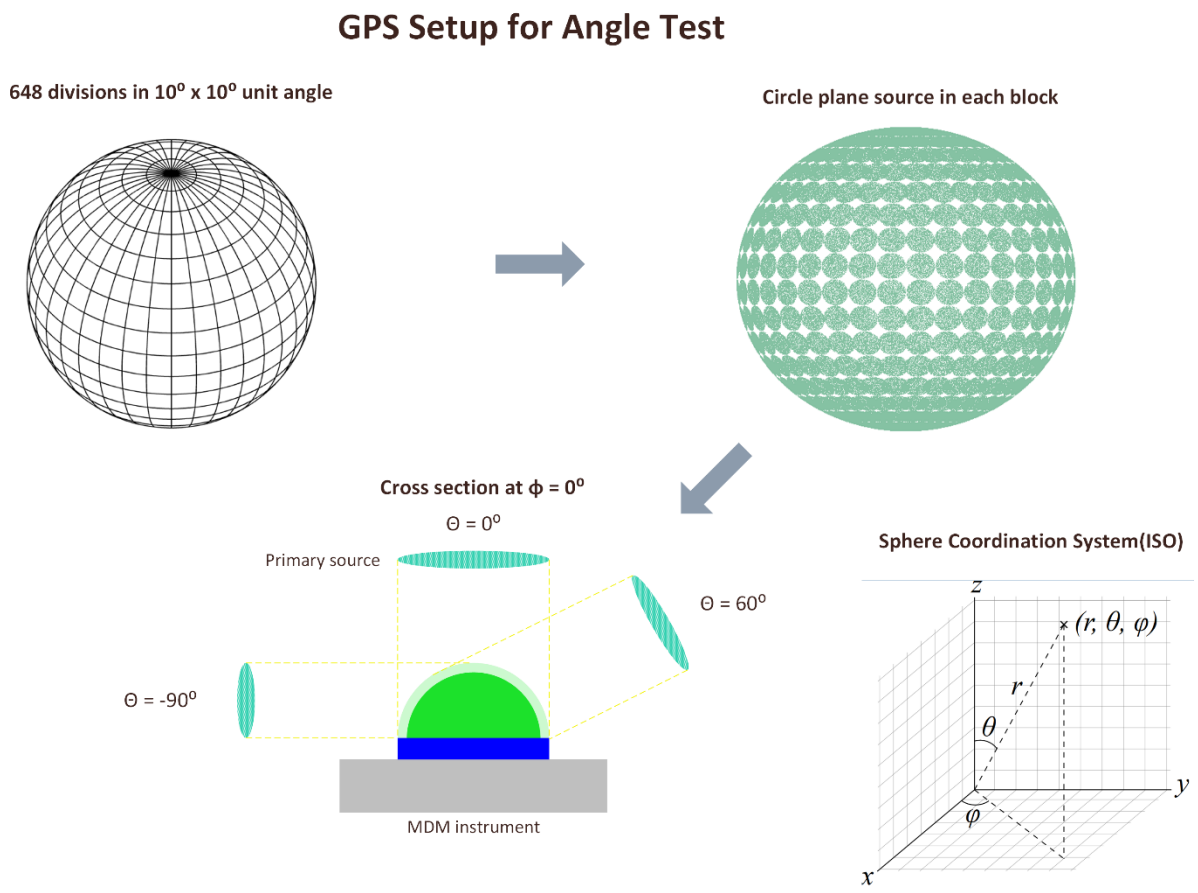


Figure 3-13 Global Particle Source (GPS) setup for angle test.

We select a few mono energies within the valid range for each particle type. Mono energy removes the response difference due to the energy difference. Table 3-8 shows the setup detail.

Table 3-8 Angle response test primary source setup

Particle	Mono Energy	Blocks	Beam On (per block)
Electron	100k,250k,500k,750k	648	1000
Proton	2M, 5M, 7.5M, 15M	324	75
Alpha	20M	324	75
Gamma	100k	648	1000

As our detector module only has 2π field of view, a 2π coverage with 324 blocks is sufficient in principle. However, around the equator, it is desirable to have more coverage. To simplify the program, we just use a 4π coverage with 648 blocks. Due to the computation time limit, we can only afford to do so for electron and gamma. For proton and alpha, we cut back to 2π coverage. The limit is so severe that we have to cut the number of primaries in each block as well.

3.6 Results

3.6.1 Initial Results

Some trial runs are conducted as a quick check for the simulation setup and Geant4 program. We use gamma as the trial source as it is quick to compute. The trial runs are manually commanded to allow more interaction. Detail trial run data can be found in Appendix C: Initial Simulation Data.

3.6.1.1 Photons Yield in CsI Scintillator

There are 34 discrete points between 5keV and 1000keV selected for the Gamma primary. For each point, the total energy deposited in the CsI scintillator and the total photon received by the SiPM sensor was recorded. The ratio of the two is the photon yield. Here we assume the scintillation photon are almost 100% detected. This can only be true in an ideal simulation configuration with almost no optics transmission loss.

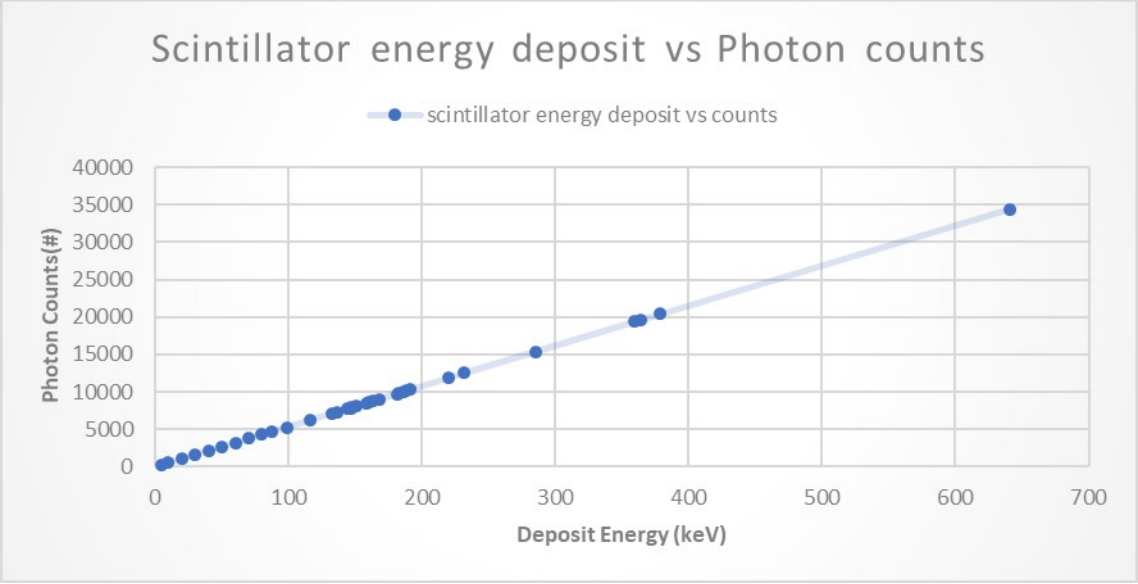


Figure 3-14 Photon yield in CsI scintillator for gamma primary in the range of 5keV to 500keV

Figure 3-14 plots the detected photon counts against the deposited energy. The slope of the fit line is the photon yield. We calculate the slope value at 54 photons/keV. This exactly matches the photon yield setting for the CsI scintillator in the Geant4 program.

3.6.1.2 Scintillator Response Timing

The scintillator response time is a critical data for particle identification and anti-coincidence algorithm. We record the timestamp of every photon at the moment it is detected by the SiPM sensor. In real experiment, it is hard to distinguish the source of photon that is detected. In the simulation, every photon is fully trackable to its origin, so we can clearly identify the source of every photon. Figure 3-15 plots the histogram of photon arrival time for both BC408 and CsI(Tl) scintillators. The histogram is generated in ROOT, which automatically calculates the statistics of each histogram and shows them in a top right box of the plot. From the statistics, we read the mean arrival time for BC408 is 1.249ns and for CsI(Tl) is 966.3ns. These figures match the simulation setting of 1ns and 1000ns well. The slightly bigger error (in terms of percentage) in BC408 timing is probably due to the extra photon transport time from BC408 through CsI(Tl) to the SiPM.

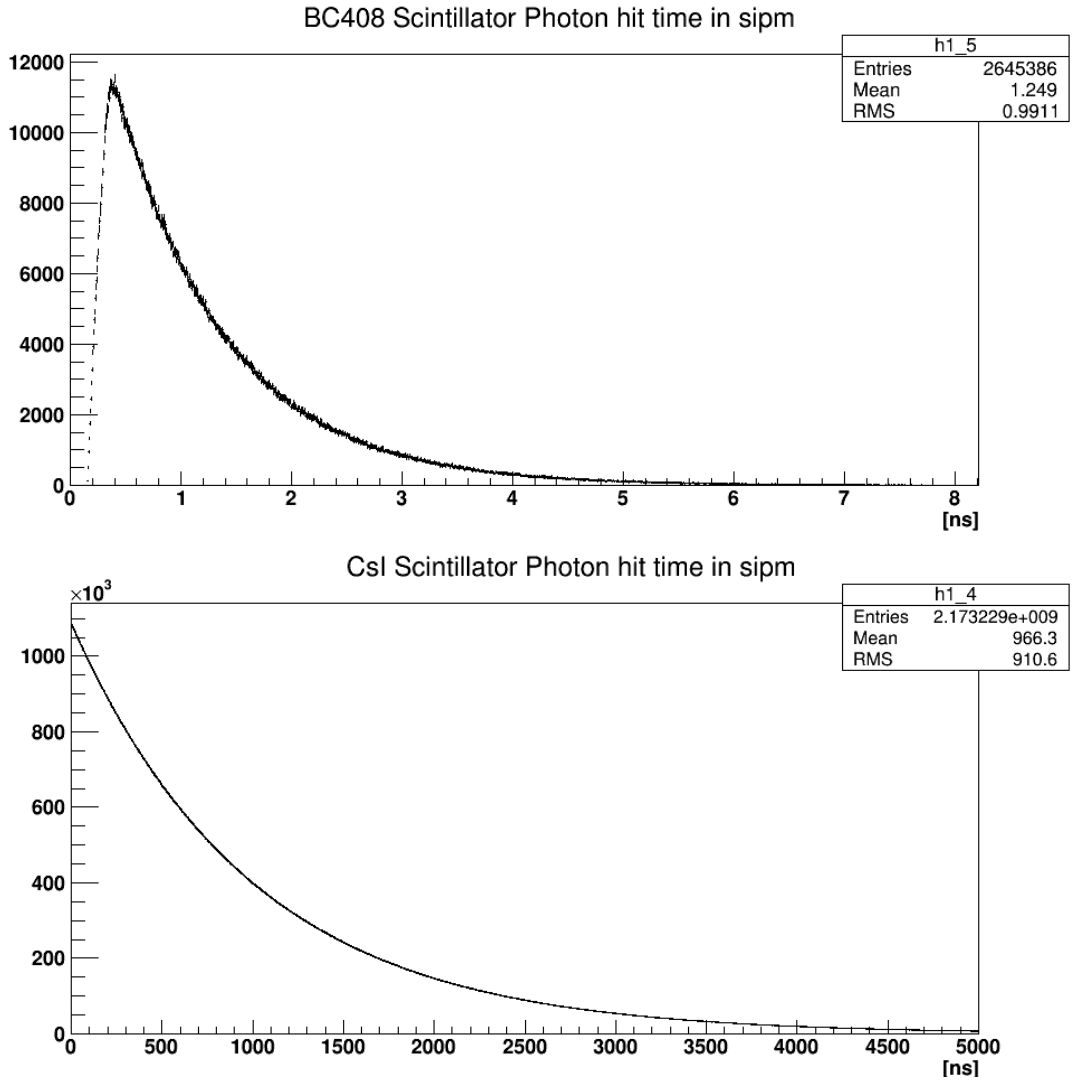


Figure 3-15 Histogram of scintillation photon arrival time. Top is BC408 and bottom is CsI (Tl)

The actual profile of the photon arrival time (≈ detector response time) matches the exponential decay distribution. The long tail can extend to multiple of the mean arrival time and the onset of the profile is much shorter than the mean arrival time. This potentially will create an overlap period when it is difficult to identify the source of photon by timing.

We also studied the relationship between arrival time and the incident energy. We found no correlation between them as shown in Figure 3-16.

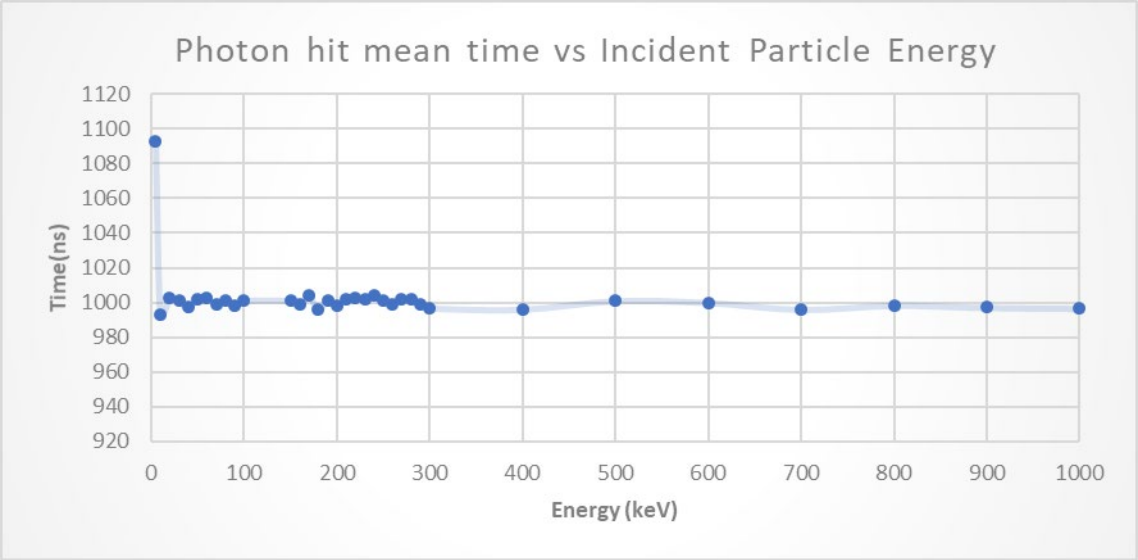


Figure 3-16 Average scintillator response time at different primary energy.

3.6.1.3 Photon Energy in the SiPM sensor

We collect the energy of the arrival photons and plot the histogram. The result is shown in the top left of Figure 3-17. The black dot is the real data and the red line is the Gaussian fit. The top right is the CsI(Tl) emission spectrum that we manually enter into the Geant4 program. We find the two plots matches very well. This proves that Geant4 generates the scintillation photon closely matching the input spectrum. The random sampling has introduced some fluctuation but over accumulated samples, the input spectrum is clearly seen. The match also proves that the scintillation photon energy/wavelength is not altered through different medium.

The bottom left of Figure 3-17 is the CsI(Tl) and CsI(Na) emission spectrum from the manufacturer datasheet. The bottom right is the spectrum we extract from the datasheet plot for CsI (Tl). The top right is just the conversion from wavelength to energy for the bottom right plot.

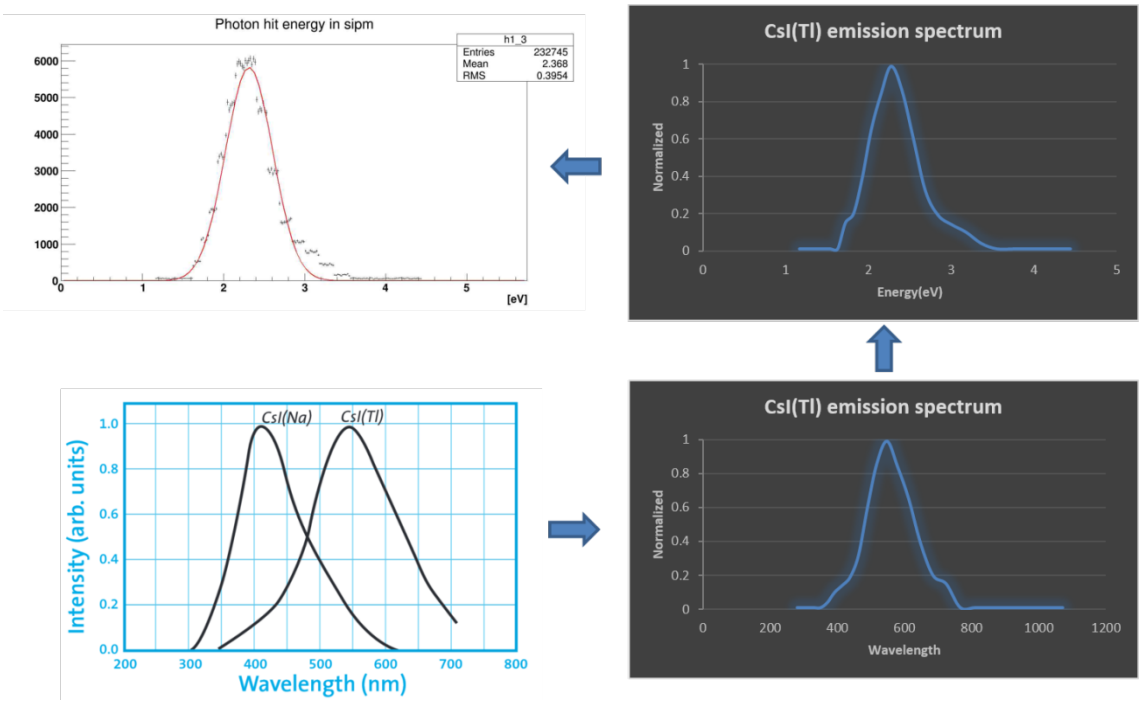


Figure 3-17 Secondary photon wavelength/energy intensity distribution from datasheet to simulation data.

3.6.1.4 Photon Hit Position at the surface of SiPM sensor

In the real experiment with one SiPM sensor, the photon hit position is not resolvable. In the simulation, we can track the hit position. This provides some interesting information. It indicates whether the photon spatial position is uniformly distributed or whether it is sensitive to the direction of the incident primary. It can also indicate what the impact of the detector geometry is on the photon spatial position. If the single sensor is replaced with a matrix of sensor, the above knowledge can help us to develop the primary direction identification algorithm.

We experiment a few discrete directions: θ = 0º, 30º, 60º, 90º for altitude and 0º for azimuth (see definition in Figure 3-13) with two types of primary: 500keV electron and 8MeV proton. The result is shown in Figure 3-18 and Figure 3-19.

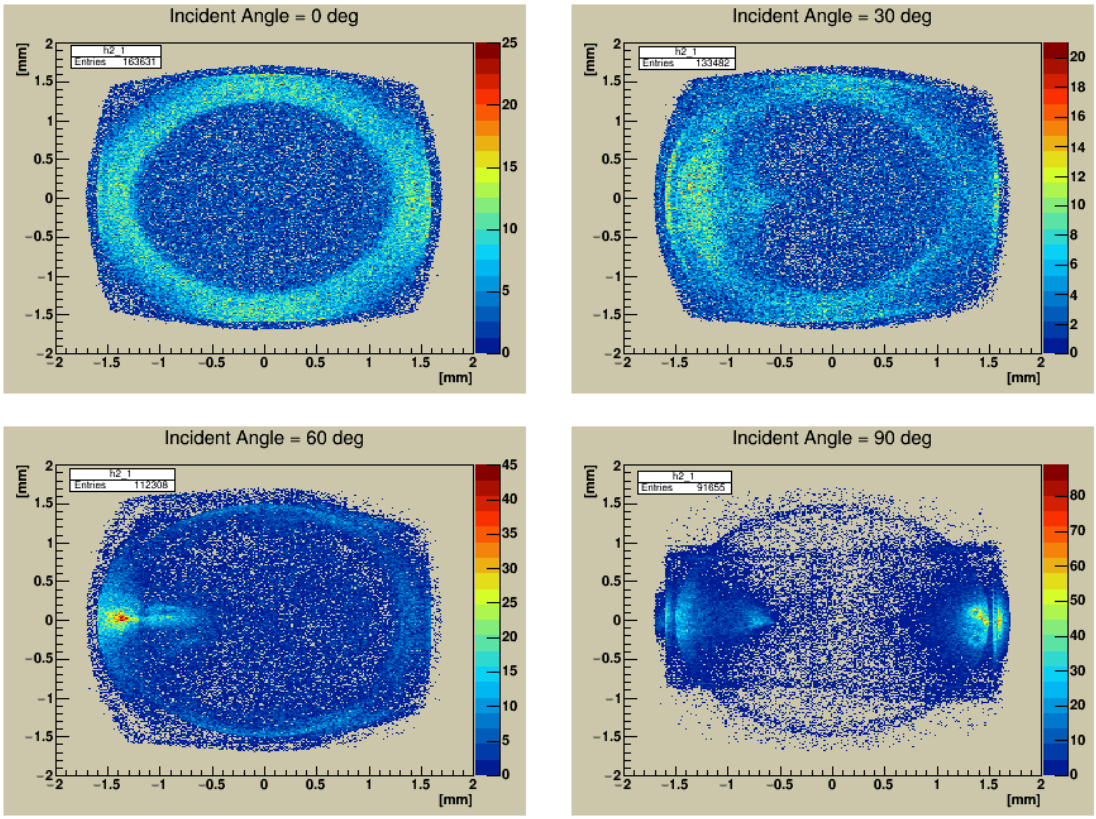


Figure 3-18 Secondary photon hit position on the SiPM sensor with a primary electron at 500keV.

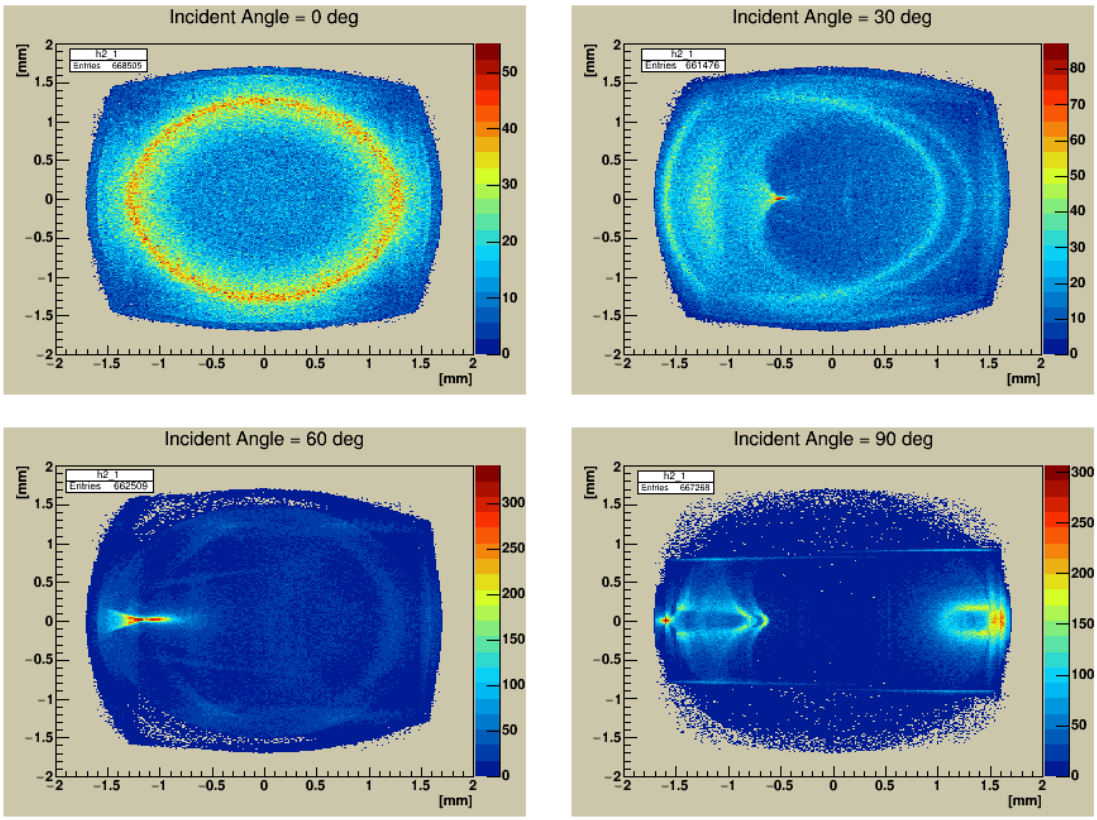


Figure 3-19 Secondary photon hit position on the SiPM sensor with a primary proton at 8MeV.

The scintillator detector geometry is a hemisphere with a radius of 1.5mm. The SiPM sensor is a square with side length the same as the hemisphere base circle diameter at 3mm. This ensures all the hemisphere base circle area is covered. This knowledge helps us to understand the photon hit position plot. The ring indicates the hemisphere base circle edge. The edge blurred square like shape indicates the SiPM sensor outline. All the plots are charted with heat map with red colour indicating high intensity.

When the primary hits from the zenith (θ = 0º) of the detector, large amount of photon hits at the edge of the hemisphere base circle. As primary direction shifts down (increasing θ), the hit position distribution shifts accordingly. These plots clearly show a link between the primary incident direction and the photon hit position. The pattern is similar for both electron and proton, with proton's photon hit centre more focused.

3.6.1.5 Mono energy response

The mono energy response shows the statistic fluctuation of the scintillation photon yield and the photon transmission loss in and across the scintillator. We launch 1MeV electron from a fixed direction (zenith) onto the detector module for 10000 times. The histogram of the total SiPM detected photons are shown in Figure 3-20.

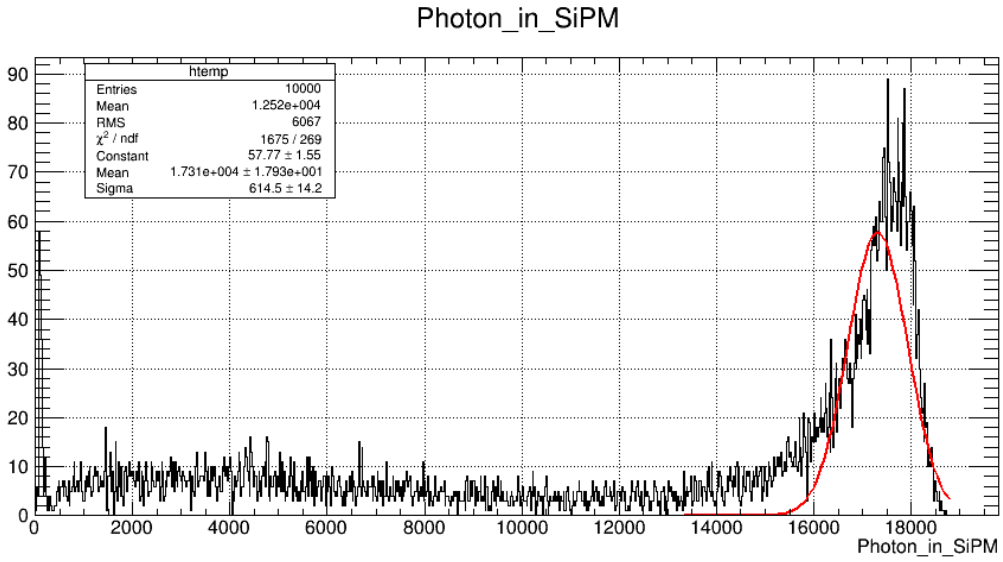


Figure 3-20 The histogram of total SiPM detected photon.

The energy deposited in BC408 scintillator and CsI(Tl) scintillator are shown in Figure 3-21.

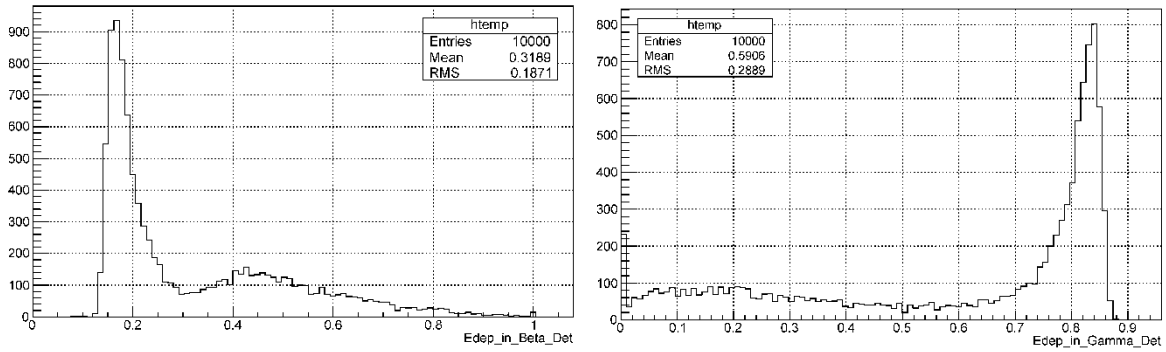


Figure 3-21 The energy deposit in each scintillator detector (left is BC408 and right is CsI(Tl)). The x-axis is MeV.

We can see the peak deposit energy in BC408 is about 0.16MeV and the peak deposit energy in CsI(Tl) is about 0.84MeV. The sum of them equals to the total primary energy of 1MeV. This demonstrates a good example of **ΔE/E** detector strategy. It might be possible to use this strategy to identify different particle with the same energy as the energy deposition share in the two scintillators will be different for different particle.

3.6.2 Energy Response Results

Once the Geant4 program was verified with good initial results, we start the full-scale simulation. This section presents the energy response simulation results. Table 3-9 lists the simulation data and setup.

Table 3-9 Simulation jobs summary for energy response test

Job ID	Particle	Beam On	Angle	Energy	Running time
2195083	e-	50,000	<0 0 1>	0.01M - 20M	23h22m
2226259	e-	50,000	<1 0 0>	0.01M - 20M	1d1h57m
2161912	Proton	5,000	<0 0 1>	0.5M – 20M	19h12m
2209211	Proton	5,000	<1 0 0>	0.5M - 20M	21h9m
2195080	Alpha	2000	<0 0 1>	10M – 80M	1d10h4m
2209212	Alpha	2,000	<1 0 0>	10M – 80M	1d13h24m
2195077	Gamma	500,000	<0 0 1>	0.001M – 1M	5h37m
2209213	Gamma	500,000	<1 0 0>	0.001M – 1M	9h7m

The energy unit is eV. The angle <0 0 1> means the position of the primary source is at coordinate <0 0 1> and its direction is <0 0 -1>. This equals to θ = 0º. Similarly, <1 0 0> means direction of <-1 0 0>. This equals to θ = 90º. Figure 3-22 illustrates the two angle directions.

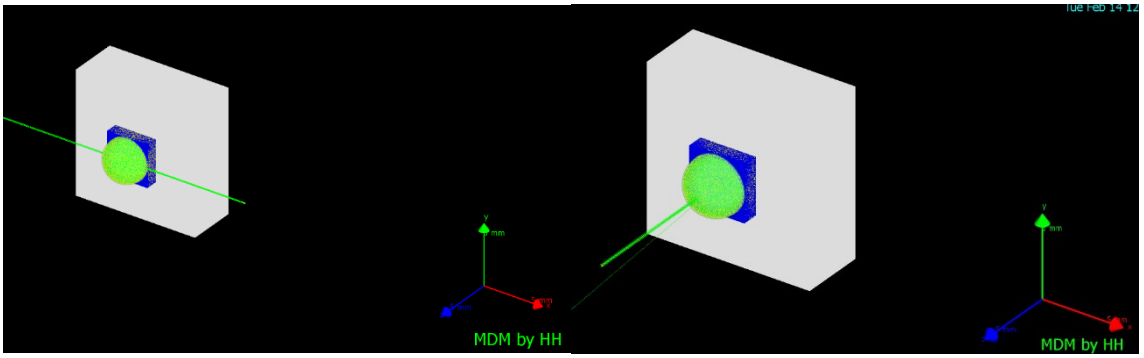


Figure 3-22 Primary injection angle <1 0 0 > (left) and <0 0 1> (right)

The running time is from the log of computation time. The length of simulation is largely driven by the primary energy and type. As the CsI scintillation photon yield is 54000 photons/MeV, it means that for each MeV of deposit energy, it will generate 54000 photons. Each photon requires tracking from birth to death (either absorbed by the scintillator or detected by the SiPM sensor). The tracking and interaction of these vast photons consumes most of the computation time. Alpha particle requires higher energy to penetrate the Teflon wrapping layer, so we have to set high energy range for alpha particle. This means the BeamOn times (=number of primary) have to be limited to cap the simulation time.

3.6.2.1 Proton Results

The proton simulation results are shown in Figure 3-23 and Figure 3-24. Figure 3-23 displays the total number of photons collected by the SiPM sensor against primary energy. The photons from both BC408 scintillator and CsI(Tl) scintillator are summed together.

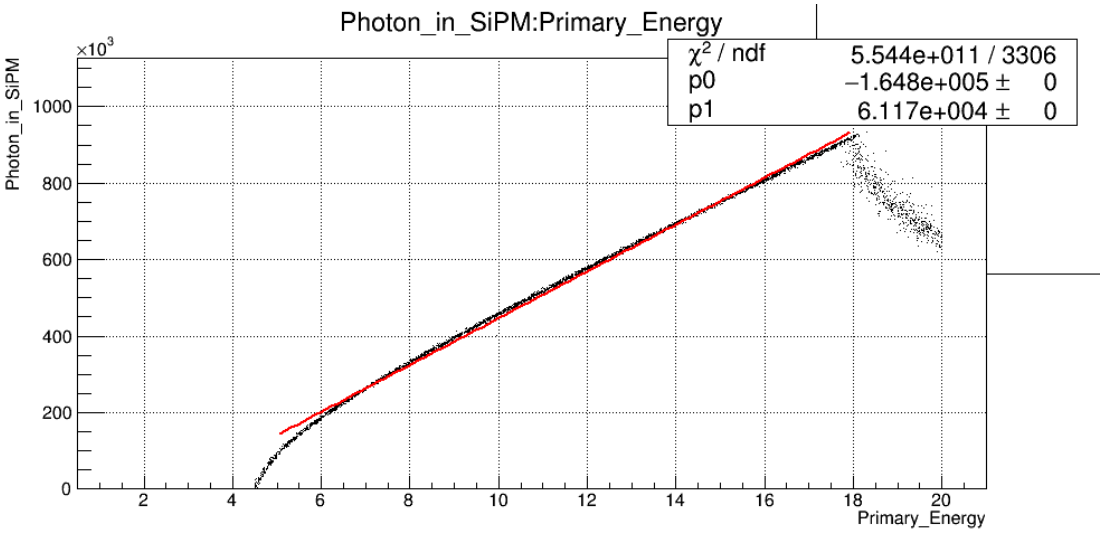


Figure 3-23 Proton energy response simulation result

With ROOT built-in tool, we fit a first order polynomial line (red) for the proton energy response. The statistics box in the top right corner shows that the fit parameter. P0 is the offset and P1 is the slope. The slope shall be close to the total photon yield of the two scintillators. The P1 reads 61170 while the sum of photon yield of BC408 and CsI is 500 + 54000 = 54500 photons/MeV.

We can see the onset of response is about 4.2MeV and response drops at about 18MeV. This means the effective detection range for our detector module for proton is 4.2MeV to 18MeV. However, the CsI(Tl) scintillation photon dominates y-scale due to its much higher photon yield. The information of BC-408 scintillation photon is not readable. Figure 3-24 splits the response to each scintillator which gives better view of the BC-408 response. From this closer view, we can see the effective range for the BC-408 is about 2MeV – 4.2MeV. And it is 4.2MeV to 18MeV for CsI(Tl). Beyond 18MeV, the deposit energy in CsI(Tl) starts dropping, which indicates proton above this energy starts leaving the CsI(Tl) scintillator. We note even the proton penetration range exceeds the scintillator depth, there is still a small amount of energy deposited in the scintillator. This result agrees with NIST data in Figure 3-10 well.

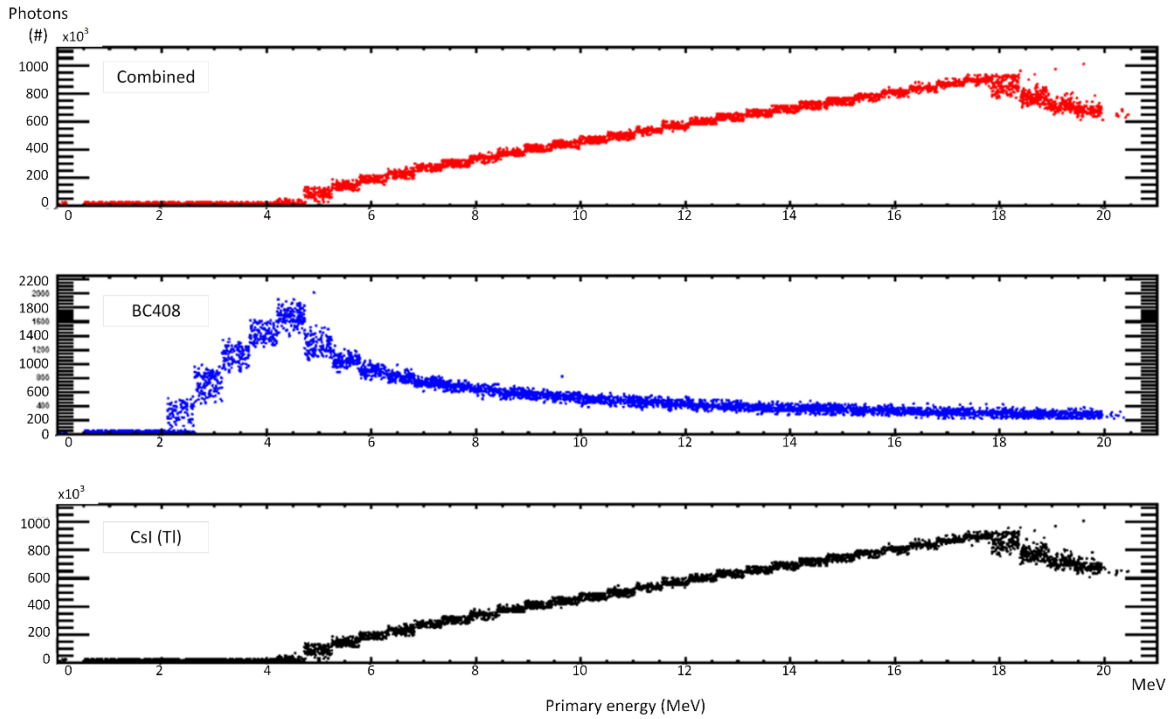


*Figure 3-24 Energy response to **proton**. Top plot is the combined response, middle plot is BC408 response, and bottom is CsI(Tl) response.*

The shape of the energy response is proportional to the shape of energy loss of charged particle in the matter. The curve matches the Landau distribution [123] well. The BC408 curve is very typical. The CsI(Tl) curve is not quite complete as the simulated energy range is not wide enough to show the long tail.

3.6.2.2 Alpha Results

The energy response to alpha is very similar to that of proton, with just different cut-off energy. Figure 3-25 and Figure 3-26 present the results.

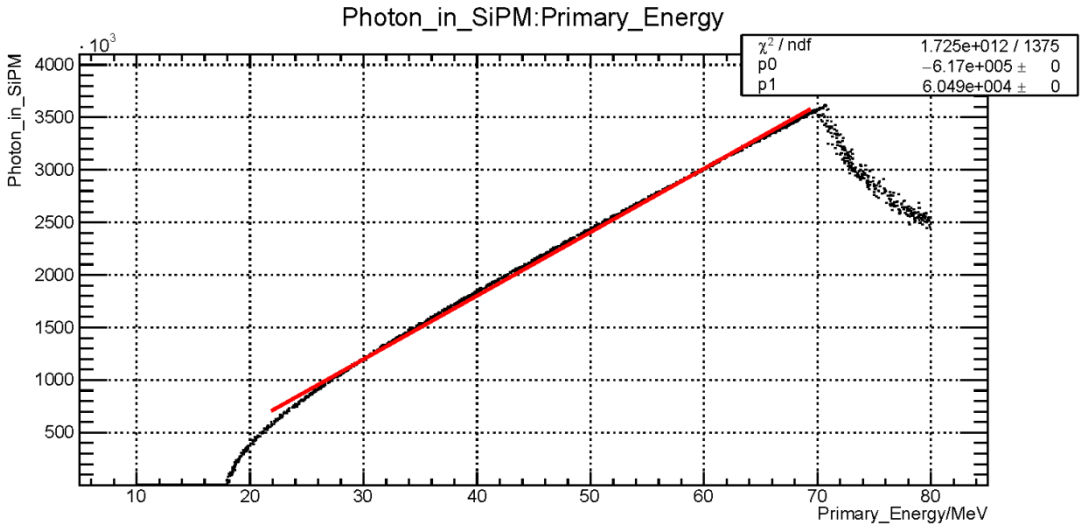


Figure 3-25 Alpha energy response result.

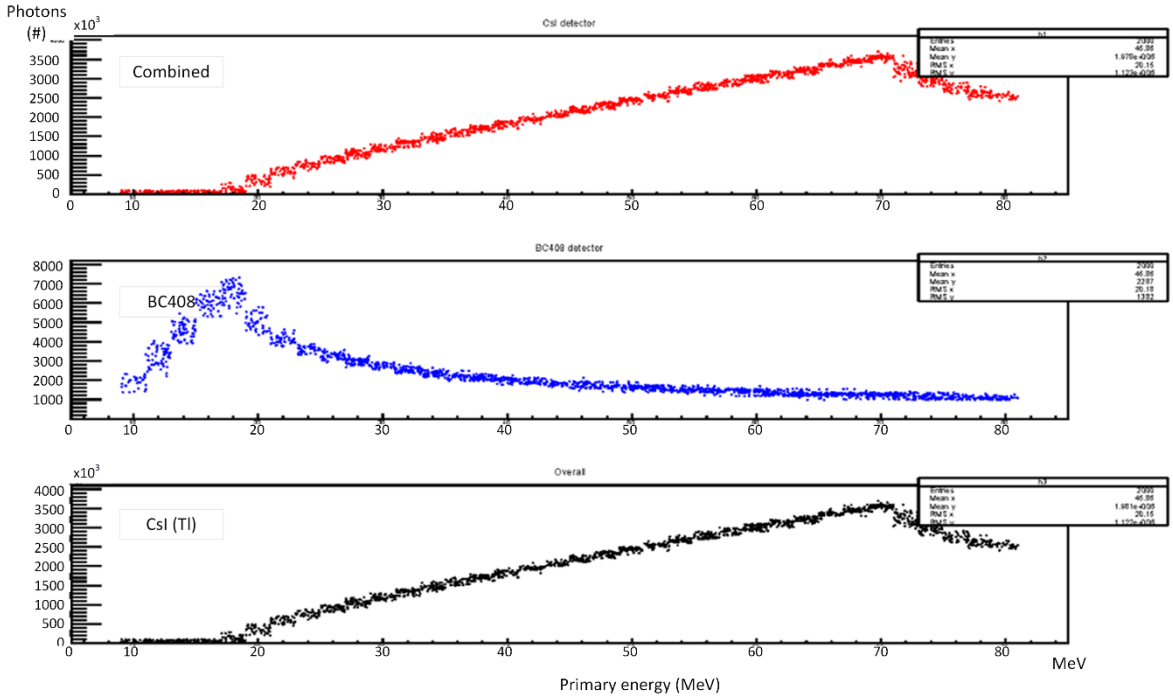


*Figure 3-26 Energy response to **alpha**. Top plot is the combined response, middle plot is BC408 response, and bottom is CsI(Tl) response.*

Similarly, the cut-off energy for the Teflon layer for alpha is about 10MeV, for the BC408 layer is about 18MeV, and for the CsI(Tl) layer is about 70MeV. This means effective energy range for BC408 is 10MeV to 18MeV and for CsI(Tl) is 18MeV to 70MeV. Beyond 70MeV, the alpha starts escaping the detector module. This result again matches the NIST data shown in Figure 3-11.

Both proton and alpha interaction with matter can be described by Bethe-Bloch formula with corrections at low energies. Within the effective range, the response shows very good linearity.

3.6.2.3 Electron Results

The energy response of electron shows different characteristics. Figure 3-27 and Figure 3-28 present the results. The number of scintillation photon in proportion to the deposit energy is widely spread from 0% to 100% of the full absorption. The red line in Figure 3-27 marks the boundary of full absorption, where all the incident energy is absorbed by the two scintillators.

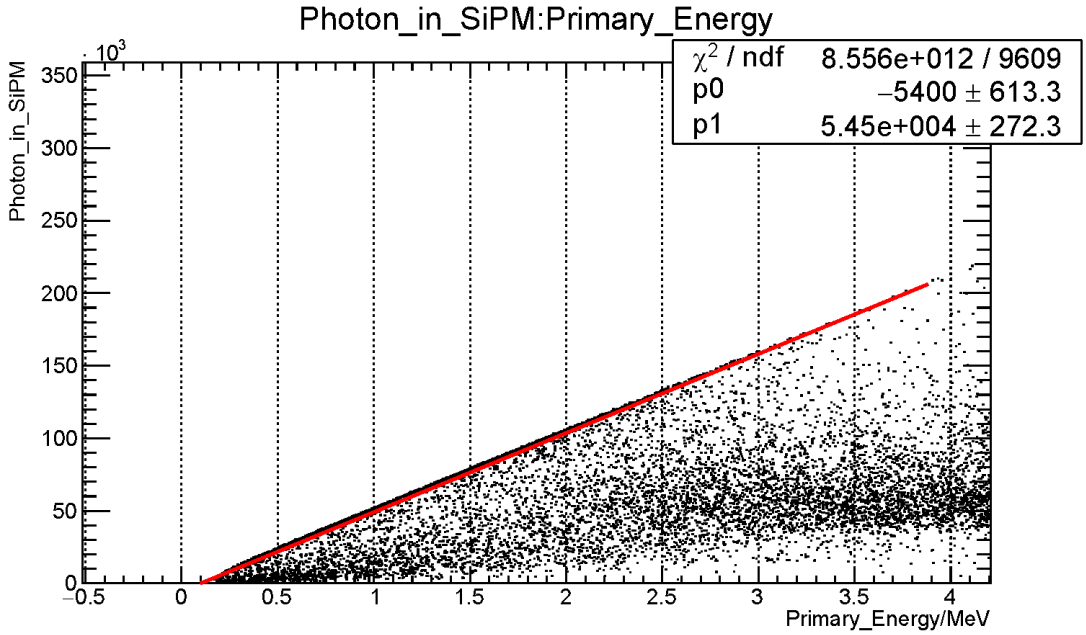


Figure 3-27 Electron energy response result.

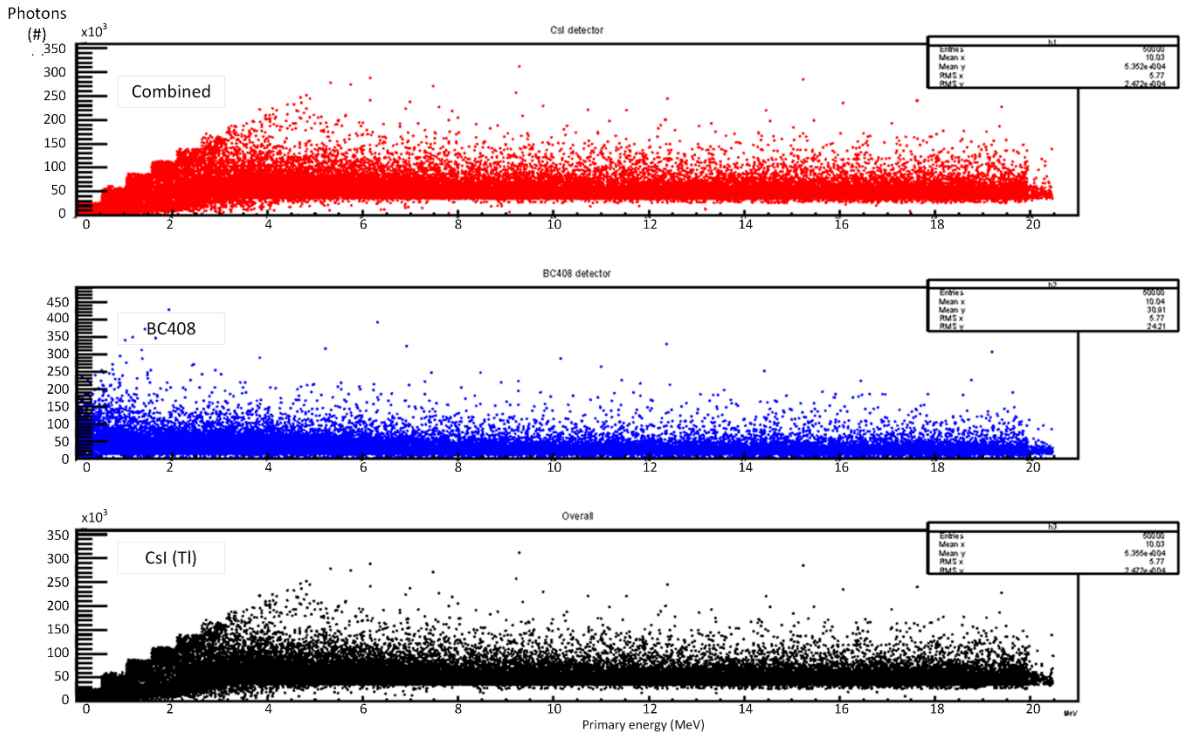


Figure 3-28 Energy response to electron. Top plot is the combined response, middle plot is BC408 response, and bottom is CsI(Tl) response.

Figure 3-29 zooms into the range of 0 to 1MeV for the same data as Figure 3-27. The response scattering is very severed. Here, scattering means for the same primary electron energy, the amount of energy being absorbed by the scintillator differs greatly from event to event.

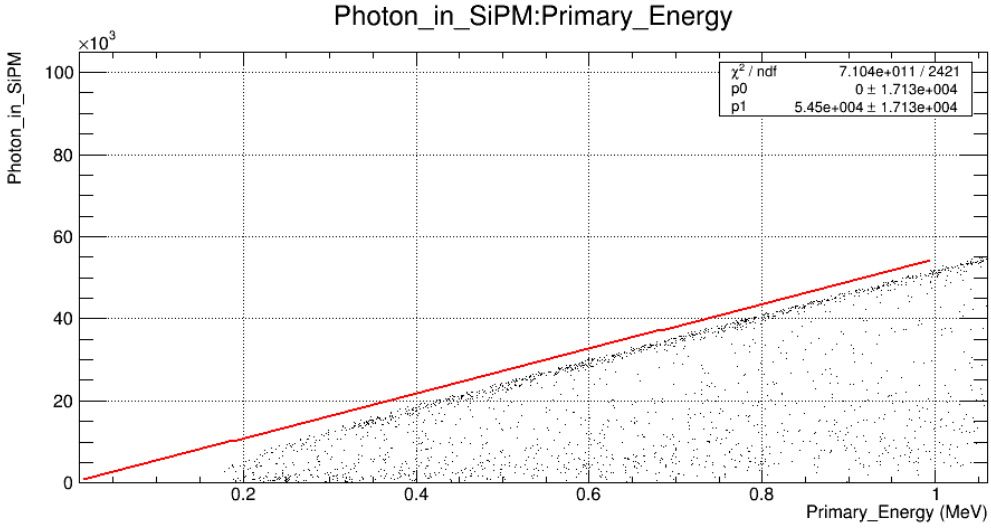


Figure 3-29 Zoom in of electron energy response into the valid detection range.

This scattering creates huge difficulty in interpreting the incident primary energy by counting the scintillation photons. For example, a 1MeV e- deposits 50% of its energy

can't be distinguished from a 500keV e- deposits 100% of its energy, as the number of the detected photons is the same.

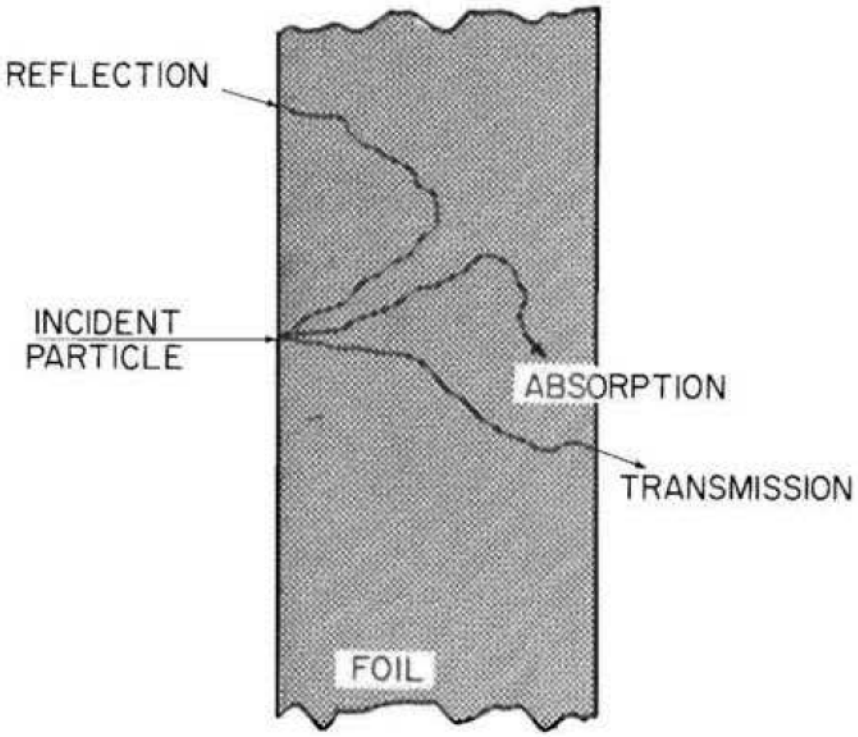


Figure 3-30 Typical electron trajectories in a foil [118]

The underlying physics explanation of the scattering is very complex. One of the explanations is that due to the light mass of electron, Coulomb interaction is the dominant mechanism for electron trajectory. As shown in Figure 3-30, at the boundary, electron can go on three types of trajectories: reflection, absorption and transmission. The resulting trajectory is influenced by the incident energy, direction and random fluctuation. Along the trajectory, electron gradually loss energy through ten- or even hundred-thousanths of Coulomb interactions/collisions. In simulation, this is modelled as continuous slowing down approximation (CSDA). The length of the trajectory is proportional to the energy loss or absorption (the energy loss of primary electron is the energy absorption of scintillator). The observed scattering of energy deposit is the reflection of fluctuation of trajectory length in the scintillator.

3.6.2.4 Gamma Results

The energy response of gamma is similar to that of electron. Figure 3-31 presents the response from 0 to 1MeV. Again, the red line marks the boundary of full absorption, which is defined by the sum of the theoretical photon yield of two scintillators. The gamma interaction with matter is dominant by photoelectric effect at low energy,

Compton scattering at middle energy and pair production at high energy. The test energy range in our simulation falls into the range that produces photoelectric effect and Compton scattering. Both processes will first transfer the photon energy to one or multiple electron. These “hot” electron will propagate further in a trajectory similar to the electron traverses in the matter. This might explain why the energy response of gamma presents similar scattering effect.

However, there is one difference in gamma response, that is between the full absorption boundary and the scattering region, there is a gap with almost no data point. Within the current knowledge, we are not clear what the underlying cause is.

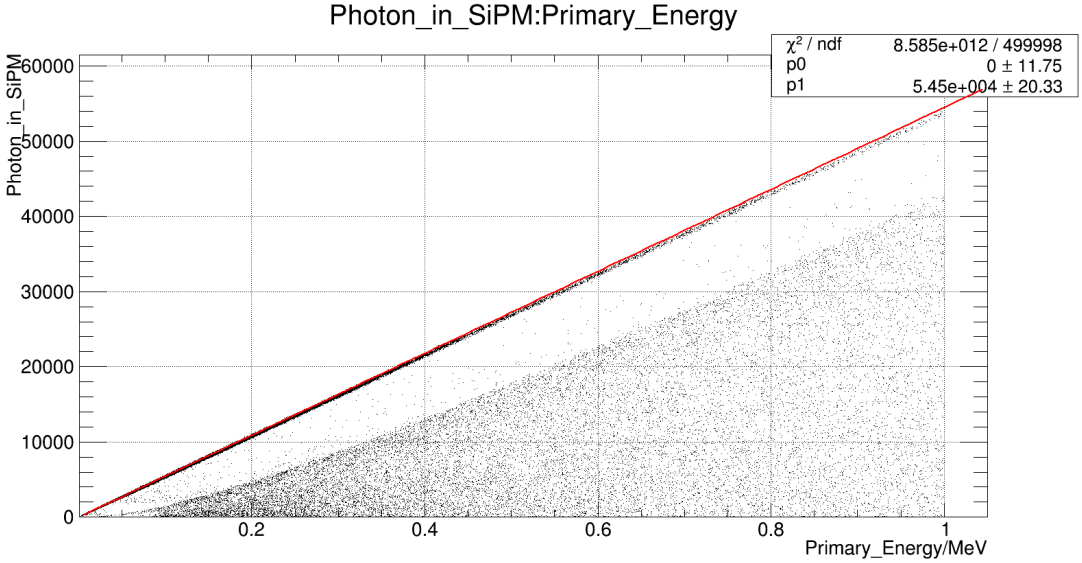


Figure 3-31 Gamma energy response result.

Figure 3-32 zooms in the energy range from 0 to 0.12MeV.

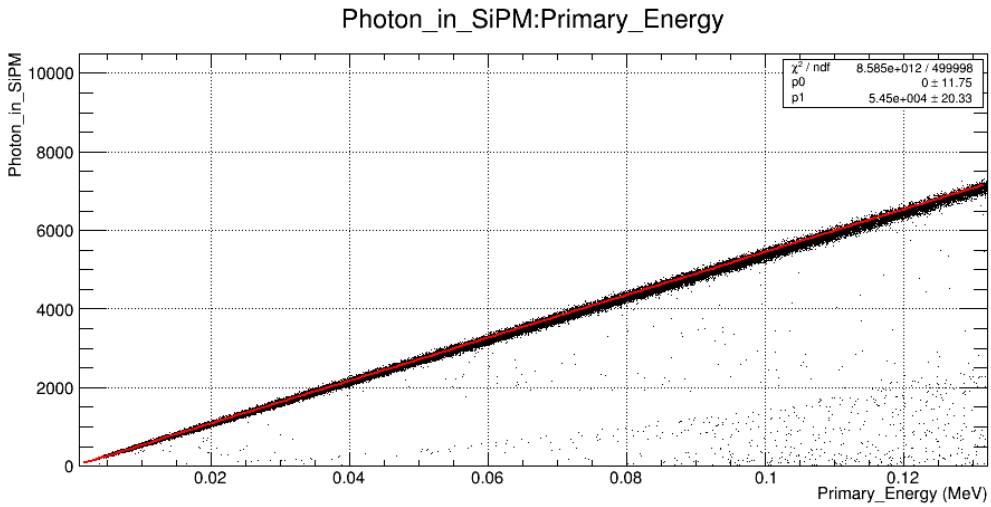


Figure 3-32 Zoom in of electron energy response into the valid detection range.

3.6.2.5 Short Summary and more results

We have the findings:

1) The proton and alpha energy response are very similar.
2) The detector acceptance range is wider for primary incident direction <1 0 0> (side injection) than <0 0 1> (top injection).
3) The electron and gamma energy response behave similarly but gamma response has an unexplained gap.

Figure 3-33 and Figure 3-34 present the results in group for easy comparison. Note the energy range (x-scale) is different.

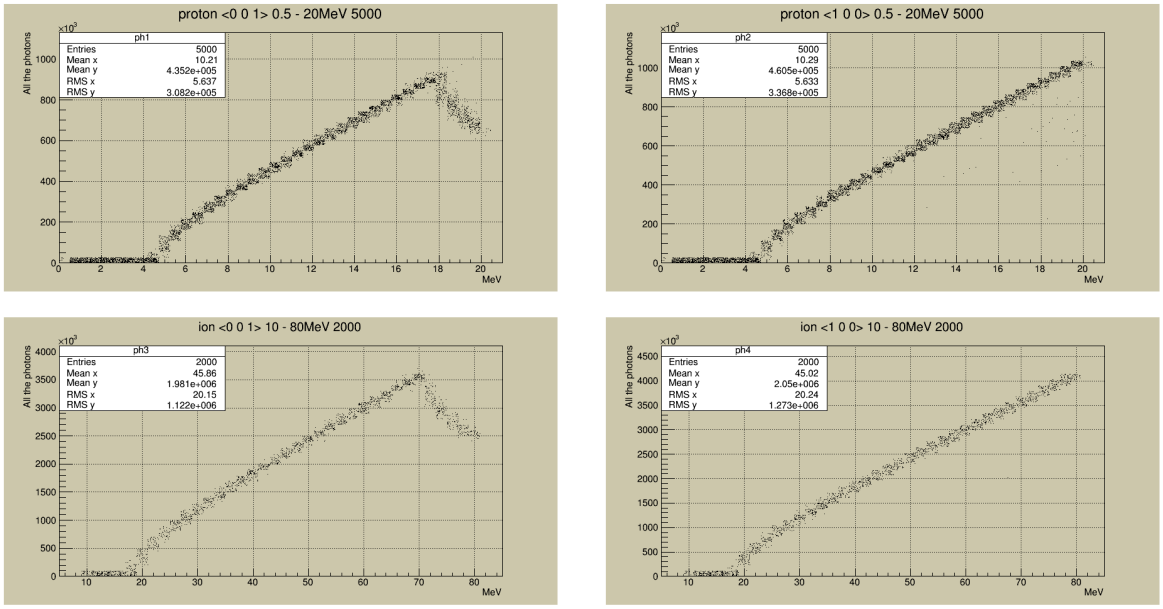


Figure 3-33 Comparison of proton and alpha (ion) energy response in two injection directions.

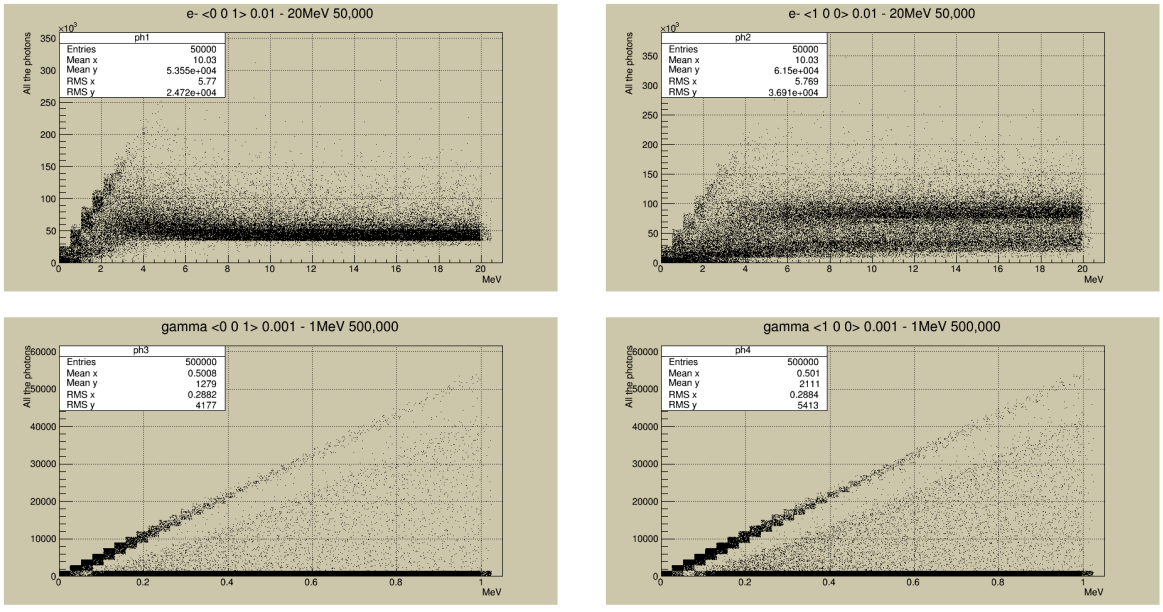


Figure 3-34 Comparison of electron and gamma energy response in two injection directions.

3.6.3 Angular Response Results

The second test is the angular response. The setup is explained in section 3.5.2. Table 3-10 lists all the test runs. A few mono-energies were tried to see the impact of energy on the angular response. High energy proton and alpha test are limited to small BeamOn times to save computation time.

Table 3-10 Job summary for angle response test

Job ID	Particle	Beam On	An-gles	Energy	Running time
2471225	Electron	1000/ang	648	100k	21m
2307579	Electron	1000/ang	648	250k	7h27m
2481819	Electron	1000/ang	648	500k	2 d 23 h 12m
2472177	Electron	1000/ang	648	750k	1 d 17 h 39 m
2471227	Electron	1000/ang	648	1M	Time out (>3d)
2471285	Proton	75/ang	324	2M	8 min
2307582	Proton	75/ang	324	5M	52m
2472180	Proton	75/ang	324	7.5M	1 d 9 h 40 m
2226257	Proton	75/ang	324	10M	2d10h39m
2308333	Proton	75/ang	324	15M	11h5m
2226258	Gamma	1000/ang	648	100k	8h45m
2308336	Alpha	75/ang	324	20M	2d11h49m

3.6.3.1 Simulation coordination system

To understand the result plots we present next, it is helpful to review the simulation setup coordination system as shown in Figure 3-35. The zenith direction is Z-axis, and X-axis, Y-axis completes the right-hand rule. The results are presented in a 2D cross section view. For example, XOZ view means view from Y-axis, XOY view means view from Z-axis and YOZ view means view from X-axis

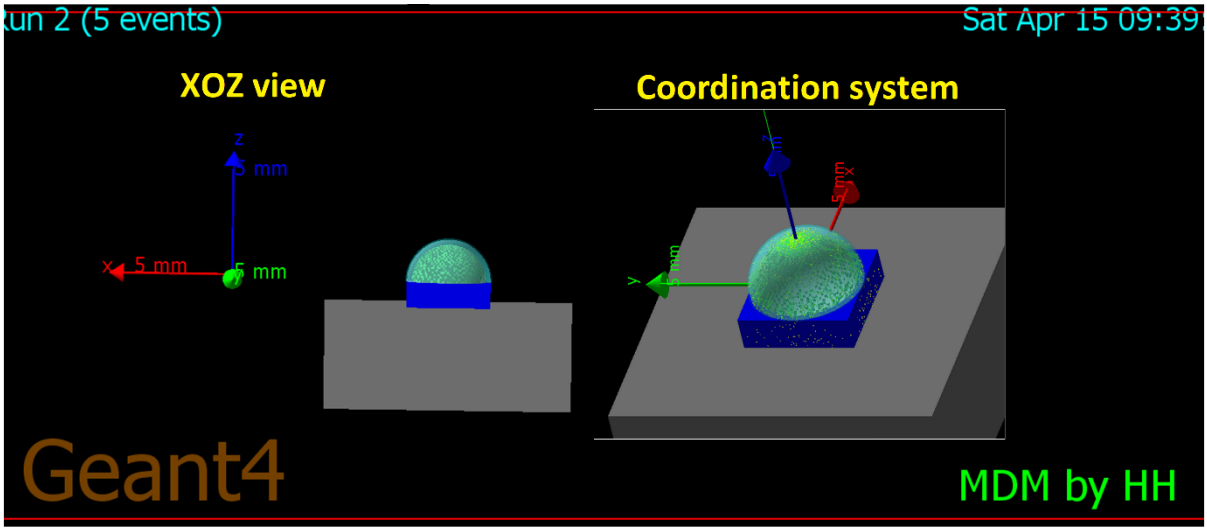


Figure 3-35 Simulation coordination system and an example XOZ view

To be noted, the angular test results are presented in a 2D projection of the 3D circular plane source sphere not the actual detector sphere. Figure 3-36 shows the projection of the 3D source sphere on three different axis view.

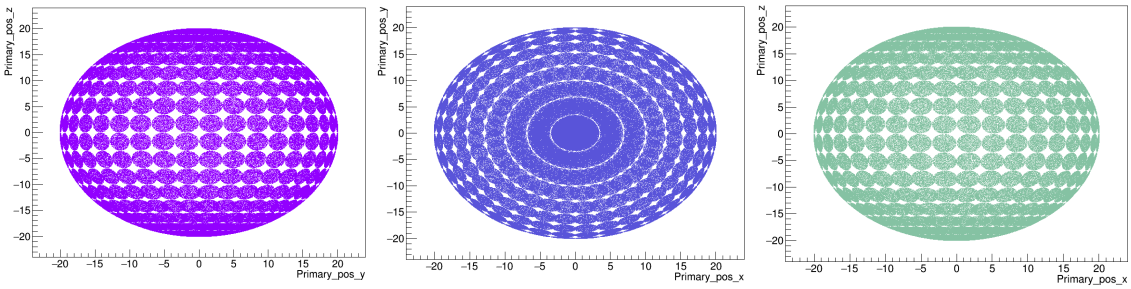


Figure 3-36 The 2D projection of the 3D source in YOZ, XOY and XOZ view (left to right)

As the hemisphere is symmetric to Z-axis, XOZ and YOZ view are identical and so we only present result from one of the two views.

3.6.3.2 Electron Results

Angle response test with electron is shown in Figure 3-37 (XOY view) and Figure 3-38 (XOZ view). Four mono energies are used: 100keV, 250keV, 500keV and 750keV. The rainbow scale represents the scintillation photon intensity.

The electron test has covered the entire 4π sphere. In XOZ view, we can clearly see the effect of the 2π FoV of our detector as the bottom half of the sphere has almost no response (cold blue). In the XOY view, it shows stronger response in the centre and weak response in the edge. This is because the effective interaction length between the source and the detector is different for the primary at different position of

the source plane as illustrated in Figure 3-39. The longer interaction length will result in better energy absorption and more photon yield in general. The impact changes as the source plane angle changes. In the XOZ view, some of the source circles show splits of responsive and non-responsive sections (red and blue sections). This might be because the source plane is slightly bigger than the cross-section of the detector module, which results in the primary in part of the source circle missing the detector.

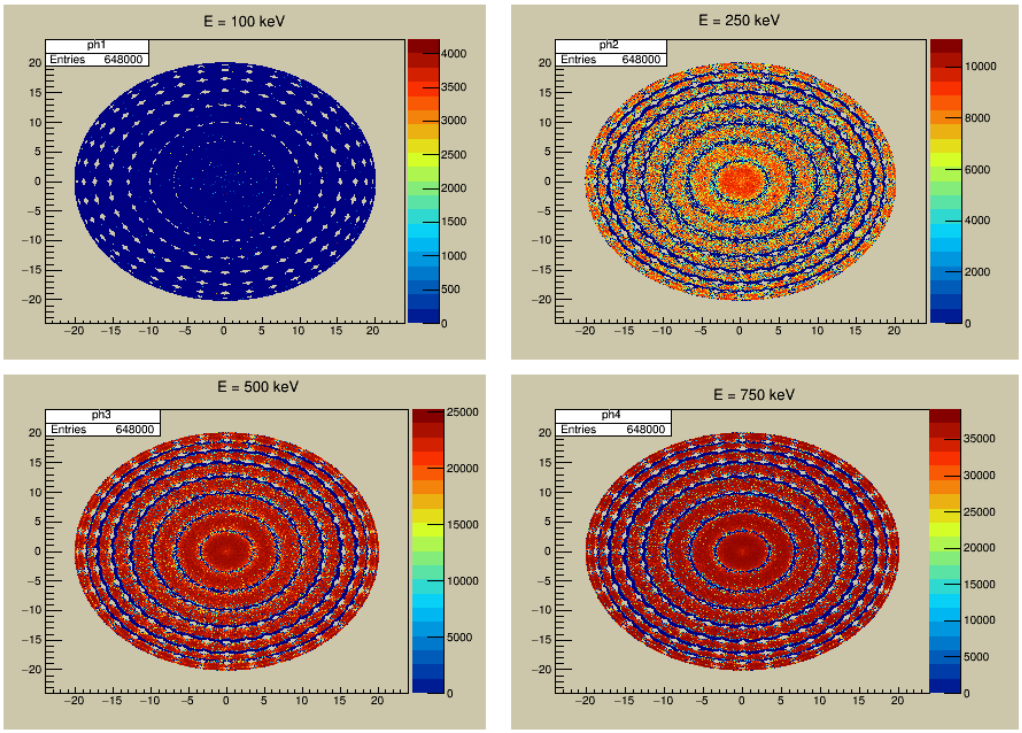


Figure 3-37 Electron angle test results in XOY view. Four mono energy is tested.

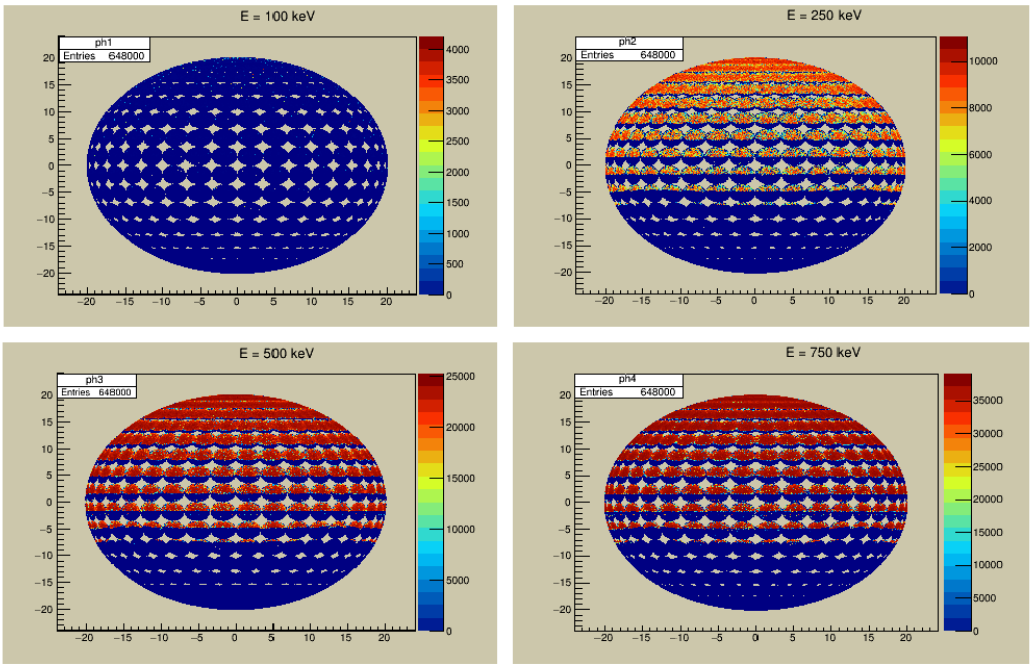


Figure 3-38 Electron angle test results in XOZ view. Four mono energy is tested.

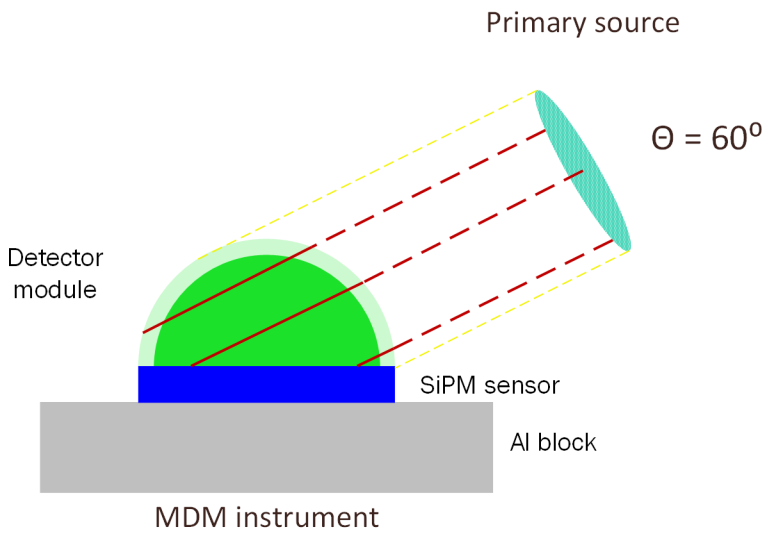


Figure 3-39 The effective interaction length (dark red solid line) between detector module and the primary at different position of the source plane

Four mono energies were tested. And we can see the energy difference has no impact on the angle response pattern.

3.6.3.3 Proton Results

For proton source, again four mono energies are chosen but with energy tailed for proton response. The results are shown in Figure 3-40 and

Figure 3-41 in the same XOY and XOZ view. Due to the computation limit, only half sphere (2π) is examined and 75 primaries in each block. As the primary number is so low, the circle shape of the source plane is not very clear. But the pattern is similar to the electron case and the explanation is the same.

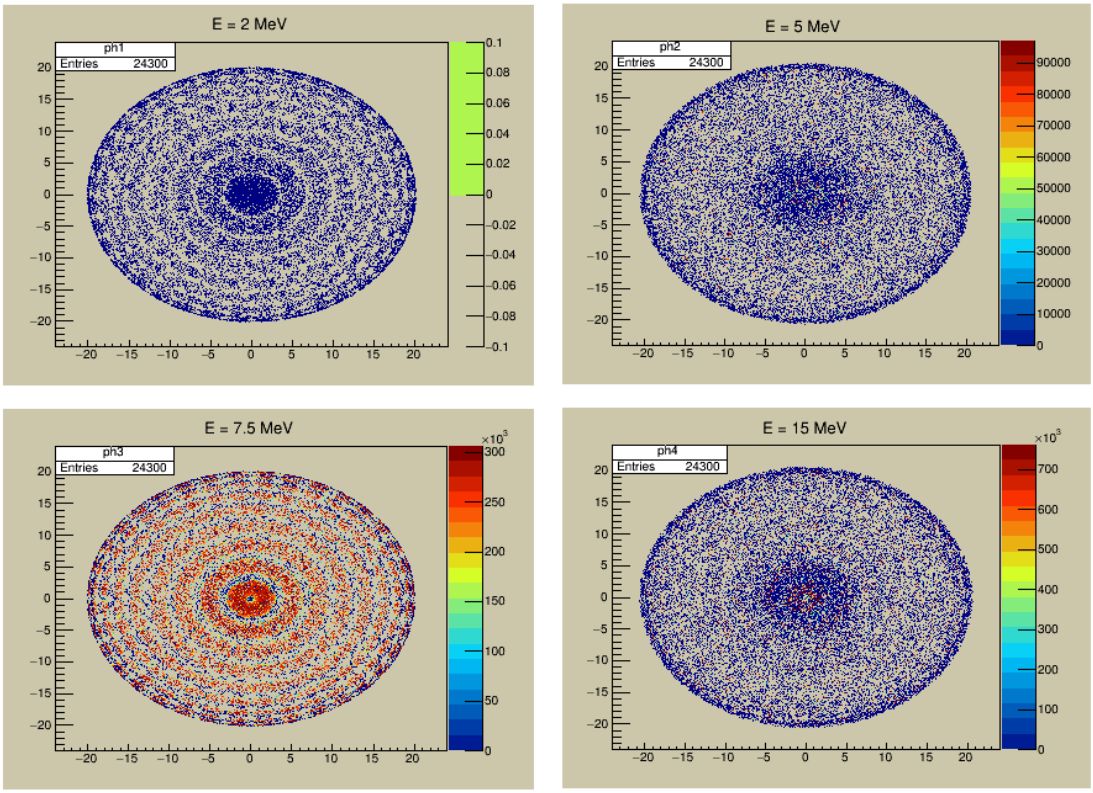


Figure 3-40 Proton angle test results in XOY view. Four mono energy is tested.

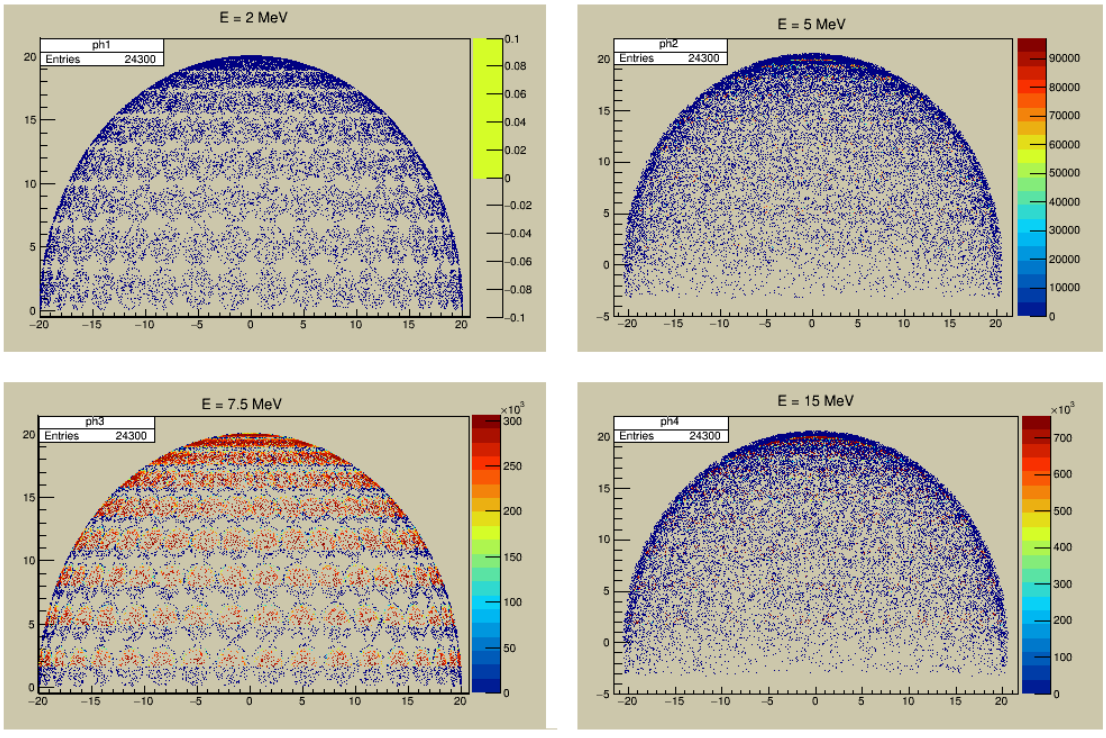


Figure 3-41 Proton angle test results in XOZ view. Four mono energy is tested.

3.6.3.4 Gamma Results

For Gamma, only one mono energy is tested in a 4π sphere. Results are shown in Figure 3-42 and Figure 3-43.

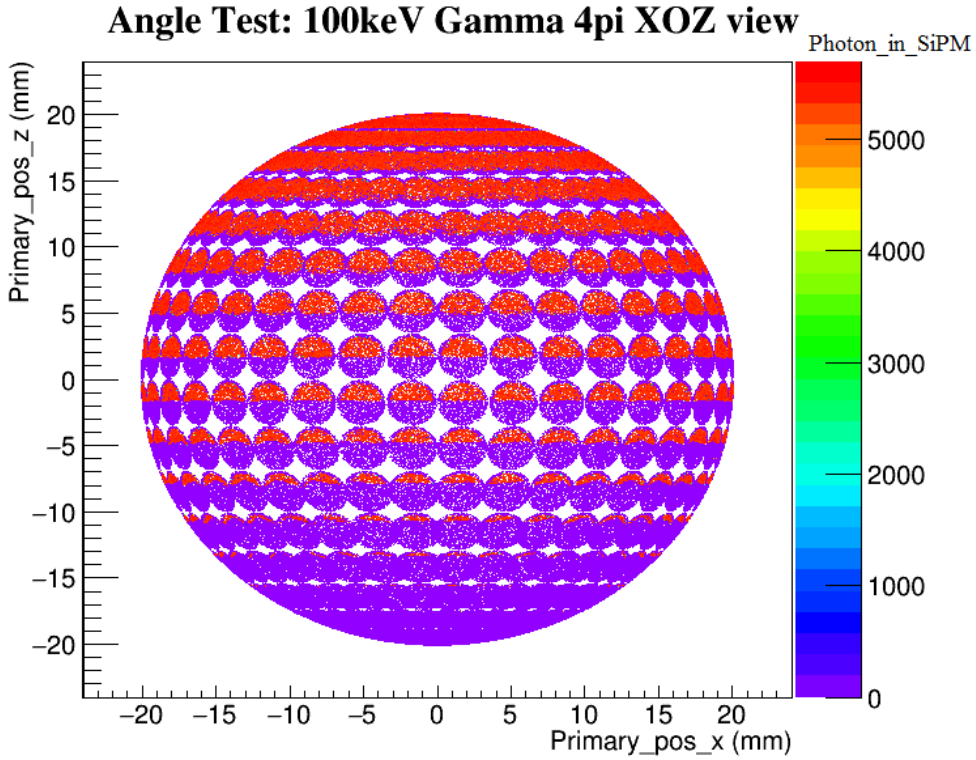


Figure 3-42 Gamma angle test results in XOZ view.

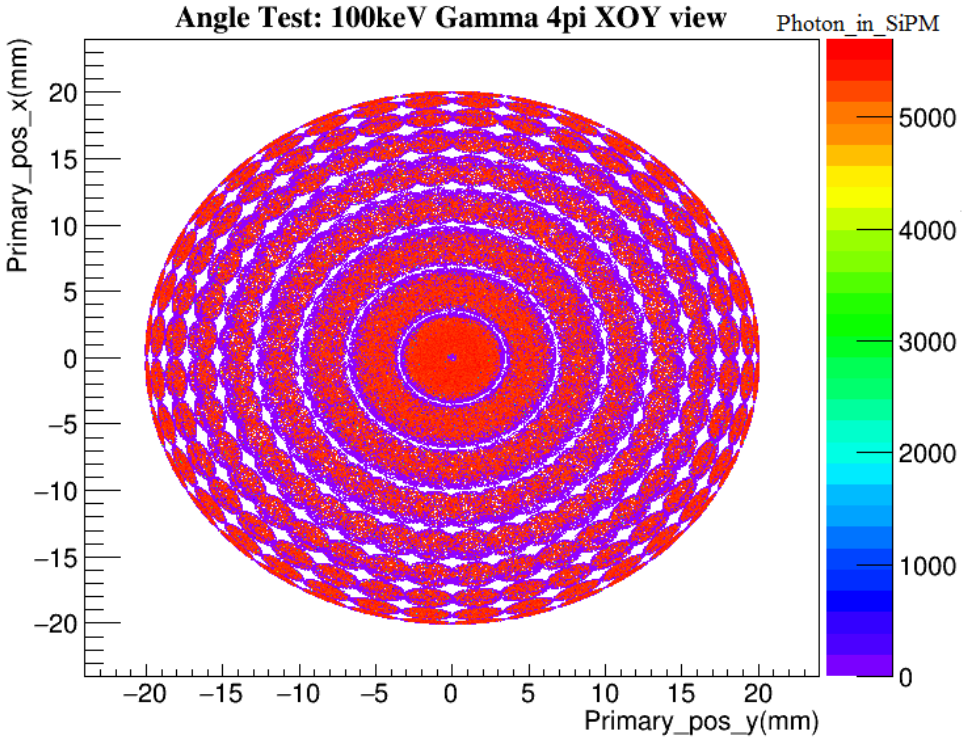


Figure 3-43 Gamma angle test results in XOY view

3.6.3.5 Alpha Results

For alpha, one energy and 2π are covered. Only 75 primaries are launched in each block due to computation limit. Results are presented in Figure 3-44 and Figure 3-45.

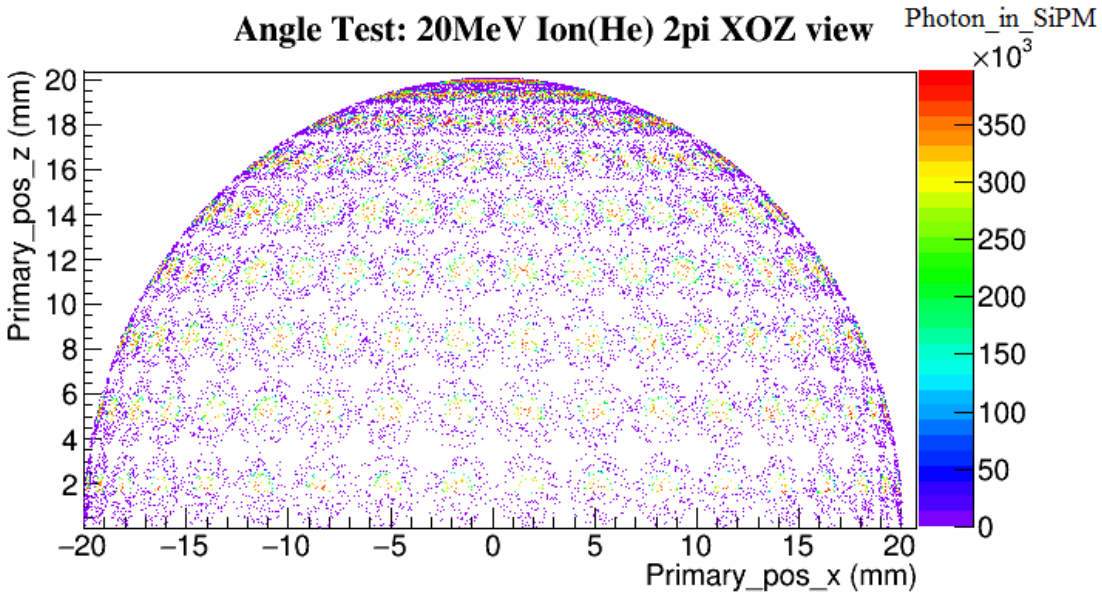


Figure 3-44 Alpha angle test results in XOZ view.

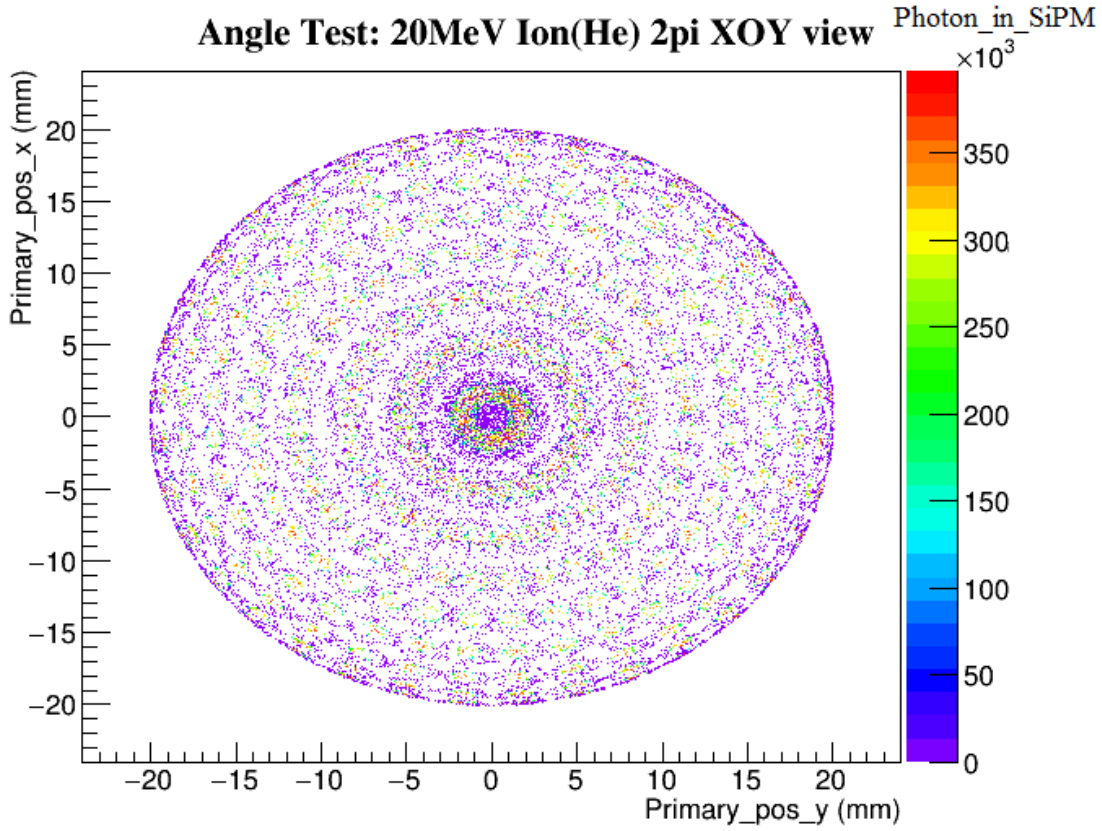


Figure 3-45 Alpha angle test results in XOY view.

3.7 Conclusion and future work

Energy Response

The simulation result of the detector module energy response matches the theoretical data from NIST database (Figure 3-9,Figure 3-10,Figure 3-11 and Figure 3-12) very well. The energy threshold results are summarized in Table 3-11. Between the min and max is the effective energy range that the corresponding scintillator can detect. The data is manually read from the plot, so we use ~ to represent estimation.

Table 3-11 Simulation results for energy response

Particle	Instrument Responsible Range				Unit
	BC408		CsI (Tl)		
	Min	Max	Min	Max	
Electron	~0.1	~0.2	~0.2	~1	MeV
Proton	~2	~4.2	~4.2	~18	MeV
Alpha	~10	~18	~18	~70	MeV
Gamma	neglect		~4	~100	KeV

Within the effective range, the energy response of proton and alpha is highly linear while the response of electron and gamma is highly scattering. This should be driven by the physics nature of the interaction between the species and matter.

Good linearity of the detector response is the essence of good instrument energy resolution. Therefore, our current detector module design can provide a good energy resolution of proton and alpha particles, but we will have difficulty to interpret the electron and gamma energy by just counting the number of scintillation photons. Additional method or different approach are required to detect the energy of single electron or gamma.

Particle with energy deposited in both scintillators demonstrates the working of ΔE/E detection method. This provides an approach to distinguish different particles as the ratio of the energy deposited in the two scintillators for a given primary energy is different for different particle specie. So, it is possible to identify the particle specie by interpreting this ratio with the total absorbed energy.

Angular Response

As expected, our detector module shows good symmetric angular response. Different particle energy, different particle species don't make a difference to this symmetric feature. The photon yield of the detector module is affected by the incident direction as the interaction length in the line of sight of the incident direction will be different for different incident direction.

The scintillation photon hit position distribution with different incident direction shows very interesting results. There is clear link between the hit position distribution and the primary incident direction. If the hit position distribution knowledge is required, it is possible to predict the incident direction.

Future work

Our simulation work has thoroughly covered the detector module energy response and angular response. Part of the particle identification work is also covered with result reported in the last chapter.

Our work has simplified the SiPM sensor response. We have assumed 100% SiPM detection efficiency. To have better simulation of SiPM, more advanced coding to represent the feature of SiPM is required. There are third party packages freely available. For example, G4SiPM[124] provides a good simulation function sets for SiPM. GODDeSS (**G**eant4 **O**bjects for **D**etailed **De**tectors with **S**cintillators and **S**iPMs)[125] is another package that covers both SiPM and scintillator. This thesis has tried out these packages briefly but without much success. Some efforts are required to integrate these packages into the existing Geant4 program.

The current work only uses simple fixed energy primary source. To simulate the real behaviour of a radioactive source, which will be used in the real system end to end test, it is desirable to introduce the radioactive source into the simulation. Geant4 provides such capability. Hence, we suggest that the radioactive source option to be included in the next simulation program.

Chapter 4 Electronics

4.1 Introduction

The key research and development of electronics for this PhD work focuses on the SiPM readout circuit. As illustrated in a typical space instrument electronics architecture (Figure 4-1), the sensor readout circuit is often referred to as front-end electronics (FEE). It usually locates next to the sensor/detector in order to minimize the electrical noise pick up. The FEE is tasked to amplifier the micro electric signal generated by the sensor, to condition the signal (including signal shaping and filtering) and to convert the signal from analogue to digital (usually via an A-D convertor). This analogue dominant circuit is critical to the instrument performance particularly in the signal to noise measure because the instrument detection sensitivity is largely determined by the noise floor of the readout circuit.

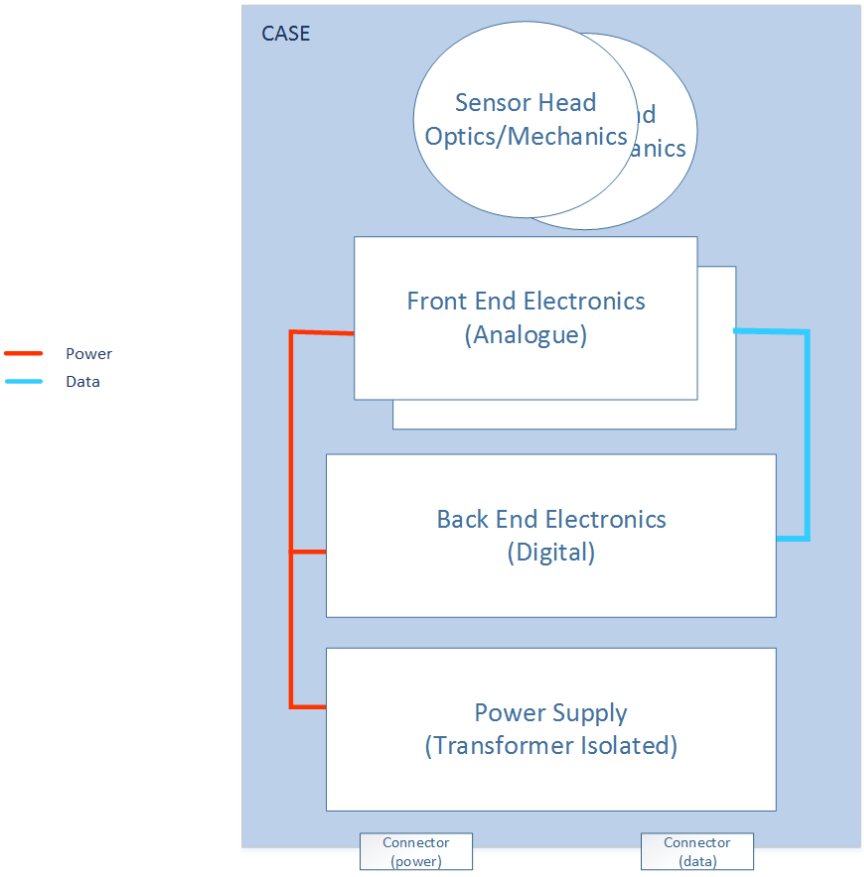


Figure 4-1 The architecture of a typical instrument

Section 2 first introduces the SiPM electric model. The work starts from Corsi's classic model [126] and moves onto Marano's improved model [127]. The circuit published in Marano's paper doesn't run very well in LTSpice (a free electric circuit simulation tool), so some practical adjustments are made to it and then the revised model is used in our project.

Section 3 discusses the methods of acquiring the electric parameters that is used in the SiPM electric model. Not all the parameters are available in the manufacturer's datasheet, actually very few are directly available. The remaining unknowns have to be acquired by either direct measurements or indirect calculation. We run through all the measurements and calculation on one real example SiPM model to examine the theory. The results are given at the end of the section.

Section 4 studies about the preamplifier circuit theory. There are two basic preamplifier circuit configurations. By studying their small signal and control stability, the current preamplifier mode proves to be a better choice for SiPM sensor. The circuit is simple in its format, but if not, enough attention is paid to the detail, the performance can be far away from satisfaction. Therefore, much is spent to study and discuss about the noise analysis of the circuit as a major part of this section. The time-domain signal analysis is also studied. And it reveals the analytic solution for the preamplifier output waveform.

Section 5, after all the theoretical preparation, we present the final hardware development work and result. Two version of the PCBs have been developed. The main difference is that version 1 is manufactured on the standard FR4 PCB material and the version 2 is on a RF PCB material with much lower dielectric constant.

Section 6 will provide a brief survey of the SiPM devices on the market and a short discussion about our choice. The radiation hardness of SiPM will also be briefly mentioned

4.2 SiPM SPICE Model

In electric circuit simulation, an electronic device is represented by a SPICE model[128] to describe its electric performance. The circuit simulation software can

read such a model and simulate the circuit behaviour. To assist the preamplifier circuit design, it is very useful to have a SiPM SPICE model to represent the circuit behaviour of SiPM. Hence, we studied such a model as the first step for the SiPM preamplifier circuit study and development.

4.2.1 Introduction to SiPM structure and its parameters

Before introducing the SPICE model, it is useful to review the SiPM structure and its electric parameters that play a role in the model. SiPM is basically a matrix of single photon avalanche diodes (SPADs) with all the anodes and cathodes connected in parallel. Each SPAD can operate independently but there is only one output for the entire SiPM device. Therefore, if multiple SPADs are triggered at almost the same time, their outputs will stack up to form one larger signal. The charge of this total signal shall equal to the sum of the charge of all the fired SPADs. When the triggering time is close enough and the total number of the triggered SPADs are relatively small, the peak of the total signal equals to the number of the triggered SPAD multiples the peak of a single SPAD signal. Hence, the operation of SiPM is straightforward while the complexity and the key parameters related to the SiPM performance mostly lay within a single SPAD.

A SPAD is basically a PN junction designed to be able to operate in Geiger-mode [129][130]. Geiger-mode means bias the PN junction reversely above its breakdown voltage. This creates an extremely high electric field in the junction region that can sustain a continuous electron-hole pair multiplication process, which is called avalanche. When an incident photon or energetic particle initializes the process, the avalanche is self-sustained until the over voltage condition is not met. This feature is utilized to stop the avalanche process at a certain point to avoid the PN junction being burned due to excessive high current. The simplest method to self-adjust the overvoltage is a high value resistor in series to the PN junction. As the current increases, the voltage drop across the resistor increases, which results in lower bias voltage across the junction given the external bias voltage is constant. This drives bias voltage below the breakdown threshold and terminates the avalanche process. This simple negative feedback mechanism is easy to implement in the SiPM construction, so it is the most common method utilized. The drawback is this resistor introduces a long recovery time with the PN junction capacitance as it limits the current that is to

restore the bias voltage across the PN junction. To address this issue, there is actively controlled termination method, for example, using a MOSFET to replace the termination resistor. This MOSFET is actively controlled to turn off to stop the avalanche and to turn on to restore the bias voltage. However, this active method requires complex implementation including independent control circuit. Both passive and active methods to terminate the avalanche is called “quenching”.

The operation principle of a SPAD can be illustrated in a simplified electrical equivalent circuit as Figure 4-2 (left) shows. The switch **S** represents the on and off status of avalanche. When **S** closes, avalanche process starts and the PN junction capacitance (**Cj**) starts discharging to the junction resistance (**Rs**) during conducting. The current flow is represented with the red arrow. When **S** opens, avalanche process ends and the $\mathbf{V_{BIAS}}$ source starts recharging **Cj** through the quenching resistor ($\mathbf{R_Q}$). The current flow is represented with the blue arrow. The current pulse is also presented in Figure 4-2 (right). The fast-rising edge represents the rapid current surging during the avalanche. The peak ($\mathbf{i_{max}}$) represents the termination point of the avalanche. The slow falling edge represents the long recovery time. **Rs** and **Cj** form the time constant for the rising edge while $\mathbf{R_Q}$ and **Cj** form the time constant for the falling edge. The peak is determined by the over voltage ($\mathbf{V_{BIAS}}$-$\mathbf{V_{BR}}$) divided by the sum of $\mathbf{R_Q}$ and **Rs**.

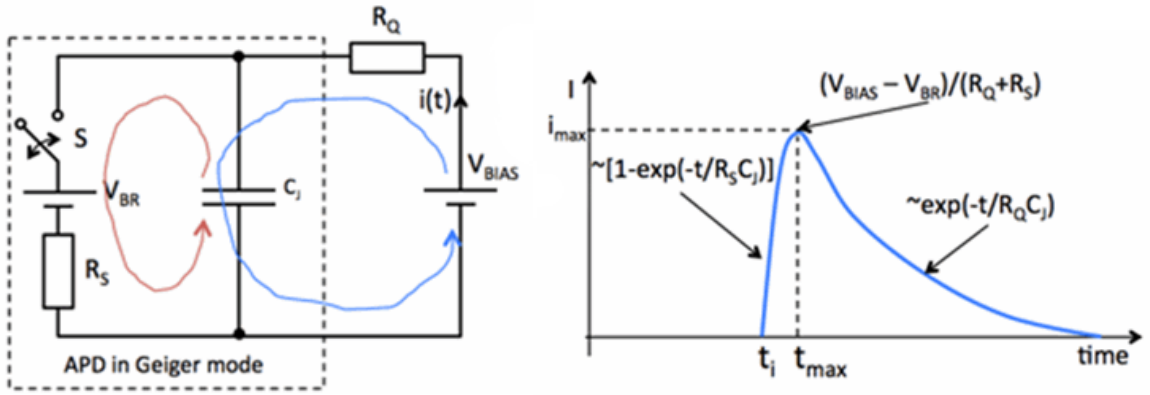


Figure 4-2 A simplified electrical equivalent circuit of a SPAD (left) and the profile of a typical current pulse (right). Both are reproduced from Hamamatsu [131]

More comprehensive overview of SiPM can be found in Acerbi’s review [132] and the excellent presentations by Collazuol [72][133].

4.2.2 Corsi's model and Marano's model

The electric model presented in Figure 4-2 doesn't include any parasitic parameters, which plays an unneglectable role in shaping the SiPM signal. Therefore, a more advanced model is required to include the impact of them.

As Shen Wei[134] argued, there is no universal SPICE model for all SiPM types due to the different parasitic effects from the different design. Of the various models, Corsi's model is so far the most widely accepted. Figure 4-3 shows the schematic description of Corsi's model.

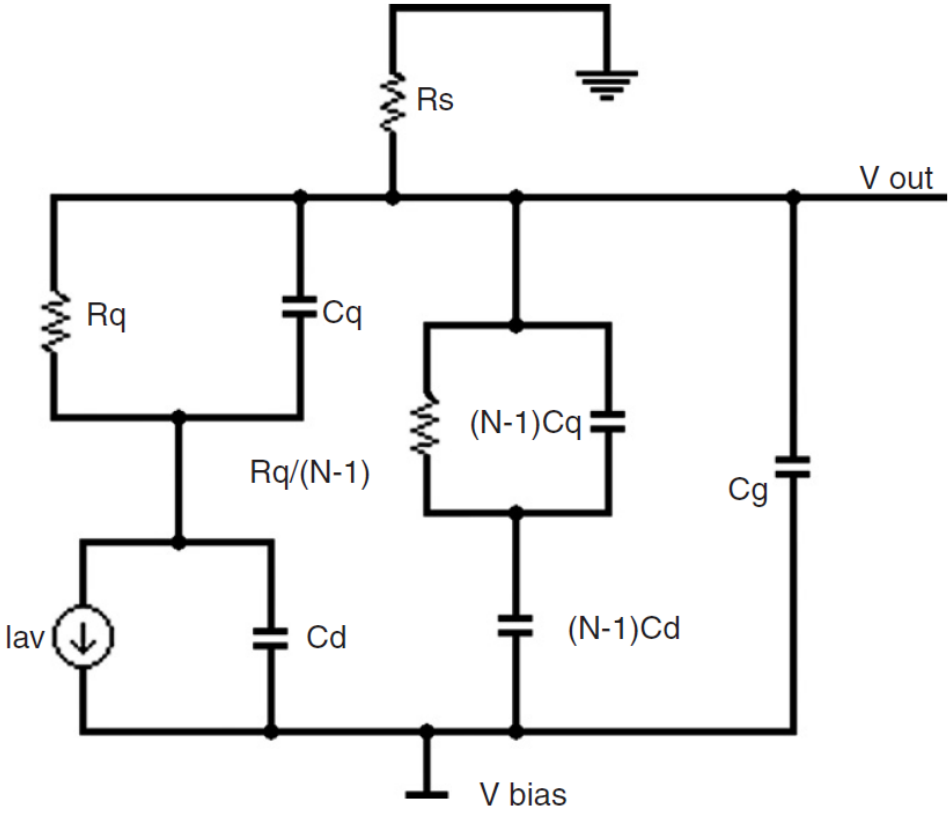


Figure 4-3 Corsi's electric model of SiPM [126]

In Corsi's model, each SiPM pixel/cell is represented as one $\mathbf{R_q}$ in series to one $\mathbf{C_d}$. $\mathbf{R_q}$ is the quenching resistor and $\mathbf{C_d}$ is the pixel (a SPAD diode) capacitance. For high frequency performance, the parasitic parameters must be considered, which are $\mathbf{C_q}$ and $\mathbf{C_g}$. $\mathbf{C_q}$ is the body capacitance of the quenching resistor $\mathbf{R_q}$. $\mathbf{C_g}$ is the SiPM total grid capacitance. The model represents the situation that one pixel is triggered, and the remaining pixels are in standby. The current source I_{av} represents the charge released by the triggered pixel. This charge can be represented as a Dirac delta pulse with a very short duration and high peak. Note this charge pulse only represents the release of avalanche current. The SiPM output current is the convolution of this Dirac

delta pulse and the RC network formed by all the parasitic parameters. Lastly, **N** represents the number of pixels.

Corsi's model is verified against the experiment SiPM electric performance. However, in practice, it is a step away to convert this theoretical model to an operable SPICE model that simulation software like LTSpice can run. Marano's improved SPICE model has filled in the gap and more. Figure 4-4 shows Marano's SPICE model in real schematic.

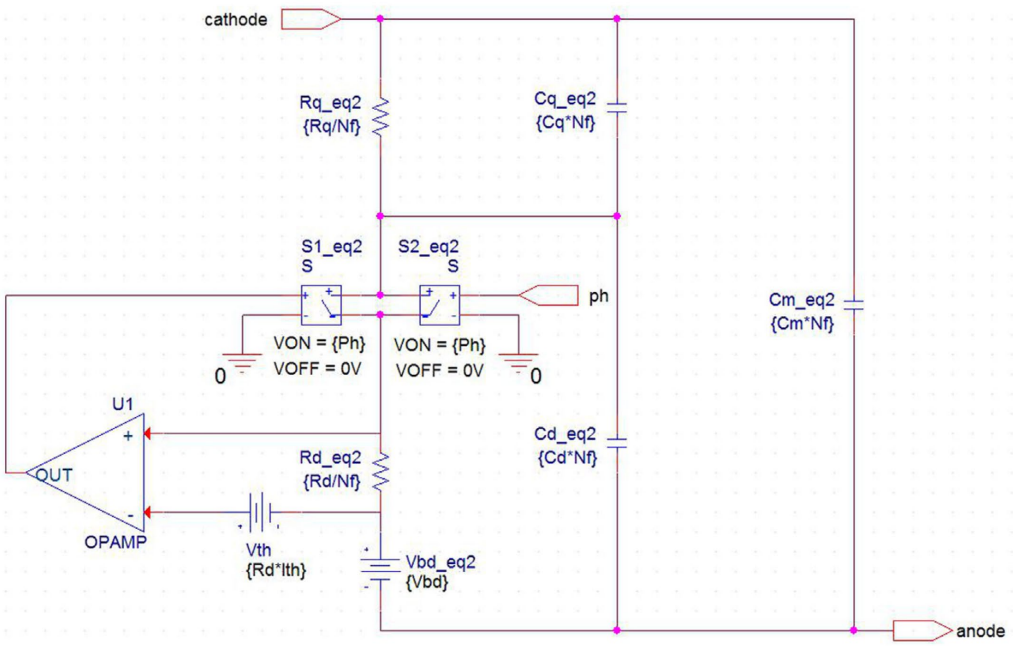


Figure 4-4 Marano's improved electric model of SiPM [127]

Compared to Corsi's model, Marano's spice model introduces one more electric parameter $\mathbf{R_d}$, which is the resistance of the diode space-charge region. This model cleverly introduces two voltage-controlled switches: **S1** and **S2**. **S2** simulates a photon strike. On striking, **S2** will close to mark the start of the avalanche event. The avalanche current through $\mathbf{R_d}$ will keep increasing until it reaches $\mathbf{I_{th}}$ (a pre-set threshold). Beyond this point, the voltage across $\mathbf{R_d}$ will be higher than $\mathbf{V_{th}}$ ($=R_d*I_{th}$). This will change the **U1** output from 0 to 1V. This will further close the switch **S1**. As the **S2** duration is very short (set to a few picosecond), switch **S1** allows the current to continue flowing from the bias supply (external connection to cathode) to the output (external connection to anode). The event is terminated when $\mathbf{V_{rd}}$ drops below the $\mathbf{V_{bd}}$ plus $\mathbf{V_{th}}$ due to the voltage loss across the $\mathbf{R_q}$ (passive quenching effect).

This dual switch design solved a practical problem that a more basic SPICE model with a single switch will present. In a single switch setup, the output signal strength (related to the total generated charge) will be dependent on the duration of the switch close time, while the dual switch design removed this dependency and it makes the output signal strength proportional to the total charge defined by equation (4.1).

$$Q = \Delta V (C_d + C_q) \tag{4.1}$$

Another advantage of Marano's model is its capability of simulating event with multiple photon strikes. By simply adjusting the $\mathbf{N_f}$ variable (define the number of photons), the model can generate the output signal with proportional amplitude.

4.2.3 Our practical improvement on Marano's model

The initial trial of the Marano's model in the LTSpice was not successful. Further study discovers that the problem was the response time of the op-amp (U1 in Figure 4-4). The role of the op-amp here is to act like a comparator. That is when the signal at the positive input is higher than the negative input, the output shall go positive. This positive signal shall turn on switch S1. Initially a general op-amp model is used in the simulation, but its response is too slow. Then parameters as slew rate, open-loop gain, are adjusted to be very large in order to achieve a fast response, but the result is still not satisfactory. What the circuit needed was almost an instantaneous response upon the two input comparison signal. We discovered a B source (an electric model in the LTSpice, with full name as arbitrary behaviour voltage source) can fit this role well. The B source can set its status on the condition of other electric signals with instant effect. The modified circuit as shown in Figure 4-5 now operates fully as expected.

Variable parameters are widely used in this circuit so that different SiPM devices can be simulated by simply updating the parameter values.

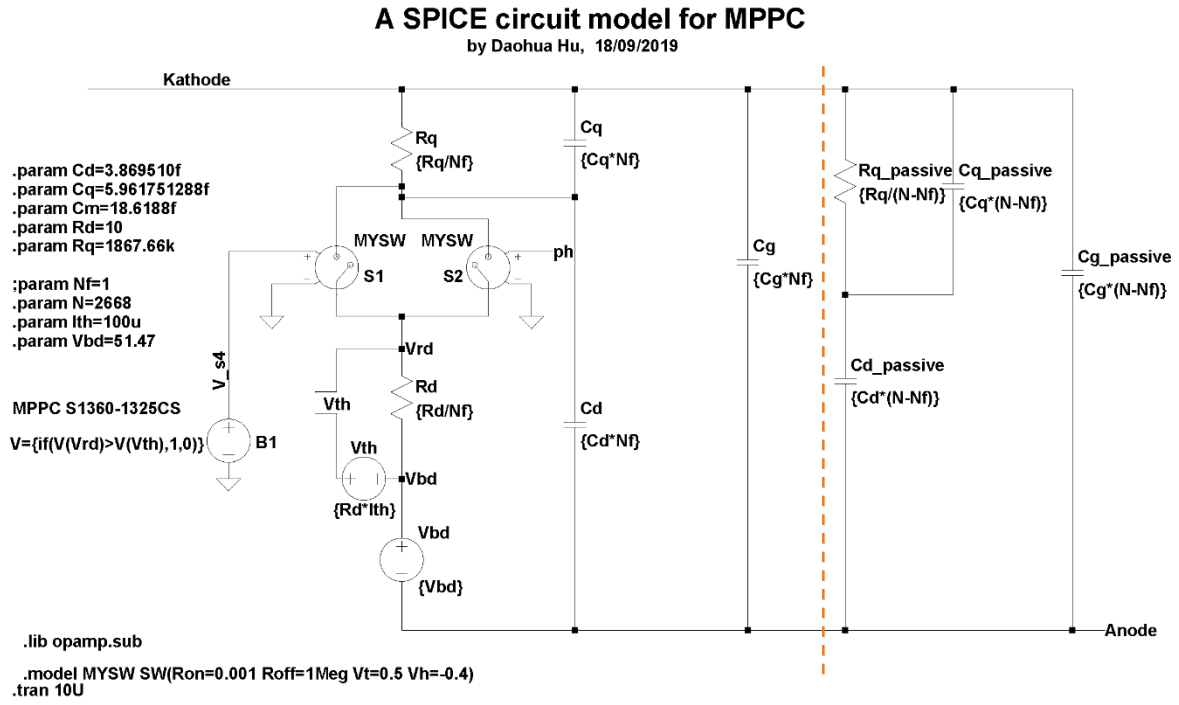


Figure 4-5 The improved dual switch SiPM spice model based on Marano's model

Note the section at the right of the dashed-line is for the untriggered pixels and the section at the left is for the triggered pixels. The passive structure of the two sections are the same. That is the quenching resistor value($\mathbf{R_q}$) and its body capacitance($\mathbf{C_q}$) in series to the pixel diode capacitance($\mathbf{C_d}$). The sum joins the pixel grid capacitance ($\mathbf{C_g}$) in parallel. Here $\mathbf{C_g}$ is defined per pixel, which is different from that in Corsi's model which defines the $\mathbf{C_g}$ per device. **S1** and **S2** are two switches for the event control and a voltage source $\mathbf{V_{bd}}$ is for setting the breakdown voltage.

To examine this SPICE model, a test circuit is developed (Figure 4-6). In this circuit, the SiPM (X1) is reverse biased. A pulse generator (V2) is set to produce the trigger signal (simulate a photon strike event). The $\mathbf{N_f}$ (number of fired pixel) variable is stepped from 1 to 3 to simulate one, two and three pixels firing events. The SiPM output current is converted to a voltage signal through R1.

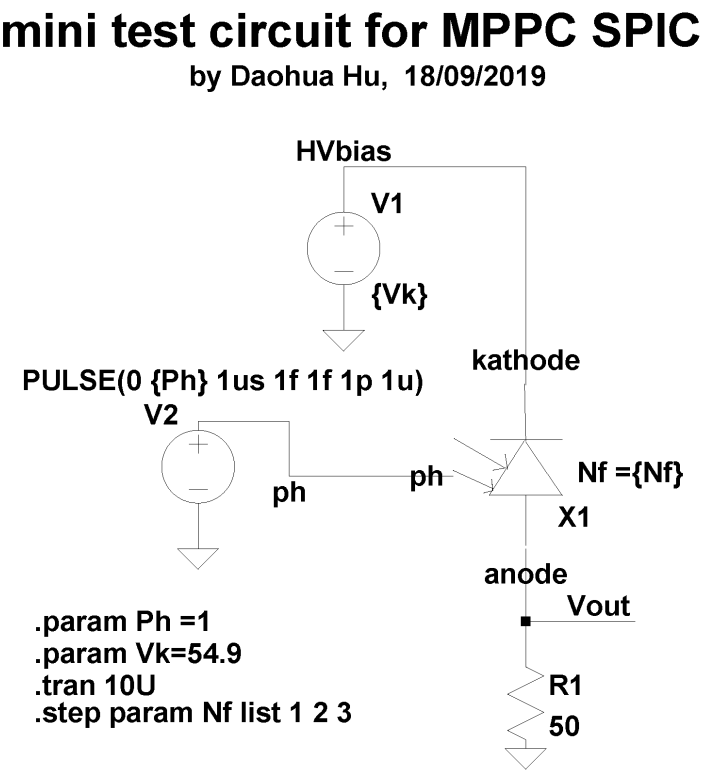


Figure 4-6 A test circuit for the improved dual switch SiPM circuit model

Figure 4-7 shows the simulation output. The top panel shows the output current and the bottom panel shows the output voltage. All three events (one photon, two photon and three photon) are stacked in the same plot. The simulation shows the output signal has a very sharp rising edge and relative slow falling edge. The rising edge builds up during the avalanche process, which is extremely fast (in the order of pico-seconds). The peak is reached when the avalanche is terminated by the quenching resistor. The falling edge is dominant by the period that the pixel capacitance ($\mathbf{C_d}$) is recharged to the external bias voltage through the quenching resistor.

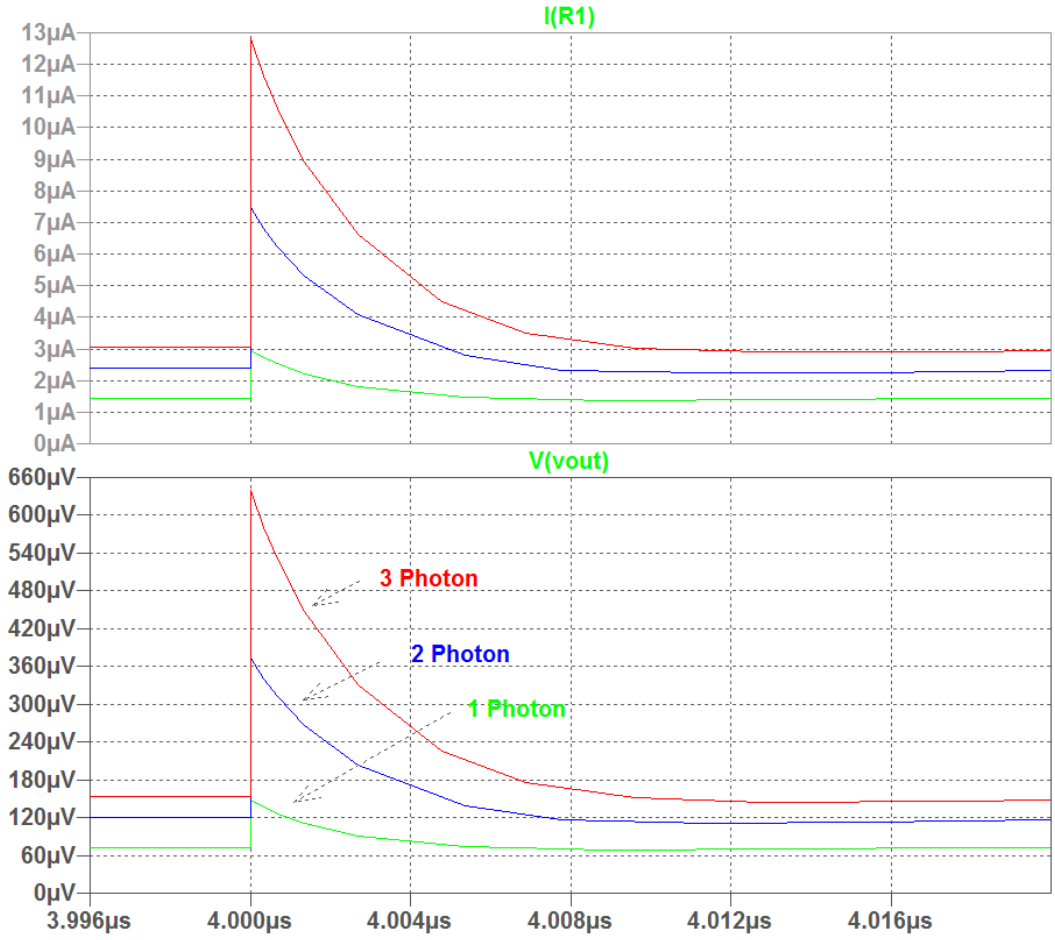


Figure 4-7 The improved SiPM SPICE model test circuit simulation output: current (top) and voltage (bottom)

We continued to simulate more photon events and record the peak current values against the number of photons. Figure 4-8shows the results with very good linearity.

No of Photons	Current peak	Differetiate
#	uA	uA
1	2.99	2.99
2	7.49	4.50
3	12.82	5.33
4	18.66	5.84
5	24.84	6.18
6	31.25	6.41
7	37.82	6.57
8	44.52	6.70
9	51.31	6.79
10	58.17	6.87
11	65.10	6.92
12	72.07	6.97
13	79.08	7.01
14	86.12	7.04
15	93.18	7.07
16	100.27	7.09
17	107.38	7.11
18	114.51	7.13
19	121.64	7.14
20	128.80	7.15

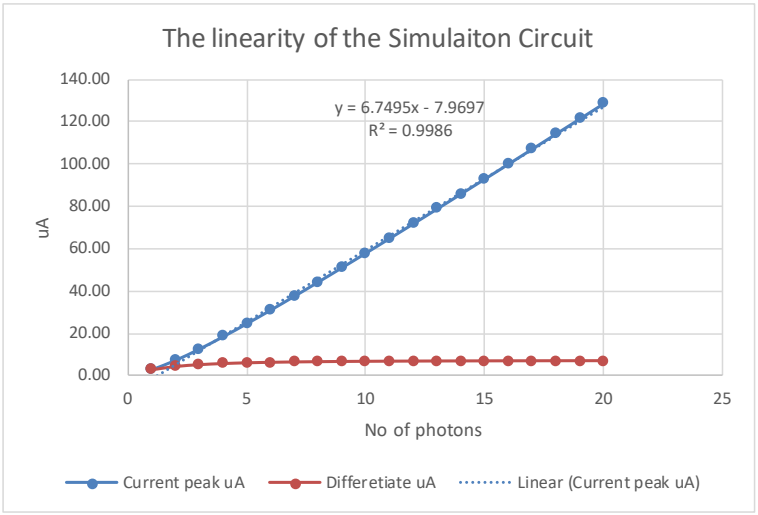


Figure 4-8 The improved SiPM SPICE model output current vs input number of photons (blue). Red line is the differentiate of blue line.

4.3 Evaluate the SiPM electric parameters

With the SPICE model defined, the next step was to determine all the electric parameters required by this model. The list of parameters is summarised in Table 4-1.

Some parameters (marked with *) can be acquired from the manufacturer datasheet. Others have to be acquired by lab measurement plus using Corsi's methods[126] to derive. Note **R_q**, **C_q**, **C_d**, **R_d** and **C_m** all refer to a single pixel. **C_g**, **V_{bd}** and **N** refer to the entire device. Corsi's formula is rewritten here to be consistent with the spice model shown in Figure 4-5.

$$C_{dTOT} = \sqrt{\frac{1+\omega^2 C_{TOT}^2 R_{qTOT}^2}{\omega^2 R_{qTOT}^2} G_m} \tag{4.2}$$

$$C_{qTOT} = C_{TOT} - C_{dTOT} \tag{4.3}$$

$$C_g = C_m - C_{dTOT} - \frac{\omega^2 C_{dTOT}^2 C_{TOT} R_{qTOT}^2}{1+\omega^2 C_{TOT}^2 R_{qTOT}^2} \tag{4.4}$$

Where $C_{qTOT} = NC_q$, $C_{dTOT} = NC_d$ and $R_{qTOT} = R_q/N$.

Table 4-1 The list of parameters used for the SiPM electric model

Symbol	Parameter Name	Method to acquire
R_q	Quenching resistor per pixel	We-V curve, forward bias
C_q	Parasitic capacitance of the quenching resistor	Corsi's formula
C_d	Diode junction capacitance per pixel	Corsi's formula
R_d	Diode conducting resistance per pixel	Estimate 1-10ohm
C_g	Parasitic capacitance between metal grid and substrate per device	Corsi's formula
C_{TOT}	A virtual parameter	Work out from single pixel charge
C_m*	Device equivalent terminal capacitance	Manufacturer data or LCR measurement
G_m*	Device equivalent terminal conductance	LCR measurement
V_{bd}*	Breakdown voltage	Manufacturer data or I-V curve, reverse biased
N*	Number of pixels	Manufacturer data

4.3.1 Measure the I-V curve of the SiPM

A typical diode I-V curve measurement circuit is setup as Figure 4-9.

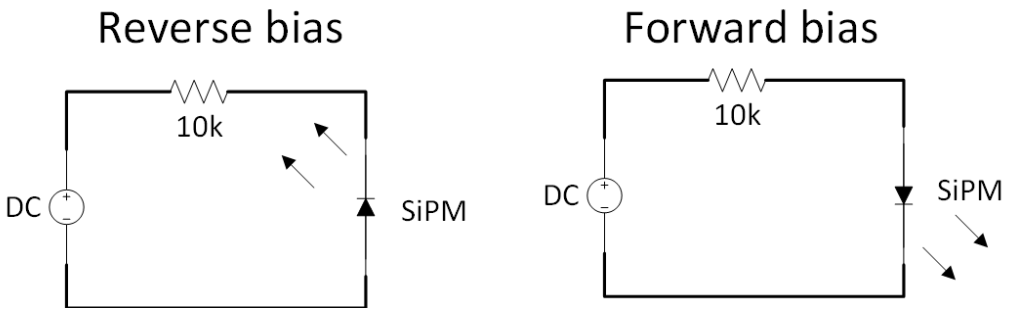


Figure 4-9 SiPM I-V curve measurement circuit

Note:

1) The current is acquired by measuring the voltage drop across the resistor and then using the ohm law to work out the current.

2) The voltage across the SiPM is acquired from the DC voltage minus the voltage drop of the resistor.

3) Sweeping the voltage and recording the current, the I-V curve can be plotted. Note under reverse bias, the voltage and current are both plotted on the negative quadrant of the XY plane. Under forward bias, both are placed on the positive quadrant.

Using the above method, the SiPM S13360-1325CS from Hamamatsu is measured. The result is plot in Figure 4-10.

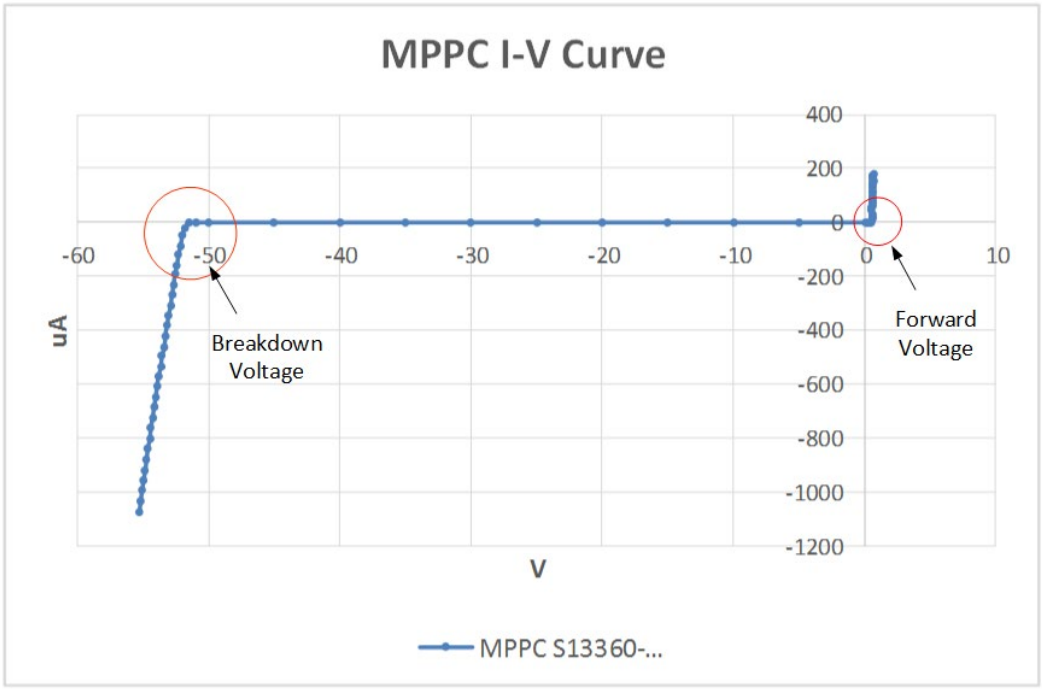


Figure 4-10 I-V curve measurement result for SiPM S13360-1325CS

The corner at the negative x-axis is the breakdown voltage, biased beyond which the avalanche will start. The corner at the positive x-axis is the forward voltage, biased beyond which the diode is forward conducted.

Reading from the I-V curve, the breakdown voltage for SiPM S13360-1325CS is 51.47V. The manufacturer test data in the product shipment package states the exact same figure. This proves that our test method is the same as the manufacturer's setup and also this I-V measurement method can provide consistent and accurate result.

Similarly, the forward voltage is read at 0.485V. Under forward bias, the diode current is determined by the (bias voltage – forward voltage) / (R_q + R_d). Since R_d is very small (0.1 to 10ohm) and R_q is very large (>100kohm), the contribution of R_d can be neglected. Hence, we can work out the R_q from the slope of the I-V curve under forward bias. The data is quite noisy, which might indicate the forward voltage of the diode is not very stable (The temperature will impact the forward voltage. As the amount of current that passes through the diode will heat up the device, the rising temperature will allow more current through at the same bias voltage.) A simple linear fit works out the slope at 0.0006Mohm (=600ohm). Since all the pixel is connected in parallel, the measured resistance is the sum of all the pixel's R_q in parallel. Given the knowledge of total number of pixels (2886) for this device, we can work out a single pixel's R_q is 600 x 2886 = 1600kohm.

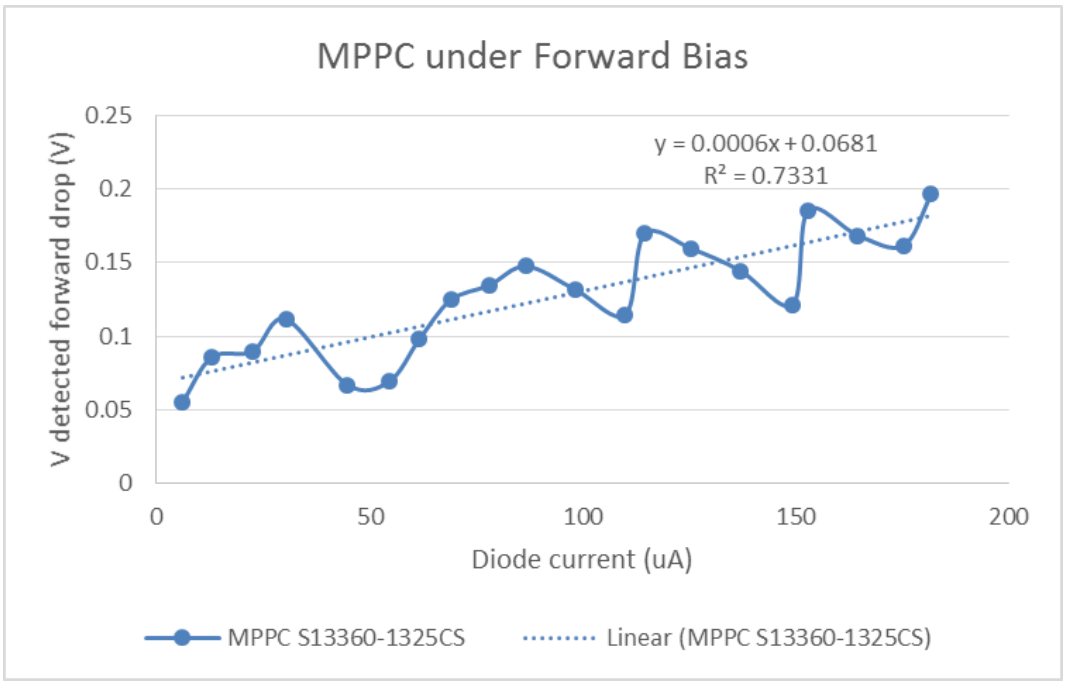
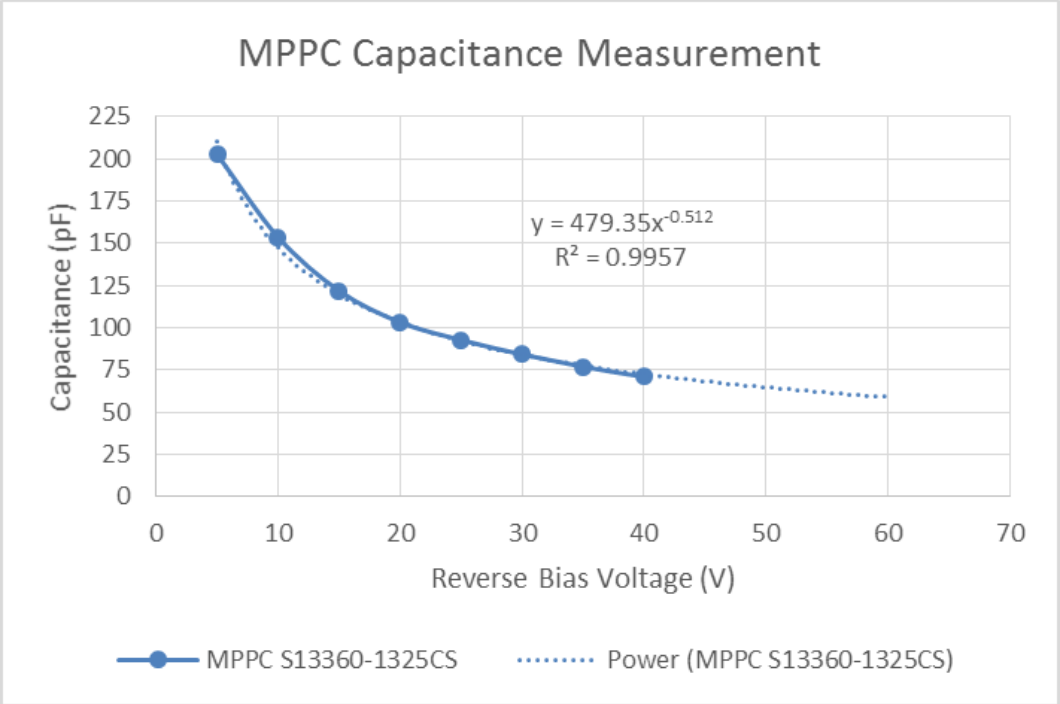


Figure 4-11 The measurement of IV curve under the forward bias

4.3.2 Measure the Terminal Capacitance and Conductance of the SiPM

The total terminal capacitance of the device can be measured with an LCR bridge instrument. The principle of the instrument is to inject a signal through the device under test at a known frequency while measuring the voltage and current across the device. The magnitude of the impedance can be calculated from the ratio of the voltage to current.

Note the SiPM terminal capacitance and conductance will change with its bias voltage. So, the measurement shall be done just under its breakdown voltage. As our LCR bridge output is limit to maximum 40V that is short of the SiPM required breakdown voltage(>50V), the out of range voltage point is extrapolated using the fit function on the measured data points from 0 to 40V. The result is shown in Figure 4-12.



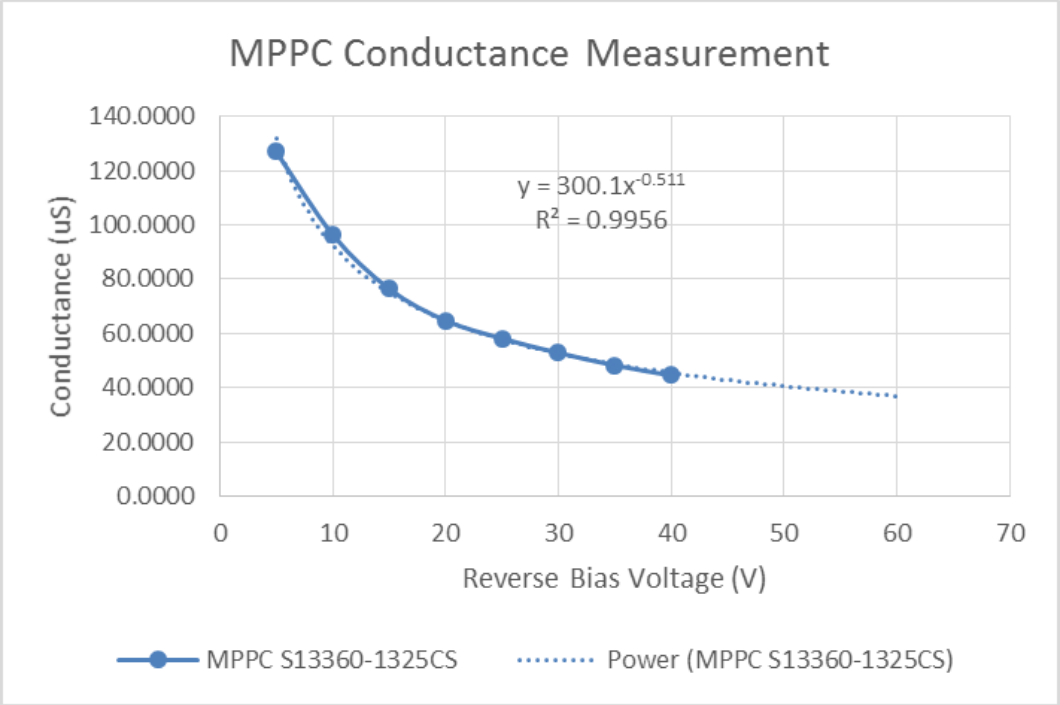


Figure 4-12 The capacitance (top) and conductance (bottom) measurement of SiPM S13360-1325CS

The acquired result is terminal capacitance (C_m) of 63.73pF and terminal conductance (G_m) of 40.055uS at the breakdown voltage of 51.47V. This again matches the manufacture datasheet value at 60pF and 40uS very well.

4.3.3 Measure the Quenching Resistor Parasitic Capacitance and the Pixel Diode Capacitance of the SiPM

These two capacitances can only be measured indirectly with the help of a preamplifier. The procedure is described below.

First, connect the SiPM device to a preamplifier with a known gain. Bias the device at a few volts above the breakdown voltage and place the system in a dark box. Next, connect an oscilloscope to the output of the preamplifier and only capture the dark count signal (ignore all those signals with pile-up effect). On the oscilloscope, the output charge can be measured as Q_{out}. With the preamp gain at A, the input charge can be worked out as $Q_{in} = Q_{out} / A$. The Q_{in} is the charge of the SiPM device with one-pixel triggered (= 1p.e. signal). Rearrange equation (4.1), we can have:

$$C_q + C_d = Q_{in} / \Delta V \tag{4.5}$$

Multiply by the pixel number N, we can work out $C_{TOT} = N \cdot (C_q + C_d)$

Here the preamp gain A is 50. The result of single pixel charge measurement result is shown in Figure 4-13. From the linear fit, we can work out the C_q+C_d = 0.0245pC. Then using Corsi's equation (4.2),(4.3) and (4.4), we can continue to work out C_q, C_d and C_g at 18.29fF, 6.29fF and 43.22pF.

DeltaV	Qin
(V)	(pC)
5.33	0.13
5.53	0.13
6.03	0.15
6.53	0.16
7.03	0.18
7.53	0.19
8.03	0.20
8.53	0.21
9.03	0.22
9.53	0.23
10.03	0.24
10.53	0.26

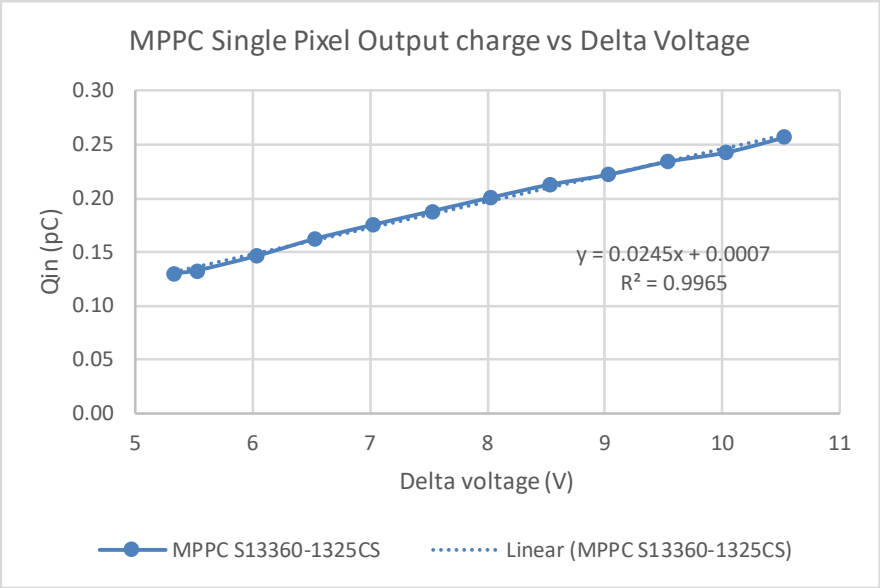


Figure 4-13 The single pixel charge measurement of SiPM S13360-1325CS

4.3.4 Summary

Now all the SiPM electric parameters are acquired. The result for the SiPM device S13360-1325CS from Hamamatsu is summarized in Table 4-2:

Table 4-2 Summary of the electric parameters of MPPC S13360-1325CS

Parameter Symbol	Value	Unit	Manufacturer specification
R_q	1600	kΩ	N/A
C_q	18.29	fF	N/A
C_d	6.29	fF	N/A
R_d	~1	Ω	N/A
C_g	43.22	pF	N/A
C_{TOT}	65.57	pF	N/A
C_m	63.73	pF	60
G_m	40	uS	40
V_{bd}	51.47	V	51.47
N	2668	#	-

N/A means Not Available

4.4 Preamplifier circuit theory

The first readout circuit stage following the sensor is called preamplifier. Generally, the role of preamplifier is to amplify the sensor signal and convert it to a suitable electrical signal format (usually a voltage signal) for the following stage of readout circuit.

Thanks to the intrinsic high gain of SiPM, its signal is relatively large (peak current at about 1-10uA). So, only mild amplification is required. The raw signal is in the form of charge. This needs to be converted to a voltage signal. Preamplifier serves for this kind of raw signal is also called charge amplifier. The charge preamplifier usually defines its gain as Volt to Charge ratio (with unit V/pC).

Another important role of preamplifier is to improve the signal to noise ratio. This is achieved by keeping the preamplifier self-generating noise low. Radeka [135] has a very thorough discussion about low-noise techniques for charge amplifier working with detectors.

4.4.1 Two basic preamplifier configurations

There are two basic configurations for SiPM detector readout. They are defined as **voltage amplifier** and **current amplifier** as illustrated in the Figure 4-14. Here the name “voltage” and “current” refer to the signal being amplified by the op-amp. Both configurations will convert the SiPM raw signal to a voltage signal. The difference is in voltage amplifier, the conversion is done by a load resistor $\mathbf{R_s}$ while in current amplifier, the conversion is done by the feedback resistor $\mathbf{R_f}$. Don’t confuse them with the specific terminology used by the semiconductor manufacturer for the two different types of operational amplifier.

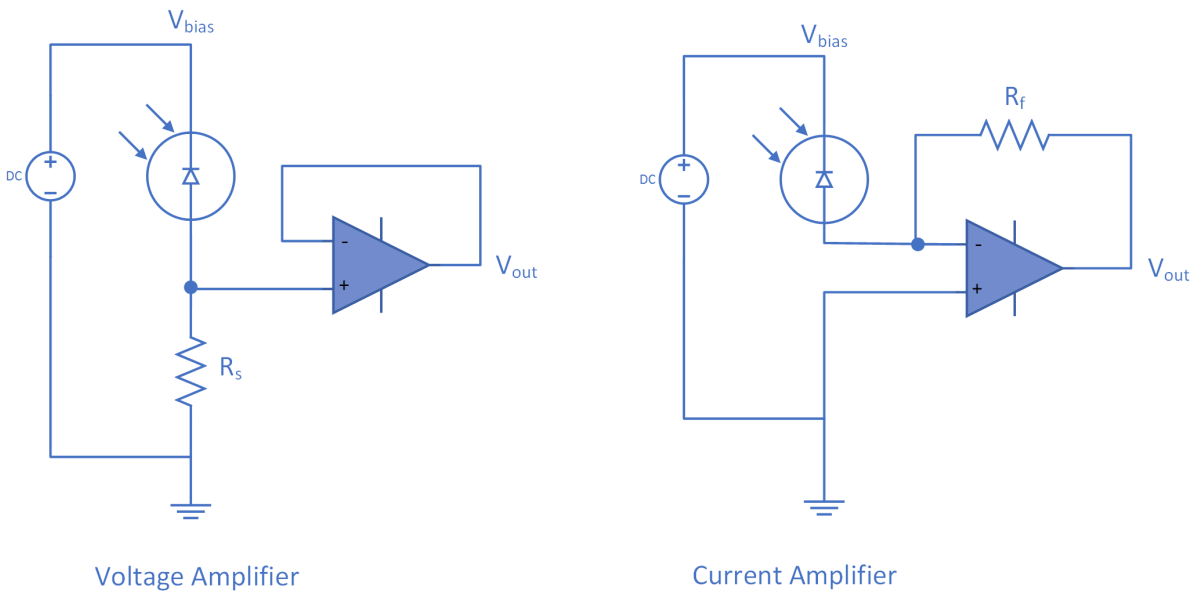


Figure 4-14 The basic circuit of two pre-amplifier configuration: voltage amplifier (left) and current amplifier (right)

4.4.1.1 Voltage amplifier

As the name indicated, in this configuration, it is a voltage signal that is amplified by the op-amp. As the op-amp has very high input impedance, the photon current I_s will only flow through the resistor R_s. By ohm's law, a voltage signal is generated on the top of R_s. The op-amp here is configured as a unit gain buffer, which simply maintains the output voltage equal to the voltage at its non-inverting input. Equation (4.6) describes the circuit mathematically.

$$V_{out} = I_s \cdot R_s \tag{4.6}$$

Sometimes, the op-amp can also be configured with a gain. This is realized by two resistors on the feedback link. This gain however will decrease the circuit bandwidth as the op-amp has a fixed gain-bandwidth product (GWP). Figure 4-15 shows the circuit with a gain. The gain is defined as $(R_2+R_1)/R_1$.

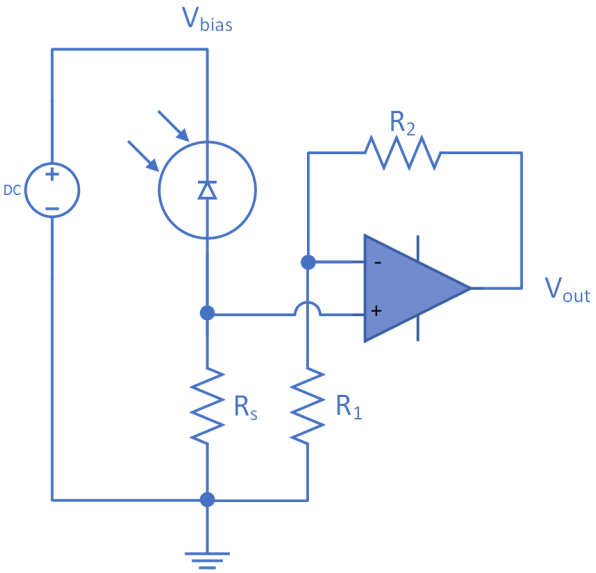


Figure 4-15 Voltage amplifier with a gain

The principle of the voltage amplifier is simple, but there are certain limits to it. The first limit is the bandwidth. The impulse photon current has very fast rising time, for example 0.5ns. To double sample this edge or more, the op-amp should have a minimum bandwidth of 4GHz. The choice of op-amp with such high bandwidth is limited and usually they come with high power consumption and high noise. More details about the bandwidth will be discussed in the next section on the small signal analysis. The second limit comes from the voltage signal generated on R_s. This voltage will effectively reduce the bias voltage across the SiPM device. This can terminate the avalanche too early. Also, the diode junction capacitance C_d is dynamically affected by the voltage across it. That capacitance increases as the bias voltage decrease. An increased C_d will extend the signal pulse falling tail, which is defined by the time constant of C_d*R_s. This means longer dead time of the circuit. This is the reason that generally R_s is kept very small (<1kΩ).

The limits of the voltage amplifier are just the advantage of the current amplifier, which will be discussed next.

4.4.1.2 Current Amplifier

In this configuration, it is the current signal being amplified. The op-amp with very high input impedance forces photon current I_s through its feedback resistor R_f. As the op-amp tends to keep its two input signals at the same potential, the inverting input signal will equal to the non-inverting input signal which is tied to ground. Hence this inverting

input is electrically at zero potential though not physically connected to ground. This feature is often referred as “virtual ground”. The output signal will be driven to the potential that helps to maintain this virtual ground potential through R_f.

We can quickly work out the formula of the gain of this circuit as:

$$V_{out} = -I_s \cdot R_f \tag{4.7}$$

The negative sign indicates the output signal has an opposite polarity to the current flow direction of I_s. That is if the current flows out of the photon sensor, the output voltage signal is negative. Vice versa if the current flows into the photon sensor, the output voltage signal is positive. We can change the flow direction of I_s by reconfiguring the circuit as Figure 4-16. This makes I_s flow into the photon sensor and hence the op-amp generates a positive V_{out}. This circuit is also called trans-impedance amplifier (TIA).

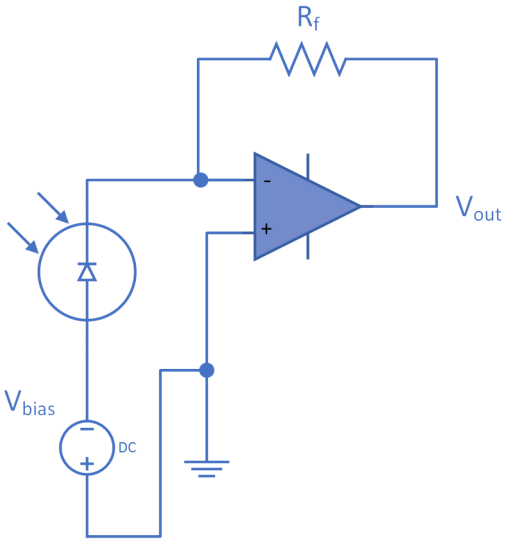


Figure 4-16 Bias the SiPM with negative Vbias to generate a positive Vout signal

In the most basic form (4.6)(4.7), the gain formula of the voltage amplifier and current amplifier is the same. The gain in both cases is defined by a single resistor. However, as mentioned in the limitations of the voltage amplifier, the choice of its resistor value is limited to relatively small value. Oppositely, the current amplifier can employ relatively large feedback resistor. This helps the circuit to achieve a relatively high gain.

For AC signal analysis, the common approach is called small signal analysis.

4.4.2 Small signal analysis

Eq. (4.6) and (4.7) are only valid for signal from DC to very low frequency. As the avalanche photon current pulse is extremely fast, the impact of all the circuit parasitic parameters becomes non-neglectable. To evaluate the circuit performance at high frequency, we convert the circuit into small signal mode to perform the AC analysis.

In small signal analysis, all the DC supply will be connected to ground and the active device will be replaced with a linear electrical model. The SiPM small signal model (as shown in Figure 4-5) is too complex for this discussion, so we simplified it to be an impulse current source in parallel to a capacitor C_d. All the op-amp's parasitic parameters are ignored for the moment, so the op-amp is just modelled as a voltage gain A_{OL}.

The small signal of the two amplifier circuits is shown in Figure 4-17.

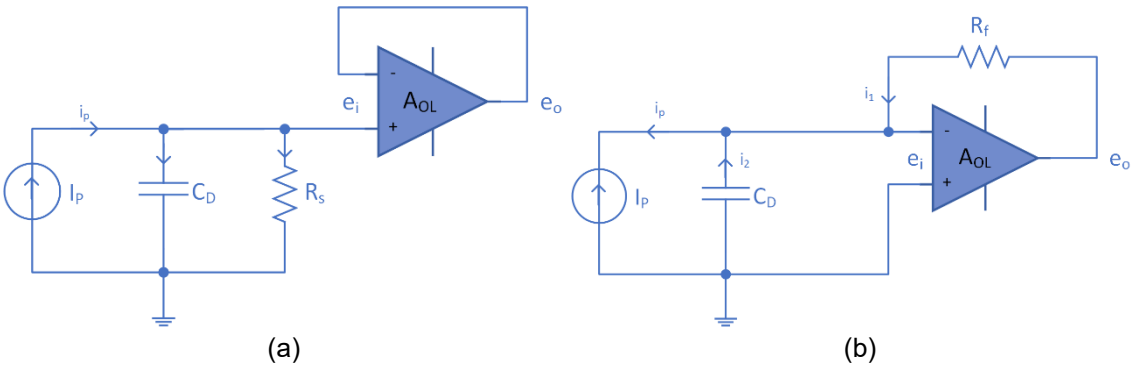


Figure 4-17 Small signal of the voltage amplifier (a) and current amplifier (b) circuit

4.4.2.1 Voltage amplifier

Let's work out the signal gain equation of the voltage amplifier circuit first.

$$e_o = e_i \times A_{OL} \tag{4.8}$$

$$e_i = i_p \times X_{RC} - e_o \tag{4.9}$$

Replace e_i in eq. (4.8) with equation (4.9) and re-organize the new equation. We can get the result of e_o as:

$$e_o = i_p \cdot \frac{X_{RC}}{1 + 1/A_{OL}} \tag{4.10}$$

Here, X_{RC} is the total impedance of R_s and C_d in parallel, which is:

$$X_{RC} = \frac{R_S}{(1 + j\omega R_S C_D)} \tag{4.11}$$

Replace X_{RC} and reorganize the equation (4.10), we can derive the final gain equation as

$$e_o = \frac{i_p \cdot R_S}{(1 + \frac{1}{A_{OL}}) + (1 + \frac{1}{A_{OL}}) \cdot j\omega R_S C_D} \tag{4.12}$$

4.4.2.2 Current Amplifier

Similar process can apply to the current amplifier circuit. First, from Figure 4-17(b), the following set of conditions is derived.

$$\begin{aligned} -e_i &= \frac{e_o}{A_{OL}} \\ i_p &= i_1 + i_2 \\ i_1 &= \frac{(e_o - e_i)}{R_f} \\ i_2 &= -\frac{e_i}{X_{C_D}} \\ i_p &= \frac{(e_o - e_i)}{R_f} + (-\frac{e_i}{X_{C_D}}) \end{aligned} \tag{4.13}$$

After cancelling i_1, i_2 and e_i, and re-organize the new equation, the gain equation is derived as:

$$e_o = \frac{i_p \cdot R_f}{1 + \frac{1}{A_{OL}} \cdot (1 + \frac{R_f}{X_{C_D}})} \tag{4.14}$$

Replace X_{C_D} with $1/j\omega C_D$, equation (4.14) can be finalized as:

$$e_o = \frac{i_p \cdot R_f}{(1 + \frac{1}{A_{OL}}) + \frac{1}{A_{OL}} \cdot j\omega R_f C_D} \tag{4.15}$$

Compare equation (4.12) and (4.15), it is found that the gain equation of the two circuits are very similar with the difference only at the denominator. That difference will

see voltage amplifier's gain rolling off quicker than the current amplifier's at increasing frequency.

At low frequency, as the gain A_{OL} is very large, the $1/A_{OL}$ is close to 0, and C_D is at the order of 10s pF, so $j\omega RC_D$ is about 0 as well. Both equation (4.12) and (4.15) can be simplified to $e_o=i_p*R$, which matches the equation (4.6) and (4.7).

From the frequency response plots of the two circuits (Figure 4-18) by LTSpice with exact the same component values, it is found that the current amplifier circuit has wider bandwidth (defined as the frequency where the gain is 3dB below the DC level) than the voltage amplifier circuit.

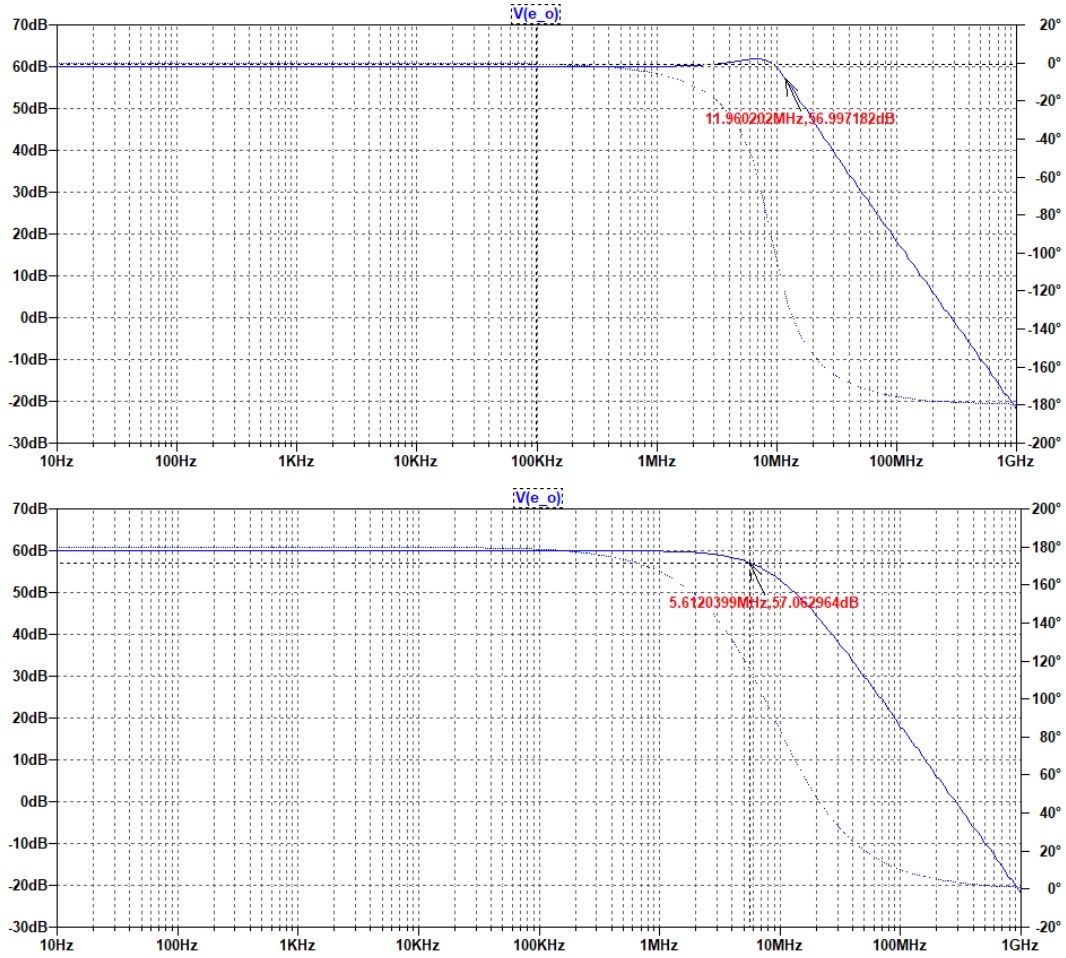


Figure 4-18 Frequency response of the two circuits by LTSPICE simulation with the same component values. Top is from current amplifier. Bottom is from voltage amplifier.

Both circuits take C_D at 20pF and R_f at 1kohm. The simulation shows for the current amplifier circuit, the -3dB point is at 12MHz while for the voltage amplifier circuit, it is at merely 5.6MHz.

With wider bandwidth and higher trans-impedance gain, the current amplifier circuit is selected as the baseline choice for the next stage development. According to [136], the current amplifier also has better linearity.

4.4.3 System stability analysis

4.4.3.1 Stability of the uncompensated system

For any circuit with feedback to close the control loop, it is crucial to understand the stability of the circuit. Because an unstable circuit will oscillate. Even the circuit appears to be designed with a negative feedback loop in the schematic, the circuit behaviour at different frequency can be very different. If at any frequency, the feedback signal turns to be 180deg phase shift with a positive gain, this negative feedback will become positive feedback and the circuit will oscillate at this frequency.

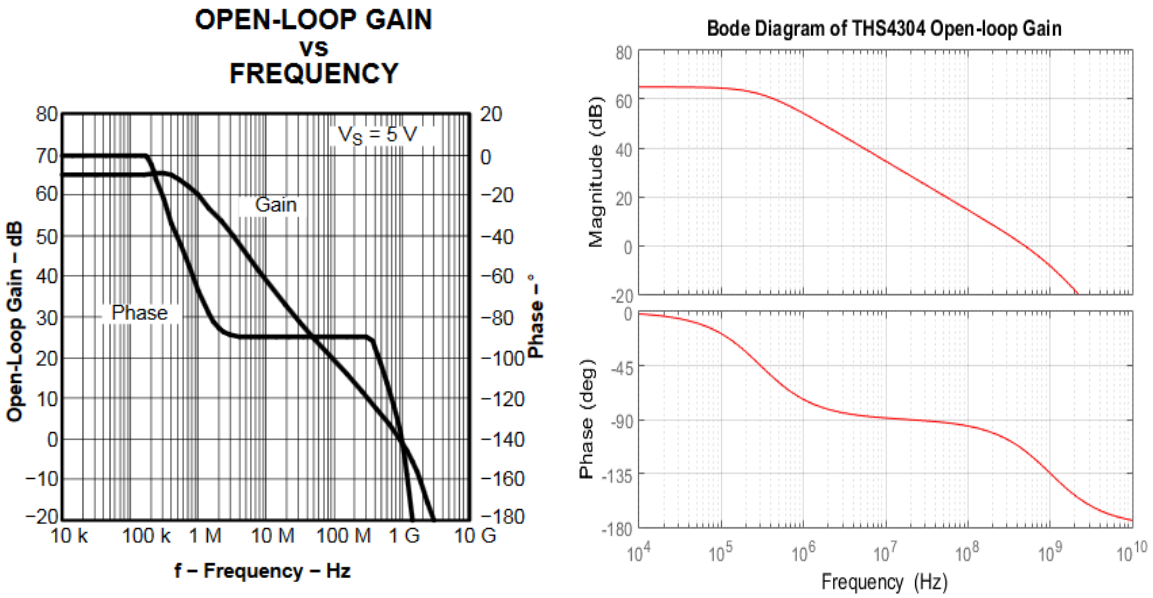
The stability of a closed loop system can be further analysed in MATLAB with the bode diagram function.

First, it is assumed the open-loop gain of the op-amp is a two-pole system with a s-function expressed as:

$$A_{OL} = \frac{a_0}{(1 + s/w_1)(1 + s/w_2)} \tag{4.16}$$

Where a_0 is the DC gain, ω_1 and ω_2 are the angular frequency of the two poles.

The wideband op-amp THS4304 (from TI) is taken as an example. Its open-loop gain (A_{OL}) frequency response plot from its datasheet is shown in Figure 4-19 (a):



(a) (b)

Figure 4-19 (a) is the open-loop gain frequency response from THS4303 datasheet. (b) is the MATLAB bode plot of the s-function of (4.16).

From Figure 4-19(a) gain curve, we can read all the parameters we need for Eq. (4.16). The DC gain a_0 is sampled at the start of the curve at 65dB. The first pole is at the first corner where the curve starts rolling off at 20dB/decade. The second pole is at the second corner where the rate of the drop doubles to 40dB/decade. Thus, f_1=300kHz, f_2=1GHz. So, the two angular frequencies are: $\omega_1=2\pi f_1$=1.885E6, $\omega_2 = 2\pi f_2$ = 6.28E9.

Using the MATLAB *bodeplot* function to plot the s-function in (4.16) with all the parameters acquired above, the bode diagram of the open-loop gain can be generated and shown in Figure 4-19 (b). We can see the plot (b) matched (a) closely well. This proves the op-amp open-loop gain frequency response can be closely approximated with a second order two-pole system.

The gain dropped to 0dB at frequency f_c. 0dB means gain of 1. So that f_c is often named as the op-amp unity-gain bandwidth. Beyond the frequency f_c, the op-amp gain will drop to below one where the op-amp is not function as an amplifier any more.

Next, the closed loop system is derived from the open loop system. First introduce two new definitions: feedback factor **β** and loop gain **L**. Feedback factor **β** is the fraction of the amplifier output signal fed back to its input signal. Loop gain **L** is the product of A_{OL} and β. These two definitions will be used below to simplify the formula.

For the current amplifier circuit defined in Figure 4-17 (b), its transfer function in (4.14) can be re-written as:

$$e_o = \frac{i_p \cdot R_f}{1+1/(A_{OL} \cdot \beta)} \tag{4.17}$$

Where $\beta = \frac{X_{C_D}}{X_{C_D} + R_f} = \frac{1}{1+sR_fC_D}$.

The trans-impedance closed-loop gain **A_{CL}** can be defined as:

$$A_{CL} = \frac{R_f}{1+1/(A_{OL} \cdot \beta)} \tag{4.18}$$

Using the same MATLAB tool, the magnitude of the frequency response of **A_{OL}**, **1/β** and **A_{CL}** are plot in Figure 4-20.

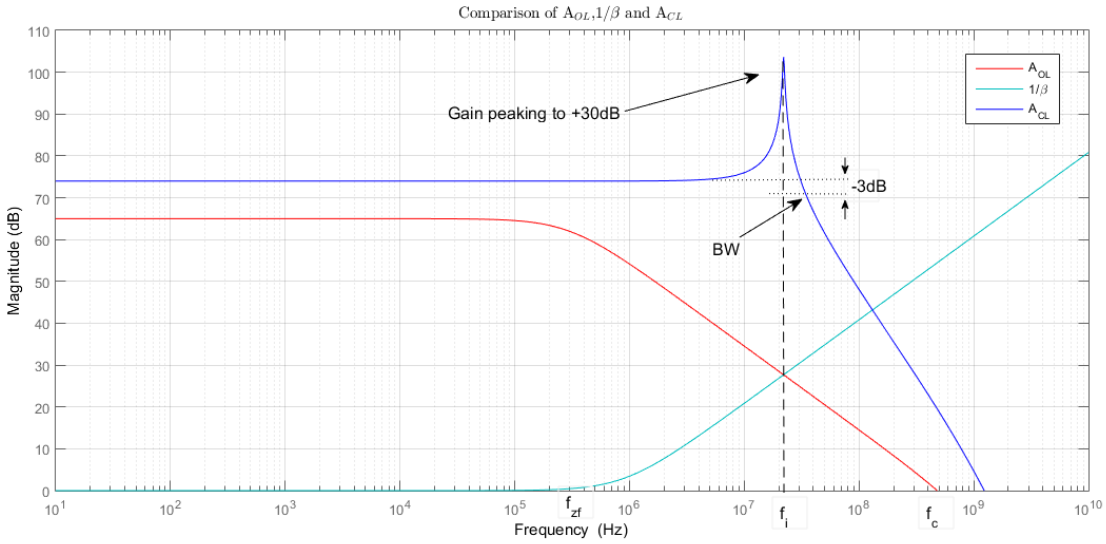


Figure 4-20 Frequency response plot of A_{OL}, 1/β and A_{CL}.

In the plot, **f_{zf}** is the pole of feedback factor **β** so it becomes the zero of **1/β**. **f_{zf}** can be defined as

$$f_{zf} = 1/2\pi R_f C_d \tag{4.19}$$

For R_f = 5kohm and C_d = 35pF, **f_{zf}** is worked out at 9.1e5 Hz.

f_c is the unit gain bandwidth of the open-loop gain. We can read its value from the plot at 5e8 Hz.

f_i is the gain peaking point of the closed-loop gain. **f_i** sits in the geometric centre of **f_{zf}** and **f_c**. As the x-axis is logarithm, **f_i** can be defined as:

$$f_i = \sqrt{f_{zf} f_c} \tag{4.20}$$

Then we can calculate the $\mathbf{f_i}$ = 2.2e7 Hz.

The system bandwidth is generally defined as:

$$BW \approx 1.4 f_i \tag{4.21}$$

Similar to $\mathbf{A_{OL}}$, the bandwidth of $\mathbf{A_{CL}}$ is defined at the point where the gain dropped by 3dB from its flat part.

The gain peaking is something that shall be limited in the system as it will contribute to the system instability. The system stability can be better viewed using MATLAB function *margin* for the loop gain **L**.

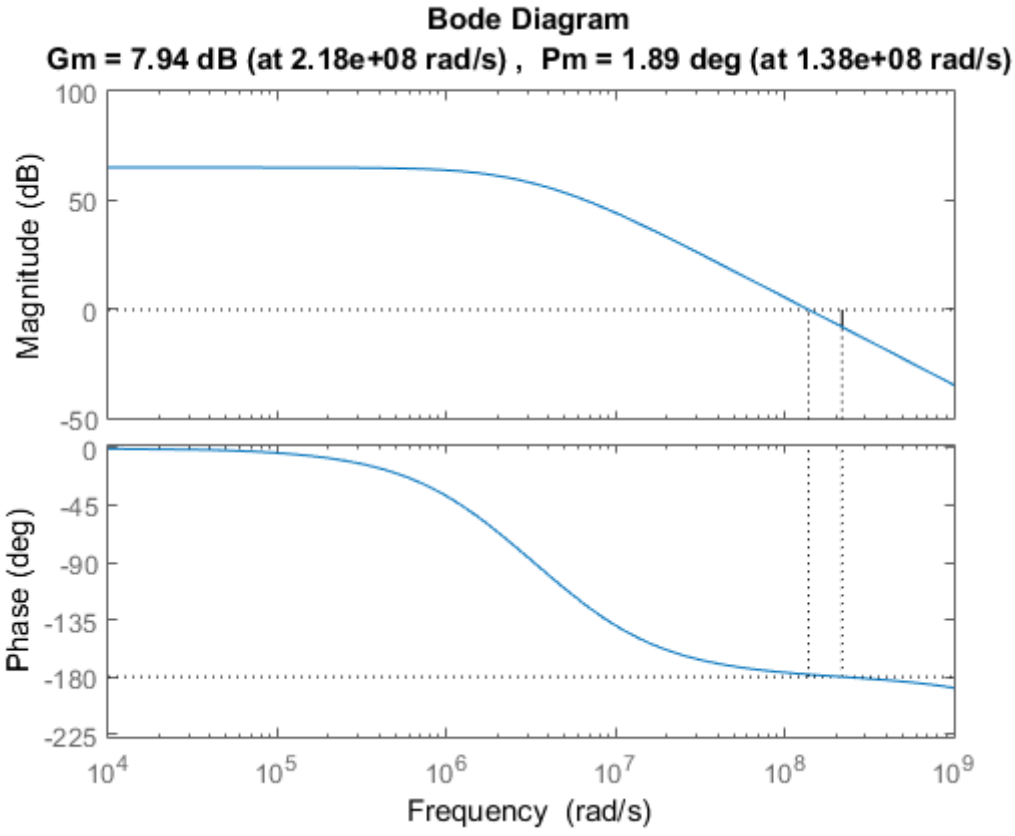


Figure 4-21 Bode diagram that shows the phase margin and gain margin of the uncompensated system

In a control system, phase margin and gain margin are the two terms often quoted to evaluate the system stability. The phase margin is defined as the phase difference to the -180deg when the gain magnitude drops to 0dB. The gain margin is defined as the distance of gain magnitude below 0dB when the phase drops to -180deg. The gain must be a negative value when phase reaches -180deg otherwise the system will oscillate. The general empirical rule of thumb for the safety design margin is over 45deg for the phase margin and over -20dB for the gain margin. Between the two, the phase

margin is more critical. For the uncompensated system shown in Figure 4-17(b), the phase margin (P_m) is only 1.89deg and gain margin (G_m) is only 7.94dB. Both are well below the safety design margin requirements.

Without enough margin, as system degrades with components aging or we encounter wider component tolerance, the system will drift into the unstable region. To improve the stability, a common method will be discussed in the next section.

4.4.3.2 Phase compensation

The method is called phase compensation. It is very simple, by just adding a capacitor C_f in parallel to the feedback resistor. This capacitor will introduce a zero to the circuit and boost the phase by 90°. Figure 4-22 shows the updated current amplifier circuit with the added feedback capacitor and the updated small signal circuit.

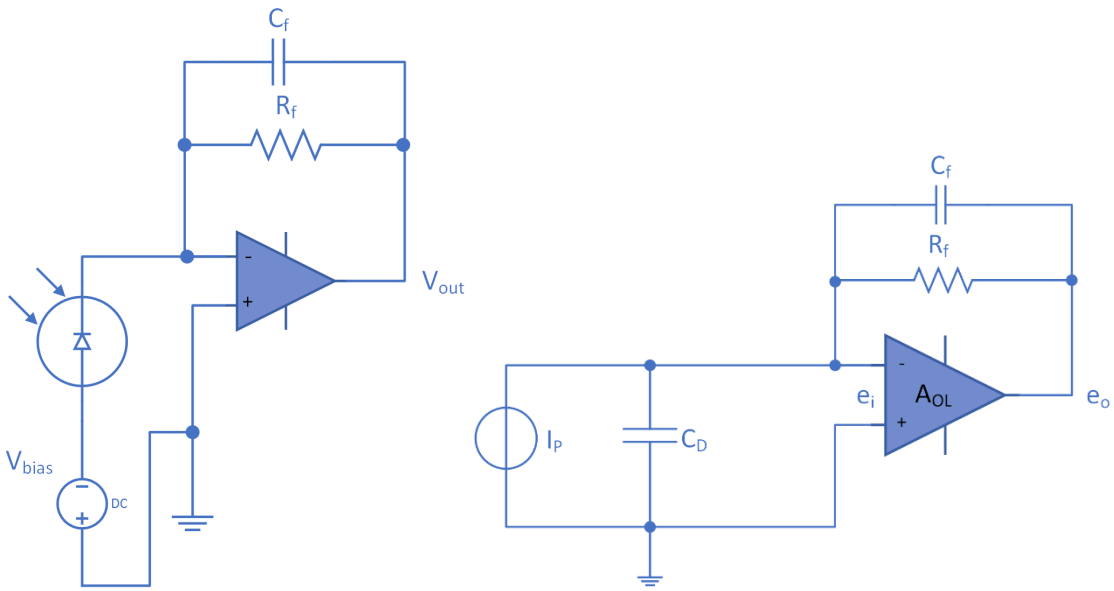


Figure 4-22 Improved circuits with the phase compensation feedback capacitor. Left is the full schematic and right is the small signal circuit.

The technique appears very simple by just using a single capacitor. However, it is important to understand the full role of this capacitor and its complication to the circuit bandwidth. The right choice of the value of this C_f is the result of the trade-off between stability and bandwidth of the circuit. To better understand the complication, the mathematic derivation of the circuit frequency response will be gone through.

The first step is to update the closed-loop gain equation (4.14). R_f is replaced with X_{RfCf}, which is the total impedance of the feedback network R_f and C_f. The same logic as the derivation of equation (4.18) is followed to derive the new **A_{CL}**.

Equation (4.13) can be updated to:

$$\begin{aligned} -e_i &= \frac{e_o}{A_{OL}} \\ i_p &= i_1 + i_2 \\ i_1 &= \frac{(e_o - e_i)}{X_{R_f C_f}} \\ i_2 &= -\frac{e_i}{X_{C_D}} \\ i_p &= \frac{(e_o - e_i)}{X_{R_f C_f}} + (-\frac{e_i}{X_{C_D}}) \end{aligned} \tag{4.22}$$

Similarly, cancel the **i_1**, **i_2** and **e_i**, then combine and rearrange the (4.22)conditions into:

$$e_o = \frac{i_p \cdot X_{R_f C_f}}{1 + \frac{1}{A_{OL}} \cdot (1 + \frac{X_{R_f C_f}}{X_{C_D}})} \tag{4.23}$$

The revised trans-impedance gain **A_{CL}** then is:

$$A_{CL\_C} = \frac{X_{R_f C_f}}{1 + \frac{1}{A_{OL}} \cdot (1 + \frac{X_{R_f C_f}}{X_{C_D}})} \tag{4.24}$$

Here a subscript **c** is added to differentiate it from the original A_{CL}.

The feedback factor can be updated to:

$$\beta_C = \frac{X_{C_D}}{X_{C_D} + X_{R_F C_F}} = \frac{1 + sR_f C_f}{1 + sR_f C_D + sR_f C_f} \tag{4.25}$$

Use β_C, equation (4.24) can be re-written as:

$$A_{CL\_C} = \frac{X_{R_f C_f}}{1 + \frac{1}{A_{OL}\beta_C}} \tag{4.26}$$

Again, plot the magnitude of the frequency response of **A_{OL}**, **$1/\beta_C$** and **A_{CL_c}** with the result shown in Figure 4-23. For the plot, the C_f is set at 1.5pF.

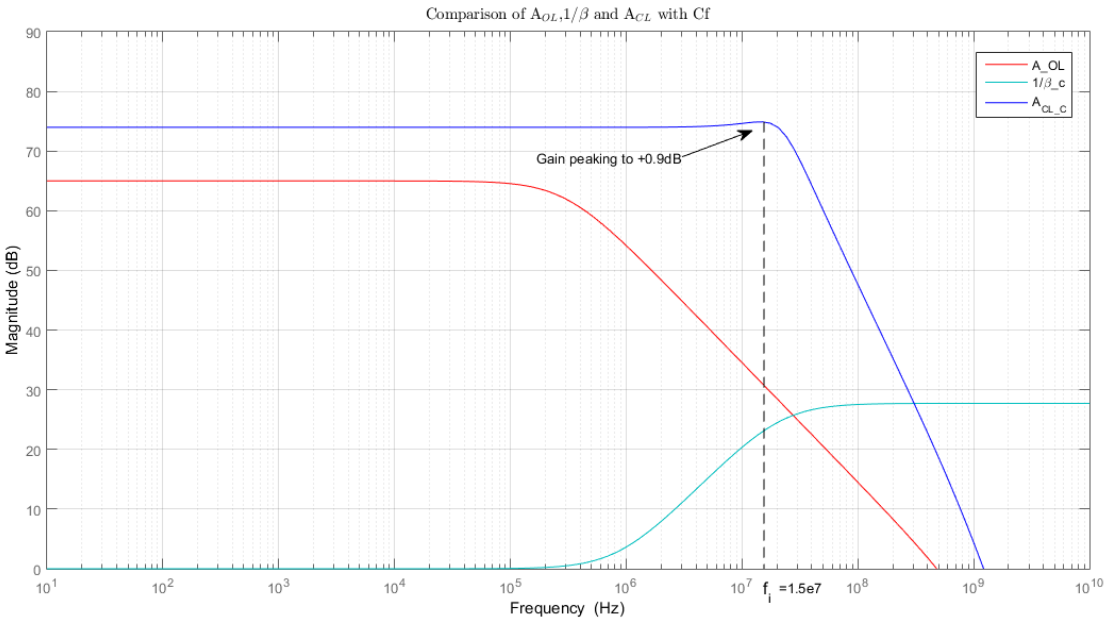


Figure 4-23 Frequency response plot of A_{OL}, $1/\beta_C$ and A_{CL_C}.

Compared to Figure 4-20, the gain peaking at $\mathbf{f_i}$ has been greatly attenuated. As a side effect, $\mathbf{f_i}$ has reduced slightly, which means the updated circuit has a narrower bandwidth. A replot the stability bode diagram of the loop gain **L_c** with *margin* function again with the result shown in Figure 4-24.

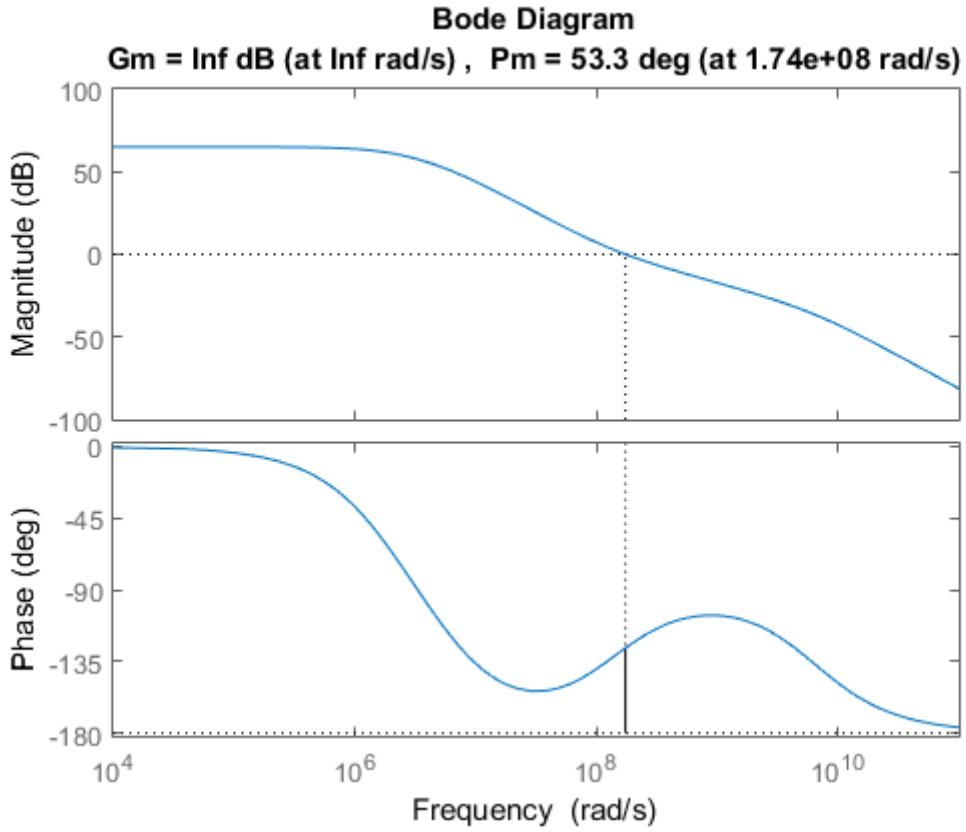


Figure 4-24 Bode diagram that shows the phase margin and gain margin of the compensated system

The improved phase margin is 53.3deg and the gain margin is Inf (it means the phase never reached -180deg). The simulation result is an ideal and may not fully reflect the real circuit performance. But clearly, it indicates the feedback capacitor makes a large improvement to the circuit stability margin.

Jerald Graeme [136] has a thorough discussion about how to work out the optimum C_f with the aim of setting the circuit phase margin at 45deg. Here, only the conclusion is presented in Eq. (4.27) and (4.28).

$$C_f = \sqrt{C_D C_c} \tag{4.27}$$

when $C_D >> C_f$ and where C_c is an artificial capacitance defined as $C_c = 1/2\pi R_f f_c$.

When the condition $C_D >> C_f$ failed, the more exact solution for C_f is

$$C_f = (C_c / 2)(1 + \sqrt{1 + \frac{4C_D}{C_c}}) \tag{4.28}$$

4.4.4 Noise model and analysis

Besides the circuit stability, circuit noise is another important issue that greatly affects the system performance. This section will discuss the major noise sources in the pre-amplifier circuit and the noise model that describes the circuit noise performance.

4.4.4.1 Common noise source

First, it is useful to understand the intrinsic noise source of a general-purpose op-amp. Figure 4-25 shows a classic op-amp noise model.

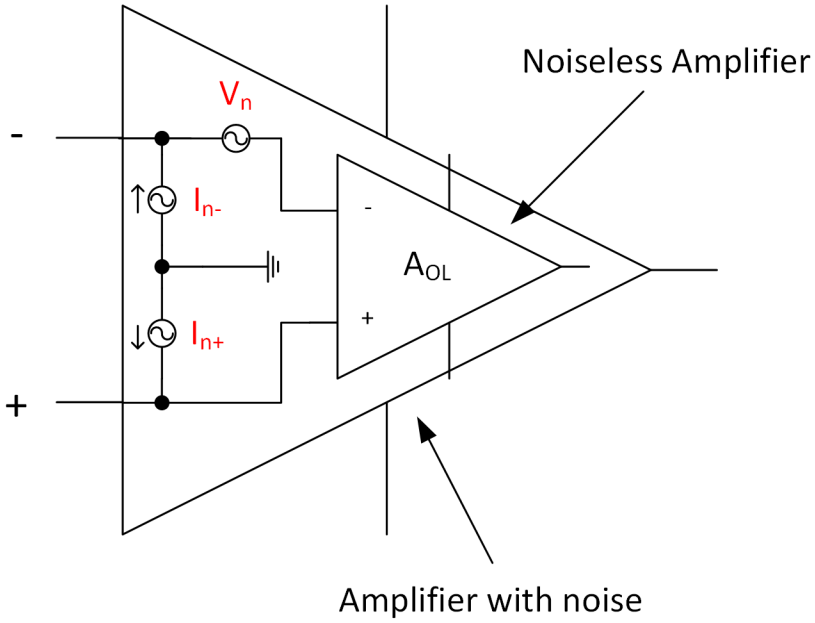


Figure 4-25 Classic operational amplifier noise model

Here **V_n** is the op-amp **input voltage noise**. It is a differential voltage across the two inputs. This noise is bandwidth dependant and measured in noise spectral density with the unit $nV/\sqrt{Hz}$. Bipolar op-amp (a type of op-amp uses bipolar junction transistor (BJT) as the first input stage) tends to have lower input voltage noise than JFET op-amp (another type of op-amp uses junction gate field-effect transistor (JFET) as the first input stage) in general. However, there are some exceptions from some specially designed op-amps.

I_{ni} is the op-amp **input current noise**. Each input has one flowing out of the device. This noise is similarly bandwidth dependant and measured as well in noise spectral density with the unit $pA/\sqrt{Hz}$ or $fA/\sqrt{Hz}$. As JFET has extremely low gate current (also called bias current), the JFET op-amp exhibits very low current noise at the level of $fA/\sqrt{Hz}$ while BJT op-amp's input current noise is much higher at the level of $pA/\sqrt{Hz}$. For simple op-amp with basic BJT or JEFT input stage, the current noise is just the shot noise of the bias current. It can be calculated with Eq. (4.29)

$$i_{ni} = \sqrt{2I_B q}/\sqrt{Hz} \tag{4.29}$$

In the voltage-feedback op-amp (Don't be confused with the voltage feedback preamplifier circuit discussed above. This voltage-feedback op-amp is a type of op-amp), the current noise in the two inputs is uncorrelated but roughly equal in magnitude.

In other more complicated op-amps with bias compensated input stages or in current-feedback op-amp (another type of op-amp), the current noise can't be calculated. Their current noise at the two inputs could be very different.

Both input voltage noise and input current noise have a 1/f frequency characteristic. This can be described by a general equation (4.30) and illustrated in Figure 4-26.

$$e_n, i_{ni} = k\sqrt{(\omega_f + s)/s} \tag{4.30}$$

Where, k is the noise floor, and ω_f represents the corner angular frequency.

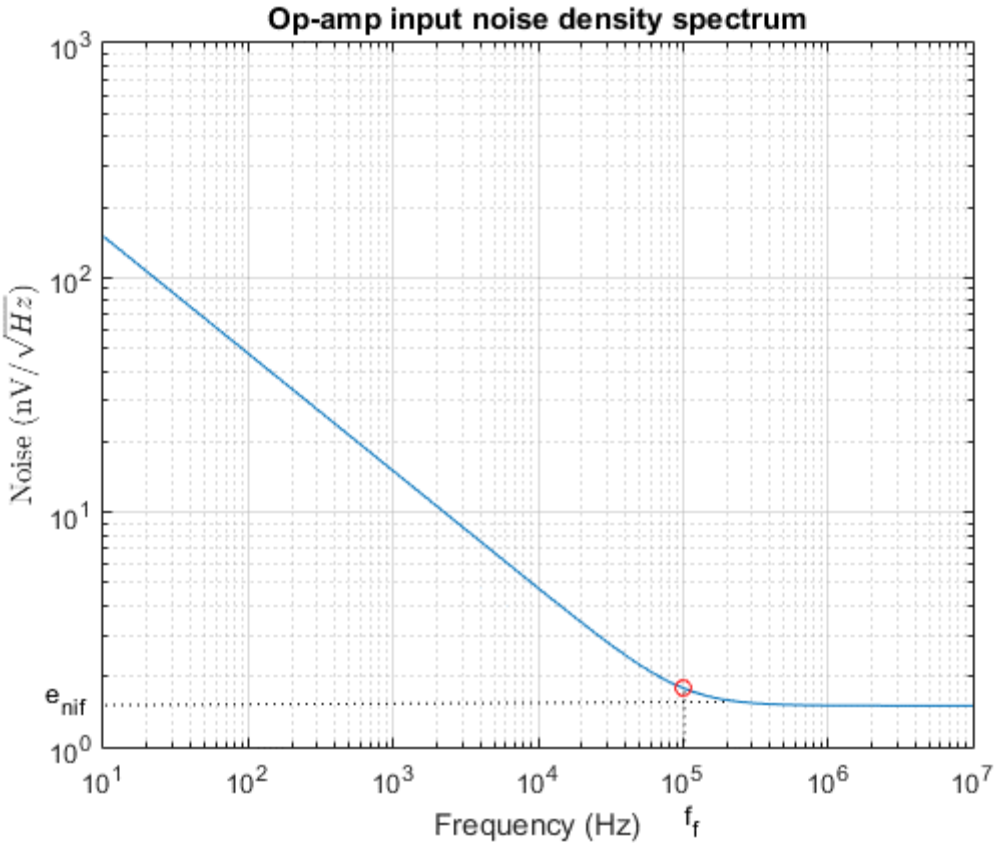


Figure 4-26 The 1/f characteristic of the op-amp input noise spectrum

Another common noise source is the thermal noise of the resistor. It is also called Johnson–Nyquist noise. It is related to the resistor value and the junction temperature. Equation (4.31) describes it clearly:

$$e_{nR} = \sqrt{4KTR_f} \tag{4.31}$$

Where K is the Boltzmann's constant (1.38×10–23J/K) and T is the absolute temperature in Kelvin.

4.4.4.2 The current amplifier circuit's noise model

After identifying the common noise source that is related to the preamplifier circuit, we can construct a noise model for the current amplifier circuit as shown in Figure 4-27.

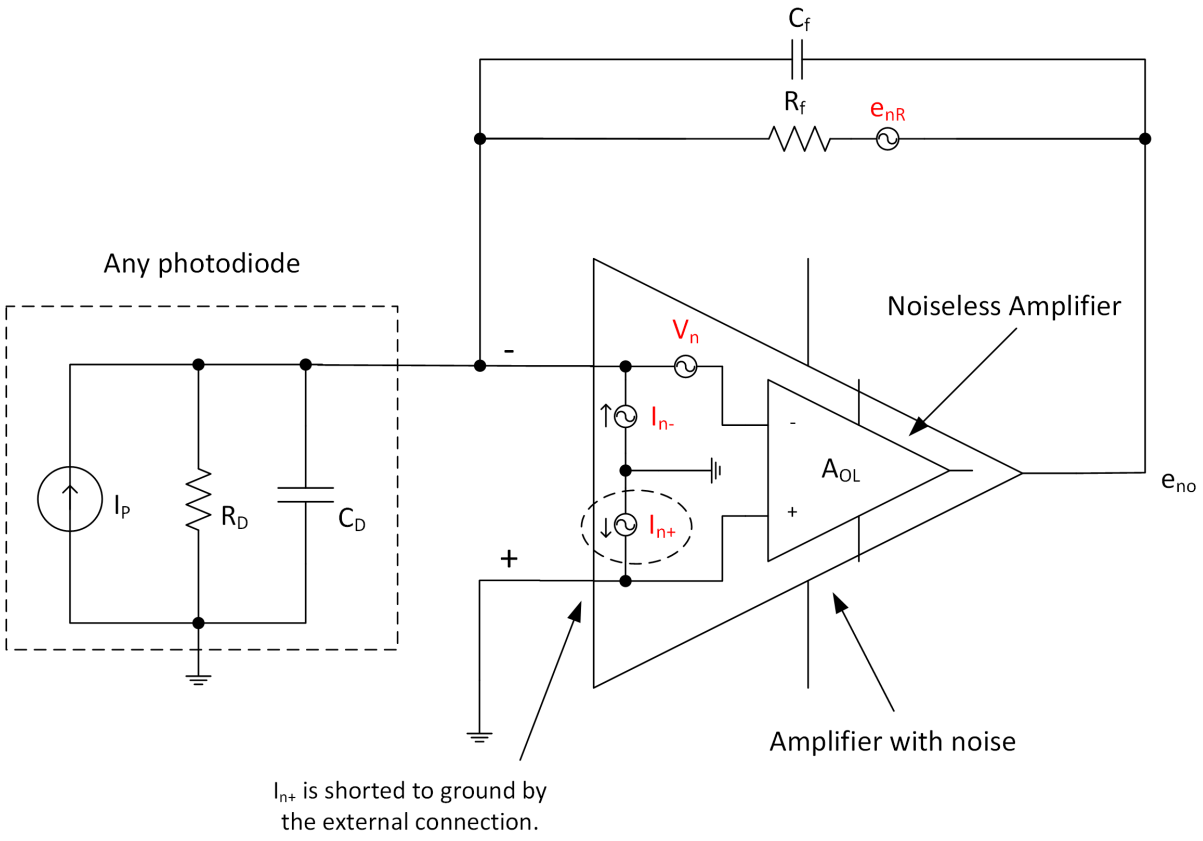


Figure 4-27 The noise model of the current amplifier circuit

Note, as the non-inverting input is connected to ground externally. Its associated current noise is shorted, so its noise can be skipped from the analysis.

For a linear circuit with multiple uncorrelated noise sources, superposition theory can be applied. The method is to analyse each noise source contribution individually and combine all the contributions by root square sum to form the total noise contribution.

Under superposition, the analysis starts with one noise source with all the other sources removed following the principle that the voltage source is replaced with a short circuit and the current source is replaced with an open circuit. From the simplified circuit, the contribution to the output signal from the one noise source can be worked out. The steps are repeated until all the noise sources are processed. Lastly, calculate the root square sum of all the contributions to get the total noise contribution.

The noise source is generally presented in spectral density. To get the output RMS noise, we need to evaluate the integral:

$$E_{no}^2 = \int_0^\infty \left| A_n e_n \right|^2 df \tag{4.32}$$

where, an is the gain of the circuit for the noise under study, e_n is the noise spectral density and E_{no} is the output RMS noise for the given bandwidth or full spectrum if we integrate from 0 to +∞.

With all the theory preparation ready, below is the application of the noise analysis on our current amplifier circuit.

4.4.4.2.1 The feedback resistor thermal noise: e_{nR}

This is the noise related to the feedback resistor. Since, for the photodiode trans-impedance circuit, this resistance provides the gain of converting the current signal to voltage signal, a high resistance is generally desired to achieve high gain. The side effect is the increased thermal noise associated with high resistance. Therefore, the choice of this feedback resistor is a balance between sufficient gain and reasonable noise.

Using the method from the superposition theory, we can update the noise model to:

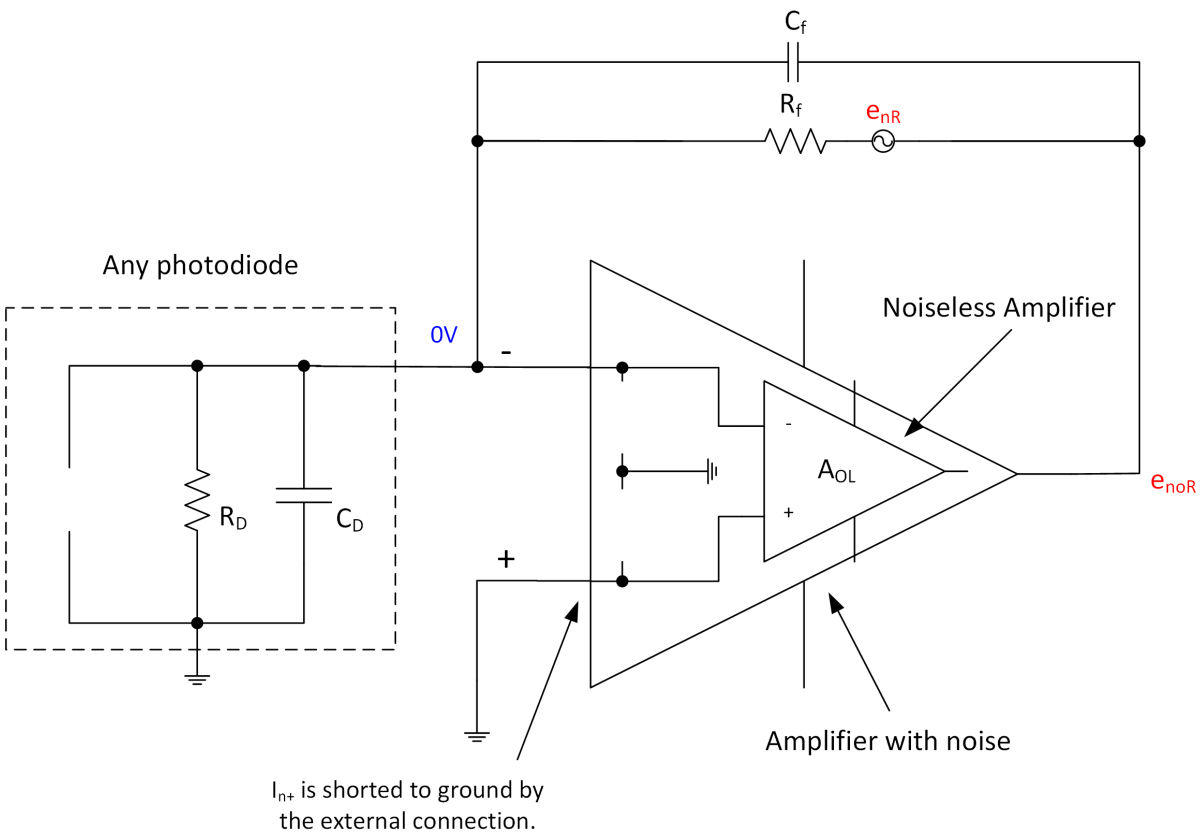


Figure 4-28 The noise model for e_{nR} noise source

Since the inverting input is held at virtual ground, we can get

$$e_{noR} = e_{nR} \tag{4.33}$$

Further with Eq. (4.31),

$$e_{noR} = e_{nR} = \sqrt{4KTR_f} \tag{4.34}$$

Then exercise the integral (4.32) to calculate its output RMS noise. As thermal noise is frequency independent “white” noise, the circuit bandwidth (BW_t) is the up limit of the integral. For this application, BW_t is limited by the pole created by C_f and R_f. Finally, E_{noR} is worked out as:

$$E_{noR} = \sqrt{2KTR_f \pi BW_t} \tag{4.35}$$

where $BW_t = 1/2\pi R_f C_f$

4.4.4.2.2 The op-amp input current noise: i_{ni}

The op-amp input current noise converts to an output voltage noise through the feedback network of C_f and R_f, see Figure 4-29. If we define the total impedance of the network as Z_f, the output voltage noise e_{noi} can be described as:

$$e_{noi} = Z_f i_{ni} \tag{4.36}$$

where $Z_f = R_f // C_f = R_f / (1 + sR_f C_f)$.

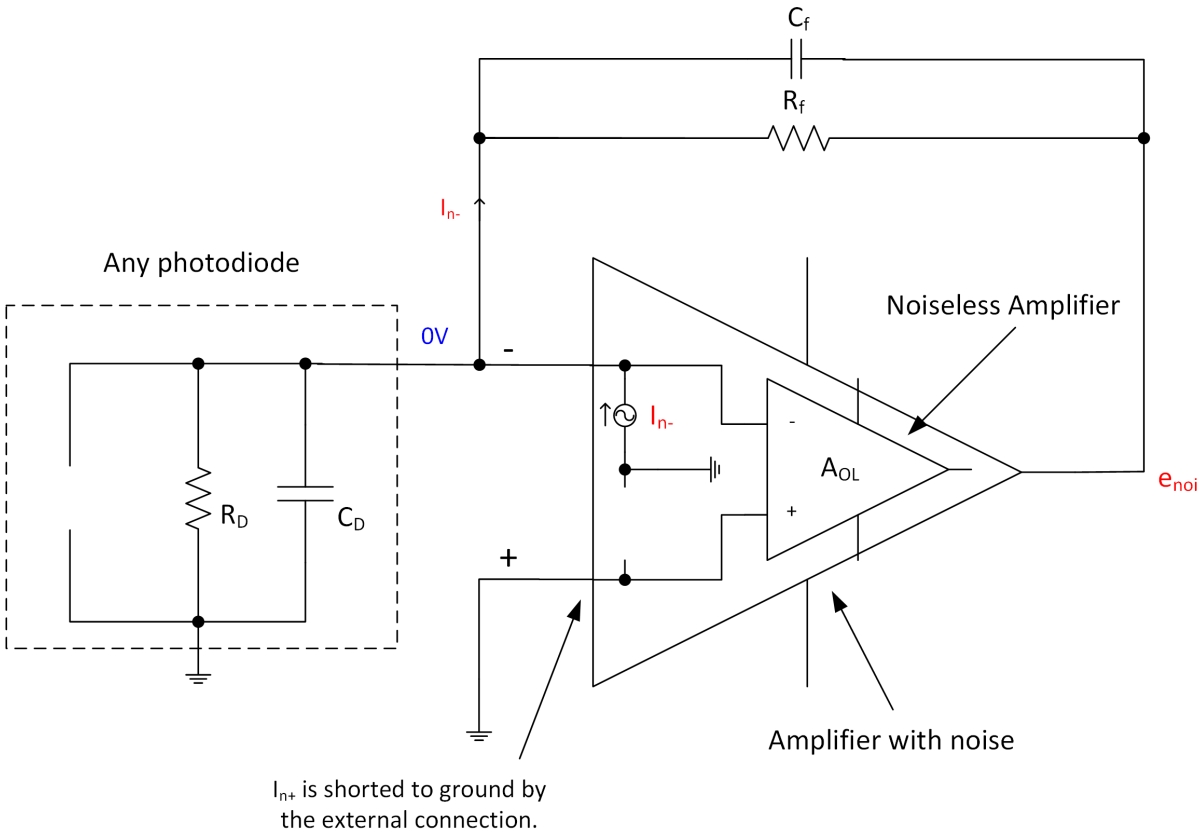


Figure 4-29 The noise model for i_{ni} noise source

Then update Eq. (4.36) with the definition of Z_f and i_{ni} (4.30),

$$e_{noi} = \frac{R_f}{(1 + sR_f C_f)} \cdot k_i \sqrt{\frac{(\omega_f + s)}{s}} \tag{4.37}$$

Work out its magnitude expression as required by the integration (4.32),

$$
\begin{aligned}
\left|e_{noi}\right| &= \left|\frac{R_f}{(1+sR_fC_f)}\cdot k_i\sqrt{\frac{(\omega_f+s)}{s}}\right| \\
&= \left|\frac{R_f}{(1+sR_fC_f)}\right|\left|k_i\sqrt{\frac{(\omega_f+s)}{s}}\right| \\
&= R_f\sqrt{\frac{1}{1+(2\pi fR_fC_f)^2}}\cdot k_i\sqrt{1+(\frac{f_f}{f})^2}
\end{aligned}
\tag{4.38}
$$

Unfortunately, this equation is too complex for the direct integration to get an analytic solution. The approach is to use numerical technique to approximate it. MATLAB is the ideal tool to do this. From the plot of the approximate frequency spectrum (Figure 4-30), we can get a quick view of the shape of the spectrum of **e_{noi}** in comparison to **i_{ni}** and **Z_f**.

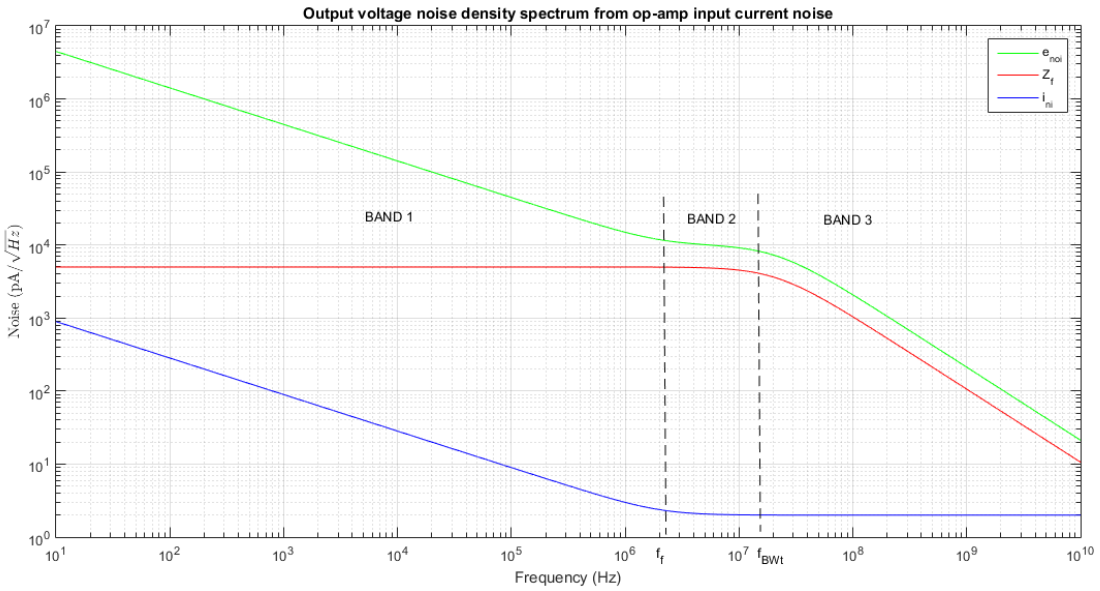


Figure 4-30 The MATLAB plots of the e_{noi} , Z_f and i_{ni} frequency response spectrum

From the spectrum plot, it is noted the curve of **e_{noi}** is modulated by a few basic curves (in this case, i_{ni} and Z_f). This suggests a way to simplify the complex curve. That is to break the curve into a few frequency bands. At each band, use the dominant basic curve to represent the final curve.

The curve of **e_{noi}** can be divided into three frequency bands. Each band is treated separately in detail.

In band 1, where $f<f_f$, the **i_{ni}**'s 1/f characteristic is dominant and the **Z_f** is at the flat region. So **e_{noi}** can be simplified to

$$\left|e_{noi}\right| = R_f \cdot k_i \sqrt{\frac{f_f}{f}} \tag{4.39}$$

In band 2, where f>f_f and f<f_{BWt}, $\mathbf{i_{ni}}$ reaches its flat plateau noise floor and $\mathbf{Z_f}$ is still at its flat region. Then $\mathbf{e_{noi}}$ can be simplified to

$$\left|e_{noi}\right| = R_f k_i \tag{4.40}$$

In band 3, where f>f_{BWt}, $\mathbf{i_{ni}}$ stays in the flat noise floor and $\mathbf{Z_f}$ starts rolling off at 20dB/decade. Then $\mathbf{e_{noi}}$ can be simplified to:

$$\left|e_{noi}\right| = R_f \sqrt{\frac{1}{1+(2\pi f R_f C_f)^2}} \cdot k_i \approx R_f k_i \cdot \frac{f_{BW_t}}{f} \tag{4.41}$$

Then do the integration with A_n = 1.

In band 1,

$$\begin{aligned} {E_{noi1}}^2 &= \int_0^\infty \left|A_n \cdot e_{noi1}\right|^2 df \\ &= \int_{0.01}^{f_f} (R_f \cdot k_i \sqrt{\frac{f_f}{f}})^2 df \\ &= \int_{0.01}^{f_f} (R_f k_i)^2 (\frac{f_f}{f}) df \\ &= (R_f k_i)^2 f_f (\ln(f_f) - \ln(0.01)) \end{aligned} \tag{4.42}$$

Here, the integration starts from 0.01Hz to avoid 0 as ln(0) is infinite. This is a theoretically unsolved problem in 1/f noise.

In band 2,

$$\begin{aligned} {E_{noi2}}^2 &= \int_{f_f}^{f_{BW_t}} (R_f \cdot k_i)^2 df \\ &= (R_f k_i)^2 (f_{BW_t} - f_f) \end{aligned} \tag{4.43}$$

In band 3,

$$
\begin{aligned}
E_{noi3}{}^2 &= \int_{f_{BW_t}}^{\infty} (R_f k_i \cdot \frac{f_{BW_t}}{f})^2 df \\
&= (R_f k_i f_{BW_t})^2 (-\frac{1}{\infty} + \frac{1}{f_{BW_t}}) \\
&= (R_f k_i f_{BW_t})^2 (\frac{1}{f_{BW_t}}) \\
&= (R_f k_i)^2 f_{BW_t}
\end{aligned}
\tag{4.44}
$$

Finally, E_{noi} is acquired by combining result (4.42), (4.43) and (4.44) in root square sum as:

$$
E_{noi} = \sqrt{E_{noi1}^2 + E_{noi2}^2 + E_{noi3}^2}
\tag{4.45}
$$

4.4.4.2.3 The op-amp input voltage noise: e_{nv}

e_{nv} is the op-amp input voltage noise. Using superposition, the noise model is updated to Figure 4-31.

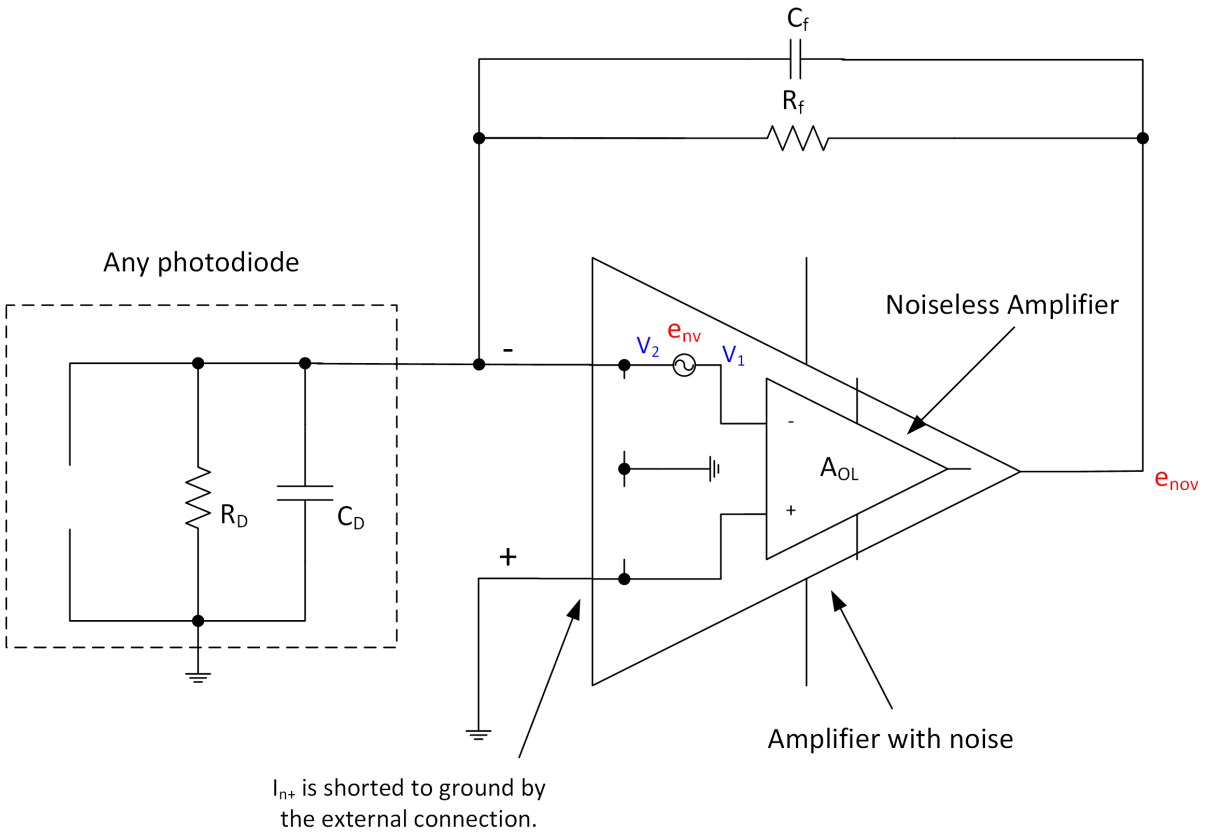


*Figure 4-31 The noise model for **e_{nv}** noise source*

Similarly, the first step is to work out the theoretical solution for **e_{nov}**. Due to the existence of **e_{nv}**, the inverting input is not at “virtual ground” potential any more. So, the effect of the photodiode input capacitance and resistance has to be considered. Using Kirchhoff’s circuit law, we can construct three equations:

$$e_{nv} = v_2 - v_1 \tag{4.46}$$

$$e_{nov} = (0 - v_1) \cdot A_{OL} = -v_1 A_{OL} \tag{4.47}$$

$$v_2 = e_{nov} \cdot \frac{z_D}{z_f + z_D} \tag{4.48}$$

where A_{OL} is the op-amp open loop gain, Z_D is the total impedance of the photodiode (Rd//Cd) and Z_f is the total impedance of the feedback loop (Rf//Cf).

After combining the three equations by removing v_2 and v_1, the relationship between **e_{nov}** and **e_{nv}** is derived as:

$$\begin{aligned} e_{nov} &= e_{nv} \cdot \frac{A_{OL}}{1 + A_{OL}\beta} \\ &= e_{nv} \cdot \frac{1}{\beta} \cdot \left(\frac{1}{1 + \frac{1}{A_{OL}\beta}}\right) \end{aligned} \tag{4.49}$$

Where β is the feedback factor. For the updated circuit as Figure 4-31, its reciprocal can be worked out as:

$$\begin{aligned} \frac{1}{\beta} &= \frac{z_D + z_f}{z_D} = 1 + \frac{z_f}{z_D} \\ &= 1 + \frac{R_f(1 + sR_D C_D)}{R_D(1 + sR_f C_f)} \end{aligned} \tag{4.50}$$

From Eq. (4.49), A_n can be defined as:

$$A_n = \frac{1}{\beta} \cdot \left(\frac{1}{1 + \frac{1}{A_{OL}\beta}}\right) \tag{4.51}$$

The reason to write A_n in this way is to allow the reader to compare it to Eq. (4.18). The two are very similar.

Eq. (4.30) can be rewritten for e_{nv} as:

$$e_{nv} = k_v \sqrt{\frac{(\omega_f + s)}{s}} \tag{4.52}$$

Here ω_f represents the corner angular frequency in e_{nv} noise spectrum. Note it has different value to the corner angular frequency in i_{ni} expression, but the same symbol is used in Eq. (4.30) to unify the formula.

With $\mathbf{A_n}$ and $\mathbf{e_{nv}}$ both defined by Eq. (4.51) and (4.52), the integral (4.32) can be called again. The result is the final equation for $\mathbf{E_{nov}}$

$$E_{nov} = \int_0^{\infty} \left| \frac{1}{\beta} \cdot \left(\frac{1}{1 + \frac{1}{A_{OL}\beta}} \right) \cdot k_v \sqrt{\frac{(\omega_f + s)}{s}} \right|^2 df \tag{4.53}$$

Now this again is very hard to solve analytically, so similar approach used for $\mathbf{E_{noi}}$ is taken. Figure 4-32 is the spectrum of the approximated frequency response calculated by MATLAB.

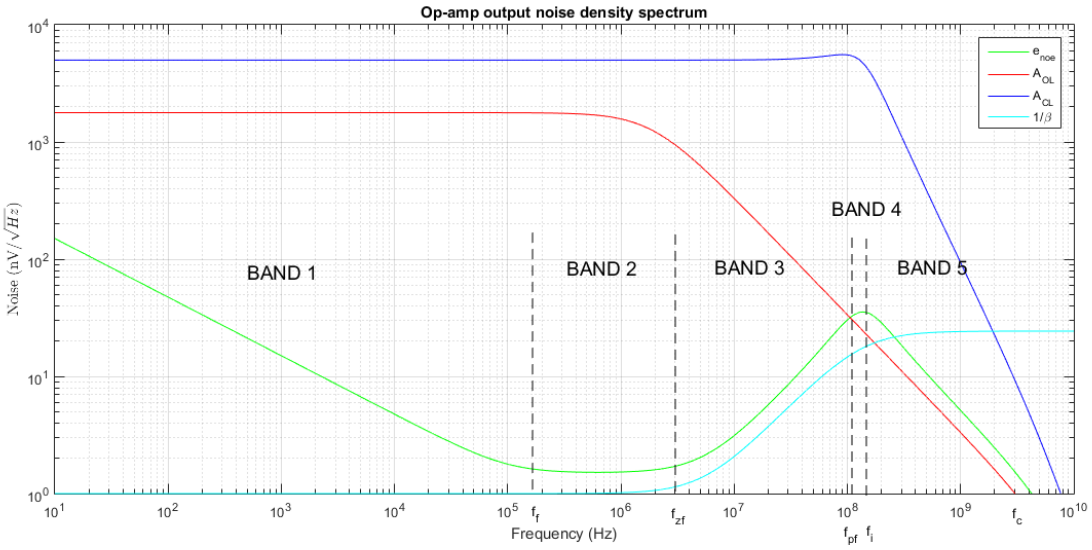


Figure 4-32 The MATLAB approximation of the e_{nov} , A_{CL}, A_{OL} and 1/β frequency response spectrum

In Figure 4-32, the green line is $\mathbf{e_{nov}}$ frequency response. It is shaped by the 1/f characteristic at the initial low frequency (band 1), then it stays constant at k_v noise floor (band 2), then it follows the 1/β curve (band 3), then it is plateaued shortly as 1/β curve is plateaued due to the feedback capacitor C_f (band 4) and lastly it is rolling off together with the A_{OL} (band 5).

As the plot suggests, $\mathbf{e_{nov}}$ curve can be divided into five frequency bands. In each band, a simpler and dominant curve can be used to approximate the $\mathbf{e_{nov}}$ curve.

In band 1, for frequency up to f_f, since 1/β = 1 and A_{OL}>>1, it is reasonable to assume $A_n \approx 1$.

As f << f_f, e_{nv} magnitude can be closely approximated as:

$$\begin{aligned} |e_{nv}| &= k_v \sqrt[4]{1+(\frac{f_f}{f})^2} \\ &\approx k_v \sqrt[4]{(\frac{f_f}{f})^2} \\ &= k_v \sqrt{\frac{f_f}{f}} \end{aligned} \tag{4.54}$$

Then, run the integration to get the output noise in rms for band 1 as e_{nov1}:

(note: below section quotes e_{nov2} for band 2, e_{nov3} for band 3, etc.)

$$\begin{aligned} E_{nov1}^2 &= \int_0^{f_f} |A_n e_{nv}|^2 \, df \\ &= \int_0^{f_f} |e_{nv}|^2 \, df \\ &= \int_0^{f_f} (k_v \sqrt{\frac{f_f}{f}})^2 \, df \\ &= k_v^2 \int_0^{f_f} \frac{f_f}{f} df \\ &= k_v^2 f_f \ln(\frac{f_f}{0}) \end{aligned} \tag{4.55}$$

Again, there is a problem of dividing by zero. The practical solution is to use a very low frequency f_1 = 0.01 to replace the zero. Hence, the final equation is derived as

$$E_{nov1}^2 = k_v^2 f_f \ln(\frac{f_f}{f_1}) \tag{4.56}$$

In band 2, for frequency between f_f and f_{zf}, e_{nv} falls to its floor k_v. The 1/β and A_{OL} hasn't changed, so A_n is still close to 1. Then E_{nov2} is:

$$\begin{aligned} E_{nov2}^2 &= \int_{f_f}^{f_{zf}} \left| A_n e_{nv} \right|^2 df \\ &= \int_{f_f}^{f_{zf}} \left| k_v \right|^2 df \\ &= k_v^2 (f_{zf} - f_f) \end{aligned} \tag{4.57}$$

where $f_{zf} = 1/2\pi R_f (C_D + C_f)$

In band 3, for frequency between f_{zf} and f_{pf}, e_{nv} remains at floor k_v. 1/β curve now becomes dominant and A_{OL} is still flat, so A_n is revised to $A_n \approx 1/\beta$. E_{nov3} is derived in detail below:

$$\begin{aligned} E_{nov3}^2 &= \int_{f_{zf}}^{f_{pf}} \left| A_n e_{nv} \right|^2 df \\ &= \int_{f_{zf}}^{f_{pf}} \left| \frac{1}{\beta} \right|^2 \left| k_v \right|^2 df \\ &= \left| k_v \right|^2 \int_{f_{zf}}^{f_{pf}} \left| 1 + \frac{R_f}{R_D} \left(\frac{1 + sR_D C_D}{1 + sR_f C_f} \right) \right|^2 df \end{aligned} \tag{4.58}$$

Rewrite Eq.(4.58) using f_{zf} and f_{pf}, where $f_{zf} = 1/2\pi R_f (C_D + C_f)$ and $f_{pf} = 1/2\pi R_f C_f$

$$E_{nov3}^2 = \left| k_v \right|^2 \int_{f_{zf}}^{f_{pf}} \left| \frac{(1 + {R_f}/{R_D}) + {s}/{2\pi f_{zf}}}{1 + {s}/{2\pi f_{pf}}} \right|^2 df \tag{4.59}$$

Since $R_D >> R_f$, $f << f_{pf}$ and $f >> f_{zf}$, Eq. (4.59) can be further simplified to

$$\begin{aligned} E_{nov3}^2 &\approx \left| k_v \right|^2 \int_{f_{zf}}^{f_{pf}} \left| \frac{f}{f_{zf}} \right|^2 df \\ &= \frac{k_v^2}{f_{zf}^2} \frac{(f_{pf}^3 - f_{zf}^3)}{3} \end{aligned} \tag{4.60}$$

In band 4, for frequency between f_{pf} and f_i (where 1/β intercepts to A_{OL}), A_n continues to follow the 1/β curve, which reaches a plateau. e_{nv} remains at its floor k_v. E_{nov4} can be derived similar to E_{nov3}:

$$
\begin{aligned}
E_{nov4}^2 &= \int_{f_{pf}}^{f_i} \left| A_n e_{nv} \right|^2 df \\
&= \int_{f_{pf}}^{f_i} \left| \frac{1}{\beta} \right|^2 \left| k_v \right|^2 df \\
&= \left| k_v \right|^2 \int_{f_{pf}}^{f_i} \left| \frac{(1 + R_f / R_D) + s / 2\pi f_{zf}}{1 + s / 2\pi f_{pf}} \right|^2 df
\end{aligned}
\tag{4.61}
$$

As f>>f_{pf}, R_f>>R_D and f>>f_{zf}, Eq. (4.61) can further be simplified to

$$
\begin{aligned}
E_{nov4}^2 &\approx \left| k_v \right|^2 \int_{f_{pf}}^{f_i} \left| \frac{s / 2\pi f_{zf}}{s / 2\pi f_{pf}} \right|^2 df \\
&= \left| k_v \right|^2 \int_{f_{pf}}^{f_i} \left| \frac{f_{pf}}{f_{zf}} \right|^2 df \\
&= \left| k_v \right|^2 \int_{f_{pf}}^{f_i} \left| \frac{1 / 2\pi R_f C_f}{1 / 2\pi R_f (C_D + C_f)} \right|^2 df \\
&= \left| k_v \right|^2 \int_{f_{pf}}^{f_i} \left| \frac{(C_D + C_f)}{C_f} \right|^2 df \\
&= \left| k_v \right|^2 \left(\frac{C_D + C_f}{C_f} \right)^2 (f_i - f_{pf})
\end{aligned}
\tag{4.62}
$$

where $f_i = f_c C_f / (C_D + C_f)$ and $f_{pf} = 1 / 2\pi R_f C_f$

In band 5, for frequency between f_i and f_c, A_n starts rolling off with the A_{OL}. e_{nv} remains at the floor k_v. For A_{OL} up to f_c, there is only one pole, after which A_{OL} starts rolling off at 20dB/decade until f_c, which is the op-amp unit-gain crossover frequency. So, the slope part of the A_{OL} can be presented as:

$$
\left| A_{OL} \right| = \frac{f_c}{f} \tag{4.63}
$$

Now we do the same integral for E_{nov5}:

$$\begin{aligned} E_{nov5}^2 &= \int_{f_i}^{f_c} \left| A_n e_{nv} \right|^2 df \\ &= \int_{f_i}^{f_c} \left| \frac{f_c}{f} \right|^2 \left| k_v \right|^2 df \\ &= \left| k_v \right|^2 \left| f_c \right|^2 \left(-\frac{1}{f} \right) \Big|_{f_i}^{f_c} \\ &= \left| k_v \right|^2 \left| f_c \right|^2 \left(\frac{1}{f_i} - \frac{1}{f_c} \right) \end{aligned} \tag{4.64}$$

Finally, we have worked out all the five bands of E_{nov}. The root square sum of them can provide the total E_{nov} as:

$$E_{nov} = \sqrt{E_{nov1}^2 + E_{nov2}^2 + E_{nov3}^2 + E_{nov4}^2 + E_{nov5}^2} \tag{4.65}$$

4.4.4.2.4 The total preamplifier circuit noise

After establishing individual noise source's contribution to the output noise and given they are independent to each other, their contributions can be combined with root square sum to get the total output noise as

$$E_{no} = \sqrt{E_{noR}^2 + E_{noi}^2 + E_{noe}^2} \tag{4.66}$$

4.4.4.3 Other simplified noise model for Transimpedance Amplifiers (TIA)

A technical note [137] from TI summarized one simpler formula for estimating the circuit noise for the application of TIA with photodiode. Figure 4-33 shows the noise model with the op-amp input voltage noise and feedback resistor thermal noise both converted into the input current noise. Combined with the op-amp input current noise, all three noise sources merge into one single equivalent input current noise $\mathbf{i_{ni}}$.

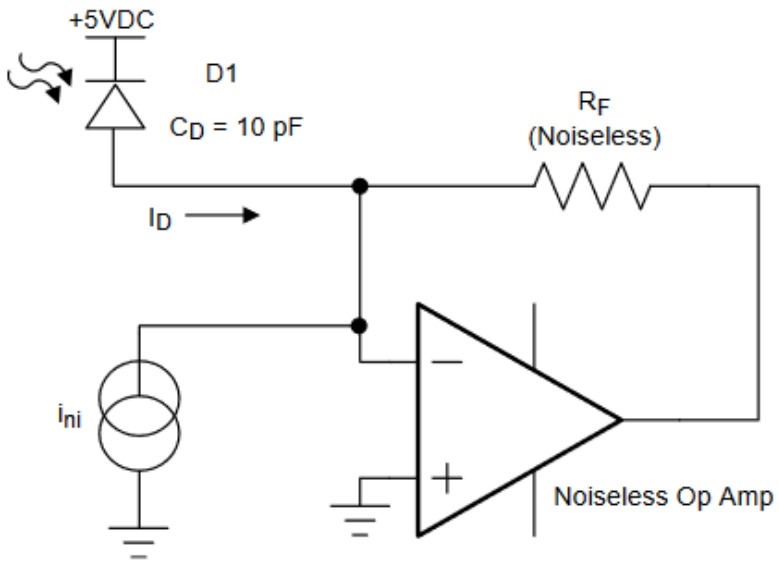


Figure 4-33 The noise model for the TIA with photodiode

The expression for i_{ni} is defined as:

$$i_{ni} = \sqrt{i_n^2 + (\frac{e_n}{R_F})^2 + \frac{4kT}{R_F} + \frac{(e_n 2\pi f_{-3dB} C_{IN})^2}{3}} \tag{4.67}$$

Where:

i_n= inverting input spot current noise

$4kT = 16.4 \times 10^{-21}$J at room temperature

R_F= feedback resistor

e_n= non-inverting input spot voltage noise

C_{IN}= Total inverting input total capacitance

f_{-3dB}= noise integration frequency limit

We haven't compared the effectiveness of this simplified noise model to our comprehensive noise model. This can be left for future study.

4.4.5 Time domain signal analysis

The frequency domain analysis is very useful to study the circuit stability and noise performance. The frequency domain analysis has wider coverage, and it is commonly used as the standard approach for checking circuit stability. However, all the circuit signal observed in real life are presented in the time domain. For example, the oscilloscope capture of the preamplifier output signal is in the time domain. Therefore, it

has practical benefit to carry out time domain signal analysis. The analytic output signal waveform can be directly checked against the oscilloscope capture from a real circuit.

The time domain analysis is generally more complicated, but it provides new insight of the circuit performance. For charge preamplifier circuit, the impact of the input capacitance on the output signal is particularly of interest. And this can be revealed by time domain analysis.

First is to re-arrange the current feedback amplifier circuit as in Figure 4-34.

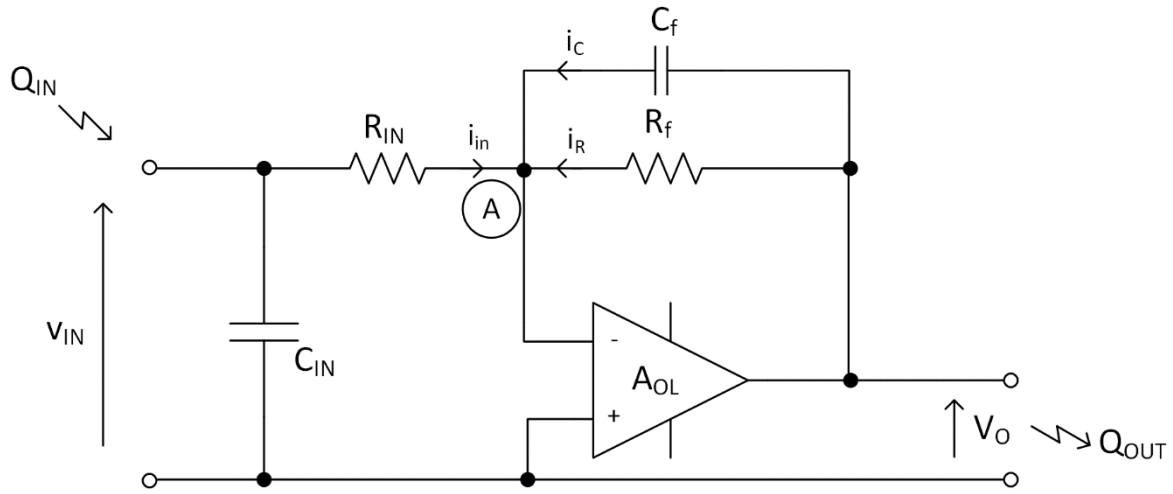


Figure 4-34 The simplified current feedback circuit for the time-domain analysis

Here the output of the SiPM is considered as a group of charge Q_{IN} (electrons cloud in very short duration). C_{IN} is the combination of SiPM capacitance and the line parasitic capacitance. R_{IN} is the combination of line resistance and any part of the series resistance inside the SiPM that is limiting the discharge current i_{IN}. The C_{IN} and R_{IN} can be linked to the parameters defined in 4.2, but it is not very straightforward.

The physical process of the event can be described as: once avalanche occurred, the group of charge Q_{IN} is released into the input capacitance C_{IN} in a very short period. Then C_{IN} is discharged through R_{IN} at the rate of i_{IN}. This i_{IN} further charges up C_f. As C_f charges up, V_o rises. At last, the charge distribution between C_{IN} and C_f will reach a balance, when V_o reaches the peak. It is important to understand C_{IN} will retain some charge and only the charge being transferred to C_f will contribute to the preamplifier output signal. The maximum possible charge transferred to C_f is half of the total charge and this occurs when C_f is equal to C_{IN}. After reaching V_o peak, the discharging of C_f through R_f becomes dominant and as a result V_o starts falling.

Point A in Figure 4-34 is the virtual ground. Use Kirchhoff's law at point A:

$$i_{IN} + i_C + i_R = 0 \tag{4.68}$$

Where, i_{IN} is the discharge current from C_{IN}, i_C is the current through C_f and i_R is the current through R_f. Note these three currents are all dynamic in time, but their sum at point A must be zero at any time.

i_R can be described as:

$$i_R = \frac{v_O}{R_f} = \frac{q_O}{C_f R_f} = \frac{q_O}{\tau_f} \tag{4.69}$$

i_C can be described as:

$$i_C = \frac{dq_O}{dt} \tag{4.70}$$

i_{IN} can be described as the discharge current of a capacitor through a resistor:

$$i_{IN} = \frac{Q_{IN}}{\tau_{IN}} e^{-t/\tau_{IN}} \tag{4.71}$$

Where $\tau_{IN} = R_{IN} C_{IN}$ and $\tau_f = R_f C_f$

Substitute (4.69),(4.70) and (4.71) into (4.68) and rearrange:

$$\frac{dq_O}{dt} + \frac{q_O}{\tau_f} = -\frac{Q_{IN}}{\tau_{IN}} e^{-t/\tau_{IN}} \tag{4.72}$$

This is a first order linear differential equation of form:

$$\frac{dy}{dx} + py = q \tag{4.73}$$

The solution is found by multiplying an integrating factor $I(x) = e^{\int p(x)dx}$ at both side of the equation, and then the solution is derived as:

$$y = \frac{\int I(x) q(x) dx + C}{I(x)} \tag{4.74}$$

From (4.72), the integrating factor $I(t)$ is worked out as

$$I(t) = e^{\int 1/\tau_f dt} = e^{t/\tau_f} \tag{4.75}$$

Multiply the $I(t)$ across (4.72), and follow the standard process of solving the first order linear differential equation, we can derive:

$$e^{t/\tau_f}\cdot\frac{dq_o}{dt}+e^{t/\tau_f}\cdot\frac{q_o}{\tau_f}=-\frac{Q_{IN}}{\tau_{IN}}e^{-t/\tau_{IN}}\cdot e^{t/\tau_f}$$

$$\int_0^t (e^{t/\tau_f}\cdot q_o)'dt=\int_0^t -\frac{Q_{IN}}{\tau_{IN}}e^{-t/\tau_{IN}}\cdot e^{t/\tau_f}dt$$

$$e^{t/\tau_f}\cdot q_o+k=-\frac{Q_{IN}}{\tau_{IN}}\cdot\frac{1}{(1/\tau_f-1/\tau_{IN})}\left[e^{t\cdot(1/\tau_f-1/\tau_{IN})}\right]_0^t$$

$$e^{t/\tau_f}\cdot q_o+k=-\frac{Q_{IN}\cdot\tau_f}{\tau_{IN}-\tau_f}\cdot\left[e^{t(\frac{\tau_{IN}-\tau_f}{\tau_{IN}\tau_f})}-1\right]$$

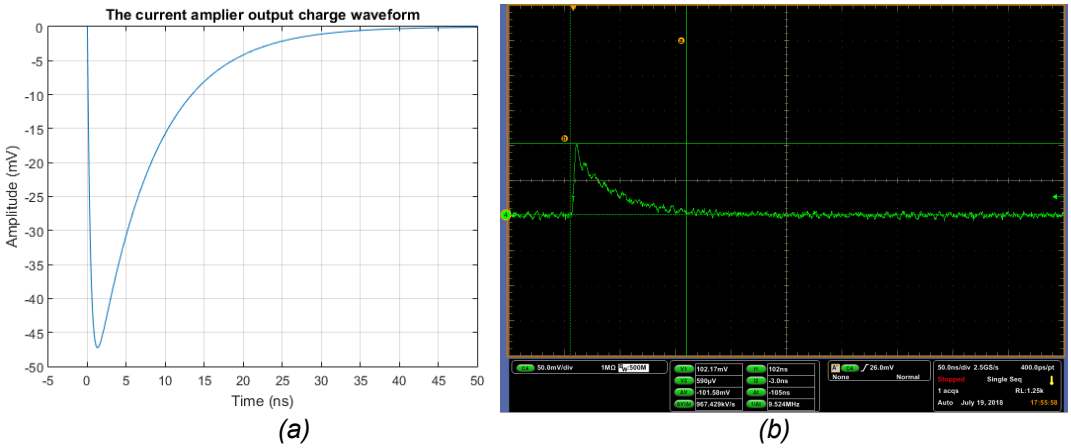
Here k is a constant. To solve k, an initial condition $q_o(0)=0$ is introduced. It is very simple to derive that $k=0$. Then continue the above solution:

$$e^{t/\tau_f}\cdot q_o+0=-\frac{Q_{IN}\cdot\tau_f}{\tau_{IN}-\tau_f}\cdot\left[e^{t(\frac{\tau_{IN}-\tau_f}{\tau_{IN}\tau_f})}-1\right]$$

$$q_o=-\frac{Q_{IN}\cdot\tau_f}{\tau_{IN}-\tau_f}\left[e^{t(\frac{\tau_{IN}-\tau_f}{\tau_{IN}\tau_f})}\cdot e^{-t/\tau_f}-e^{-t/\tau_f}\right] \qquad (4.76)$$

$$q_o=-\frac{Q_{IN}\cdot\tau_f}{\tau_{IN}-\tau_f}\left[e^{-t/\tau_{IN}}-e^{-t/\tau_f}\right]$$

Equation (4.76) is quite significant. It describes the output charge waveform in the time domain. To find out the output voltage waveform, simply divide Eq. (4.76) by C_f. Remember, Q_{IN} is defined in Eq. (4.1).



(a) (b)

Figure 4-35 (a) the calculated output waveform in the time domain for the current feedback amplifier circuit with a SiPM; (b)the oscilloscope captures of a real SiPM preamplifier output waveform.

Figure 4-35 shows the output waveform (a) calculated from Eq. (4.76) in comparison of a real SiPM preamplifier output (b).

The theoretical output waveform has the same characteristics as the real waveform with a very sharp rise time and a slow and exponential decay time. Note the different polarity is because in real circuit, SiPM is negatively biased to make the preamplifier output a positive pulse. This is to tackle the 180deg phase shift introduced by the preamplifier circuit. And generally, a positive pulse is preferred for circuit processing and the later stage analogue to digital conversion. This polarity difference doesn't affect the discussion here.

Further, the time for the peak q_0 can be worked out from the fact that the derivative of the waveform is zero at the peak as $dq_o/dt = 0$.

The derivation process of Eq. (4.76) is skipped and the result is directly given as:

$$t_{peak} = \frac{\ln(\frac{\tau_f}{\tau_{IN}}) \cdot \tau_f \tau_{IN}}{(\tau_f - \tau_{IN})} \tag{4.77}$$

4.5 Development of the preamplifier board

4.5.1 Schematic capture

With all the theory established, we are ready to engineer an actual preamplifier board. The Electronic Design Automation (EDA) tool that was chosen for the schematic capture and board layout is DesignSpark [138] from RSComponents. It is simple and free to use, which is ideal to create a fast turn-around prototype. One good feature of this tool is its seamless link to the RSComponents' product database. This not only provides the access to a very large electronic components library, but also makes the process of generating the bill of material a simple click. This EDA tool also has a growing large community, where extra component libraries can be sourced. Even for those components that still can't be found, it is quite straight-forward to build a customized component with the built-in utility. Of course, this is a relatively basic EDA tool and for more advanced projects, other large commercial EDA tools like DxDesigners from MentorGraphics, SolidWorks PCB etc. can provide much more functionality.

We chose the current amplifier circuit shown in Figure 4-1 as the base form of our preamplifier board. The work focuses on selecting the wideband op-amp(IC1) and determining the right feedback resistor (R1) and capacitor(C3) for the selected SiPM model.

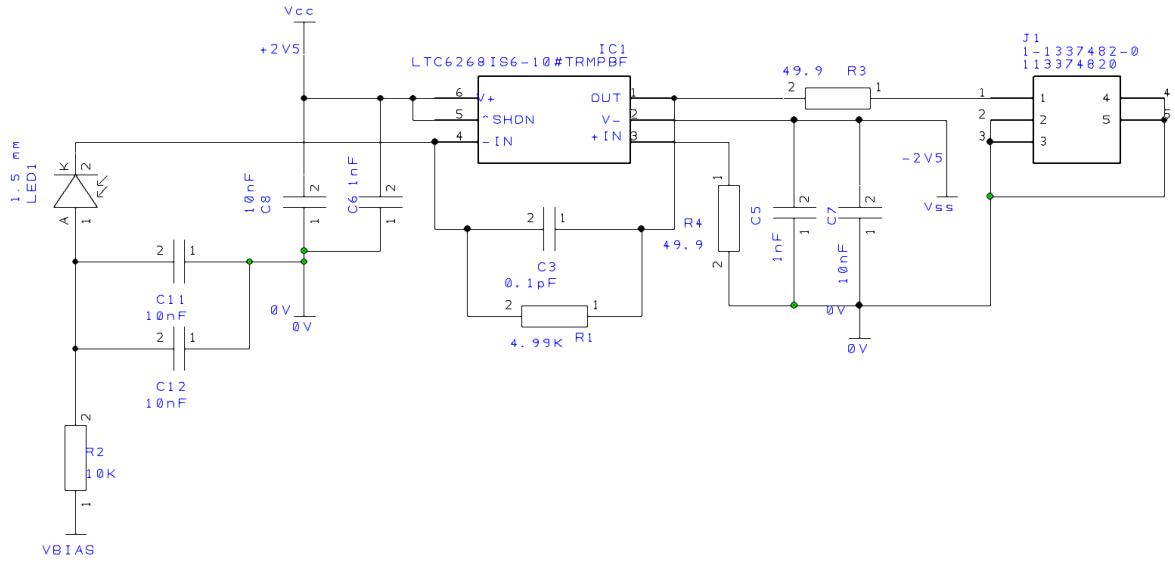


Figure 4-36 Schematic of the pre-amp board

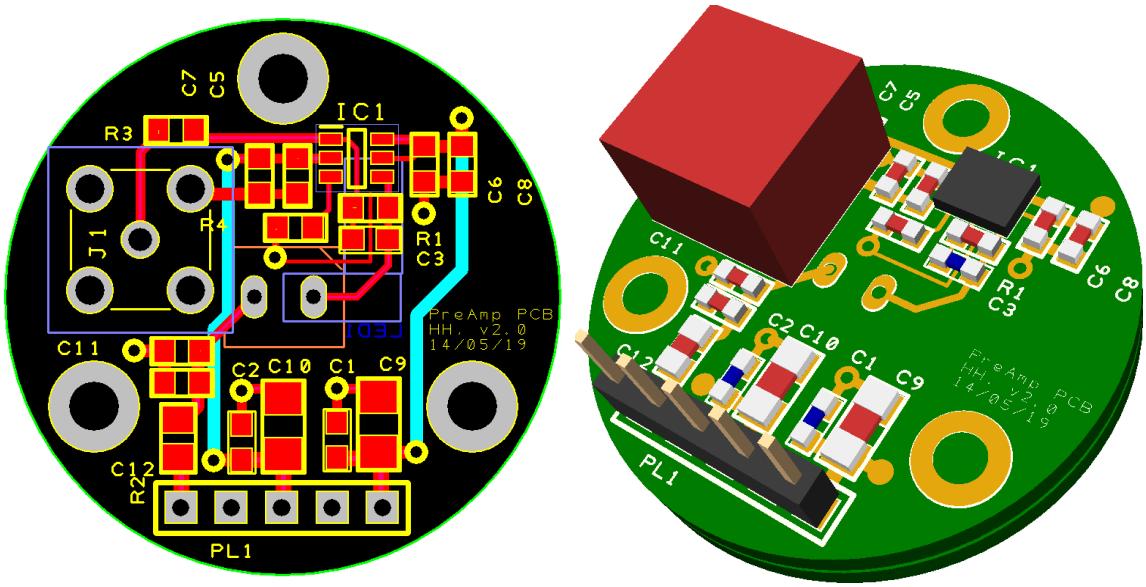
R2 and C12, C11 filters the V_{BIAS} supply voltage for the SiPM. The SiPM is biased negatively at the anode and its cathode is directly connected to the inverting input of the op-amp (IC1). This is explained before as the need to compensate the 180deg phase shift. The non-inverting input of the op-amp is connected to ground (0V) via a 49.9ohm resistor R4. R4 aims to counter balance the input bias current impact at the inverting input. R3 sets the op-amp output impedance at 50ohm (49.9ohm is just an engineer practical value which is close enough to 50ohm), which allows the impedance matching with the common 50-ohm co-axial cable that takes the signal to the next stage. The op-amp is biased with ±2.5V. The output of the op-amp is limited to its supply voltage. Hence, the selection of the supply voltage is to match the input range of DRS4 data processing module that we use as the next stage processing. The choice of the op-amp and the feedback network value will be discussed in detail in 4.5.4.

4.5.2 PCB layout

Once the schematic is complete, the next step is the layout of PCB, shown in Figure 4-37 (a). The PCB profile is set to match an existing PCB mechanic holder. The PCB

design opts for a four-layer solution to comfortably route all the tracks with one layer dedicated to the ground plane. For high speed application, it is vital to have at least one ground plane. Layer 1 (top) is for mounting the SiPM sensor only. The remaining space shall be kept clear to avoid clashing with the mechanical holder. Layer 2 is set to a ground plane (not shown). Layer 3 is for signal and power tracks (coloured in cyan). And Layer 4(bottom, coloured in red) is for fitting all the remaining circuit components and for some signal/power tracks.

One nice feature of DesignSpark is the integration of the 3D view tool. RSComponents provide 3D models for many popular components. If the component's 3D model is not available, by setting the component height parameter, the tool will just visualize the part as a cube to represent its volume. This is very useful to check for any potential confliction when the PCB is integrated into the mechanical holder. Figure 4-37 (b) shows the 3D view of the bottom side of the PCB.



(a) (b)

Figure 4-37 (a) The PCB layout of the pre-amp board; (b) the PCB 3D show

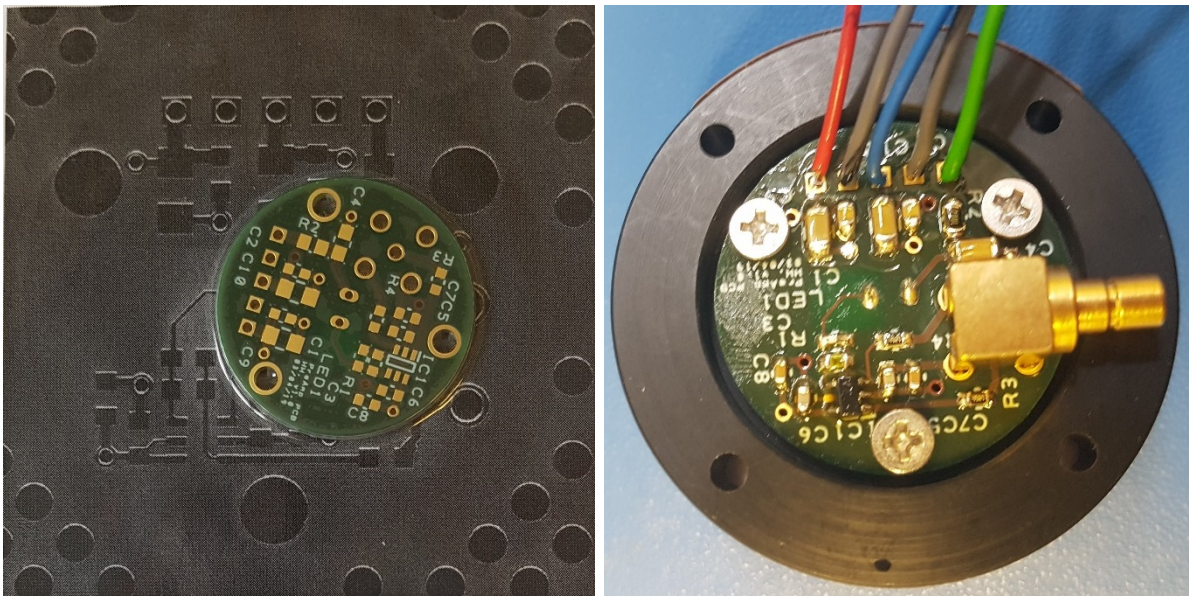
4.5.3 PCB manufacturing

Two versions of PCB have been manufactured. Version 1 was manufactured with common PCB material FR4 and version 2 was changed to the special RF material RO4350B. Compared to FR4, RO4350B has a lower dielectric constant. This is key for high frequency performance, because the low dielectric constant can reduce the parasitic capacitance formed between any PCB conducting surface e.g. tracks, planes,

chassis. Despite their very small value (usually at the order of few pf or even less), these parasitic capacitors can play a key role in altering the circuit signal path at very high frequency (say above a few hundred MHz, the impact starts surfacing.). For many circuit applications, the schematic is only a part of the design. The layout and the selection of PCB material are also crucial to the circuit performance. The latter is to ensure the signal travels through the designed path as the schematic specifies across the frequency spectrum of interest.

RO4305B material also exhibits other benefits such as better thermal stability. Its thermal coefficient of expansion (CTE) is closer to that of copper, which reduces the mismatch of the thermal expansion of the two material under extreme temperature environment. For space application where extreme temperature cycle is expected to experience often, two contacting material with very close CTE is important to minimize the structure failure caused by the CTE mismatch.

The PCB is manufactured by PCB-POOL, from Beta-layout. This company is famous for low cost and quick turn-around for PCB prototype. Figure 4-38 shows the manufactured PCB of prototype v1 and its bottom view post fitting.



(a) (b)

Figure 4-38 (a) The blank v1 PCB; (b) the populated v1 PCB bottom view

4.5.4 Key component selection

The preamplifier circuit is relatively simple. The key to the superb performance is to understand the principle of operation and the theory of the noise model. Based on such knowledge, the right op-amp can be selected. Table 4-3 shows the comparison of system noise with different commercial op-amps.

Table 4-3 System Noise Calculation with different op-amps

Parameter	LMH6629	THS4304	MAX4223	AD8000	LTC6268-10	Unit
SiPM – S13360-1325CS						
Device capacitance (C_g)	43.22	43.22	43.22	43.22	43.22	pF
Number of pixels	2668	2668	2668	2668	2668	#
Grid capacitance per pixel	16.20	16.20	16.20	16.20	16.20	fF
Parasitic capacitance of pixel quenching resistor (C_q)	18.29	18.29	18.29	18.29	18.29	fF
Diode junction capacitance per pixel (C_d)	6.29	6.29	6.29	6.29	6.29	fF
Total SiPM output capacitance per pixel	40.78	40.78	40.78	40.78	40.78	fF
Pixel quenching resistor (R_q)	1600	1600	1600	1600	1600	kΩ
Circuit Setting						
Feedback resistor (R_f)	4.99	4.99	4.99	4.99	4.99	kΩ
Feedback capacitor (C_f)	0.60	0.60	0.60	0.60	0.60	pF
Op-amp						
Opamp input capacitance	4	1.5	0.8	3.6	0.1	pF
Current noise floor (k_i)	2.5	1	3	30	12	pA/√Hz
Current noise corner (f_{f_i})	200	2000	10	5	200	kHz
Voltage noise floor (k_v)	0.7	2.4	2	0.8	4	nV/√Hz
Voltage noise corner (f_{f_v})	10	30	10	20	500	kHz
Open-loop unit gain cross point frequency (f_c)	4000	1000	1000	1500	4000	MHz
Circuit Characteristics						

Parameter	LMH6629	THS4304	MAX4223	AD8000	LTC6268-10	Unit
Total input capacitance (C_D)	44.78	42.28	41.58	44.38	40.88	pF
Circuit Bandwidth (BW_t)	53.16	53.16	53.16	53.16	53.16	MHz
Circuit first zero frequency (f_{zf})	0.70	0.74	0.76	0.71	0.77	MHz
Circuit first pole frequency (f_{pf})	53.16	53.16	53.16	53.16	53.16	MHz
Circuit peaking point frequency (f_i)	52.89	13.99	14.22	20.01	57.86	MHz
Thermal noise						
Boltzmann constant	1.4E-23	1.4E-23	1.4E-23	1.4E-23	1.4E-23	J/k
T at ambient	298	298	298	298	298	k
Thermal noise (EnoR)	82.81	82.81	82.81	82.81	82.81	uVrms
Input Current Noise						
Band 1 noise (E_{noi1})	22.87	30.85	5.56	38.35	109.80	uVrms
Band 2 noise (E_{noi2})	90.78	35.69	109.14	1091.4	435.76	uVrms
Band 3 noise (E_{noi3})	90.95	36.38	109.15	1091.4	436.58	uVrms
Total input current noise (E_{noi})	130.53	59.58	154.45	1543.9	626.54	uVrms
Input Voltage Noise						
Band 1 noise (E_{nov1})	0.26	1.61	0.74	0.43	11.91	uVrms
Band 2 noise (E_{nov2})	0.58	2.03	1.73	0.66	2.07	uVrms
Band 3 noise (E_{nov3})	222.86	721.99	591.84	252.45	1164.0	uVrms
Band 4 noise (E_{nov4})	0.00	0.00	0.00	0.00	599.63	uVrms
Band 5 noise (E_{nov5})	382.46	637.09	526.49	266.47	2088.2	uVrms
Total input voltage noise (E_{nov})	442.65	962.89	792.13	367.06	2464.8	uVrms
Output						
Total circuit noise at output	**468.87**	**968.28**	**811.28**	**1589.2**	**2544.5**	uVrms

Note:

LTC6268-10 and MAX4223 are current-feedback op-amp, which has different current noise characteristic curve than that of the voltage-feedback op-amp, which are the case for the other three.

This table is designed into eight sections. The first two sections are the characteristics of the SiPM device S13360-1325CS and the circuit setting for R_f and C_f. These are fixed once selected and kept unchanged for the comparison of op-amps. The third section is the characteristics of op-amps. The data is quoted from the manufacturer's datasheet. However, for the current and voltage noise spectrum, most datasheets only publish a plot. So, that data is extracted from the plot by sight reading with limited accuracy. With the information from the top three sections, the circuit characteristics then can be calculated with results in the fourth section. Note here C_D refers to the total input capacitance, which is the sum of the SiPM output capacitance and the op-amp input capacitance. In a real circuit, the PCB parasitic capacitance will also contribute to this sum. The next three sections calculate three dominant noise sources using the equations developed in 4.4.4.

From the comparison, LMH6629 seems to be the best choice for its lowest total noise.

4.5.5 Electric bench test result

Theoretical calculation provides a direction of component choice. But it cannot replace the real circuit test. So, we have tried all the op-amps under comparison for our circuit. Here, the initial result with op-amp LMH6629 and SiPM S13360-1325CS is presented in Figure 4-39 and Figure 4-40.

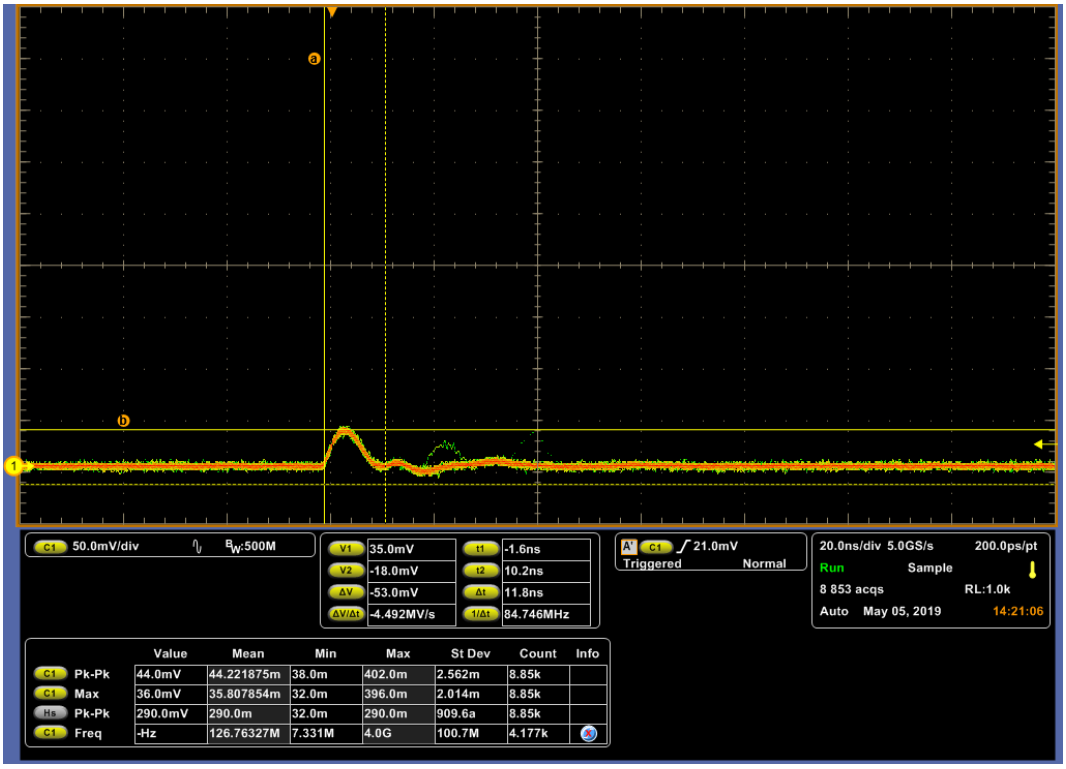


Figure 4-39 Oscilloscope capture of a single dark count pulse with LED off. The oscilloscope is set to persistent mode.

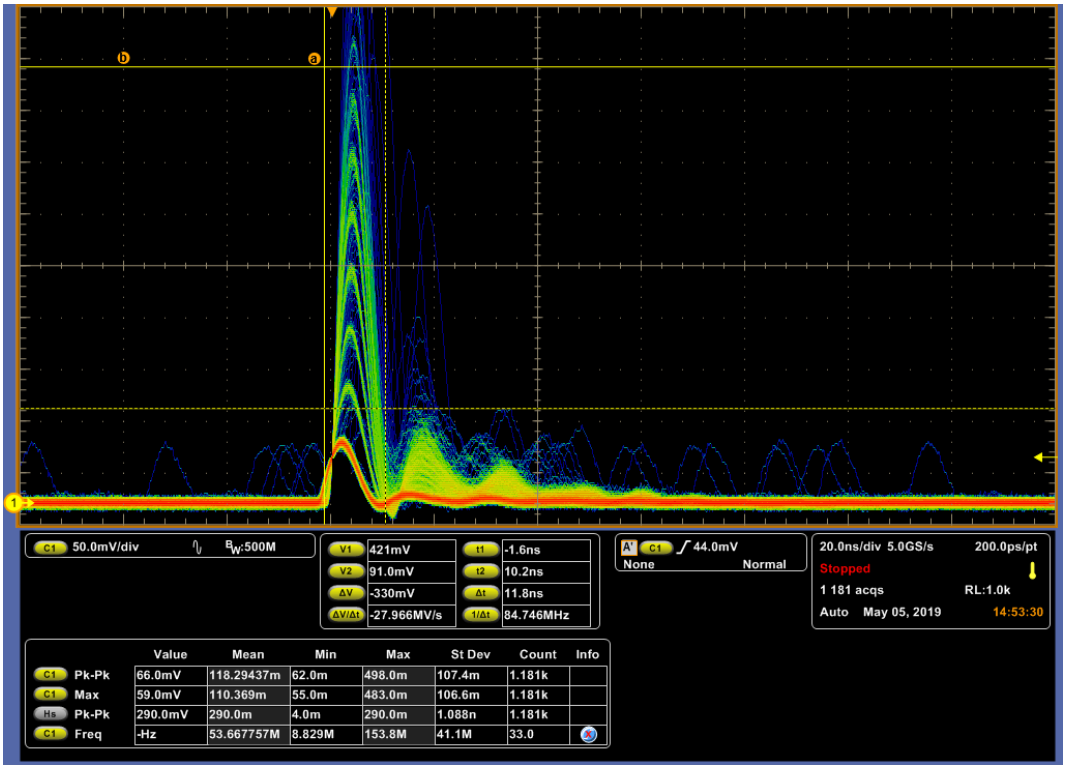


Figure 4-40 Oscilloscope capture of multiple photons events with LED on. The oscilloscope is set to persistent mode.

The preamplifier output captured directly by the oscilloscope is very clean and smooth. The peaks of different p.e. signal in Figure 4-40 are clearly separated. This suggests the preamplifier circuit preserves the SiPM signal waveform very well while amplifying

it. There is a small ringing after the leading primary pulse. The amplitude of the ringing increases as the primary pulse amplitude increases. This suggests the control loop is a bit underdamped. A little bit higher feedback capacitance can help to improve the damping.

In the multiple photon events plot (Figure 4-40), the pulse peaks correspond to different number of triggered pixels. If the peaks are plotted in a linear chart, a straight line is presented as shown in Figure 4-41. This line can be fit with a first order polynomial. The slope of that is the gain of the preamplifier circuit and the SiPM together in unit mV/pixel. The y-axis zero crossing value of the fit is the preamplifier circuit DC offset, which is a constant output signal level when the input signal is zero.

More about the preamplifier circuit performance (including the SiPM) analysis can be found in Chapter 5 and Chapter 6.

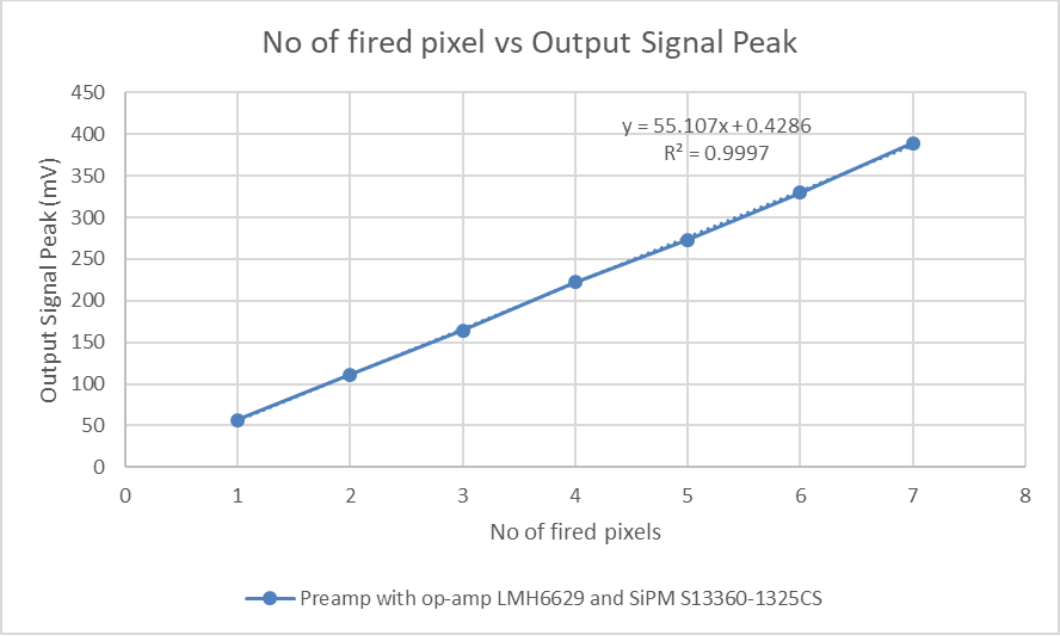


Figure 4-41 The plot of preamplifier circuit output pulse peak vs the number of firing pixels from the prototype

4.6 The selection of SiPM and its radiation hardness

At the early stage of this research project, we conducted a market survey of SiPM sensor covering the manufacturer, typical product number, and the sensor parameters. The aim is to build an overview of SiPM available on the market and the status of mainstream performance. This survey is presented in Appendix G. This survey was

done in 2015 and the market offer has progressed significantly since, therefore, some data might be out of date.

Follow the MPPC technique note from Hamamatsu [139], we calculate the performance of four Hamamatsu SiPMs as presented in Table 4-4.

Table 4-4 Comparison of four Hamamatsu SiPMs

MPPC comparison	S12572-010C	S12572-015C	S13360-025C	S13360-050C	
Pixel size	10u	15u	25u	50u	
MPPC terminal capacitance	320	320	320	320	pF
Npixel	90000	40000	14400	3600	
Pixel capacitance	0.0036	0.0080	0.0222	0.0889	pF
Pixel quench resistor	1000	1000	300	150	kohm
Recovery time	16.4	36.8	30.7	61.4	ns
PDE	0.1	0.25	0.25	0.4	
Dark current	1000	1000	400	500	kcps
Nphotons (a 35MeV proton)	39855	39855	39855	39855	
Nfired	3899	8820	7191	3557	
Pulse width	42.0	42.0	42.0	42.0	ns
Dark count per ns	0.0010	0.0010	0.0004	0.0005	
S/N ratio	62.4	93.9	84.8	59.6	
Non-Linearity	2.2	11.5	27.8	77.7	%
Bandwidth	44.76	19.89	23.87	11.94	MHz
f = 1/recovery time	61.07	27.14	32.57	16.29	MHz

The yellow section is the electrical parameters of each SiPM. The blue cell is an assumed scintillator photon emission and it is kept the same for all the four SiPM for comparison. Then we use the formula in the MPPC technical note to calculate the expected number of fired pixels, the S/N ratio and the non-linearity. We can find that the pixel number of the SiPM shall be selected to maintain a good linearity at the target maximum scintillation photon. The increasing pixel number comes at the cost of increasing dark count rate. This will reduce the signal to noise ratio. The calculation shows that S12572-015C is the best choice which balances the linearity and S/N ratio well for the assumed scintillation photon.

Common to other silicon based electronic device, SiPM is sensitive to radiation damage. The most sensitive parameters are dark count rate. Other parameters like signal amplitude and shape, I-V curve shape, timing (timing resolution, rise and fall time) are relatively stable after mild exposure e.g. 30krad dose. This level of radiation hardness is sufficient for near-Earth orbit environment for a 2-3 year mission.

We don’t expand the discussion here. More details can be found in these publications: [140][141][142].

4.7 Conclusion and future work

A complete preamplifier circuit development from theory to realization is covered in this chapter. Though the SiPM signal is relatively large due to the amplification provided by the avalanche process, the preamplifier circuit still needs to be carefully designed to achieve the targets: (1) good signal to noise ratio, (2) good stability and (3) enough bandwidth to keep the original signal not distorted.

The theory study in this chapter can help to guide the real circuit design and reduce the number of try-and-error practices.

Two types of preamplifier circuit (voltage and current) are introduced and compared with small signal analysis. We found the current preamplifier circuit had better bandwidth with the same circuit setting.

The stability analysis shows the importance of phase compensation technique to stabilize the preamplifier circuit.

Focussing on the current preamplifier circuit, we constructed a noise model of the circuit. Three major noise sources feedback resistor thermal noise, op-amp input current noise and input voltage noise are identified and their spectrum distribution are analysed. The total noise projected to the preamplifier output are the square root sum of all the individual noise projected independently.

We also derived the equations for the time domain analysis. This is very useful to help to understand the output waveform profile and its timing. This analysis also proves the rising edge and falling edge of the output is associated with the two time constants formed by the SiPM and preamplifier circuit together.

Lastly, we designed and built a preamplifier circuit board. We used the knowledge built in the theory study to calculate and select the proper op-amp for the chosen SiPM. The initial test with an oscilloscope directly connected to the preamplifier board output shows good system linearity. More system characterization with more advanced DRS4 module as the next stage data sampling and digitizing will be discussed in Chapter 6.

For future work, more advanced preamplifier with temperature monitor and bias voltage compensation can be added. The preamplifier can be integrated with the SiPM coupling to the scintillator detector or centralized into a mother board located nearby. The choice shall be application dependant.

Chapter 5 Signal Processing

5.1 Introduction

The last chapter introduced the electronic design of the preamplifier board as the first stage of the readout electronics. The remaining stages of the readout electronics is made up of a discrete electronic module called the DRS4 module from PSI[143]. Later, we also refer the preamplifier circuit plus the DRS4 module as the readout system. More about the full readout system is covered in Chapter 6.

The output of the readout system is a waveform in digital format, like the oscilloscope capture. This chapter will discuss about the data processing of this waveform. A series of Python scripts are developed to batch process the waveforms and to acquire the parameters like the peak, charge, duration, rise time of the waveform. These are presented in 5.4.

During the analysis of the digital waveform, a phenomenon called signal pile-up is often observed. This occurs when two consecutive signals are too close to each other. This led to some research into how to tackle this issue with a data processing algorithm. The adopted algorithms are Deconvolution and Wiener filter. This is discussed in depth in 5.3.

5.2 Typical digital waveform of the SiPM readout system

The output of the SiPM photon sensor is a very short pulse with extreme fast rising and falling time (depends on the SiPM design, the falling time can be almost as fast as the rising time.) This pulse is amplified by the preamplifier circuit board, but the pulse waveform shape is largely conserved, except both rising time and falling time are extended with falling time extending more significantly. This amplified pulse waveform is sampled and digitized by the DRS4 module. The data is transferred to a PC where it becomes available for display and further processing. Figure 5-1 shows a typical amplified pulse waveform sampled at 5GHz frequency. When one pixel is triggered, its output pulse is called 1p.e. signal (p.e. is short of photoelectron, which is commonly used as a unit for SiPM pixel event.). The pulse for two-pixel event is called 2p.e. and so on. The number not only indicates how many pixels are triggered, but also suggests the amplitude of the pulse peak as that is linearly proportional to the number of pixels within the readout system dynamic range.

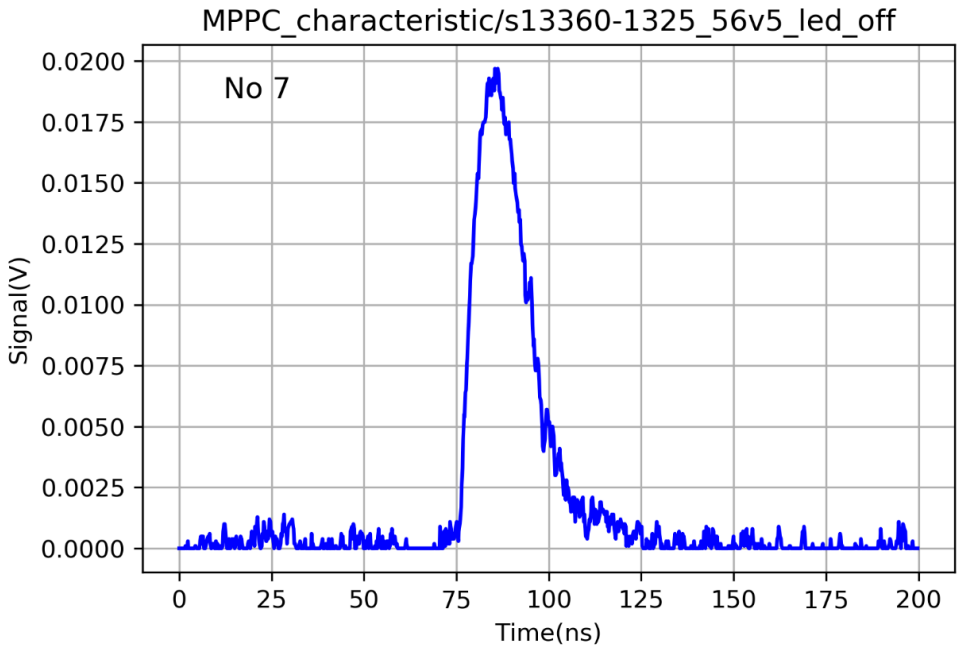


Figure 5-1 A typical amplified SiPM pulse waveform for a 1p.e. signal.

But some events will look like Figure 5-2, where two pulses occur so close that their pulse waveforms overlap with each other. This phenomenon is called pulse pile-up. Pile-up on SiPM sensor is quite common due to its design. A SiPM has a matrix of pixels. Each pixel operates independently with its own bias voltage electrode, quench-

ing resistor and the photon-sensitive p-n junction. But the output of each pixel is electrically connected together to form one single output. This creates a problem that if the second pixel is firing while the first pixel event is not completed, the two signals will pile-up.

There could be many causes for the second pixel occurring immediately after the first pixel. One cause is the incident photon flux is too high, i.e. under a strong light source. Another cause could be the after-pulse signal, which is triggered by the delayed electron in the trapping centre or defect of the silicon. Another cause could be the coincidence with a self-generated thermal triggered dark count signal. This occurs by chance and the probability is very low.

The preamplifier circuit extends the pulse duration due to the inevitable delay introduced by the circuit. This makes the pile-up situation worse. Therefore, the preamplifier circuit is desired to be as fast as possible from this point of view.

Figure 5-2 shows an example of two pulses piling up. From the plot, it suggests that the second pulse is not necessarily smaller than the first pulse.

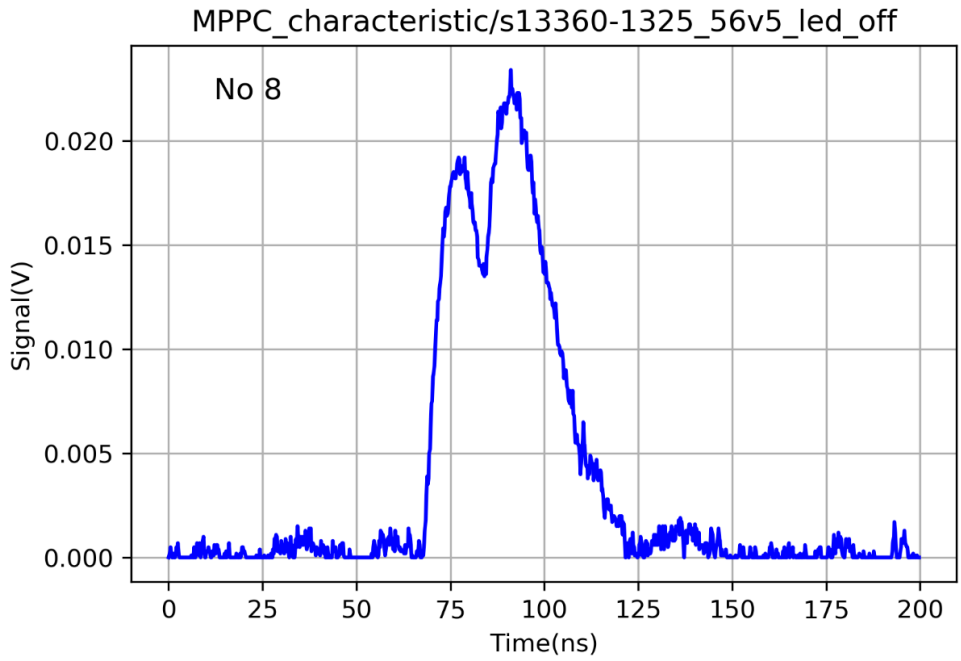


Figure 5-2 An example of two pulses pile up

The pile-up increases the error in the signal processing. Often, the overlapped multiple pulses are treated as a single pulse. When the system is operating in counting mode, this means loss of counts. When the system is in pulse height analysis (PHA) mode, this makes the peak searching algorithm miss pulses. Figure 5-3 shows an example

of three peaks with only two peaks are identified. The middle peak is too close to the leading pulse.

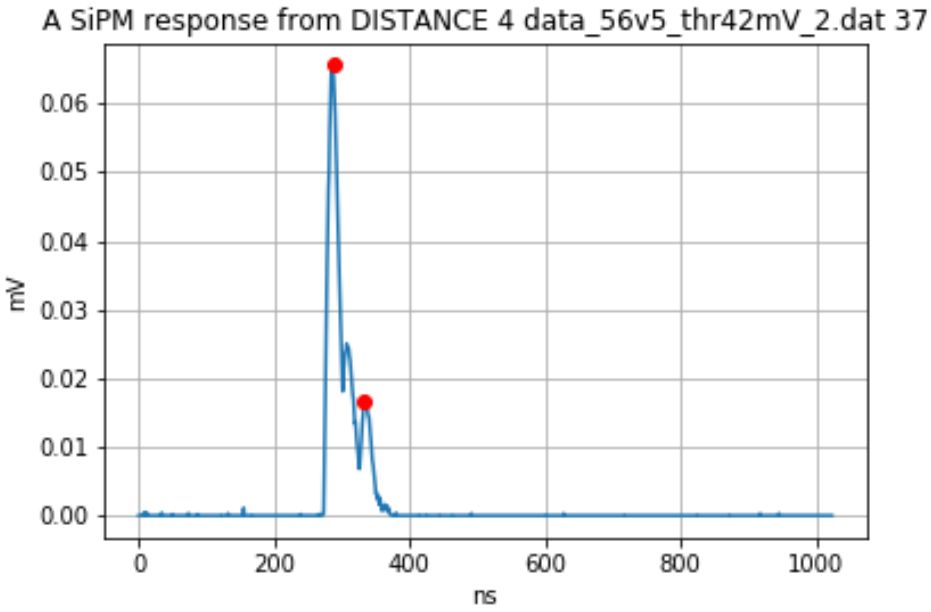


Figure 5-3 The example of pulse pile-up causing missing peak search.

There are ways to tackle or relief the pile-up issue at both hardware and software level. A common hardware method is called CR-RC shaping network[144]. As indicated in Figure 5-4 (left), this network is typically composed of a high pass filter (CR), a buffer op-amp and a low pass filter (RC). The function of the buffer is to isolate the CR and RC network to avoid interaction between them. CR forms a high pass filter, which can sharpen the long decay time. RC forms a low pass filter, which improves the signal to noise ratio. The Figure 5-4 (right) shows the signal waveform from input to output. We can see the prolonged input waveform (green) is reshaped to a gaussian like short waveform (blue) after the network.

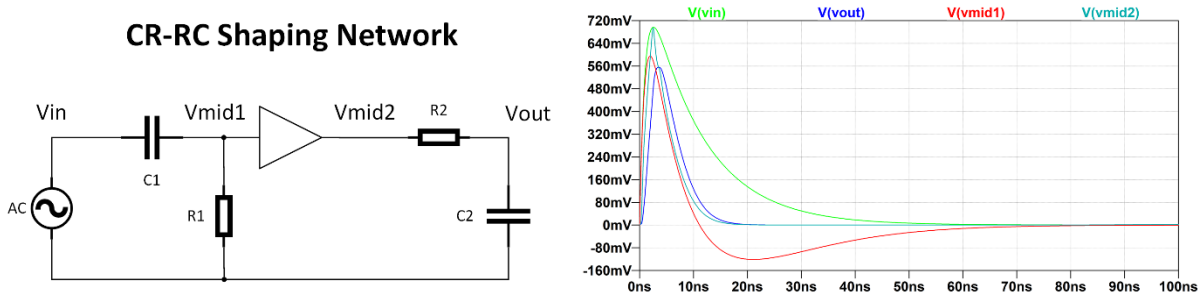


Figure 5-4 The typical circuit of a CR-RC shaping network (left) and its typical signal waveform (right)

This method is widely used in nuclear detector application. The CR-RC values can be adjusted to fit the detector signal characteristics and the flux rate. The big drawback of the hardware method is the inflexibility. Once the hardware is manufactured, it is not possible or very costly to change the CR-RC values. On the other hand, the dynamic signal flux rate will require more flexible configuration of the CR-RC shaping network.

This drawback is the advantage of the software-based solution. A popular choice to mimic the performance of the CR-RC network with a software solution is the Deconvolution filter and Wiener filter. Deconvolution filter is a variant of high pass digital filter and Wiener filter is a variant of low pass digital filter. The combination of the two filters can provide similar signal shaping function as the CR-RC network.

Next section will discuss the Deconvolution filter and Wiener filter in detail.

5.3 Deconvolution and Wiener filter

5.3.1 Background knowledge

It is necessary to prepare some background knowledge first. In signal processing, a continuous function **f(t)** in time domain can be transformed into frequency domain by Fourier Transform (FT) (4.78). Reversely, a function **F(ω)** can be transformed back to time domain by Inverse Fourier Transform(IFT) (4.79).

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \tag{4.78}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{+i\omega t} d\omega \tag{4.79}$$

For discrete signal, the Discrete Fourier Transform(DFT) pairs are defined as:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{\frac{-i2\pi}{N} kn} \tag{4.80}$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{\frac{+i2\pi}{N} kn} \tag{4.81}$$

DFT is widely used in the digital signal processing.

Figure 5-5 shows a classic signal processing system.

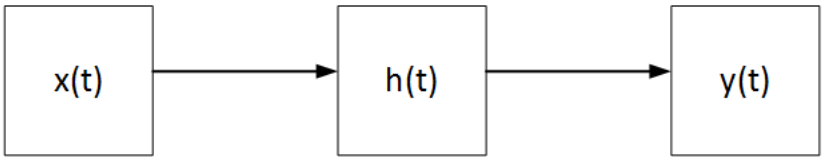


Figure 5-5 A classic signal processing system.

x(t) represents the input signal, h(t) represents the processing function and y(t) represents the output signal. All the three functions are prepared in the time domain.

Through Fourier transform defined by equation (4.78), the three functions can be transformed to frequency domain as X(ω), H(ω) and Y(ω). The system can be expressed as:

$$X(\omega)\cdot H(\omega)=Y(\omega) \tag{4.82}$$

Equivalent but in the time domain, Eq. (4.82) can be rewritten as:

$$x(t)*h(t)=y(t) \tag{4.83}$$

The * operation is called convolution. Mathematically, convolution is defined as:

$$y(t)=\int_{-\infty}^{\infty}x(\tau)h(t-\tau)d\tau \tag{4.84}$$

Equation (4.82) and (4.83) are equivalent in describing a linear system. They are interchangeable through Fourier Transform or Inverse Fourier Transform.

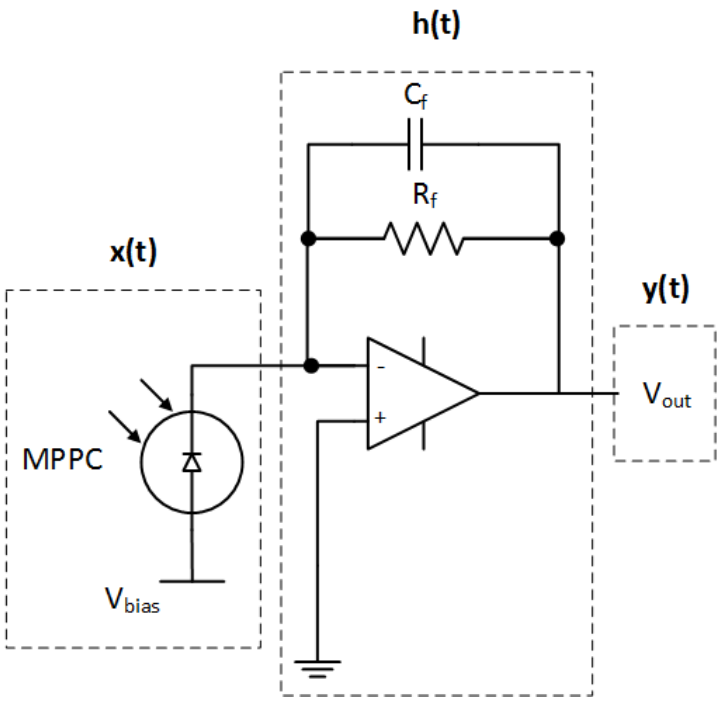
Deconvolution means given the knowledge of h(t) and y(t), find the solution of x(t). The basic approach is:

1. Compute the Fourier transform of h(t) to get H(ω) and of y(t) to get Y(ω).
2. Obtain X(ω) from $X(\omega)=Y(\omega)/H(\omega)$.
3. Compute the Inverse Fourier Transform (4.79) of X(ω) to get x(t).

This seems very simple and straightforward, but in the real application, quite often we find y(t) is mixed with noise, and h(t) is difficult to identify. These all make the problem difficult to solve.

5.3.2 Deconvolution filter

The SiPM readout system that is discussed in Chapter 4 can be constructed to match the classic system in Figure 5-5 as:



$$x(t) * h(t) = y(t)$$

$$X(\omega) \cdot H(\omega) = Y(\omega)$$

Figure 5-6 Construct the SiPM readout system into the classic signal processing system

Here, the SiPM raw signal is set as the system input x(t), the preamplifier circuit that processes the input signal is h(t) and the output signal is y(t).

y(t) is known from the measurement result. If h(t) can be found, then x(t) can be derived following the approach defined previously. It is easier to discuss about this problem mathematically and in frequency domain. What needs to be found is such a H(ω) that can help to work out the X(ω) from

$$X(\omega) = \frac{1}{H(\omega)} \cdot Y(\omega) \tag{4.85}$$

Define $H_D(\omega) = 1/H(\omega)$ as the deconvolution filter. So equation (4.85) can be re-written as:

$$X(\omega) = Y(\omega) \cdot H_D(\omega) \tag{4.86}$$

The way to find $H_D(\omega)$ is to define a desired x'(t), then estimate h'(t) from some training data y'(t) as:

$$H_D(\omega) = \frac{X'(\omega)}{Y'(\omega)} \tag{4.87}$$

Then with this acquired deconvolution filter, we can recover the original signal x(t) using equation (4.86) followed by a IFT.

The choice of x'(t) is arbitrary, but a pulse-like function with a smaller pulse width than the original signal x(t) is a good choice. The widely used x'(t) is Blackman window function:

$$BW(t;\tau_{BW}) = \begin{cases} 0.42 - 0.5\cos(2\pi t/\tau_{BW}) + 0.08\cos(4\pi t/\tau_{BW}) & (0 < t < \tau_{BW}) \\ 0 & (otherwise) \end{cases} \quad (4.88)$$

The parameter τ_{BW} determines the width of the Blackman window function. Its value will affect the time resolution and the signal to noise ratio(S/N) of the filtered waveform result. More detail about the choice of τ_{BW} will be discussed later.

Figure 5-7 shows the Blackman window function in time domain and its frequency response.

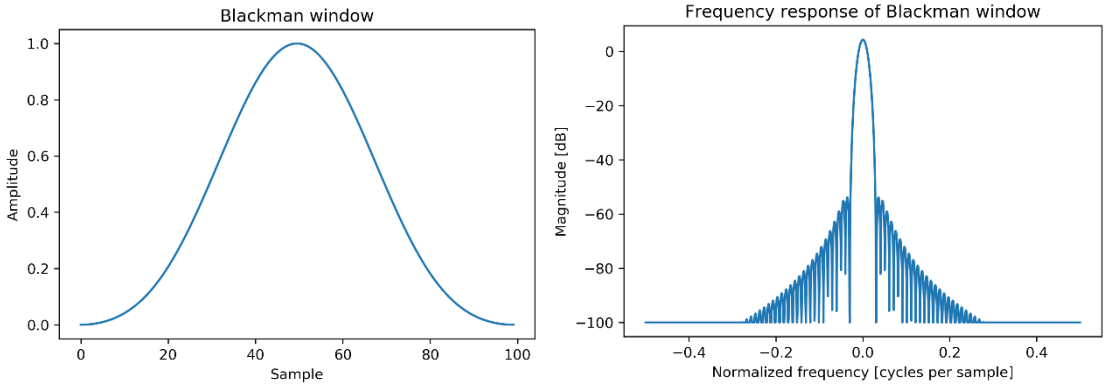


Figure 5-7 Blackman window function waveform in time domain (left) and its frequency response (right)

In real system, noise will be inevitably added into the signal through the processing function h(t). The Deconvolution filter as a type of high pass filter is quite poor in tempting the noise. So, another low pass filter is needed to reduce the system noise and hence compensate for the degradation of signal to noise ratio. The Wiener filter is a common choice and it often works together with the Deconvolution filter as a pair.

Adding the Deconvolution filter and the Wiener filter, system described in Figure 5-5 can be extended to:

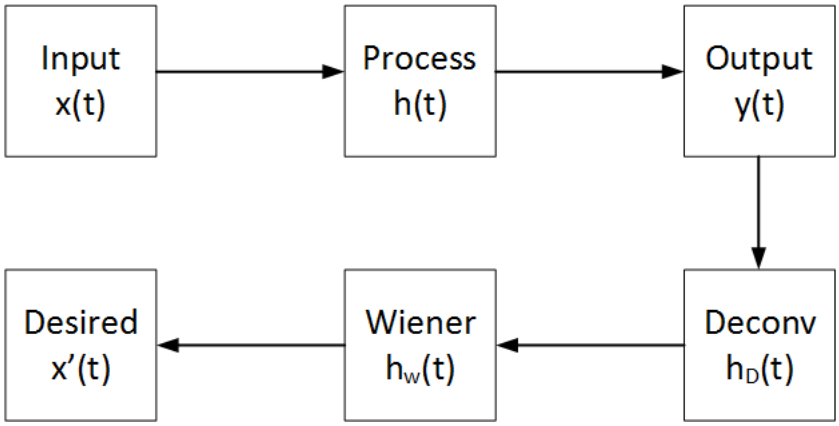


Figure 5-8 Extended signal processing system that includes the Deconvolution filter and Wiener filter

The Deconvolution filter and Wiener filter can be cascaded together to form a single filter H_{dw} (= H_dH_w) in practice. This can save the computer processing power as only one FFT and one IFFT calculation is required to switch the combined filter between time and frequency domain.

5.3.3 Wiener filter

To include both signal and noise item, the input signal can be redefined as

$$x_{s+n}(t) = x_s(t) + x_n(t) \tag{4.89}$$

Then equation (4.83) can be rewritten as:

$$y_{s+n}(t) = \left(x_s(t) + x_n(t)\right) * h(t) \tag{4.90}$$

Apply the Fourier transform, and Eq. (4.90) becomes:

$$\begin{aligned} Y_{s+n}(\omega) &= [X_s(\omega) + X_n(\omega)] \cdot H(\omega) \\ &= X_s(\omega) \cdot H(\omega) + X_n(\omega) \cdot H(\omega) \\ &= Y_s(\omega) + Y_n(\omega) \end{aligned} \tag{4.91}$$

Apply the Inverse Fourier transform, and Eq. (4.91) becomes:

$$y_{s+n}(t) = y_s(t) + y_n(t) \tag{4.92}$$

This suggests the output signal is also composed of two parts: signal and noise. According to Figure 5-8, write the signal processing system equation:

$$Y_{s+n}(\omega) \cdot H_D(\omega) \cdot H_w(\omega) = X'(\omega) \tag{4.93}$$

And assume the Deconvolution filter $H_d(\omega)$ is perfect, then:

$$Y_s(\omega)\cdot H_D(\omega)=X(\omega) \tag{4.94}$$

Here, $H_w(\omega)$ is the Wiener filter. To find out what is the best form for this filter, the judgement criteria is set to achieve the minimum error between x(t) and x'(t).

First define a function L as:

$$L=\int_{-\infty}^{+\infty}\left|x'(t)-x(t)\right|^2 dt \tag{4.95}$$

Using Parseval's theorem, equation (4.95) can be converted from time domain to frequency domain as

$$L=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|X'(\omega)-X(\omega)\right|^2 d\omega \tag{4.96}$$

Then expand L with the definition from (4.93) and (4.94) to get:

$$\begin{aligned} L&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|Y_{s+n}(\omega)\cdot H_D(\omega)\cdot H_w(\omega)-Y_s(\omega)\cdot H_D(\omega)\right|^2 d\omega \\ &=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|(Y_s(\omega)+Y_n(\omega))\cdot H_D(\omega)\cdot H_w(\omega)-Y_s(\omega)\cdot H_D(\omega)\right|^2 d\omega \end{aligned} \tag{4.97}$$

Since there is no correlation between $y_s(t)$ and $y_n(t)$ (this means $Y_s(\omega)Y_D(\omega)=0$), further re-organize (4.97) into:

$$\begin{aligned} L&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|H_D(\omega)\right|^2\left|(H_w(\omega)-1)Y_s(\omega)+H_w(\omega)Y_n(\omega)\right|^2 d\omega \\ &=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|H_D(\omega)\right|^2\left[\left|H_w(\omega)-1\right|^2\left|Y_s(\omega)\right|^2+\left|H_w(\omega)\right|^2\left|Y_n(\omega)\right|^2\right]d\omega \end{aligned} \tag{4.98}$$

To find the $H_\omega(\omega)$ which makes L minimum, need to solve this partial differential equation:

$$\frac{\partial L}{\partial H_w(\omega)}=0 \tag{4.99}$$

The mathematic derivation of the solution is quite simple as:

$$\frac{\partial L}{\partial H_D(\omega)} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |H_D(\omega)|^2 \left[2|H_w(\omega)-1||Y_s(\omega)|^2 + 2|H_w(\omega)||Y_n(\omega)|^2 \right] d\omega = 0$$

$$\rightarrow 2|H_w(\omega)-1||Y_s(\omega)|^2 + 2|H_w(\omega)||Y_n(\omega)|^2 = 0 \qquad (4.100)$$

$$\rightarrow H_w(\omega) = \frac{|Y_s(\omega)|^2}{|Y_s(\omega)|^2 + |Y_n(\omega)|^2}$$

This is the form of the Wiener filter. What this means is that the filter is constructed as the ratio of the power spectrum of signal and the sum of the power spectrum of both signal and noise. The power spectrum is defined as the square of the magnitude of the complex number. This form also suggests that Wiener filter is a pure scalar with no phase element. So, applying the filter won't add any phase shift to the original signal. With Wiener filter, the recovered signal x'(t) has the smallest absolute error to the original signal x(t).

5.3.4 Impact of normalization for the Wiener filter

It worth noticing that it is important to normalize the signal y_s and noise y_n before applying equation (4.100) in practice. To understand the reason, first examine this Figure 5-9.

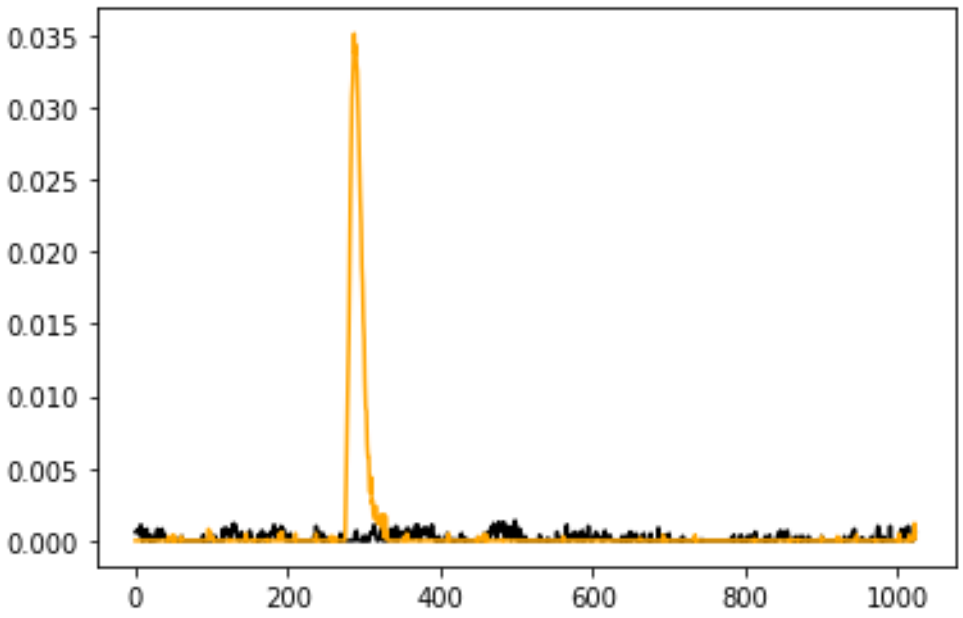


Figure 5-9 The waveform of a 1p.e. signal(yellow) and the noise floor(black)

The pulse in yellow is the waveform of a 1p.e. signal. The white noise like signal in black is the electronic background noise acquired by biasing the SiPM just under its

breakdown voltage. Given no correlation between the signal and noise, rewrite equation (4.97) as

$$\begin{aligned} L &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|(Y_s(\omega)+Y_n(\omega))\cdot H_D(\omega)\cdot H_w(\omega) - Y_s(\omega)\cdot H_D(\omega)\right|^2 d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|(Y_s(\omega)\cdot[H_D(\omega)\cdot H_w(\omega) - H_D(\omega)] + Y_n(\omega))\cdot H_D(\omega)\cdot H_w(\omega)\right|^2 d\omega \\ &= \frac{1}{2\pi}\left[\int_{-\infty}^{+\infty} \left|(Y_s(\omega)\cdot[H_D(\omega)\cdot H_w(\omega) - H_D(\omega)]\right|^2 d\omega + \int_{-\infty}^{+\infty} \left|Y_n(\omega))\cdot H_D(\omega)\cdot H_w(\omega)\right|^2 d\omega\right] \end{aligned}$$

(4.101)

Equation (4.101) is composed of two integration terms, one for the signal and the other for the noise. Refer to Figure 5-9, it is corresponding to the integration of the 1p.e. signal and the electronic noise over the entire time window. Clearly the 1p.e. signal only exists for a very short period while the electronics noise occupies the full time window. If both are integrated over the full time window, the result of the noise term is amplified compared to the signal term. The problem gets worse if the window gets larger.

Therefore, it is necessary to localize the noise integration only over the time window of signal. Update the equation (4.100) as:

$$H_w(\omega) = \frac{\left|Y_s(\omega)\right|^2}{\left|Y_s(\omega)\right|^2 + c_n^2\left|Y_n^{'}(\omega)\right|^2} \tag{4.102}$$

Where $Y_n^{'}(\omega) = \frac{T_{sig}}{T} Y_n(\omega)$ and **C_n** is a normalization constant with size of ~1. The factor T_{sig}/T is the ratio of the signal duration over the full window. To simplify the calculation, the ratio can be regarded as a constant and be merged with C_n as a single correction constant.

5.3.5 Impact of T_{BW} for the Deconvolution filter

The shape of the desired signal x'(t) after the deconvolution filter is determined by the Blackman window function. And the width of this window function is determined by the parameter T_{BW}. How to select the right value for this parameter depends on the pulse width of the original signal being processed. If the generated signal x'(t) has the equal

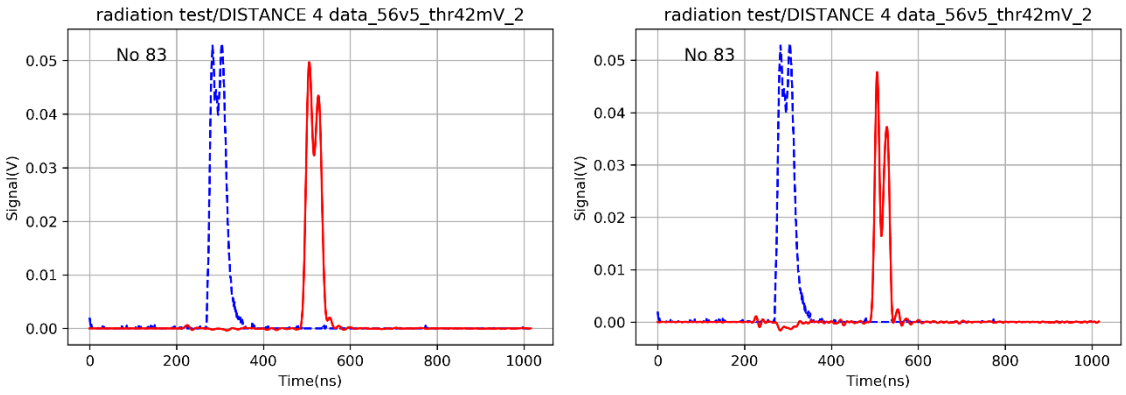
or higher pulse width to the original signal x(t), the piling-up issue will be worse. So, the upper limit for τ_{BW} is that it shall ensure the output x'(t) has shorter pulse width than x(t). However, we can't make τ_{BW} too small, as a smaller τ_{BW} will reduce the signal to noise ratio (S/N).

The lower limit is determined by the minimum signal level that should be maintained after the filter. To distinguish signal from noise, we will generally set a threshold at a certain amplitude. Above the threshold, the data is discerned as part of the signal and run the necessary processing like peak searching algorithm and below that, we regard the data as noise and simply discard them. A good figure for such threshold is 5 times σ (amplitude of noise). For the after-pulse signal after the primary pulse, its amplitude usually is lower than the primary pulse. Let's allow the after-pulse to be as low as 25% of the primary pulse. This means 25% of the full signal level (defined as S) shall be above 5 times σ (defined as N) threshold. A quick calculation can conclude:

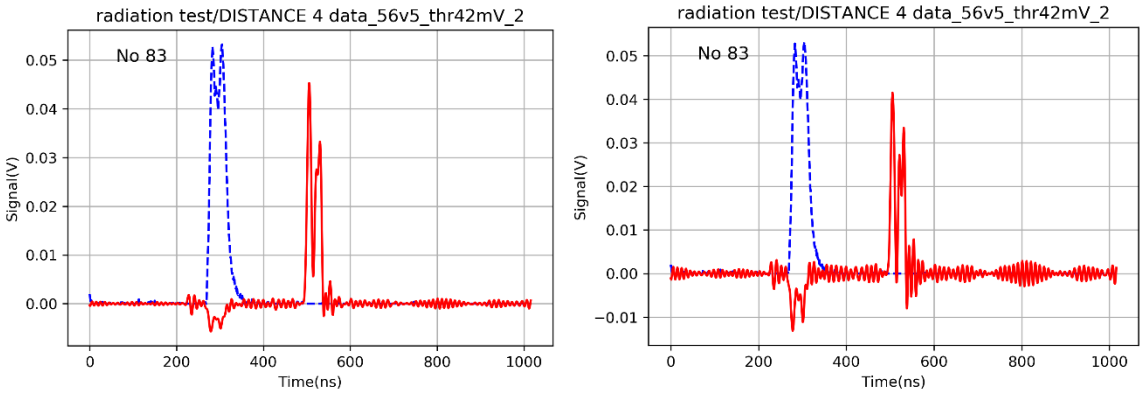
$$\begin{aligned} 25\%S &> 5N \\ \frac{S}{N} &> 20 \end{aligned} \tag{4.103}$$

Therefore, the bottom limit of τ_{BW} is that it shall ensure the S/N is above 20.

The mathematic definition of the link between τ_{BW} and x(t) pulse width and the link between τ_{BW} and S/N is not very clear. So, it is hard to work out the best value by the defined calculation. Therefore, in practice, the old-fashioned try and error method is applied to find the right τ_{BW} that can achieve a good balance. An example of the effect of different value of τ_{BW} is shown in Figure 5-10:



τ_{BW}=8ns τ_{BW}=6ns



τ_{BW}=4.8ns τ_{BW}=4ns

Figure 5-10 The filter performance impact with different setting of τ_{BW}.

Note the dotted blue waveform is the original signal and the red waveform is after the filter. Four different values of τ_{BW} have been experimented.

Clearly the blue waveform is the result of two individual signals piled up. From the results, we can see τ_{BW} = 4.8ns has the best balance of the four values, that it can separate the two pulses with clear gap while still maintain enough S/N ratio. τ_{BW} = 4ns generates too much noise and the other two values fail to separate the two pulses wide enough.

The performance of the Deconvolution+Wiener combined filter is not only determined by the choice of τ_{BW}, but also by the choice of c_n. For example, although selecting a very small τ_{BW} leads to higher noise, a proper choice of c_n can improve the performance of Wiener filter to damp the noise. This improves the signal to noise ratio and compensate the negative effect of Deconvolution filter. Thus, the combined effect could achieve a better performance. See Figure 5-11 demonstrates a better outcome with the choice of a large c_n while adopting the overly filtering τ_{BW} = 4ns setting.

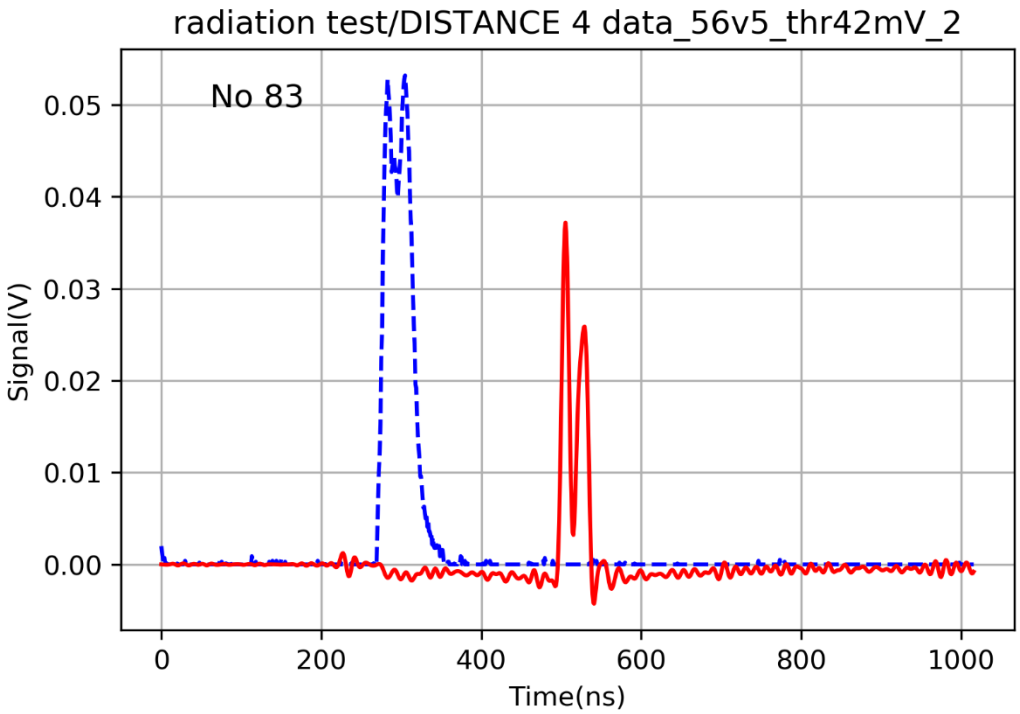


Figure 5-11 The performance of the combined filter with the choice of τ_{BW} = 4ns and c_n=500

Therefore, in a practical system, the two parameters can be evaluated together to achieve the best outcome of the combined filter. The theoretical method in the form of one or a set of mathematic formula that can work out the optimized parameters combination of τ_{BW} and c_n require further study.

5.4 Python program

The data processing scripts in this project are developed in **Python**[145]. As an open source, platform independent language, Python has one of the largest communities in the world. This brings the benefit of enormous third-party packages. Most of the functions that are required by the data processing in this thesis study are readily available from the well-established packages. For example, the Discrete Fourier Transfer (DFT) function can be found from **NumPy** package[146]. This advantage can greatly shorten the script development time. Python as an interpretive language is not great in computing efficiency compared to the compiling language like C or C++. However, computing time of data processing is not the primary concern for this project. The focus is to demonstrate the effectiveness of the filter algorithm and explore the influence of different filter parameters.

The Python scripts are developed in **Jupyter Notebook**[147]. The notebook can run on any standard browser as a web page. The Jupyter console running in the background provides all the interactive interpretation. Jupyter Notebook features what you see is what you get. This makes it perfect to try out different ideas quickly and debug the scripts in real time.

5.4.1 Deconvolution + Wiener Filter implementation

One Python script is developed to demonstrate the implementation of the Wiener + Deconvolution filter. Figure 5-12 shows the principal data flow of the script.

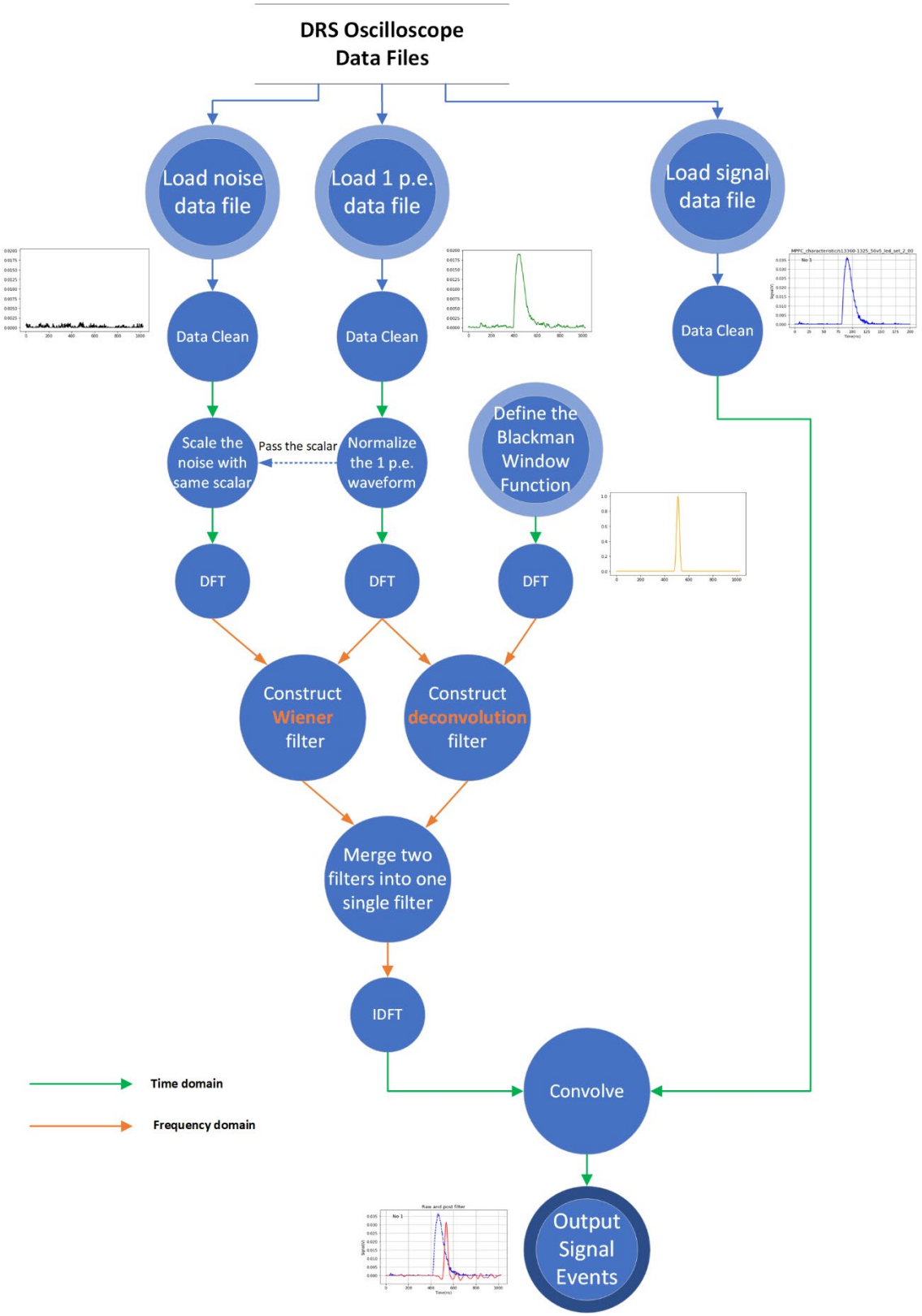
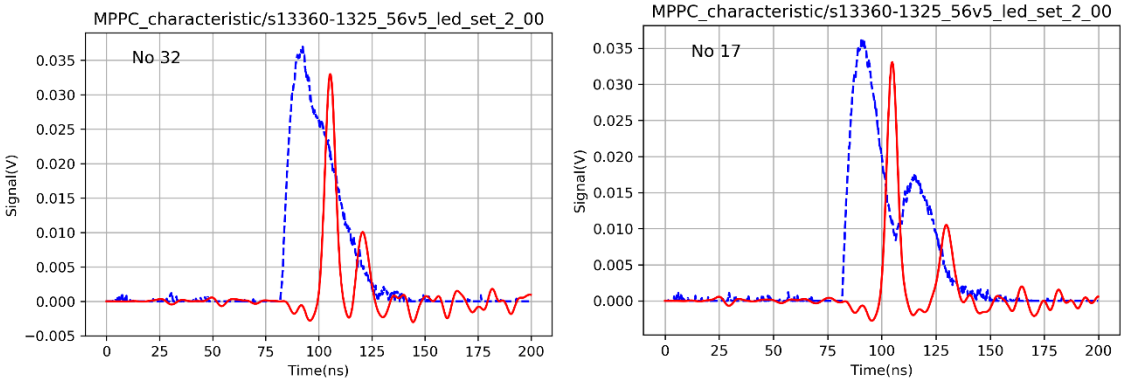


Figure 5-12 Data flow of the Python script for Deconvolution and Wiener filter

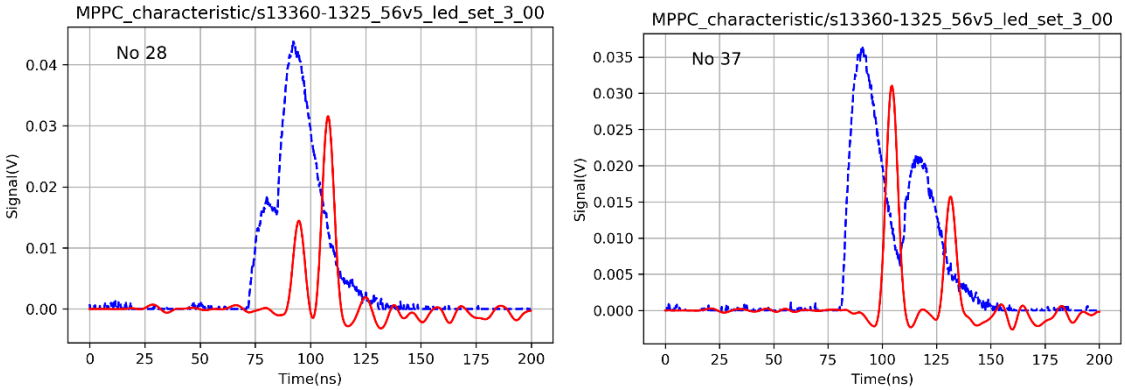
The script first loads three data files containing different waveforms. The first is the noise data file. This is acquired when biasing the SiPM just under its breakdown voltage. This file shall contain no signal but only white noise (a spectrum independent

noise). The second is the 1p.e. signal data file. This can be acquired when the LED signal is off and SiPM is biased at normal operating voltage. Signal acquired in this condition is mostly 1p.e. signal but occasionally 2p.e. or higher signal can occur. Since only one 1p.e. waveform is required to construct the filter, it is feasible to manually pick up one correct 1p.e. waveform from the signal data file. The third is the normal signal data file. This contains the working signal waveforms to be processed. The signal file in this chapter study is acquired under a LED illumination.

Figure 5-13 shows some sample results of the Deconvolution+Wiener filter on working signal waveform. These waveforms are sampled at different LED setting (brightness).



LED set = 2.00



LED set = 3.00

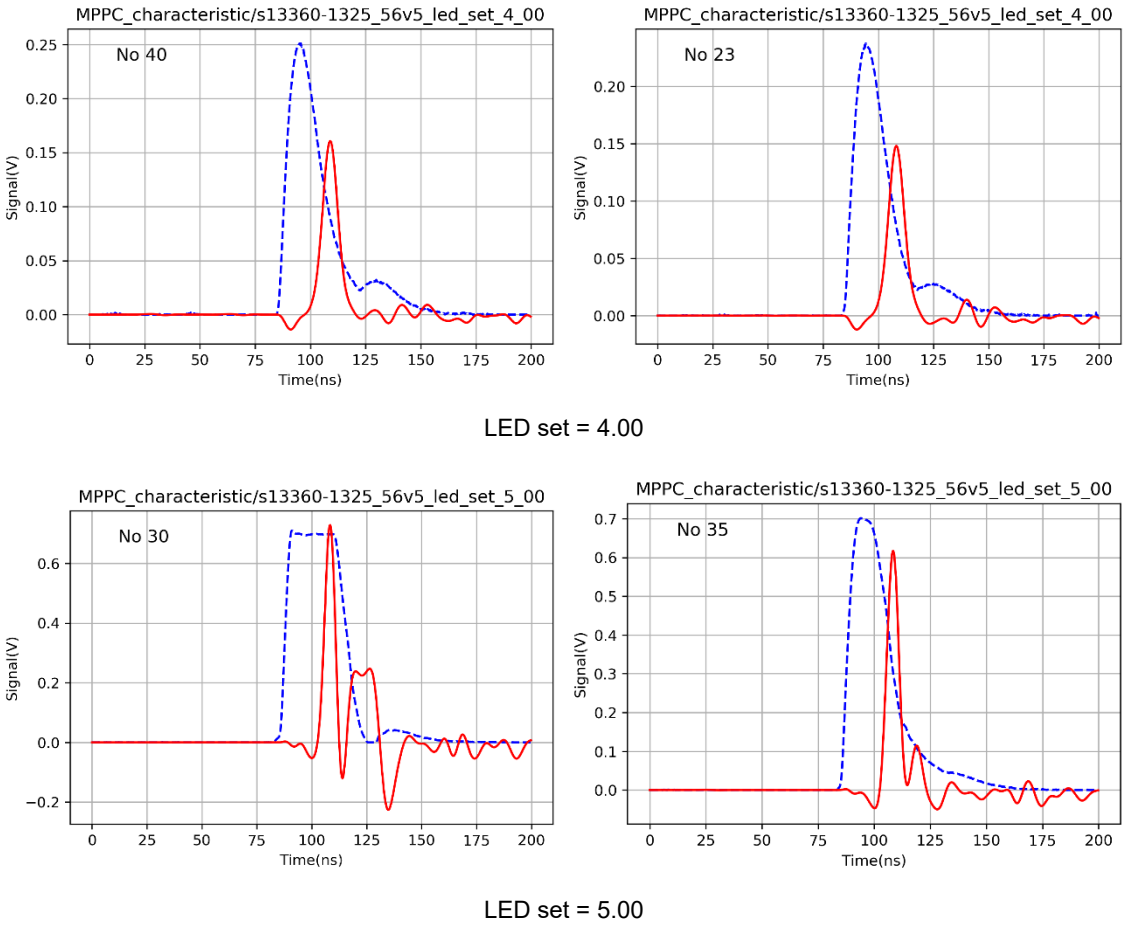


Figure 5-13 Waveform examples at different LED setting. The dotted blue waveform is the original signal. The red solid waveform is after the Deconvolution+Wiener filter.

The filter parameter setting is kept unchanged for all the signal waveforms while the signal amplitude is changing under different LED setting. The results show that this filter parameter setting works well for low amplitude signal but not equally well for the high amplitude signal. This finding may suggest that the filter parameters shall be adjusted according to the signal amplitude. Exactly how to adjust the filter parameters in relation to the signal amplitude are not explored in this project. This is left for a future study.

5.4.2 Pulse waveform parameters and its histogram

Another Python script is developed to help to understand the system linearity and dynamic range. This script can automatically scan through all the events stored in a data file and extract key pulse parameters like peak amplitude, rise time, falling time and total charge from each event. One event might contain more than one pulse. Each pulse will be analysed. Then, the script can create the histogram of each parameters

with all the pulse entry. Figure 5-14 shows how an event looks like and the calculated parameters are shown on the same graph. Note, an event is presented and processed as one waveform frame that the DRS4 module can capture at a time.

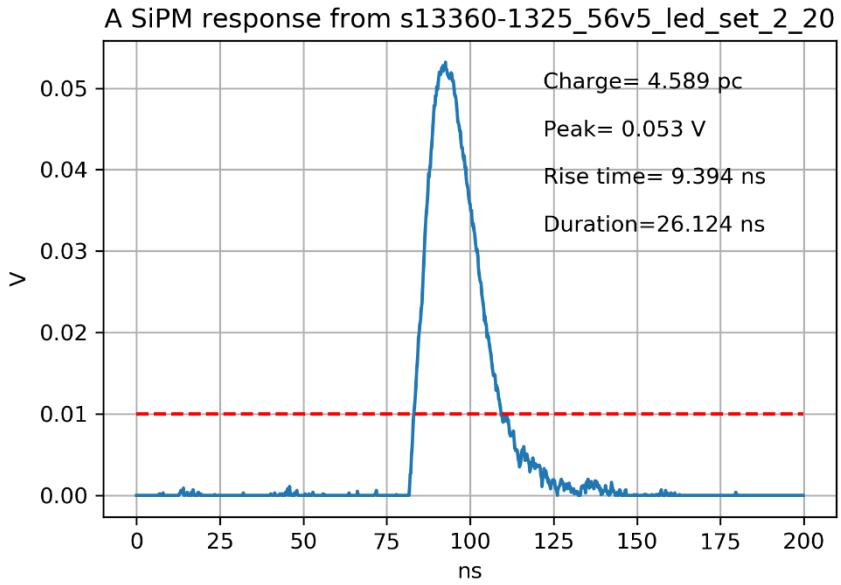


Figure 5-14 Example waveform with the prime pulse parameter calculation information

From the pulse waveform, initially four parameters are extracted. They are pulse **charge**, **peak amplitude**, **rise time** and **duration**. A threshold (red dotted line) is chosen so that only the waveforms above the threshold are counted. The threshold is chosen at about half of 1p.e. signal peak. Our first version of parameter extraction algorithm is very basic.

The assumption is each signal pulse will intersect with the threshold line twice. The waveform between the two intersection points is defined as the pulse window. In this window, the maximum point is defined as **peak amplitude**. The time between peak point and the window starting point is defined as **rise time**. The length of the window is called **duration**. The area formed by the waveform and the x-axis within the window can be calculated by simply sum all the waveform points. The sum divided by 50ohm is defined as **charge**. Note the pulse under study here is the output of the preamplifier circuit, not the original SiPM signal. So, the **charge** is not the represent of SiPM signal charge, however they are related. The definitions of the parameters are visually shown in Figure 5-15.

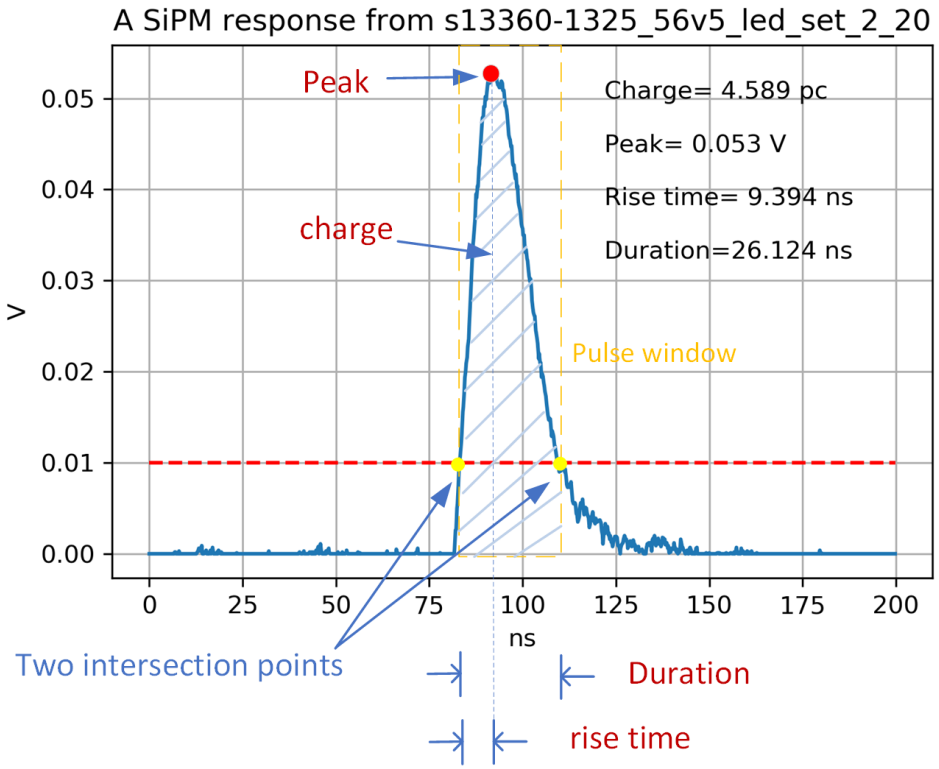


Figure 5-15 Illustration of the definition of peak, charge, rise time and duration for a typical pulse

The algorithm of calculating the four parameters just follow from their definitions. It works well for most of the pulses with the typical shape as the example in Figure 5-15. It is vital to ensure the two intersection points are correctly picked up. If there is glitch in the waveform just around the intersection points, mis-judgement by the algorithm will occur. This raises the demand to smooth the waveform to remove the glitches before the calculation. The smooth filter we used is the simple **Median filter** [148], with the filter window length set at 3 or 5.

One data set file often contains thousands of events. For each pulse in each event, the four parameters are calculated and saved. This is done in batch processing. Once all the parameters of all the pulses have been extracted, the collected parameter data sets are plotted in histogram. The histogram provides a visual interpretation of the system response statistically. Figure 5-16,Figure 5-17 and Figure 5-18 shows the example of the histogram of the peak, charge and rise time parameter for one experiment data file.

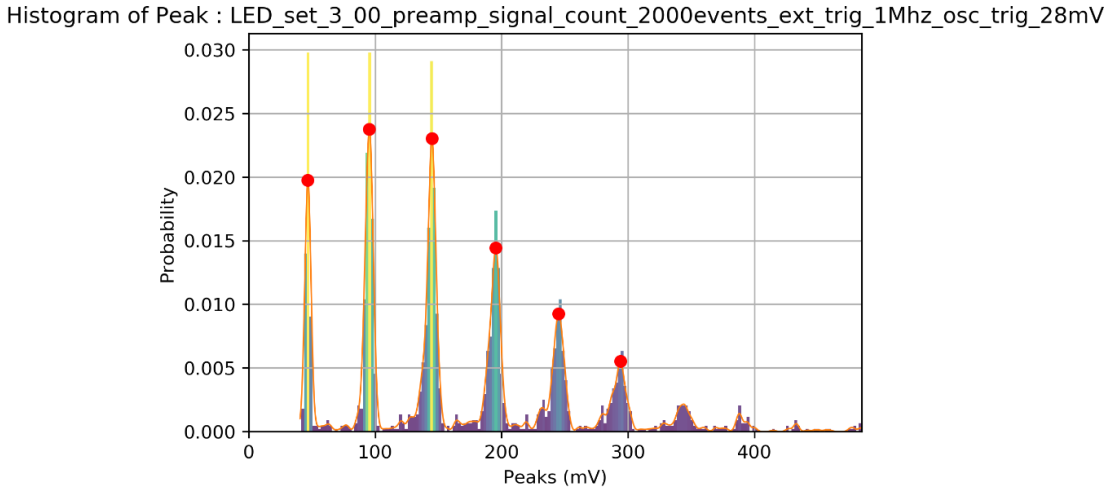


Figure 5-16 An example of histogram of pulse peak

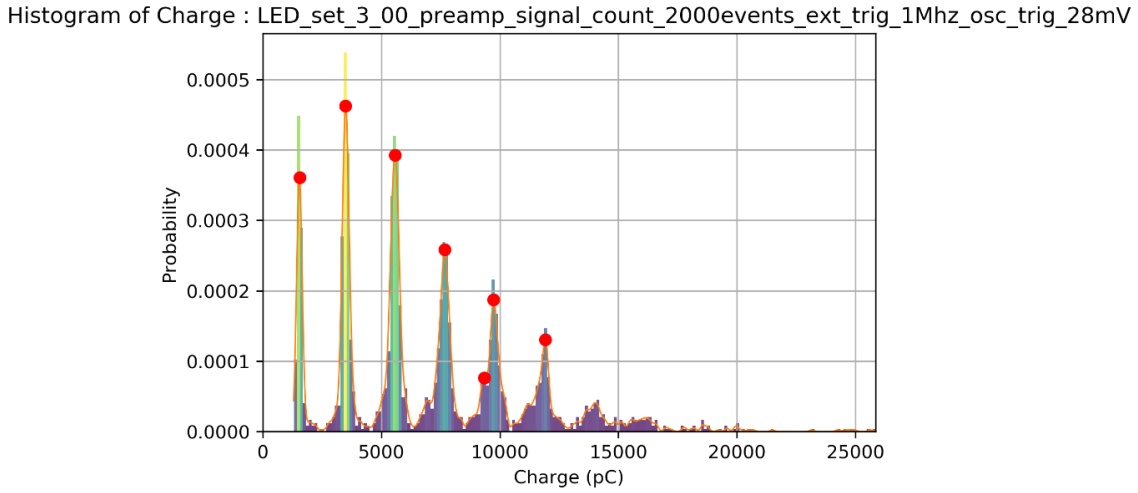


Figure 5-17 An example of histogram of pulse charge

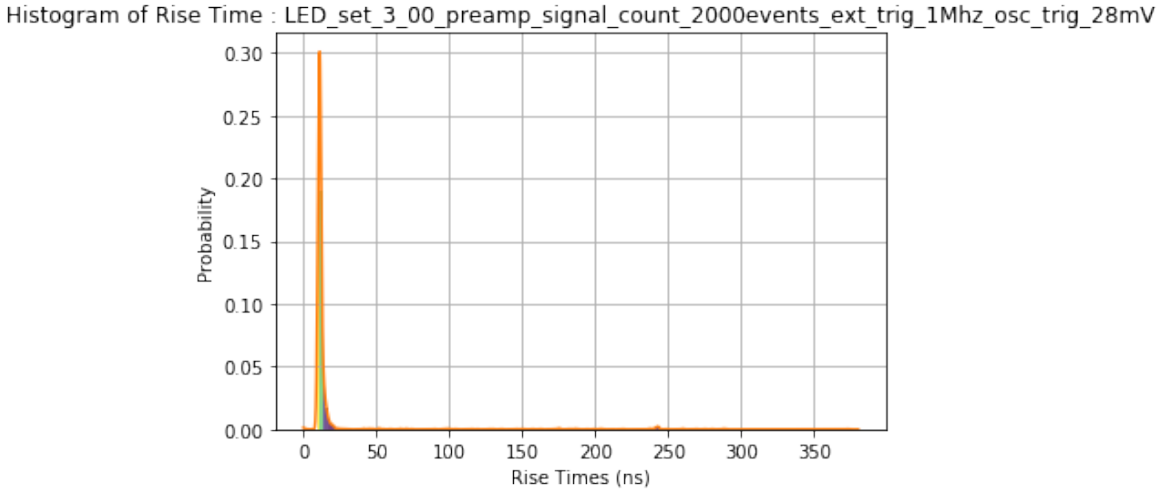


Figure 5-18 An example of histogram of pulse rise time

The histogram is constructed using the **hist()** function from the **Matplotlib** package [149] for Python. Each bin is colour coded corresponding to its height. Note, the y-axis of each histogram is normalized and displayed as probability, not as frequency.

The key parameter for the histogram construction is the bin size (this determines the number of equal-width bins for the entire sample population). Choosing the right bin size to recover the true data distribution can be quite technical. If the chosen bin size is too small, the structure of the data distribution will be distorted. If it is too big, the height of the individual bin will be impacted by sampling error as each bin doesn't contain enough data sample to smooth out the individual sample's error. There are two popular methods named as Scott rule[150] and Freedman&Diaconis rule[151] for tuning the optimal number of bins. Here by practice, a bin size at a fraction of about 1/20 – 1/50 of the total number of sample population is found to work well.

5.4.3 Histogram Analysis

The histogram can reveal the feature of the data that can't be clearly seen from the time series data (waveform). In Figure 5-16, it is observed the waveform amplitudes concentrate in a few discrete Gaussian shape pulses with almost equal gap between them. This observation matches the nature of the SiPM device, which sums all the pixels signal as a single output. Each pixel signal has an almost identical amplitude. Hence, the final output signal waveform has an amplitude proportional to the number of pixels triggered.

With this knowledge, the histogram peak can be easily interpreted as the first peak (from the left) represents 1p.e. output, the second peak as 2p.e. output, the third peak as 3p.e. and so on. The average gap between the histogram peaks represents the system gain in mV/p.e. (Note, this is briefly touched in 4.5.5, but here we use the histogram method.)

The y-axis of a histogram plot means the probability of the data that lands in the corresponding bin at x-axis. Therefore, the height of each histogram peak means the likelihood of that p.e. signal in the data set. The breadth of the histogram peak suggests the fluctuation of the p.e. signal amplitude. As the preamplifier circuit and the digitization circuit have very good linearity, the fluctuation most likely originates from the SiPM pixel signal introduced by the internal avalanche process.

As the x-axis position of each histogram peak represents the most likely amplitude for the corresponding p.e. signal, it worth to find the x-axis position of these peaks.

5.4.3.1 Kernel Density Estimation (KDE)

In the histogram, the information that is of interest for further study is the x-axis value. For y-axis value, the absolute number is not very important. The Y-axis value is only needed to find the histogram peaks. Therefore, some data processing technique can be applied to smooth out the samples to assist the peak searching but at the cost of losing the absolute y-axis value. The technique used here is called **Kernel Density Estimation** (KDE) [152]. This is a way to estimate the Probability Density Function (PDF) of the variable that the data sample represents. The general mathematic definition of KDE method is:

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h}) \tag{4.104}$$

Where *K* is the kernel function and *h* > 0 is called the bandwidth, which is a parameter that can strongly influence the estimation results. The optimization of the selection of *h* is beyond the discussion of the thesis. The method chosen in the implementation is called Scott rule [150] offered by the **gaussian_kde()** function from Python **SciPy** package[153].

The orange smooth curves shown in Figure 5-16,Figure 5-17 and Figure 5-18 are the outcome of the KDE method on the raw data samples. The KDE curve removes the small glitches and randomness of the raw data samples. It should be noted that the KDE method is fundamentally different from those noise filtering methods. The noise filtering methods can only reduce the noise at the cost of losing some of the signal at the same time. The KDE method creates a linear distribution function that is closely matching the distribution of the data sample. The created function is perfectly smooth and free of noise.

Then the KDE curve is ready for next step of processing: peak search.

5.4.3.2 Peak search and linear fit

The peak search algorithm selected here is from the Python **PeakUtils** package [154]. This package is based on the first-order difference method to find all the peaks. This method is better than the max searching method used in 5.4.2. The max searching method simply finds the maximum point within a specified window. Within that window,

only a single highest point will be identified. The first-order difference method is based on the fact that the first-order difference of the waveform at the peak point is zero. This method allows the detection of multiple peaks. While the method also allows the detection of waveform valley, this error can be excluded by checking the signs of the first-order difference before and after the zero point. This correction is implemented in the package. Another source of error is the wrong picking of noise spike particularly around the base of the histogram peak. This can be mitigated by some PeakUtils package built-in functions, like minimum peak threshold, minimum distance between two peaks, the compensation for the baseline distortion etc.

The red dots in Figure 5-16,Figure 5-17 and Figure 5-18 represent the peak search results. The peak search detection is on the KDE smoothed curve. So, most of the search results are accurate. There is still some small errors as missing peak or false peak. This is because the data is highly dynamic, which causes some KDE mis-filtering and the consequent peak search failures. The static search setting (minimum threshold and minimum distance) can hardly be one fit for all cases. Maybe some dynamic setting approach can be explored. Such task will be left for the future research.

The collected peak data points are plotted in a 1D series graph. Figure 5-19 illustrates the principle of how to map the histogram peak x-axis values to the 1D series graph.

The dashed line is the linear regression fit for the data points. The work uses the **PolyFit()** function from the **NumPy** package. The better the fitting result suggests the better the linearity of the data acquisition system is. The **slope** of the fitting line is the **gain** of the system. As the gain of the analogue to digital conversion (DRS4 module) is 1, the gain of the preamplifier circuit represents the gain of the full data acquisition system. The **cross** of the fitting line means the y-axis value when the input signal is zero (=0 p.e.). This is the **offset** of the full data acquisition system. The DRS4 module has very small contribution to the offset, so the offset is dominant by the preamplifier circuit. In the top figure of Figure 5-19, it shows the slope at 0.015765V per p.e. and cross at 0.001017V. The merely 1mV offset is very small compared to the amplitude of a typical 1p.e. at around 15.7mV.

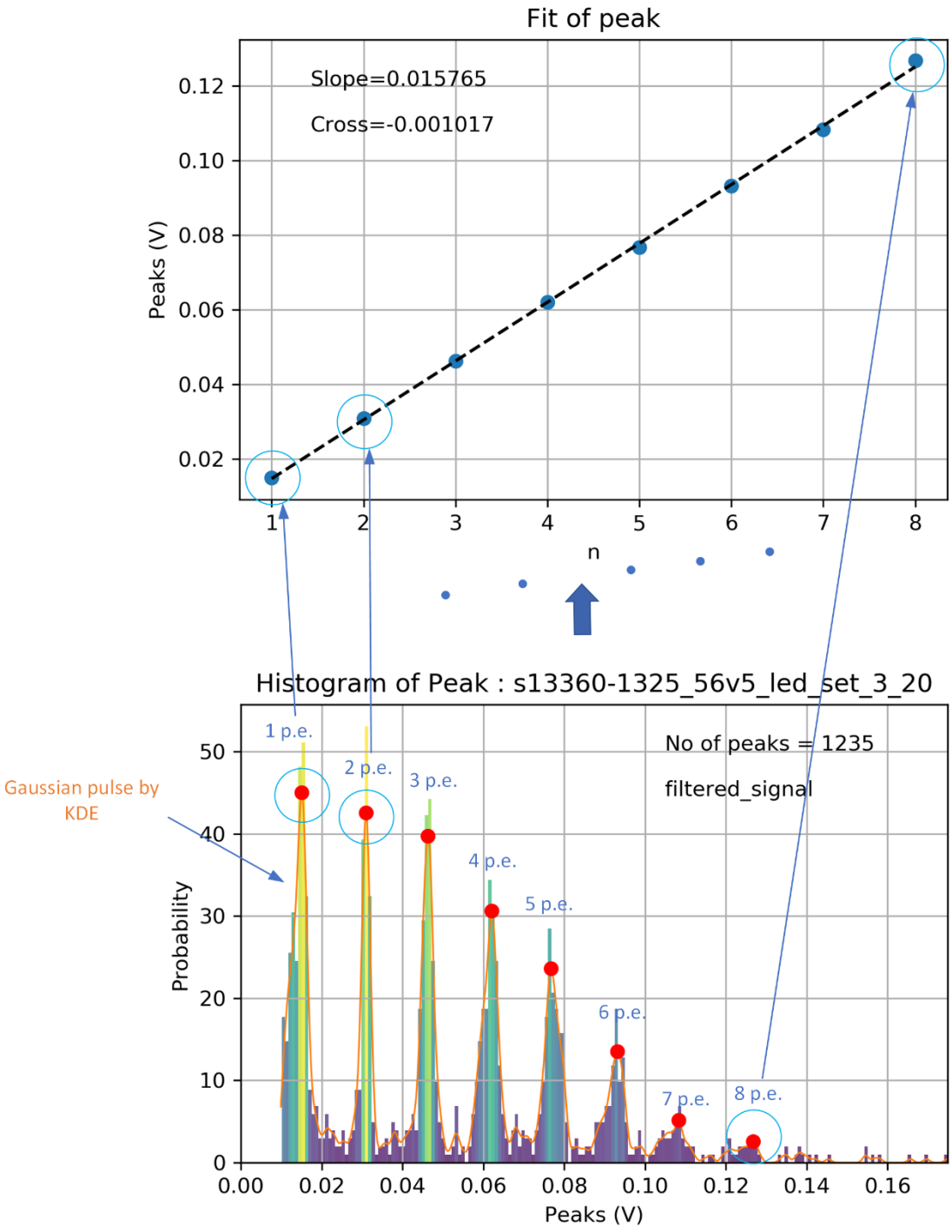


Figure 5-19 Plot the Gaussian pulse peaks in the histogram as a 1-D line

Besides the histogram of pulse amplitude, the study also covers the histogram of the charge parameter. Figure 5-20 shows the similar fit of histogram peaks from the charge histogram in Figure 5-17. In this case, the linear regression fit shows poorer result. This suggests it is less accurate to use the charge parameter to estimate the readout circuit gain. The reason can be the method used to calculate the charge in section 5.2 is too coarse, particularly when there is signal pile-up, the result of the

calculated pulse charge is hugely misleading as an example shown in Figure 5-21. In this example, the two peaks can still be identified, but the calculated charge is the merge of two pulses body which does not equal the sum of two pulse charge if the pulses were separate.

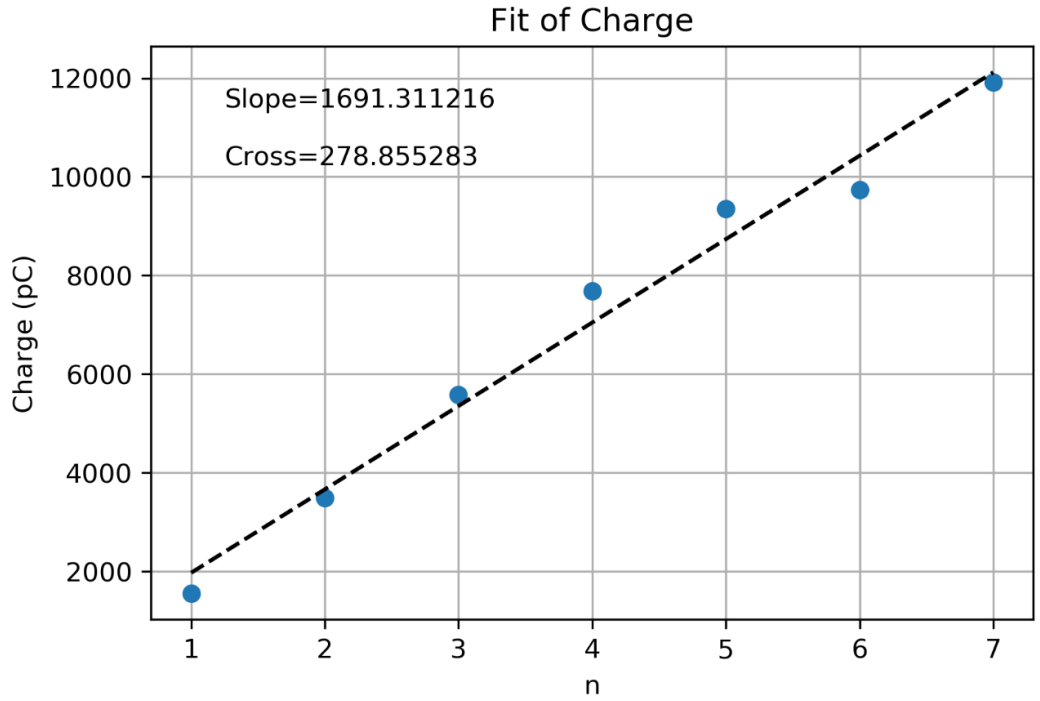


Figure 5-20 The fit curve of the histogram peak points [x-axis value] for the signal pulse charge

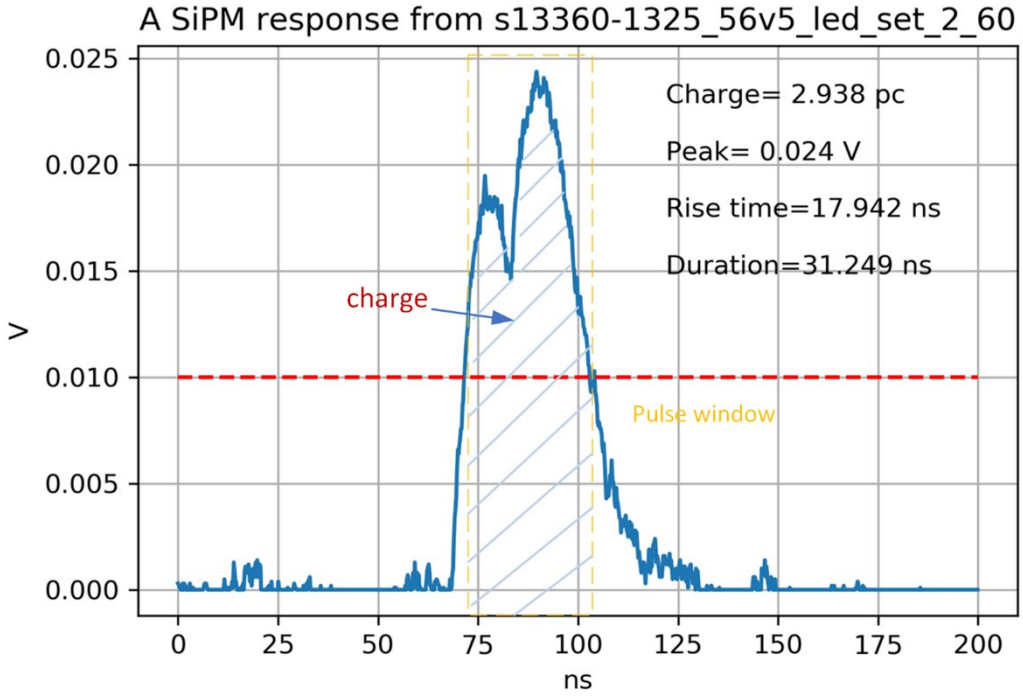


Figure 5-21 An example of misleading charge parameter in a pile-up event.

Although not as accurate as the peak parameter, the charge parameter is more useful at higher photon flux condition, where the pulse peak saturates and stops responding to further incident photon flux increment. While, the charge parameter can continue to respond but not in a linear way.

As an independent check, an experiment is set to use the DRS4 module PC software to calculate the output signal waveform peak and charge parameters. The experiment sweeps the LED driver setting from step 0 to 100 and records the signal peak and charge at each step. The results are plotted in Figure 5-22.

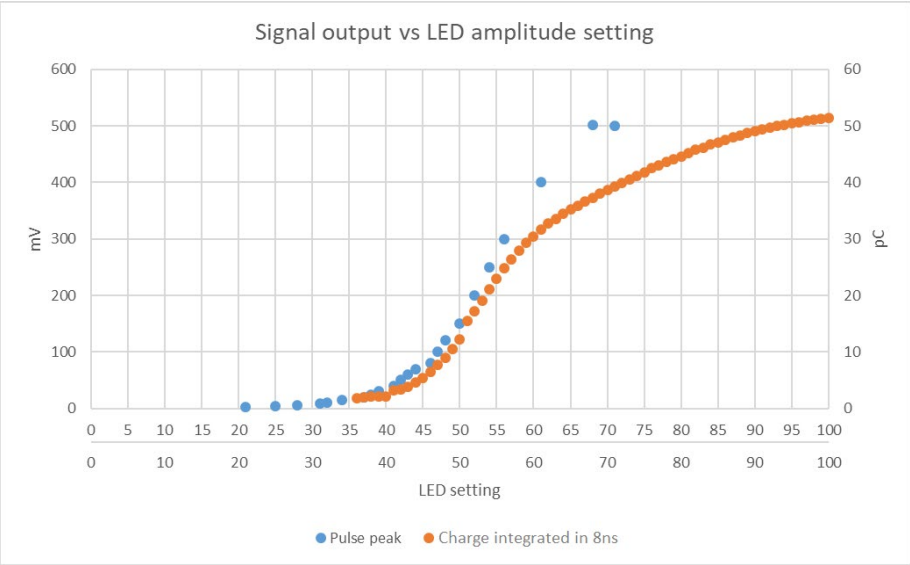


Figure 5-22 the plot of system output signal peak and charge against the LED brightness

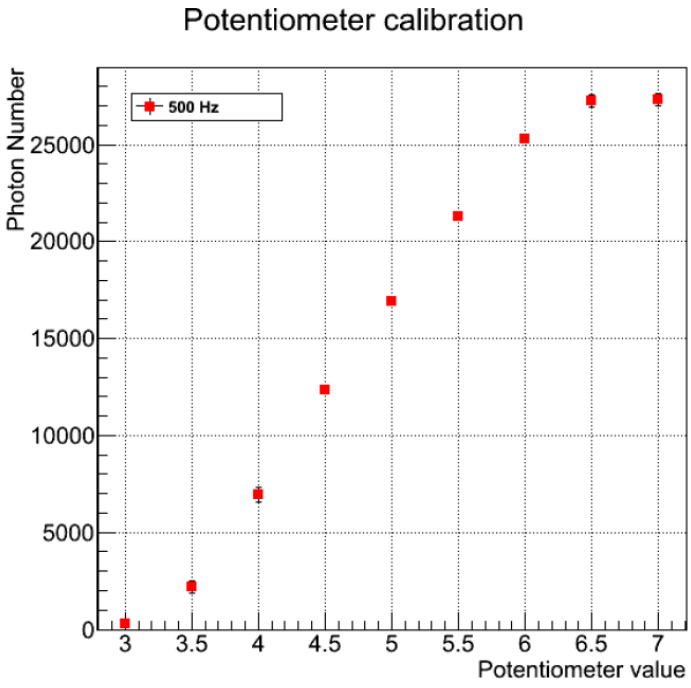


Figure 5-23 the plot of output photon flux against the LED setting from the LED driver datasheet

In Figure 5-22, the blue curve is the response of peak parameter and the orange curve is the response of charge parameter. Two curves track each other during the low LED setting range up to setting of 55. After that, two curves start separating. The peak curve stops increasing after setting of 70, but the charge curve continues to increase. However, the increase rate is slowing down.

Figure 5-23 is a calibration plot from the LED source manufacturer. The plot shows the relationship between LED setting and the output photon flux. The x-axis of Figure 5-22 and Figure 5-23 are the same but expressed in different scale with Figure 5-22's value 10 times that of Figure 5-23.

It is interesting to compare Figure 5-22 and Figure 5-23. It is found the peak curve in Figure 5-22 is very similar to the photon number curve in Figure 5-23. The close match suggests the good linearity between the SiPM readout system output and the LED driver output photon (the input to the SiPM system) when measured with peak parameter.

Although the dynamic range is wider when the SiPM system output is measured with charge parameter, the linearity is quite poor after the photon flux reaches a certain level. Therefore, for the system characterization discussed in the next chapter, peak parameter is our first choice.

5.4.4 Mapping signal amplitude peak to p.e. number

Due to the nature fluctuation of the avalanche process, the amplitude of the SiPM signal follows a gaussian distribution. The good linearity of the preamplifier can prevent from adding further distortion but the intrinsic SiPM signal fluctuation can't be reduced. Hence, the amplitude is not very suitable for pulse height analysis. It is best used to interpret the number of p.e.

With the proper avalanche gain of SiPM (by setting the deltaV) and a proper preamplifier gain, the amplitude of different p.e. signal can be clearly separated even with the avalanche fluctuation as shown in Figure 5-19. Figure 5-20 shows the relationship between the signal amplitude to p.e. number is a simple first order linear function. The function shall have crossed the origin (assume the small system offset is neglectable.).

So, the conversion of signal peak to p.e. number can be a very simple ladder comparison as shown in Figure 5-24. The procedure is first to acquire the 1p.e. signal peak from some training data. Then use this 1p.e. signal to construct a ladder with each step as high as 1p.e. amplitude. All the thresholds are set at half of 1p.e. position. If the signal falls between two neighbouring thresholds, the output is the p.e. number between the threshold. For example, if a signal falls between 0.5p.e. and 1.5p.e. , the output is 1p.e. The signal below 0.5p.e. is counted as noise and recorded as 0p.e. This method is crude and simple, but also effective. It is important to keep the SiPM gain and preamplifier gain stable across the operation temperature range, because any shift in the system gain will distort the constructed comparison ladder.

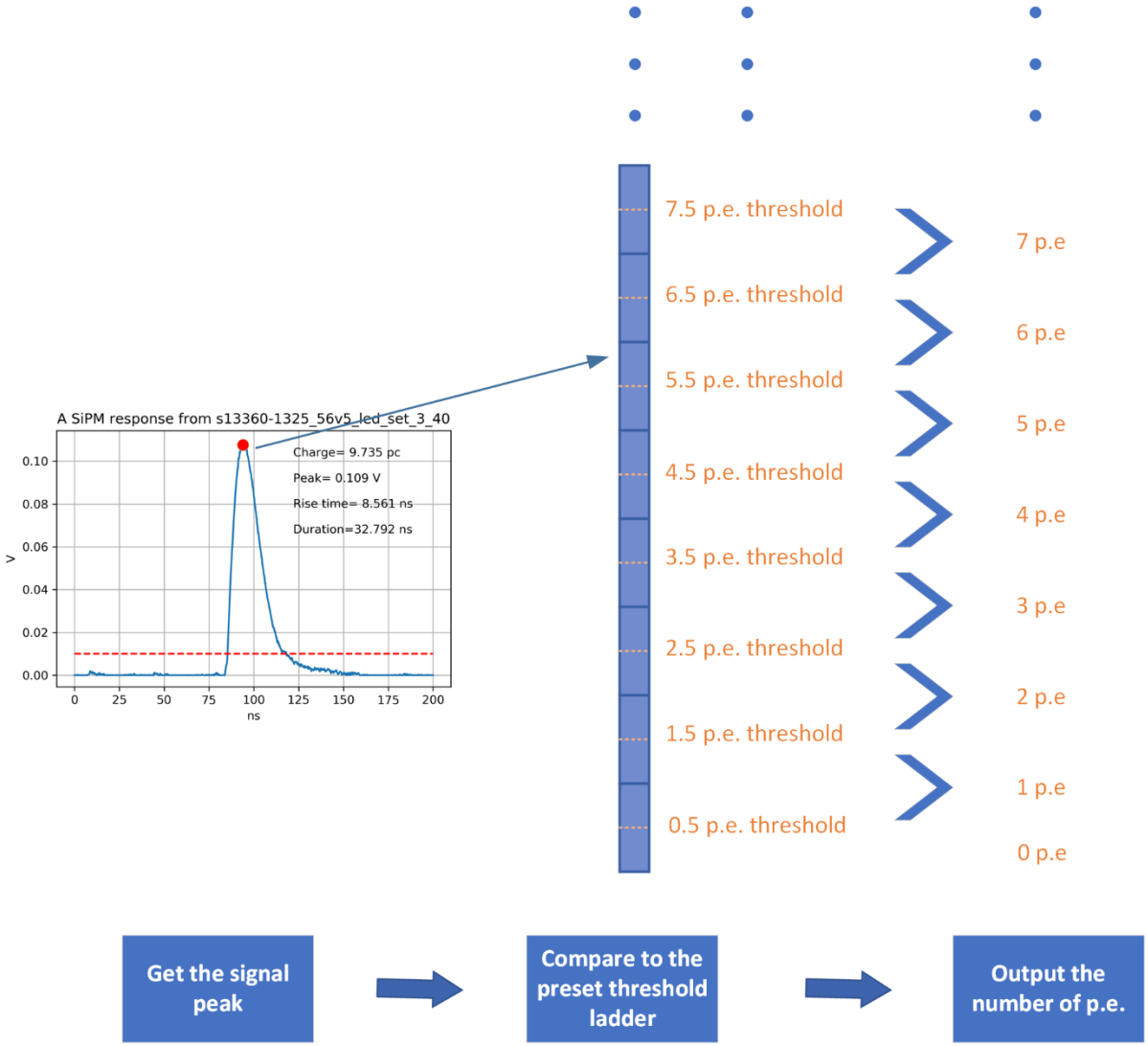


Figure 5-24 the illustration of process and principle of converting the signal peak to the p.e. number

Use this method, the difficulty to process analogue histogram is addressed by conversion to a discrete digital histogram. Figure 5-25 shows some conversion examples:

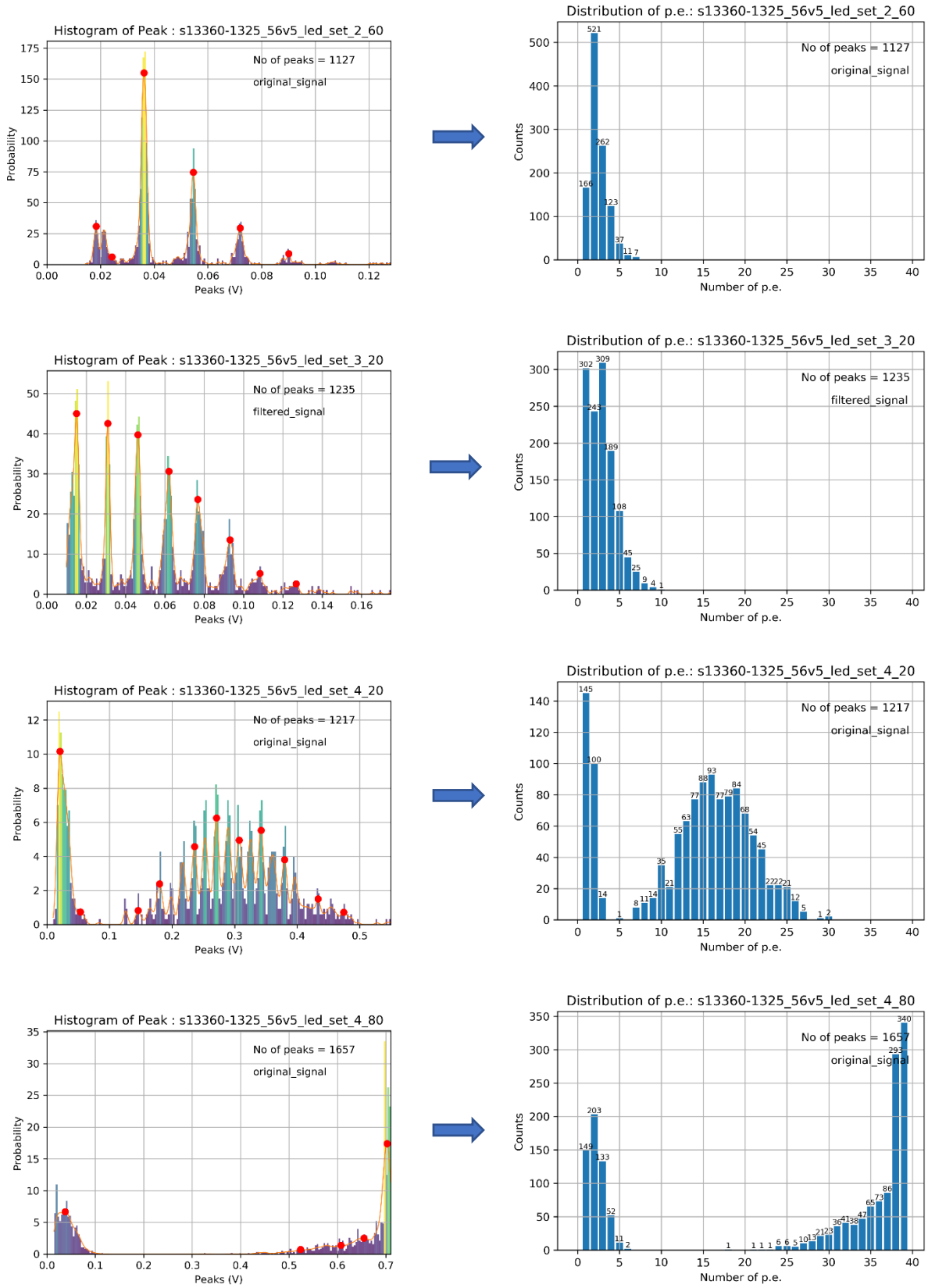
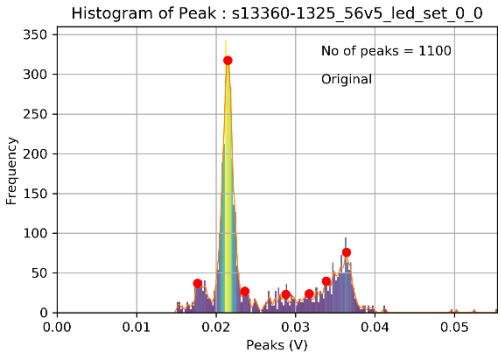
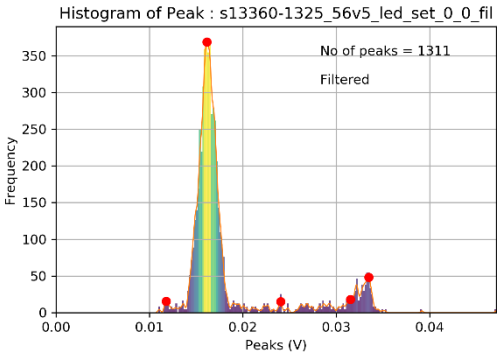


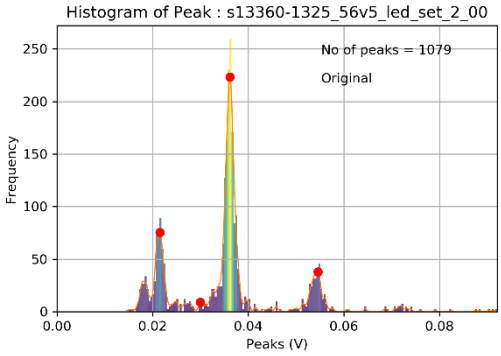
Figure 5-25 Examples of mapping analogue peak histogram to the discrete digital histogram in p.e. number

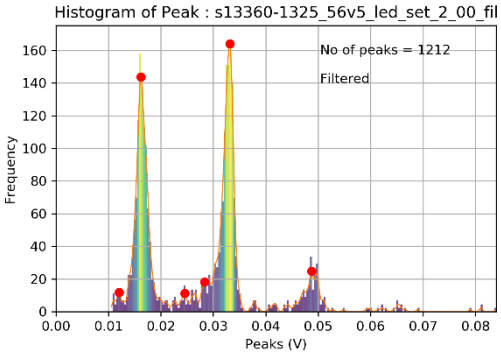
The histogram of p.e. number is much clearer than the histogram of the signal peak. The method has a digitization error close to each threshold. A signal just under a

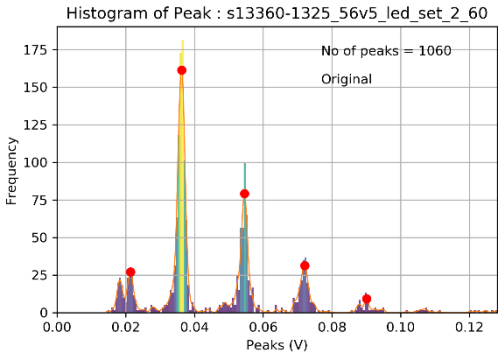
threshold might be flipped above the threshold by random noise. In this case, the output p.e. number is one digit above its true value. We called this digitization error and it shall be no more than 1p.e. as long as the maximum noise is kept below 50% of 1p.e. signal. The frequency of such error is related to how wide the gap is between SiPM signal amplitude. In Figure 5-25, it is found that at low light condition, the gap between histogram pulse is very clear but at intense light condition, the gap is much less clear. There might be a few reasons behind it. First is the fixed bin numbers for different amplitude scale. At intense light, the SiPM signal amplitude is much larger compared to the low light condition. With the same bin number, this means the resolution per bin for large signal is poorer. This will result in reduced gap. A second reason could be for each light condition, the same amount of SiPM events are acquired. The fixed amount might be sufficient to show the histogram structure for low light but not sufficient to show that of intense light, where the range of signal amplitude is much wider. The third reason could be the baseline shifting at the intensive light. This shift is due to the readout system can't reset back to the default baseline before the next signal is triggered. This is explained at the beginning of the chapter for signal pile-up effect. Another interesting observation of Figure 5-25 is for intense light, the 1p.e. and 2p.e. signal is blurred together and appeared to be much broader than the corresponding signal at low light condition. The first and third reasons discussed above shall apply to this phenomenon. The third reason might be more dominant as the shifted baseline, which varies randomly depending on the signal intensity, will lead to shifting 1p.e. and 2p.e. signal amplitudes. A signal pile-up removal algorithm discussed in this chapter shall help to address this issue. However, our first attempt at filtering software works well for low light, but for intense light, the effect is not that obvious. The histogram display at intense light is largely distorted by the dominant 1p.e. pulse. One positive outcome is very clear that after applying the filters, more SiPM signals are detected (see the no of peaks statistics in the plot). Some results are presented in Figure 5-26 with left column represents the original signal and the right column represents the filtered signal.

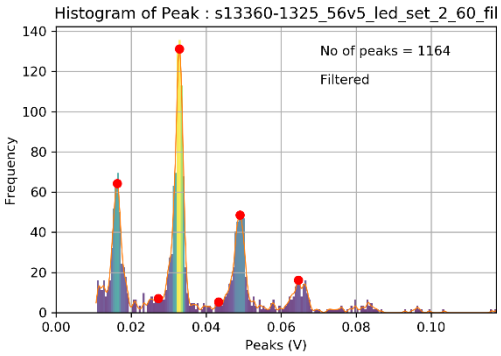

Histogram of Peak : s13360-1325_56v5_led_set_0_0
No of peaks = 1100
Original
Frequency
Peaks (V)

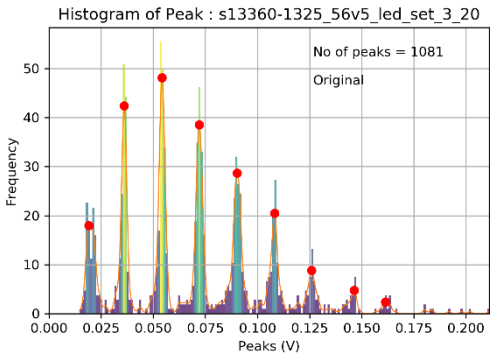

Histogram of Peak : s13360-1325_56v5_led_set_0_0_fil
No of peaks = 1311
Filtered
Frequency
Peaks (V)

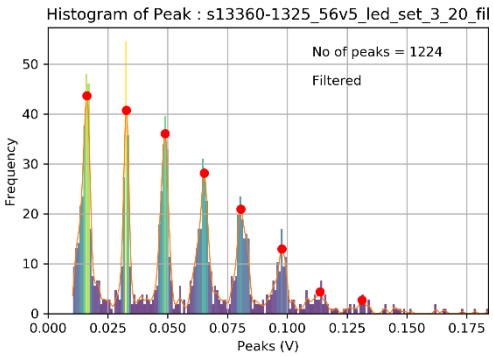

Histogram of Peak : s13360-1325_56v5_led_set_2_00
No of peaks = 1079
Original
Frequency
Peaks (V)


Histogram of Peak : s13360-1325_56v5_led_set_2_00_fil
No of peaks = 1212
Filtered
Frequency
Peaks (V)


Histogram of Peak : s13360-1325_56v5_led_set_2_60
No of peaks = 1060
Original
Frequency
Peaks (V)


Histogram of Peak : s13360-1325_56v5_led_set_2_60_fil
No of peaks = 1164
Filtered
Frequency
Peaks (V)


Histogram of Peak : s13360-1325_56v5_led_set_3_20
No of peaks = 1081
Original
Frequency
Peaks (V)


Histogram of Peak : s13360-1325_56v5_led_set_3_20_fil
No of peaks = 1224
Filtered
Frequency
Peaks (V)

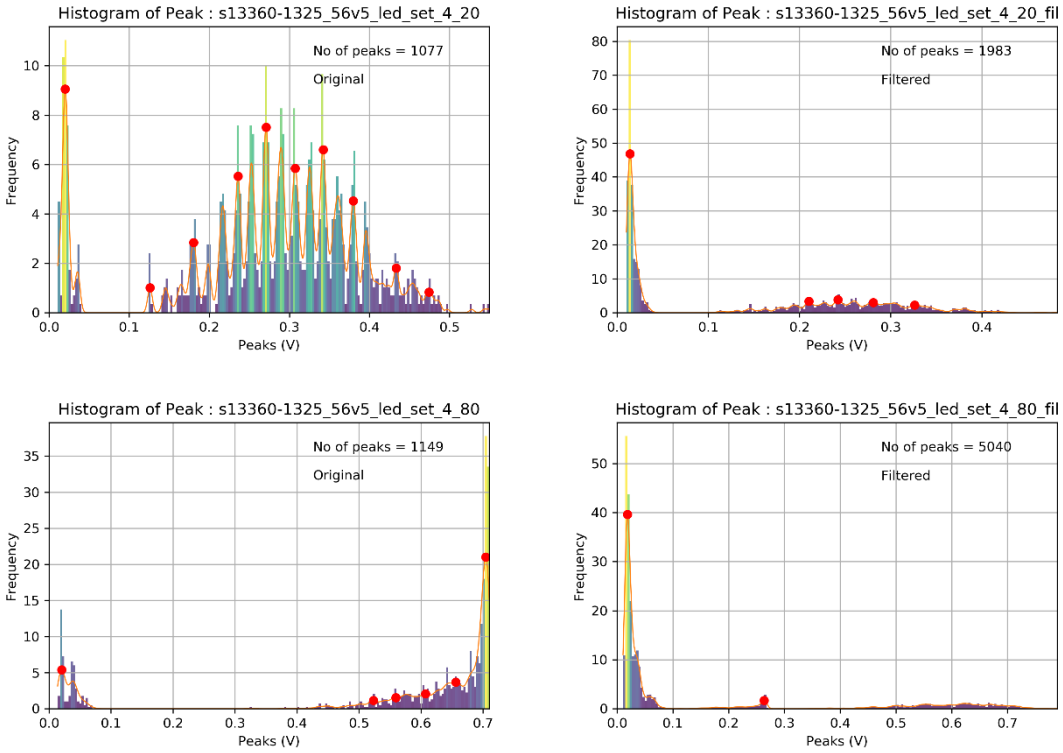


Figure 5-26 The comparison of SiPM signal peak histogram before the Deconvolution+Wiener filter processing (left) and after (right).

5.5 Conclusion and future work

This chapter discussed the system data processing (software aspect of this research work). This includes the Deconvolution and Wiener filter, the waveform parameters extraction algorithm, the histogram analysis tools, the KDE pulse fitting algorithm and pulse peak search algorithm. All the processing scripts are developed in Python language on Jupyter console. All the scrips are developed for concept demonstration purpose only with no performance consideration in mind. So, they are only suitable for post data processing on the PC but are not suitable for the real time processing in the onboard computer. Therefore, the next step will be to optimize the code to fit into the onboard real time processing environment. Besides, the key parameters of the filters which greatly determines the filter performance should be studied further to understand how to dynamically change them to fit the dynamic data.

Chapter 6 System Characterization

6.1 Introduction

This chapter describes the work based on the previous chapters. With all the tools (both hardware and software) ready, it was possible to make a major characterization of the SiPM detector system. The system is excited with a LED source powered by an adjustable LED driver.

This chapter starts with a review of the hardware and software setup and the limits these setup present particularly in timing. After the dataset overview, the system characterization mainly driven by the SiPM sensor characteristics are thoroughly discussed. Limited by the availability of equipment and the simplicity of test setup, not all the SiPM parameters can be measured very accurately. However, many characteristics have been successfully measured with good accuracy.

The key aspect of the system characteristics or performance is the dark count rate. This is studied in great depth. A statistic model is built to describe the dark count timing distribution based on a Poisson distribution. A software simulator is built on this statistical model. The simulator is useful to generate a large amount of data for signal timing analysis

Our system in normal configuration is not feasible to measure the dark count rate due to the timing limit. To correct this, we replace the DRS4 data acquisition module with a fast counter. This way, the dark count rate can be measured accurately.

6.2 Test setup and system limitation

6.2.1 Review of the data acquisition system hardware setup

In this research, a full chain data acquisition and processing system has been developed from scratch. Figure 6-1 shows the system architecture:

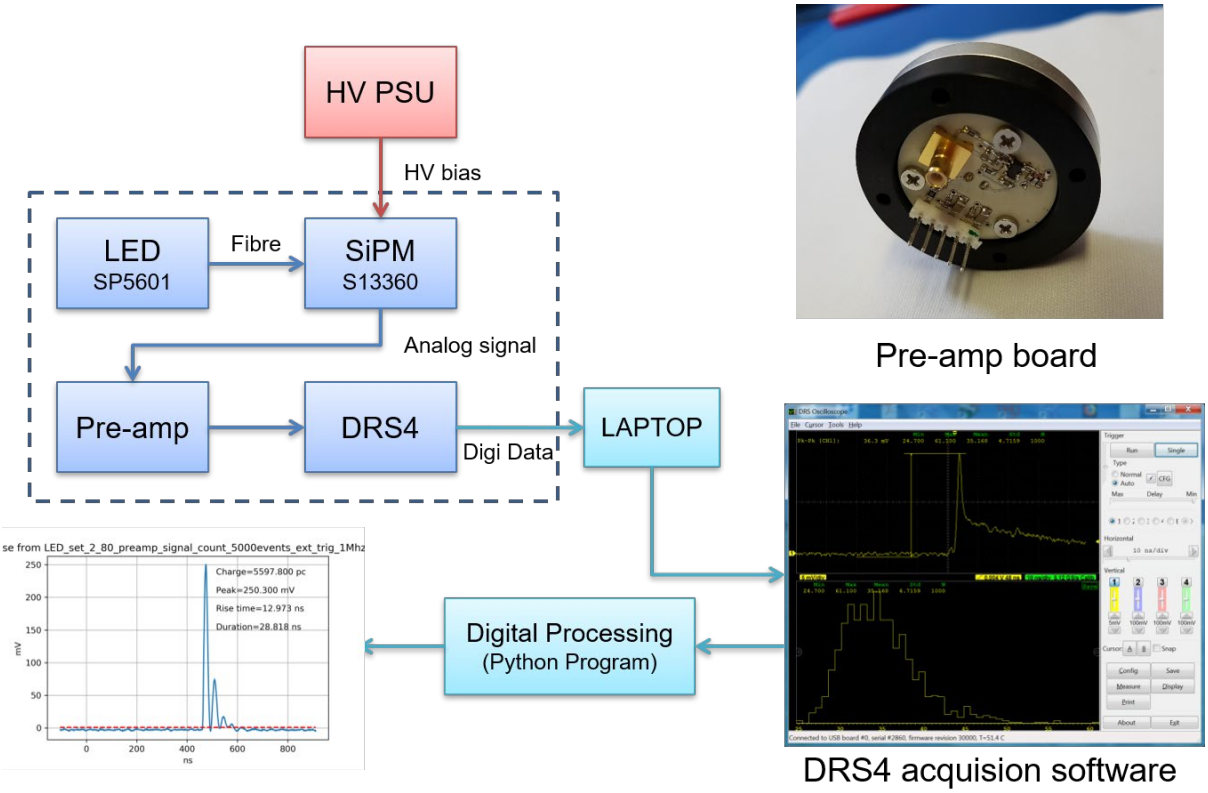


Figure 6-1 The full chain data acquisition and processing system for SiPM sensor

The high voltage power supply unit (HV-PSU) is a bench instrument, which provides the bias voltage for the SiPM. The temperature compensation for the breakdown voltage drift of the SiPM is not built into the HV-PSU. For the next system revision, a Hamamatsu power supply C11204-11 with temperature compensation function can replace this bench PSU.

The LED driver SP5601 from CAEN shown in Figure 6-2 including a LED source is used as stimulus signal. The driver is equipped with a 5mm super violet LED (part number OSSV5111A from OptoSupply), which has the peak wavelength of 405nm. The driver can clock the LED from 500Hz to 5MHz with a typical 8ns pulse width. This frequency can be externally synchronized. This LED driver provides an alternative test signal to the radioactive source. Unlike the radioactive source, a LED driver can be easily controlled and synchronized. It also has less safety hassle for operation. The

wavelength of the LED is chosen to be as close to the SiPM absorption spectrum peak as possible. The 8ns pulse width is not adjustable. The short duration limits the number of photons the LED produces in each pulse, but it is not short enough for more timing critical system testing such as coincidence function or time of flight application. For those more demanding application, a laser driver like the NPL series from THORLABS are more suitable, albeit 10x as expensive.

Figure 6-2 SP5601 LED driver with the SP5650 sensor holder (the round piece on the left). The left is the back of the SP5601 and the right is the front.

SP5601 can adjust the amplitude (which sets the output pulse light intensity) with a 10-turn dial. Each turn contains further 50 steps. So, in total, there are 500 steps of amplitude level. The SP5601 driver adjusts the LED current to control the output photon flux. This way it can achieve better control linearity than adjusting the LED voltage.

Figure 6-3 shows the output photons vs LED amplitude setting. Figure 6-4 shows the response of the LED output to various synchronization frequency. Both plots are provided by the manufacturer.

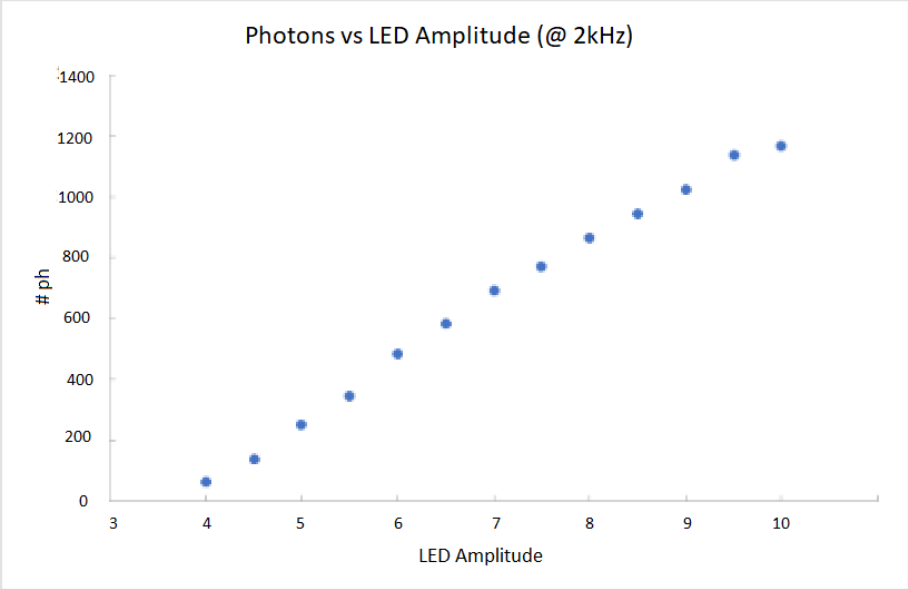


Figure 6-3 The manufacturer calibration of the photons output vs LED amplitude setting [155]

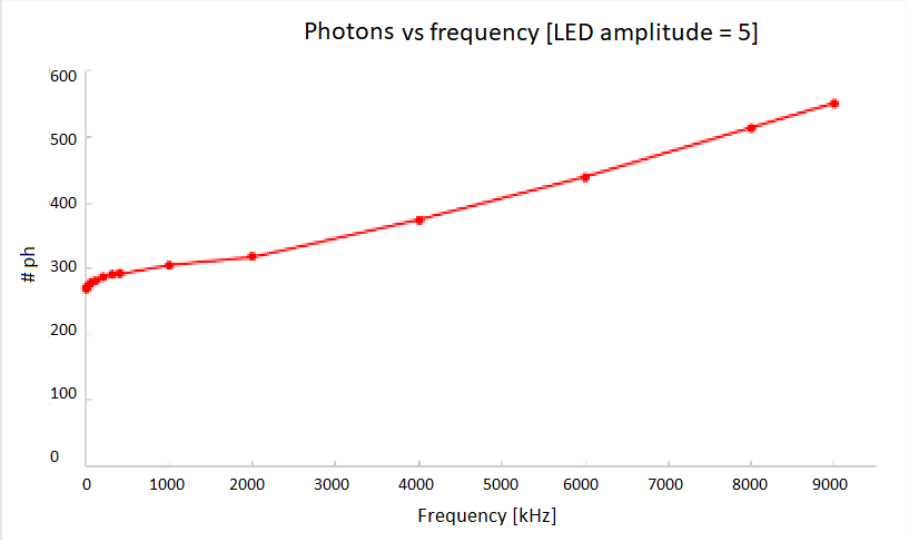


Figure 6-4 The manufacturer calibration of the influence of synchronization frequency to LED output photon number for SP5601 LED driver [155]

The output light is focused onto the SiPM by an optic fibre cable (model type JTFLH400430730 from Polymicro) as shown in Figure 6-5. A percentage of photon will be lost during the transmission. This is usually defined as the attenuation in dB/km.

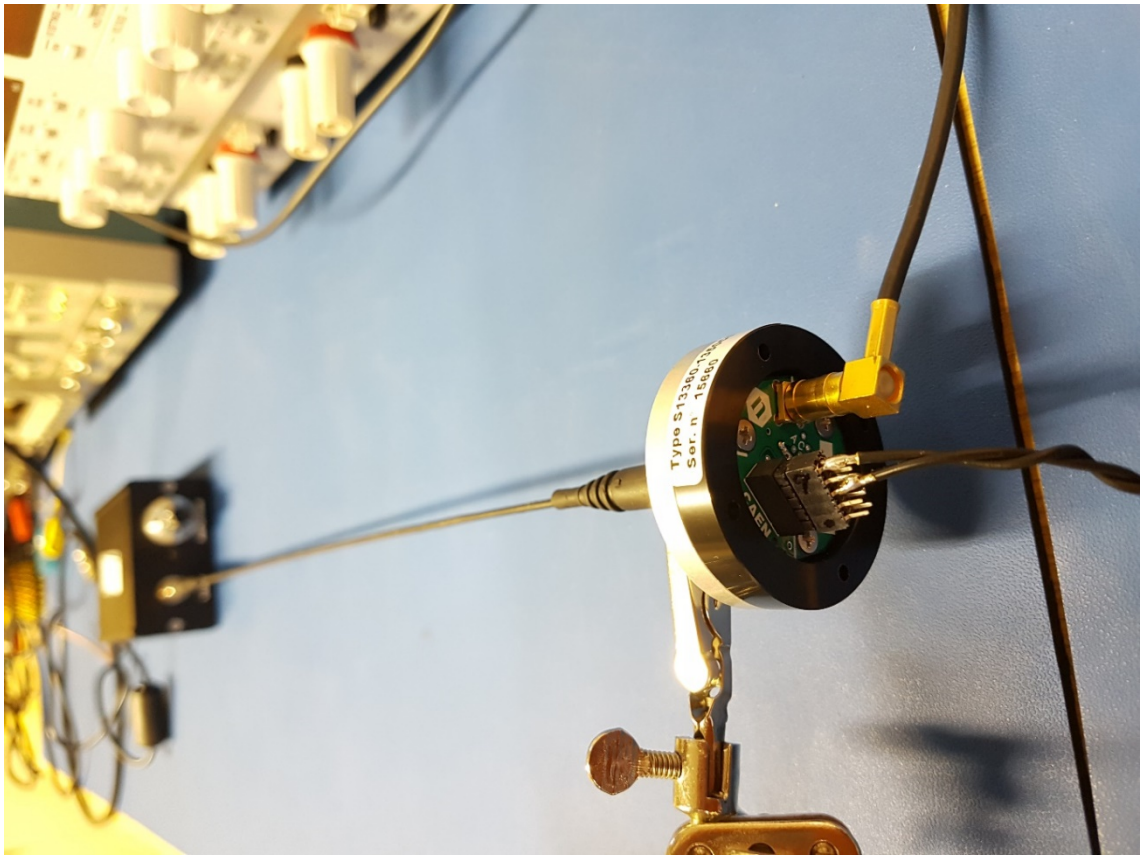


Figure 6-5 The optic fibre cable connects the SP5601 LED output to the SiPM sensor holder input.

The CAEN SP5650 sensor holder (Figure 6-6) can support a wide range of SiPM sensor. A few Hamamatsu SiPMs have been tried and tested, from the old S10362 series to the latest S13360 series. The SiPM is mounted on a round PCB fit in a metal mechanical case. The original PCB (shown in Figure 6-6) from CAEN can only provide basic function such as biasing the SiPM. This is replaced with a more advanced PCB that includes an active preamplifier circuit (discussed in Chapter 4). All the tests and data collected and analysed in this thesis use MSSL preamplifier circuit board.

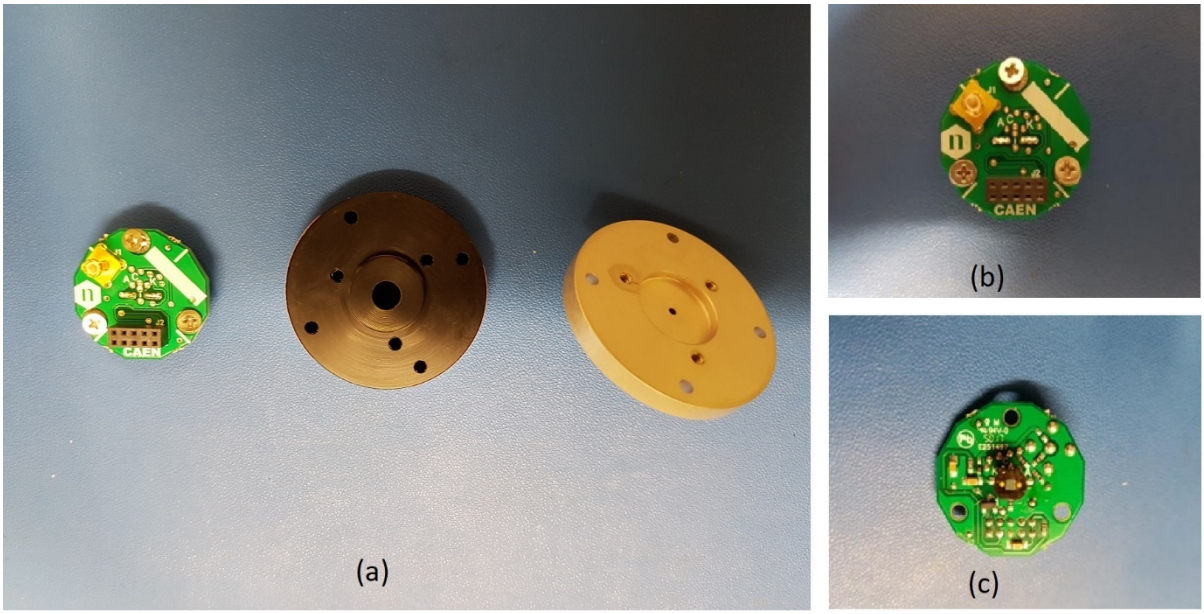


Figure 6-6 CAEN SP5650 sensor holder assembly break down (a). The original circuit board (b)(c) is simply a SiPM sensor host with no preamp circuit. (b) shows bottom side and (c) shows bottom side

The preamplifier output signal is collected and digitized by the DRS4 evaluation board (Figure 6-7, also called DRS4 module in this thesis.), which contains a Domino Ring Sampler (DRS) chip developed by PSI. The DRS4 has 9 independent channels. Each channel contains 1024 sampling cells that can store the analogue waveform at a sampling rate between 700MHz and 5GHz. The DRS4 chip only samples and holds the analogue signal. To digitize the signal, an external ADC is required. In DRS4 chip, the optimized readout time for each cell is 30ns. This means an ADC operated at 33MHz (~30ns) can provide a continuous and smooth output data stream. This slow ADC data rate is the major bottle neck of the readout system processing rate. More about the timing will be discussed in 6.2.3. The DRS4 module is designed around DRS4 chip and it includes the ADC for digitization and a FPGA for data handling and communication to the external interface (USB2.0). This module is effectively a mini oscilloscope. In the future instrument design, the design of the DRS module can be a very good reference design for the central processing and control board.

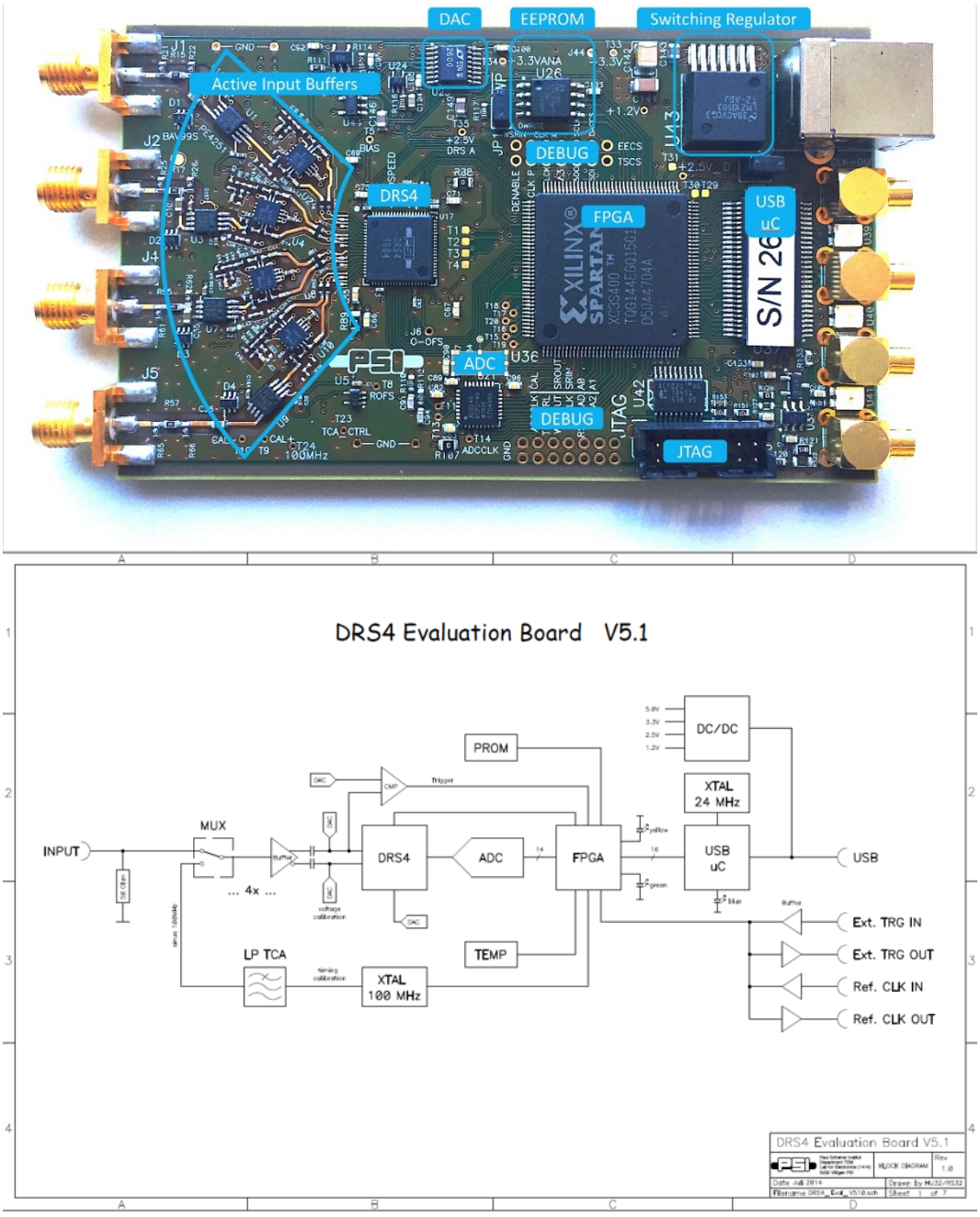


Figure 6-7 DRS4 evaluation board from PSI. The top picture shows the evaluation board PCB. The bottom picture shows the DRS4 evaluation board architecture [156].

With the DRS4 evaluation board comes a PC based application software called DRS Oscilloscope (Figure 6-8, in short DRSOsc), which serves to control the module and collect data from the evaluation board. The software allows user to save the digitalized data to the PC in either XML or BIN(ary) format. The XML format is easier to decode (as it is better structured, and the data is saved in ASCII text), but the file is much bulkier. It takes longer for DRSOsc to generate an XML data file. This will directly result in much lower event processing rate. The BIN format is much faster to process. In this

project, BIN format is selected as the standard data format. With BIN format, the processing rate can reach about 442 events per second. Each event is a complete waveform with 1024 data point. The BIN format does require more effort to write the decode function, but it is just a one-off development effort. This decode function is integrated in the overall waveform processing Python script.

Figure 6-8 DRS Oscilloscope software screen capture

The full hardware setup is shown in Figure 6-9 with each subsystem marked out.

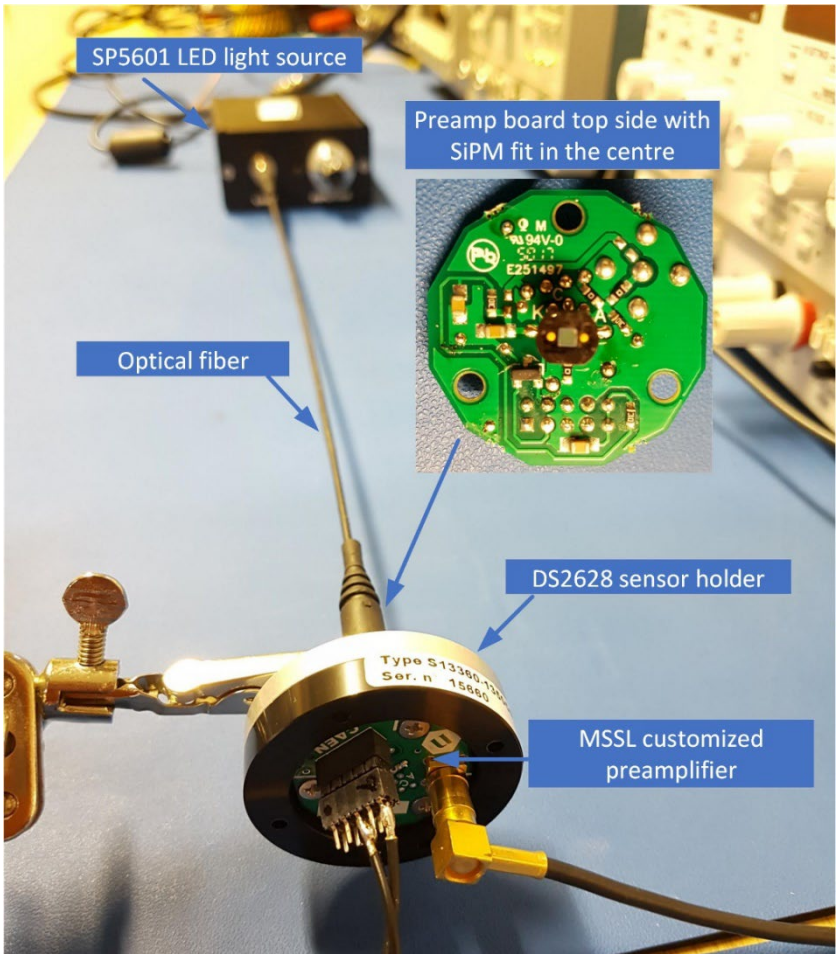


Figure 6-9 The bench test setup with each subsystem marked out

6.2.2 Review of the data acquisition system software setup

In 5.4, a set of Python scripts developed for this project are discussed. These scripts are combined into a single master script to handle the full data processing chain as shown in Figure 6-10.

The processing flow is firstly to decode and load an experiment data file, which contains 2000 or more events(also called frame in the later discussion) in BIN format, into the PC memory. Next, each waveform is cleaned with the mean filter. Next, construct the Deconvolution+Wiener filter using the reference data acquired in the same setup including the electronics noise waveform and the 1p.e. signal waveform. The filter is applied to each waveform to reduce the signal pile-up. Next, run the SiPM signal parameters calculation function. Each waveform might contain more than one signal. So, each signal will be calculated. The extracted parameters of all the signals are stored for the histograms construction. Last is the histogram analysis. KDE algorithm is used

to fit the histogram Gaussian like distribution. Then apply the peak search algorithm to find all the distribution peaks (each distribution represents a p.e. number). Then apply a linear fit to all the collected peaks. The system gain ,offset and other information can be extracted from the fit line. Save all the SiPM signal parameters, histogram plots, peaks plots as various PC files. Further process and deep data analysis can be added for the future work.

The master script can automatically loop through all the data files within a folder. This helps to automate the post experiment data processing, because generally an experiment will generate a batch of data files with each specified for different experiment setup. The file name is used to identify the key test setup. And all the generated plots use the file name as the base name for the plot title. This is just for the convenience of the data interpretation.

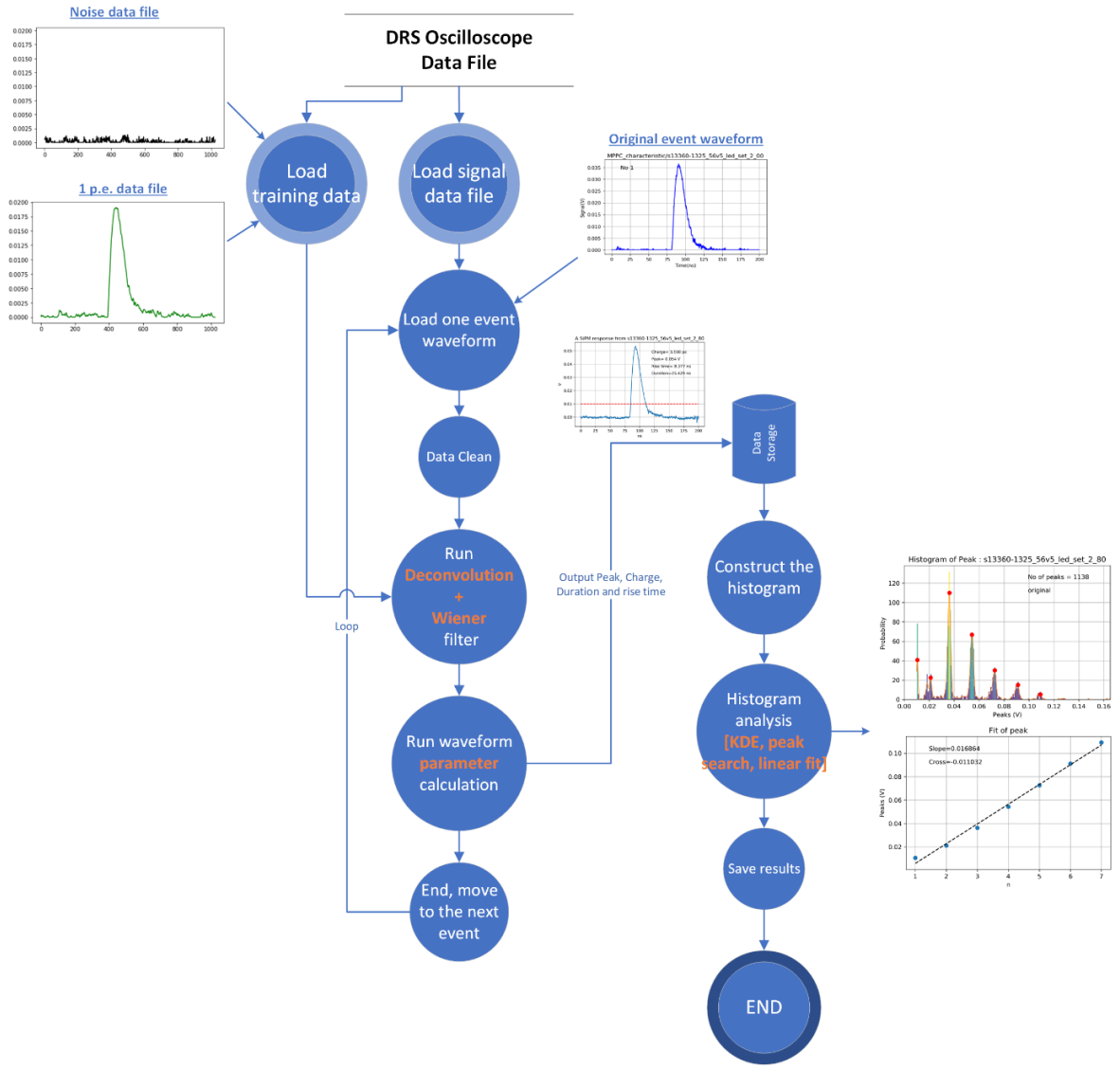


Figure 6-10 The data flow of SiPM data acquisition system software

6.2.3 The system time limit

6.2.3.1 System data acquisition timeline

The current data acquisition system has couple of bottlenecks limiting the data rate. There is no public data about the dead time of the DRS4 module and the DRSOsc software processing time. The DRS4 module user manual and the DRS4 datasheet suggests the module operates to the following timeline shown in Figure 6-11.

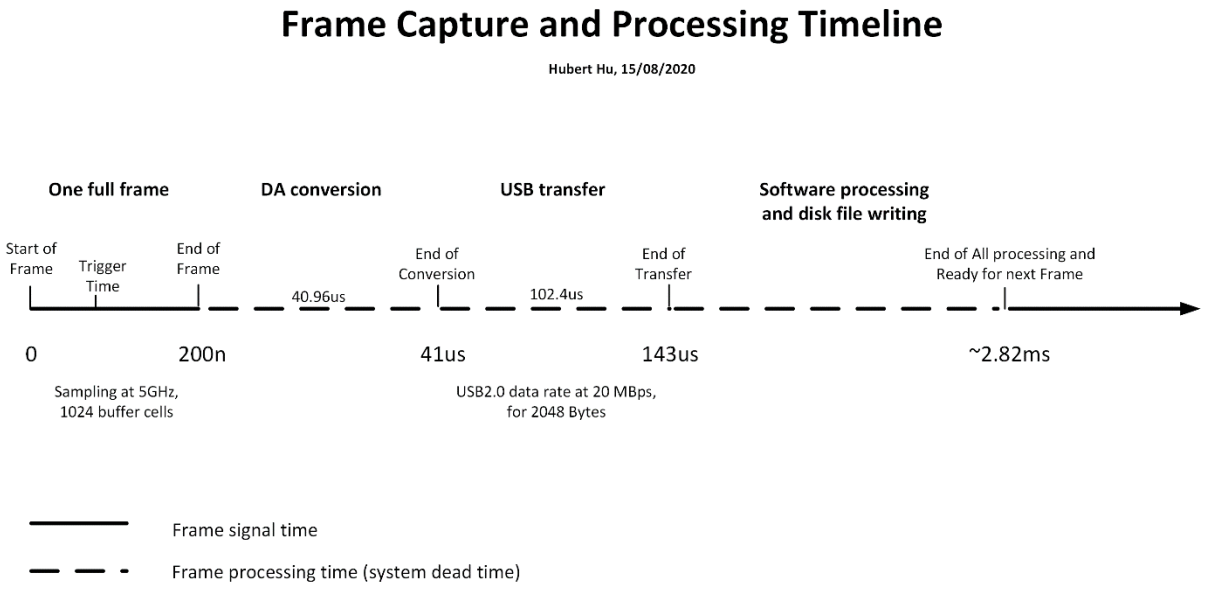


Figure 6-11 The timeline of the full data acquisition system based on the theoretical calculation and some data observation

The DRS4 chip has 9 independent sampling channels. Each channel contains 1024 sampling cells. The cell sampling rate can be set to 5GHz at maximum. At 5GHz, 1024 cells can hold a period of analogue signal for ~200ns. Such a 200ns window is defined as a full frame. In the previous software processing flow discussion, a frame is defined as an event or waveform. At the data level, they refer to the same thing, which is the full amount of data points one DRS4 channel can hold at one trigger.

Each channel of the DRS4 chip can be connected to one external ADC. Thus, all the channels can be read out in parallel. Or to save resources, all the channels can be multiplexed into one single output and then to be readout by just one ADC. The DRS4 module takes the latter approach.

The cell shifting time is 30ns for optimized performance. The multiplex will add 10ns delay. The DAC conversion rate is set to 33MHz (=30ns per sample) to match the cell shifting time. This means to readout and digitize all the cells in a single channel, it will take 1024 * (30 + 10) = 40.96us.

The DRS4 module uses a 14-bit ADC (AD9245 from Analog Device). To simplify the data format, two-byte (=16-bit) is assigned for one cell. This means the full frame has a total data volume of 1024*2 = 2048 bytes. The onboard USB2.0 bus module has the capability of data transfer at 20MBps (Mega Byte per second). This works out that to transfer a full frame data from the module to the host PC, it will take 2048 / 20 = 102.4us. There is a small overhead for the USB2.0 protocol, which is ignored here for simplicity.

Once the data is collected by the PC, the PC software DRSOsc will pack the frame data (defined as an event) into a structured format with additional information like timestamp, voltage and time calibration data. The timestamp is from a PC system clock, so its accuracy is only down to one millisecond. Then the structured data will be written to hard disk as either an XML file or binary file under user choice. As discussed already, we adopted the binary format for the shorter PC processing time. The exact processing time depends on the PC hardware performance. For the PC used in the experiment, the average frame rate is about 354fps. This is calculated from 27 data file containing total 28,000 frames.

By close examining of the timestamp of each frame package (the timestamp is added by the DRSOsc software with accuracy of 1ms) , it is found for those photon intensive test data files, there is often a continuous set of 6 frames (also called events in the thesis) at the same timestamp. This suggests the DRS4 module can buffer 6 frames onboard and send to the PC in one USB data packet. Following the 6 frames, the timestamp of the next frame often shows a large delay of 15 or 16ms. The reason is likely that the DRSOsc takes much longer time to process a large USB packet. Another case is for the lower photon intensive test data files, the gap between consecutive frames is usually 2,3 or 4ms. This probably is the time for the DRSOsc to process a USB data packet containing just one frame. Of all the data files, apart from those cases of sharing the same timestamp, there is no observation for a gap less than 2ms. This probably suggests the full system deadtime is around 2ms. If inverse the average frame rate 354fps, the average delay is 1/354 = 2.82ms. Given the accuracy of the software timestamp is just 1ms, this can match the observation of 2, 3 or 4ms gap.

In summary, if the system operates in a real time mode, the system deadtime after processing a frame (200ns at 5GHz sampling rate) is 2.82ms. This greatly limits the capability to characterize the SiPM detector system in real time. To mitigate this, the

SiPM detector system characterization experiment data is acquired in trigger mode, and the experiment focuses more on the signal amplitude related performance characterization rather than the timing related performance.

6.2.3.2 The deadtime of SiPM and the preamplifier

From the above analysis, the dominant deadtime of the system is the PC processing and the data transfer from DRS4 module to the PC. The deadtime of the SiPM and the preamplifier are much smaller. However, it is useful to understand these deadtime as that can help to estimate how much photon signal will be lost due to the deadtime.

The deadtime of the front-end electronics can be discussed in three different areas: **the pixel deadtime** (a pixel of the SiPM sensor), **the device deadtime** (the whole SiPM sensor) and **the readout deadtime** (limit to the preamplifier).

The **pixel deadtime** is the period that a triggered pixel takes to recharge its body capacitance. This time is estimated at 5 times the time constant formed by the quenching resistor and the pixel body capacitor. Using the SiPM parameter calculated from Table 4-2 in section 4.3, the pixel deadtime of SiPM S13360-1325 is calculated as:

$$t_{pixel\_deadtime} = 5R_qC_d = 5 \times 1600k\Omega \times 6.69fF = 50ns$$

The **device deadtime** is the period for the device grid (parasitic) capacitance to discharge. The similar rule of thumb as 5 times of the time constant by the grid resistance and total effective capacitance can be used to estimate this deadtime. The grid resistor should not exceed a few ohms. The grid capacitance is around tens of pF. Using the figures from Table 4-2, the device deadtime of SiPM S13360-1325 can be estimated as:

$$t_{device\_deadtime} = 5R_gC_g = 5 \times 2\Omega \times 43.2pF \approx 432ps$$

This device deadtime is neglectable small.

The **readout deadtime** (limit to the preamplifier) is determined by the recovery time of the preamplifier circuit, which is mainly the reset time for the integration capacitor. The capacitor is discharged by the feedback resistor in parallel to it. Again, take 5 times of the time constant as a figure in merit. Using the component values from the preamp board design (4.5.4), we can estimate this deadtime as:

$$t_{preamp\_deadtime} = 5R_fC_f = 5 \times 4.99k\Omega \times 0.6pF \approx 15ns$$

The device deadtime and the readout deadtime both locates at the serial section of the front-end electronics, therefore this two deadtime can be combined. Between them, the readout deadtime is dominant. The pixel deadtime locates at the parallel section of the front end, so its situation is different. While one pixel is inactive during the recovery period, the other pixels are still active and are ready to trigger. At very low light intensity or dark count only cases, the effective pixel deadtime can be relaxed to the single pixel deadtime divided by the number of pixels, as in these cases, the chance of more than two pixels being triggered simultaneously is extremely low. This feature has some analogy to MCP. For S13360-1325CS SiPM with 2668 pixels, the effective pixel deadtime is calculated as:

$$t_{pixel\_deadtime} / n_{pixels} = 50ns / 2668 \approx 18.7ps$$

So, the **total deadtime** of the SiPM readout front end electronics can be calculated as:

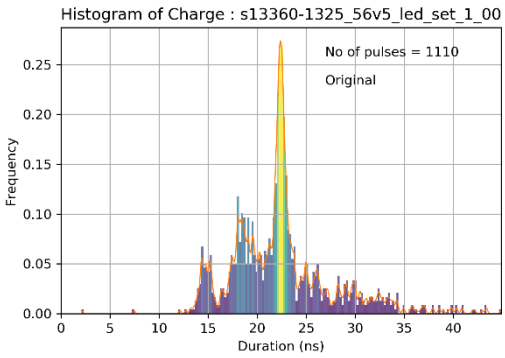
$$t_{total\_deadtime} = t_{pixel\_deadtime} / n_{pixels} + t_{device\_deadtime} + t_{preamp\_deadtime} \approx 15.5ns$$

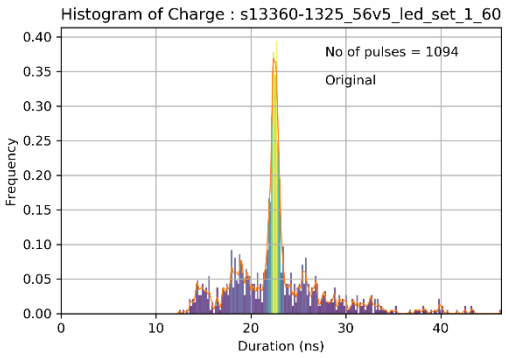
To check the total deadtime in the real experiment data, the duration (see Figure 5-15 for the definition of the duration) of 3679 dark count pulses from a data set are collected. The duration represents the total deadtime including the pixel deadtime, SiPM device deadtime and preamp deadtime. The calculated average is **15.4ns**. This matches the theoretical calculation of **15.5ns** very well.

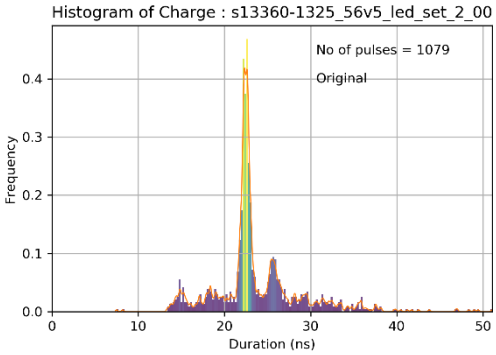
To eliminate the noise interference, the duration is defined as the period that the signal is above the 10mV threshold line (see Figure 5-15). It can be argued that the 10mV threshold line has cut off part of the signal that is below the threshold, which shall be counted as deadtime as well. This is true, but this missed extra duration is only a very small fraction of the total duration. Therefore, skip of such extra only introduces a very small error that won't impact our conclusion.

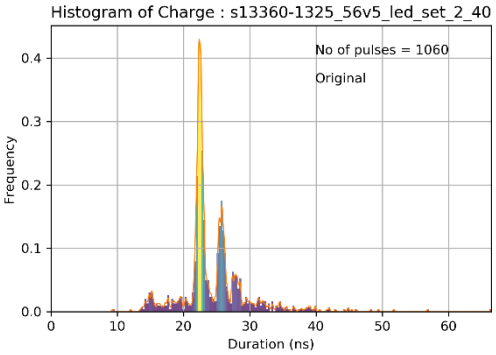
The front end deadtime changes as the signal amplitude changes. The larger the signal is, the longer the deadtime is. The above theoretical dead time calculation and the real experiment data (dark count) dead time measurements are all based on 1p.e. signal. For 2p.e. signal or bigger, the theory needs to be adjusted. As signal amplitude

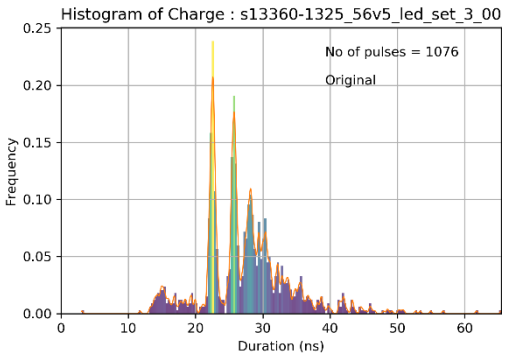
increases, the total deadtime will increase but not linearly. Figure 6-12 presents the signal duration against different light intensity (leads to different p.e. number)

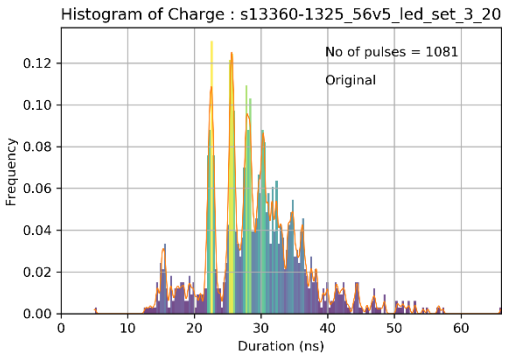


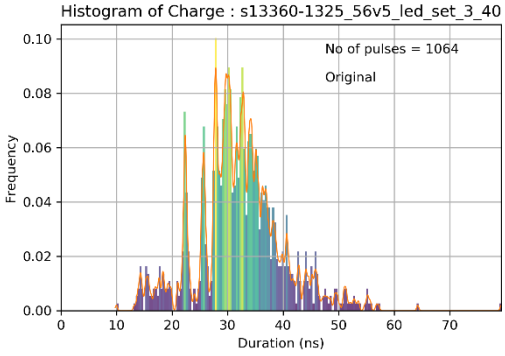


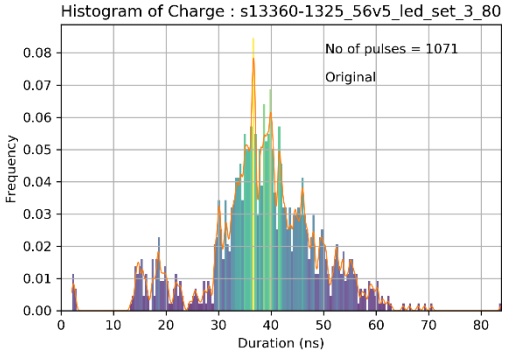












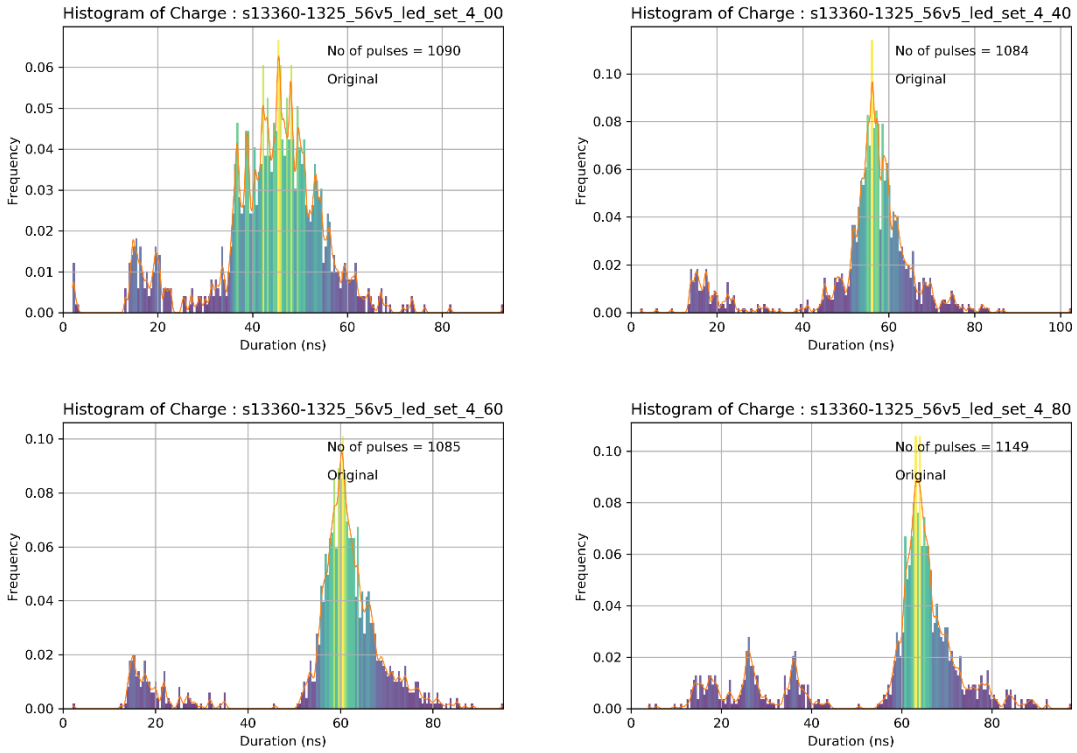


*Figure 6-12 The histogram of the **duration** of SiPM preamplifier output pulse at different LED condition. The parameter is based on the original waveform, i.e. no Deconvolution+Wiener filter is applied*

We can see at low light condition, the structure of different duration corresponding to different p.e. number is clearly visible. As light intensity increases, the structure is not clear any more. The relationship between the signal amplitude and the duration is left for future study.

Another observation on real data is that our dead time theory based on the pulse duration might have overestimated the real system deadtime if some degree of signal pile-up is allowed. For example, in Figure 6-13, a second signal immediately appears before the first signal is fully reset to zero potential. This suggests the preamplifier can accept new event before its integration capacitor is fully discharged. Though it is not perfect, duration of the output signal is still a close estimation of the front end deadtime.

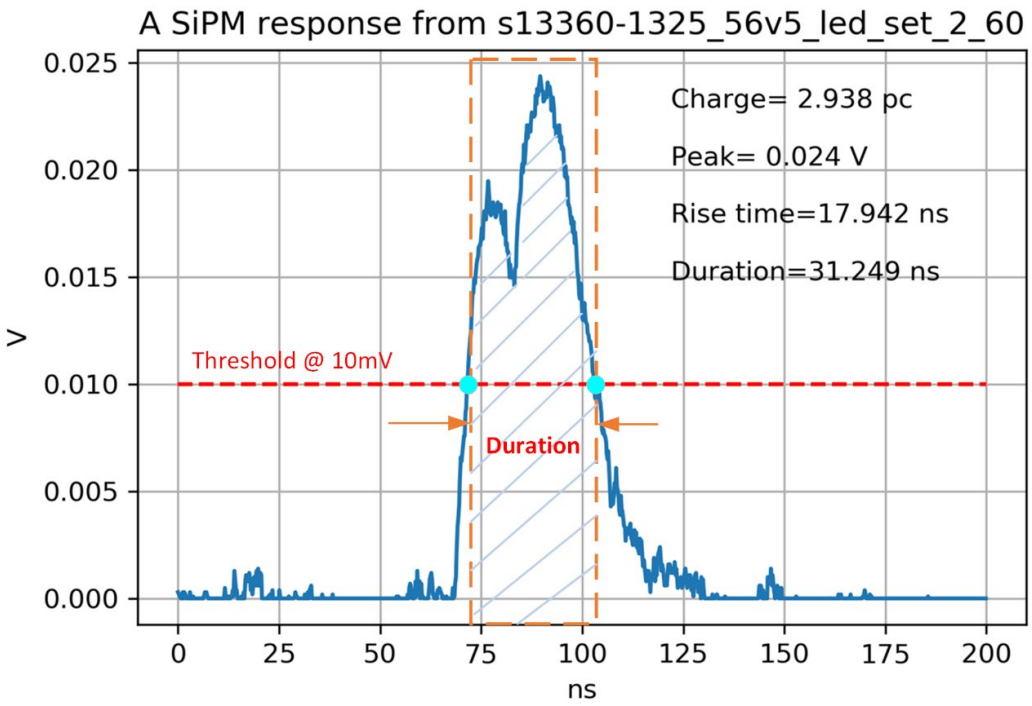


Figure 6-13 The overestimate of the system total deadtime when signal pile-up appears.

6.3 SiPM detector system characterization data overview

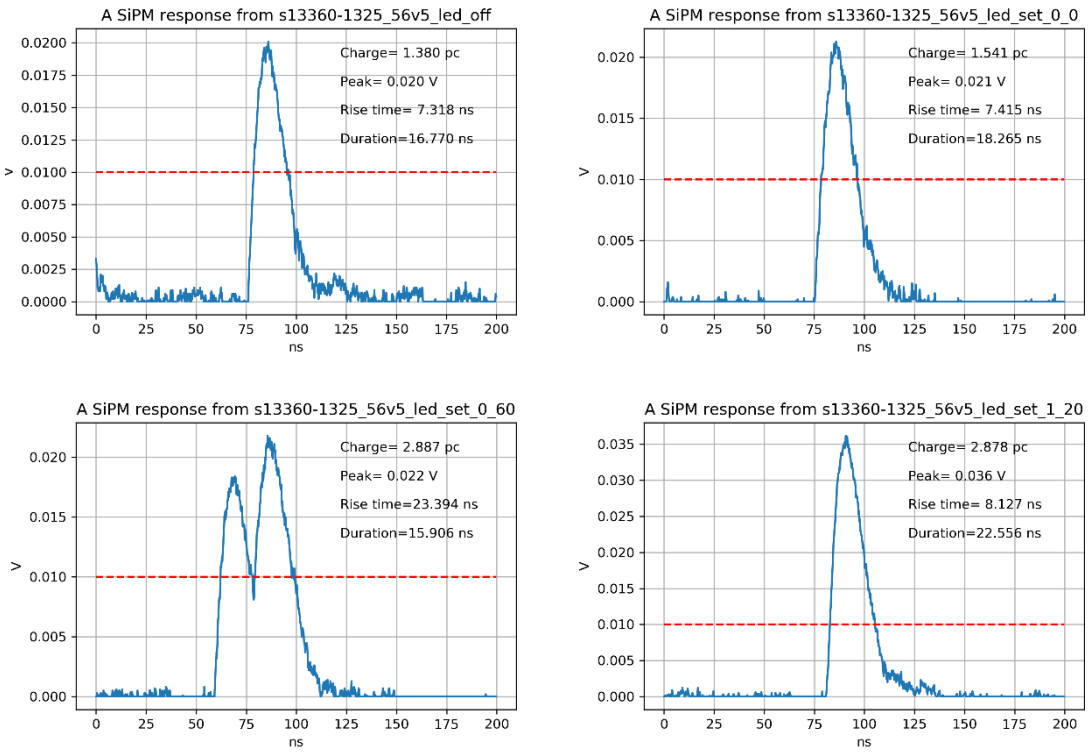
As introduced above, the SiPM detector system characterization experiment is done using an adjustable LED source on the bench at room temperature. The LED intensity adjusting knob has two levels of setting. The course level is from 0 to 9. At each step of the course level, the fine level ranges from 0 to 100 in a step size of 2. This means the total adjustable range contains 10x50=500 fine steps.

In this characterization experiment, the LED intensity is swept from setting 0.00 to 5.00 in a step of 0.20. The first digit represents course level. The two digits after decimal represent the fine step. At each step, we acquired 1000 events in trigger mode with a threshold at 0.5p.e. amplitude. This produces a dataset containing 26 different LED intensity level. The reason that the intensity beyond setting 5.00 is not covered is because the pre-scan tests show the SiPM system deeply saturates beyond setting 5.00.

In addition, a dark count (made up of majority 1p.e. signal) data file is captured with LED completely off (power removed, not the intensity setting to 0.00). An electronic noise background data file is also captured with LED off and SiPM biased below the breakdown voltage. These two files are used for the Deconvolution+Wiener filter construction. And the dark count data file is also used for the dark count analysis work.

Figure 6-14 SP5601 LED driver (right) and the DS2628 sensor holder (left)

During the data acquisition, the DRS4 module is configured at 5GHz sampling rate. So, for an event frame made up of 1024 samples, frame can cover 1024*0.2ps≈200ns. Figure 6-15 shows some randomly selected, typical event waveforms at different LED setting. These are random chosen waveforms so that their peak amplitude not necessary reflect the increasing LED setting.



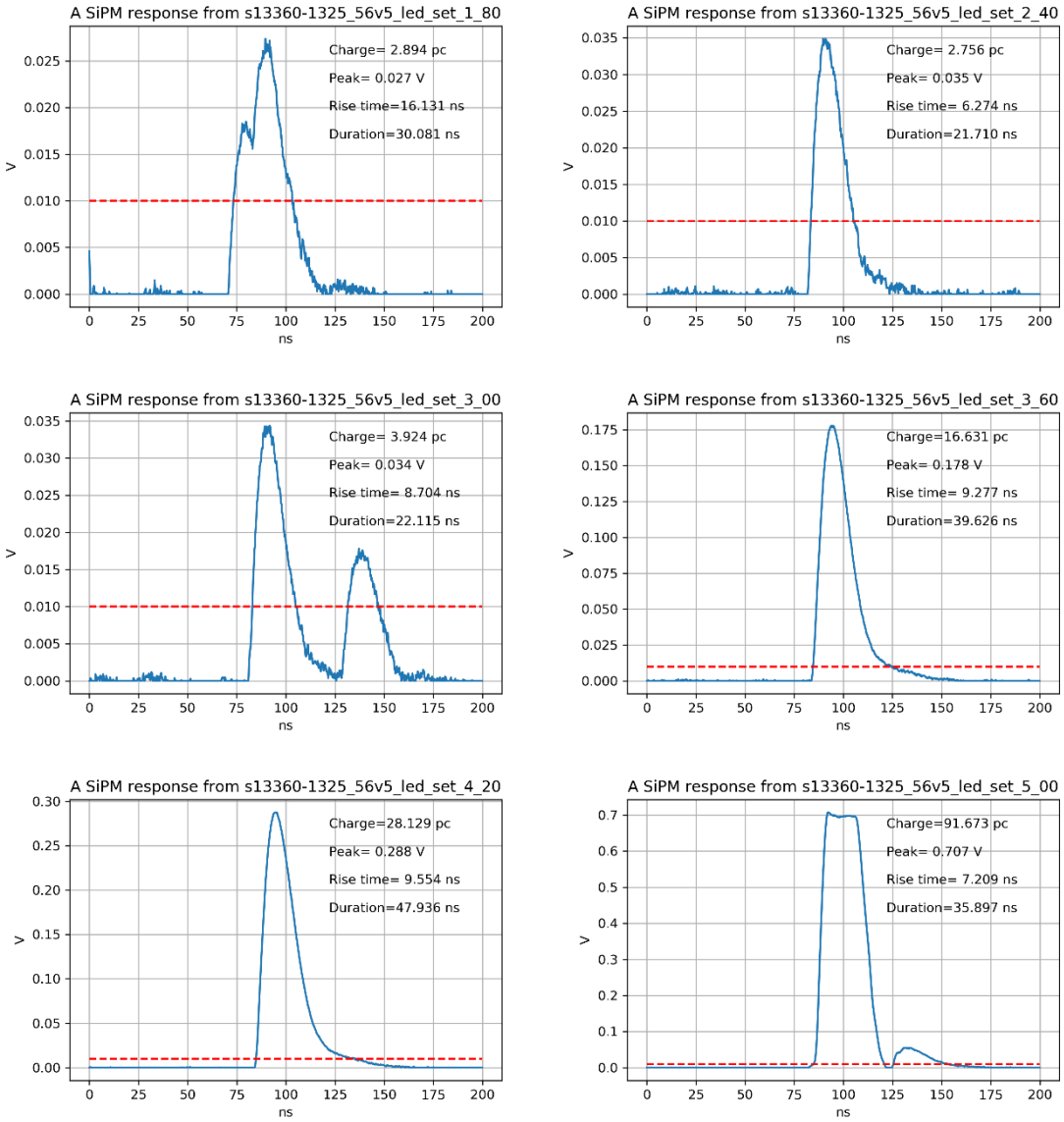


Figure 6-15 Random event waveform examples at different LED setting

The last sample is at LED setting of 5.00. It shows the system output signal reaches saturation with a sign of flat plateau of the signal pulse top. The saturation voltage is 0.7V, which is the up limit of the DRS4 module intake signal.

The acquired data set contains 27 files in total (1 LED off and 26 LED on files). The file detail is listed in Table 6-1.The *find pulses* column means how many signal pulses are found over all the captured events in that file. This figure is higher than the total events because some events contain more than one signal pulse. The minimum and maximum counted pulse peaks for each data file are recorded here. Due to the fluctuation, for the same LED setting, the signal pulse peak will swing in a wide range. Duration is the time that it takes to collect the 1000/2000 events. This is calculated using

the first and last event's time stamp. The events rate is then calculated using the number of events divided by the duration.

Table 6-1 The overview of the full data set acquired for the SiPM system characterization experiment

#	LED setting	Events	Find Pulses	Pulse Peak (V)		Duration	Events Rate
		#	#	min	max	ms	fps
1	LED = off	2000	2131	0.01538	0.03980	5564	359
2	LED=0.00	1000	1100	0.01509	0.05539	2995	334
3	LED=0.20	1000	1109	0.01419	0.05390	2964	337
4	LED=0.40	1000	1103	0.01429	0.05599	2995	334
5	LED=0.60	1000	1107	0.01090	0.05689	3011	332
6	LED=0.80	1000	1105	0.01470	0.05500	2933	341
7	LED=1.00	1000	1110	0.01200	0.09049	2918	343
8	LED=1.20	1000	1112	0.01380	0.07439	2902	345
9	LED=1.40	1000	1089	0.01429	0.06159	2902	345
10	LED=1.60	1000	1093	0.01509	0.09040	2855	350
11	LED=1.80	1000	1079	0.01499	0.09069	2871	348
12	LED=2.00	1000	1079	0.01490	0.09209	2746	364
13	LED=2.20	1000	1079	0.01409	0.11080	2714	368
14	LED=2.40	1000	1060	0.01659	0.14069	2730	366
15	LED=2.60	1000	1060	0.01470	0.12809	2699	371
16	LED=2.80	1000	1063	0.01589	0.16519	2730	366
17	LED=3.00	1000	1076	0.01389	0.19838	2761	362
18	LED=3.20	1000	1081	0.01479	0.21118	2715	368
19	LED=3.40	1000	1064	0.01549	0.26529	3105	322
20	LED=3.60	1000	1071	0.01110	0.30959	2730	366
21	LED=3.80	1000	1071	0.01090	0.40099	2715	368
22	LED=4.00	1000	1090	0.01090	0.43529	2730	366

#	LED setting	Events	Find Pulses	Pulse Peak (V)		Duration	Events Rate
23	LED=4.20	1000	1077	0.01110	0.54949	2714	368
24	LED=4.40	1000	1084	0.01189	0.68839	2683	373
25	LED=4.60	1000	1085	0.01119	0.70919	2714	368
26	LED=4.80	1000	1149	0.01409	0.71009	2684	373
27	LED=5.00	1000	1704	0.01490	0.71128	2714	368

Note:

The number of pulses found in LED setting 4.80 and 5.00 are not accurate due to the primary pulse secondary ringing bouncing above the 10mV threshold and being counted as a separate pulse. The ringing is related to the preamplifier board design and in general, the higher the primary pulse is, the bigger the ringing will be. Therefore, the ringing is caught as a new pulse in error only in those two highest signal cases.

The max pulse peak found at each LED setting can be used as a rough guide for system response, but it is not very accurate because the max includes the sum of nominal response and the uncertainty. Later in the system linearity analysis, the system response will be better presented with the nominal response which can be acquired from the pulse peak histogram distribution. Here, for initial view, the max pulse peak is plotted against LED setting in Figure 6-16. This plot is only partially linear for the middle range because: first the LED intensity vs LED setting is not fully linear (as shown in Figure 6-3), second the SiPM system saturates after LED setting above 4.60.

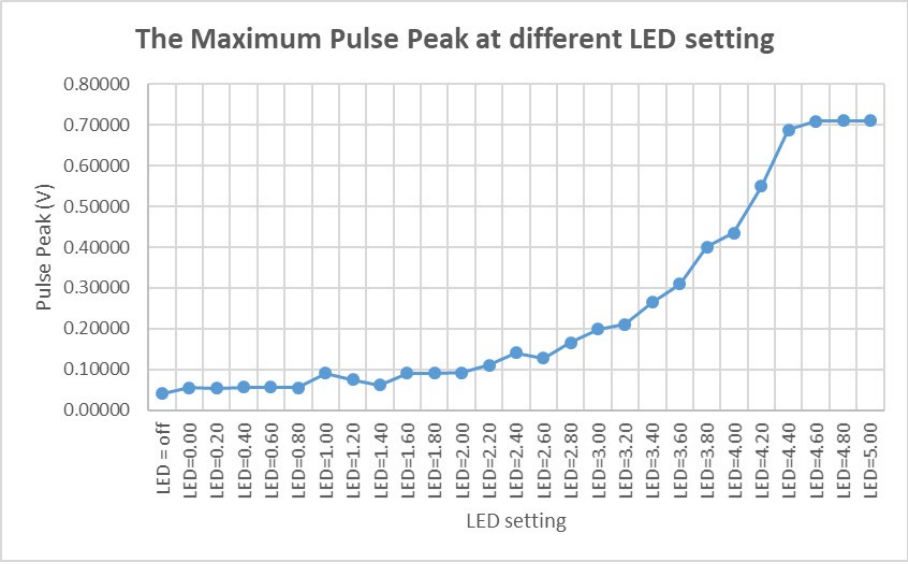


Figure 6-16 The max pulse peak recorded for each LED setting

6.4 SiPM detector system characterization

6.4.1 SiPM after-pulse

During the avalanche process, a small portion of the avalanche carriers get trapped in defects in the silicon but are released after short delays (typically on the order of a few ns to many tens of ns). If the release occurs during the cell recovery time, its amplitude will be lower than that of a normal (1p.e) pulse. If it occurs after the cell's recovery, it will develop into a full avalanche process and appear the same as a genuine pulse. These are called after-pulses. After-pulse is an unwanted signal and a contamination to the true signal. The partially developed after-pulse can be discriminated by the pulse height. At some specific conditions, the fully developed after-pulse can be excluded by a time-delay filtering algorithm. Such conditions can be the timing of the pulse is known when the incoming signal is driven by a trigger, or in very low pulse rate application where the time difference between consecutive pulses is very large.

Since the amplitude of the after-pulse generally is equal to or below the amplitude of 1p.e. signal, it can be filtered out together with most of the dark count together by setting the discrimination threshold at 1.5p.e. Figure 6-17 shows the full-grown after-pulses captured by an oscilloscope set in hysteresis mode, which means stacking up multiple waveforms. The waveforms are coloured in heat mode, with red standing for high intensity and blue for low intensity. The leading/triggering pulse is the primary SiPM signal. The trailing pulse is the full-grown after-pulse. This plot is acquired with no LED signal, so only dark count presents. As dark count hardly reoccurs within the same screen frame (1us width), these blue (blue means occur less frequently) pulses are mostly after-pulse.

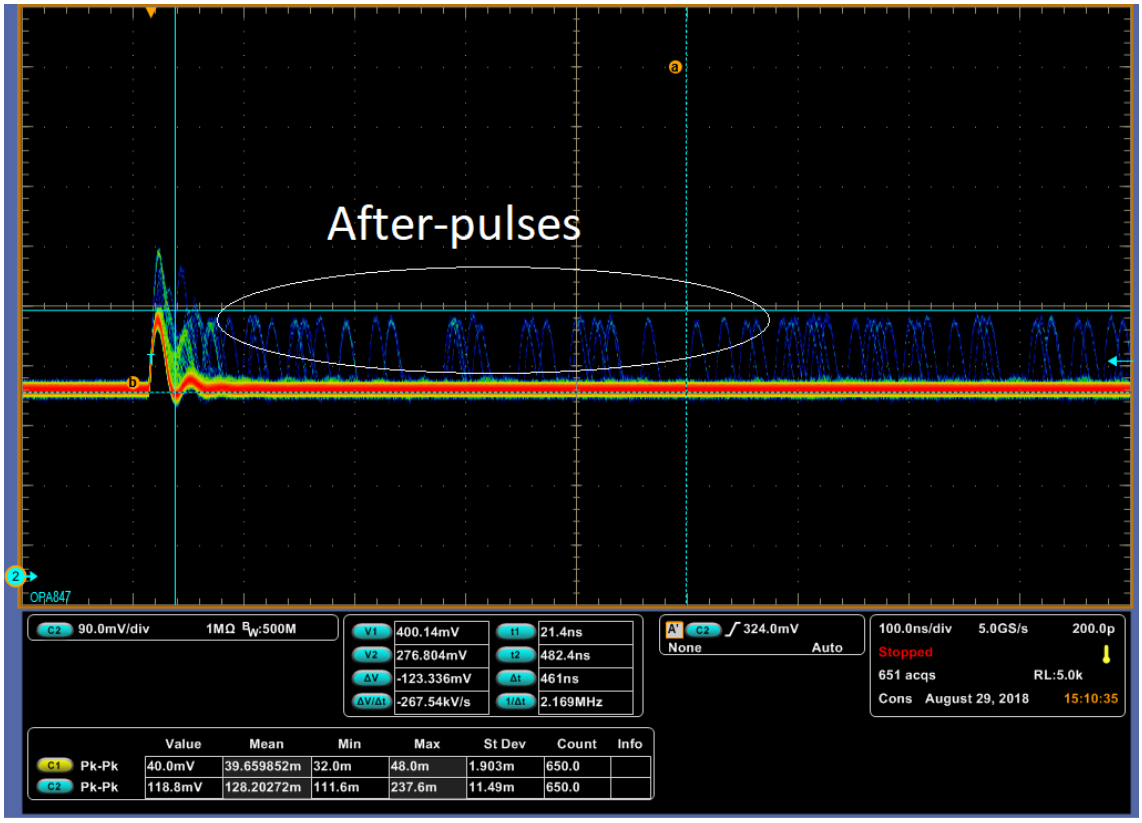


Figure 6-17 The oscilloscope captures of after-pulse in hysteresis mode in dark count measurements

From the SiPM characterization dataset, we picked up two examples in Figure 6-18 to illustrate the partially developed and fully developed after-pulse.

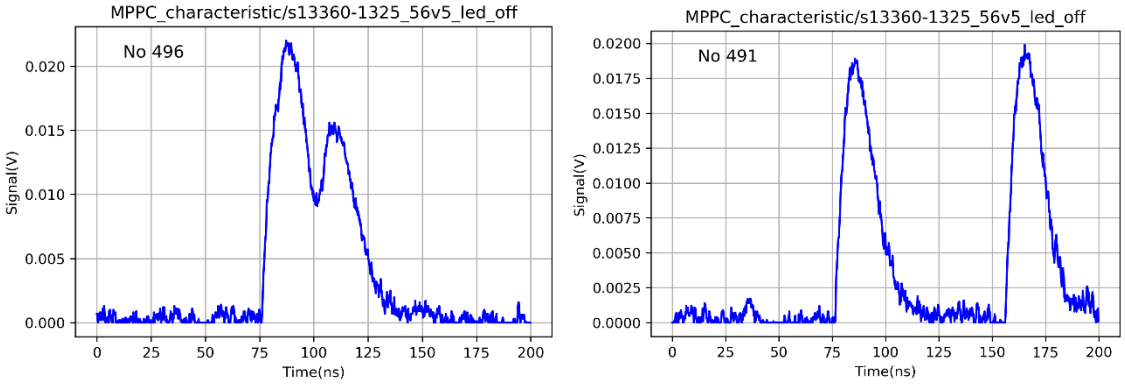


Figure 6-18 Example of partially developed after-pulse (left) and fully developed after-pulse (right)

6.4.1.1 Calculation of after-pulse ratio

The data file #1 with light source switched off is used to calculate the after-pulse ratio. As there are two types of after-pulse: partially developed and fully developed, the discussion will be separated into two parts.

Contribution from partially-developed after-pulses

Before the discussion, let's explain some background information about the data file. This data file is made up of 2000 events. Each event is a 200ns long frame. As each event is acquired on trigger, there is at least one pulse in each frame. This triggered pulse is defined as the primary. In some frames, there is one or more pulses after the primary. These pulses are defined as the secondary. A customized algorithm has been developed to detect and analyse all the pulses.

From Figure 6-18 left plot, we notice the immediate after-pulse is so close to the primary pulse, that it often causes signal pile-up. To detect the overlapped pulses properly, we applied the Wiener-Deconvolution filter on this data file. Before applying the filter, our algorithm picks up 131 secondary pulses over the entire 2000 events. After the filter, our algorithm picks up 186 secondary pulses. This shows the success of the filter. And it can be quite certain the delta (186-131=55) number of pulses are the partially-developed after-pulse.

Then plot all the pulses (2000 primary pulses + 186 secondary pulses) as a time series taking the relative time to the first pulse as the timestamp of each pulse. The result in Figure 6-19 shows most of the peaks lay around the 1p.e. line and a few around the 2p.e. line. This reflects the fact that with no light source, the majority of the signal is the 1p.e. dark count. The 2p.e. signal can be a mixture of a 1p.e. dark count triggering one optical cross-talk or two dark counts from different pixel at the same time. Those peaks lay below the 1p.e. line and between 1p.e. and 2p.e. line are the under-developed after-pulse.

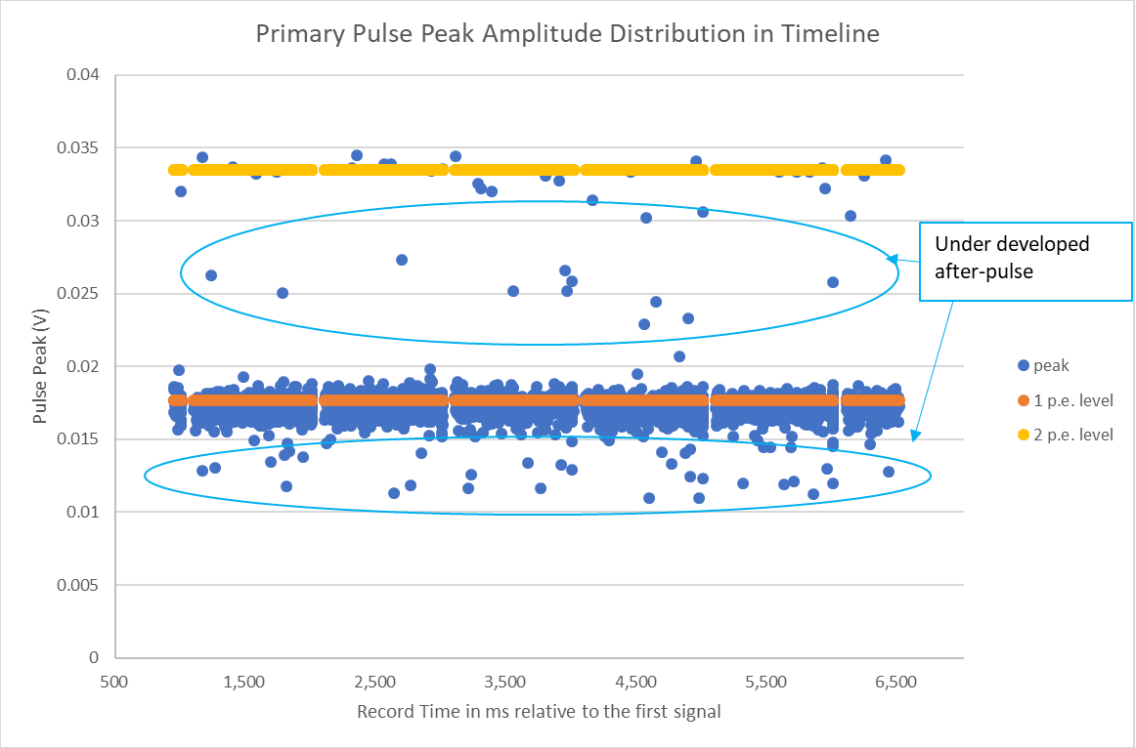


Figure 6-19 Pulse peak against the pulse arrival time (relative to the first pulse)

Then the after-pulse ratio contributed by the under-developed pulses can be estimated by counting those partially developed secondary pulse to the total number of primary signals. Here the data is calculated by reading the plot Figure 6-19. With pulse peak under 0.015V counted as less than 1p.e. pulse, with peak between 0.02V and 0.03V counted as less than 2p.e but above 1p.e., we derived the results shown in Table 6-2. The thresholds used here are a bit arbitrary. Because the fluctuation of a nominal 1p.e. or 2p.e. signal amplitude partially overlaps with the amplitude of the under-developed after-pulses, this causes the thresholds very hard to set precisely.

Table 6-2 Statistics for After-Pulse calculation

less than 1 p.e.	45
less than 2 p.e. but above 1 p.e.	18
Sum of partially developed after-pulse	63
Total number of signals	2000
Ratio	3.15%

Contribution from fully-developed after-pulses

This type of after-pulse can't be distinguished from a normal dark count pulse. So, we took a statistic approach to estimate how many fully-developed secondary pulses are genuine dark counts and the remaining shall be after-pulses.

It is reasonable to assume the average dark count rate is uniform across all the 200ns event frame. The data file #1 includes 2000 events. This gives a total 2000*200ns long time window. With the average dark count rate at 70kHz, the expected dark counts for the total time window is:

$$Total\_dark\_counts = 70k \times 200ns \times 2000 = 28$$

This means 28 pulses out of the 186 counted secondary pulses are dark count. This calculation is valid assuming each dark count is an independent event. Then after removing 63 under-developed secondary pulse from Table 6-2, the balance is mostly the fully developed after-pulses. We then can calculate the portion of fully developed after-pulse as:

$$R_{after\_pulse\_\sec} = \frac{186 - 28 - 63}{2000} = 4.75\%$$

Combining this result with the result in the Table 6-2, we can acquire the final after-pulse ratio as:

$$\begin{aligned} R_{after\_pulse\_final} &= R_{under\_developed} + R_{fully\_developed} \\ &= 3.15\% + 4.75\% = 7.9\% \end{aligned}$$

This calculation provides a very good approximation for the SiPM after-pulse performance. However, there is inaccuracy in the process. For example, the amplitude of the pulse peak varies even if it is fully developed. This is governed by a Gaussian distribution. So, at extreme cases (whose probability is quite low), some pulse will fall into the low amplitude region, where it can be counted as an under-developed after-pulse. Another example is the statistic estimation for the number of dark counts. The events are captured in triggering mode while the triggered pulse is mostly like already a dark count. This creates a bias as its timing (of the 200ns window) is selected rather than a random section in the timeline.

Nevertheless, the methodology here provides an easy and very practical (it requires very little computing resources.) approach to estimate the after-pulse ratio with a reasonably good accuracy.

6.4.1.2 The study of the delay time between the primary and secondary pulse in a single capture frame

The parameter that is closely studied is the delay time between the primary pulse and the leading secondary pulse (in data file #1, only the delay time between the primary and the first secondary is counted. The time between the primary to the second or more secondaries is ignored. There are very few such cases, so this cut-off shall not affect the results. Plus, for event with more than 2 secondary pulses, it is hard to judge the nature of the secondary, which could be a genuine after-pulse, or a coincident dark count, or another after-pulse triggered by the first after-pulse.) The delay is defined as the peak time of the primary pulse to the peak time of the next secondary pulse as shown in Figure 6-20.

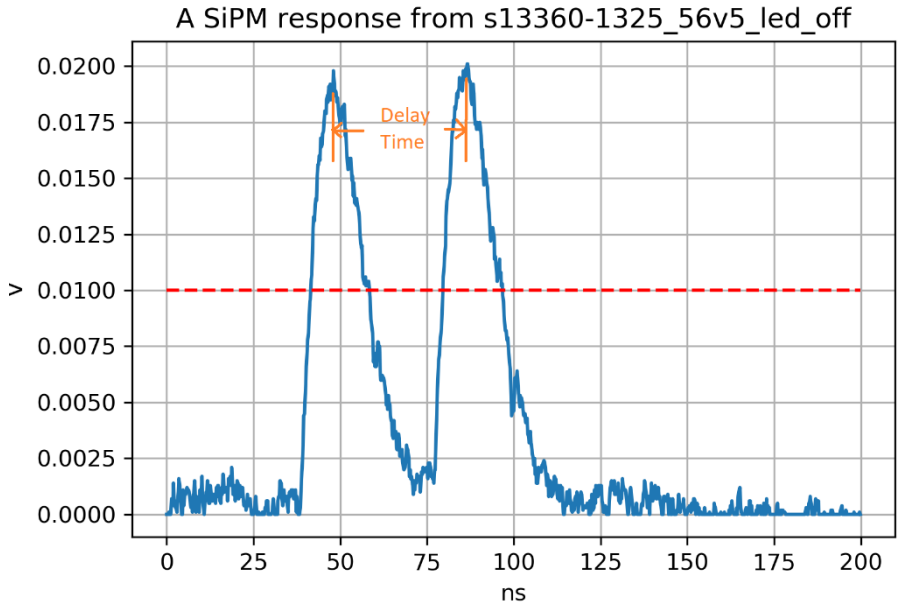


Figure 6-20 The definition of delay time between the primary(triggering) pulse and the first secondary pulse

In this way, the delay time for all those frames is calculated and collected. Then plot the histogram of all the delay time in 24 bins ranging from 5ns to 120ns. The result is shown in Figure 6-21. Note, the data used here is the original data before the filtering because we would like to study the nature of the occurrence of the after-pulse excluding the impact of the filter.

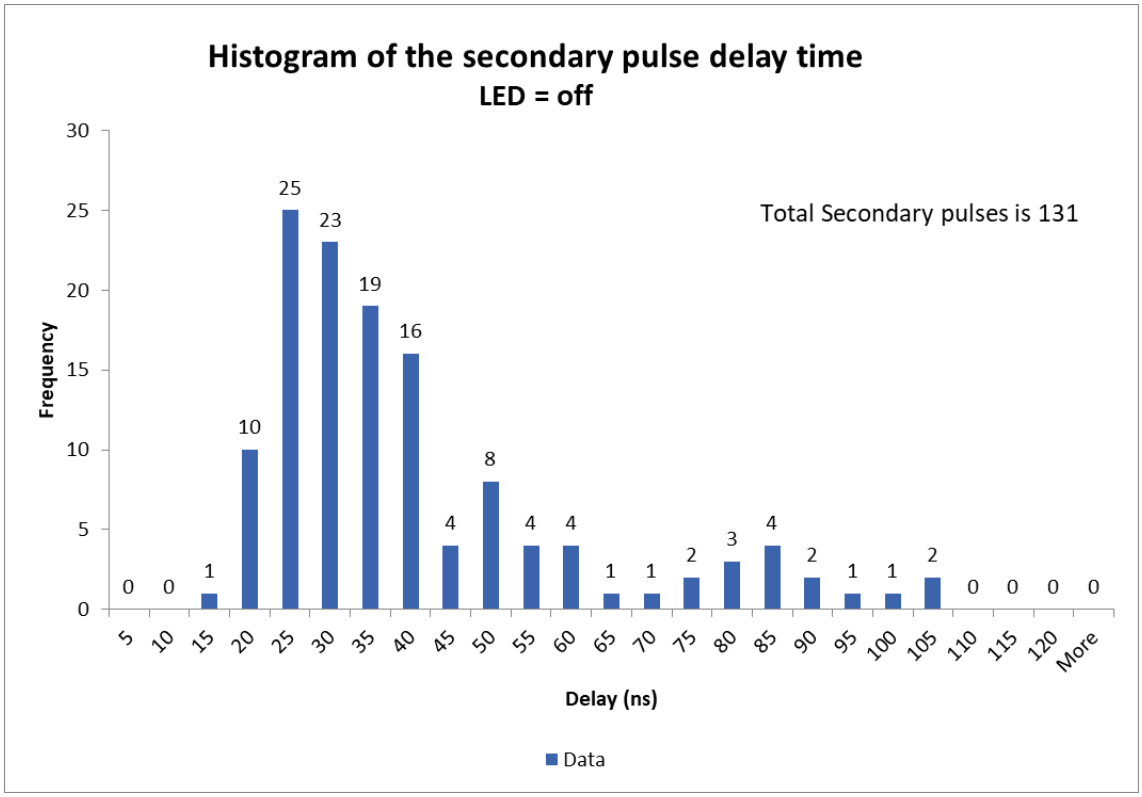


Figure 6-21 Histogram of the delay time between the primary pulse and the leading secondary pulse

Firstly, this distribution is interpreted as the sum of several Gaussian pulses. The fitting formula for each Gaussian pulse is constructed as:

$$y = B + A \cdot \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \tag{5.1}$$

Where,

y represents the frequency,

x represents the delay,

B is the offset,

A is the amplitude of the peak of the Gaussian pulse,

μ is average of delay, and

σ is the standard deviation of the delay, which determines the width of the pulse.

By trial and error of the four parameters B, A, μ and σ, a set of Gaussian pulses that can closely fit to the plot are found. Four Gaussian pulses are created and the sum of them represents the final fit function. The quality of fitting is examined with Chi-square calculation. With Gaussian pulses, the best fitting score we can get has a Chi-square error of 41.5. The result is shown in Figure 6-22. In the figure, four

Gaussian pulses are coloured differently. The leading pulse dominants. The dotted curve is the sum of all the four pulses.

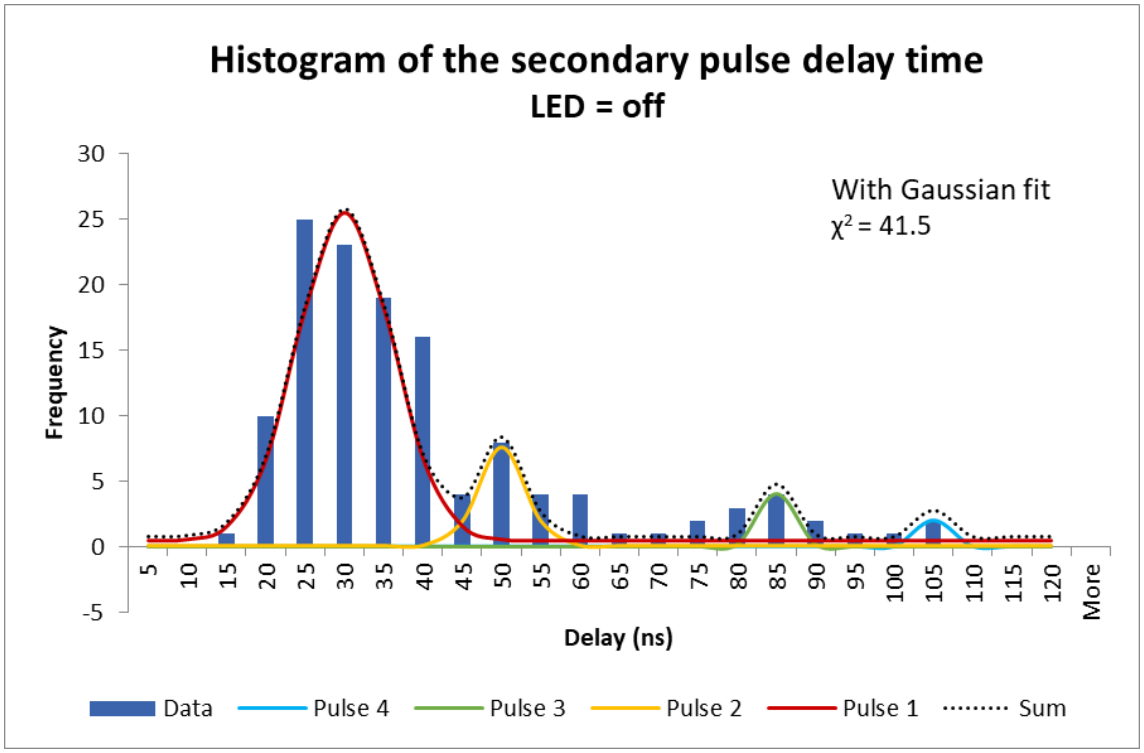


Figure 6-22 The fit function using Gaussian pulse

The Gaussian pulse doesn't look like the best fit to the data as the pulse appear in the data shows a bit asymmetric. This suggests that a Poisson fit might be better. The standard Poisson distribution function is slightly modified to include a x-axis shift and y-axis offset. The revised function is shown as:

$$y = B + A \cdot \frac{\mu^{(x-C)} e^{-\mu}}{(x-C)!} \tag{5.2}$$

Where,

y is the frequency,

x is the delay,

μ is the mean and standard deviation of delay,

B is the offset,

C is the shift of peak in x-axis to compensate the non-origin start point and

A is the amplitude of the peak of the Poisson pulse.

Again, Chi-square score is used to check the quality of the fitting and find the best fit function parameters. The final fit function is similarly made up of four independent

Poisson pulses. With the Poisson pulse, a better Chi-square score at 16.5 is obtained. The new result is presented in Figure 6-23.

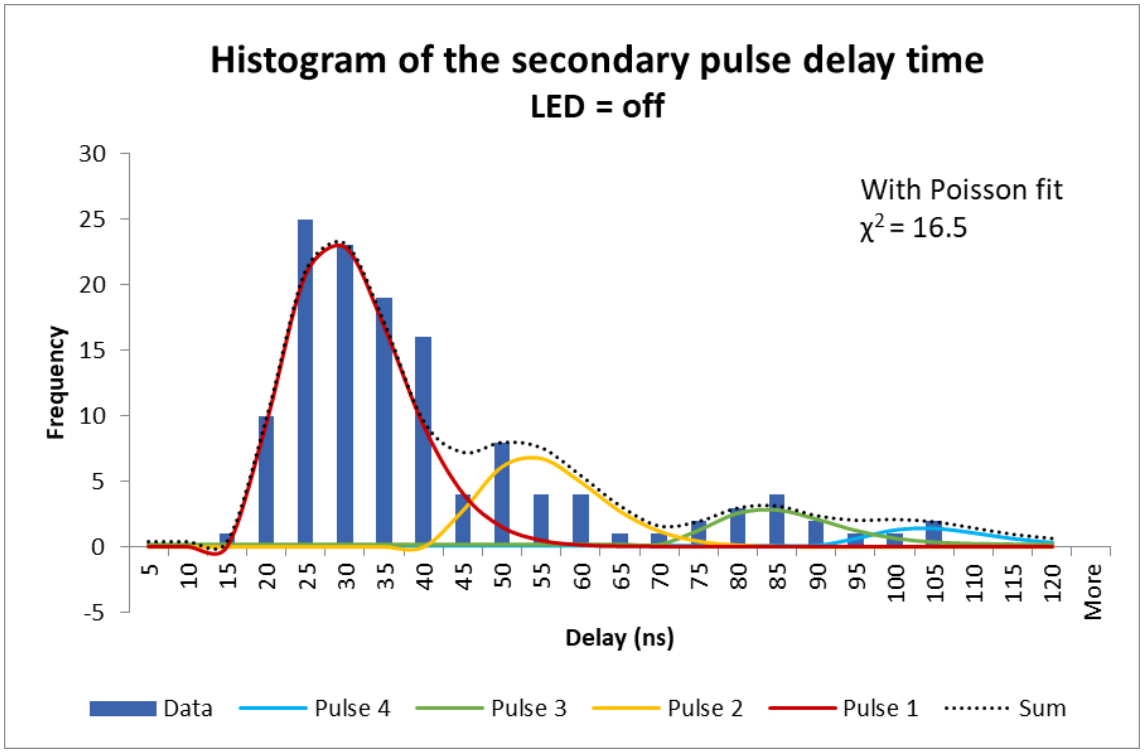


Figure 6-23 The fit function using Poisson pulse

The statistic modelling of the delay time of the secondary pulse doesn't directly contribute to the after-pulse ratio calculation, but it indicates an important feature of the delayed after-pulse. That is the timing of the occurrence of the after-pulse seems to obey the superposition of a number of Poisson distribution.

At the time of thesis writing, we discovered a study about the timing of the occurrence of after-pulse and crosstalk by J.Rosado [157] . We haven't examined the outcome of that study. It will be interesting to compare his study to our independent study finding in future work.

6.4.2 SiPM optical crosstalk

In the avalanche process, millions of carriers will be generated in a very thin depletion region, where the electric field can reach $10^5 – 10^6$V/cm. Of all the carriers, some might not contribute to the carrier multiplication, but instead they lose their energy to emit new photons. These photons with energy above the silicon band gap (~1.14eV) can travel to the neighbouring pixels and trigger new avalanches in them. This phenomenon is known as optical crosstalk and first observed by Newman[158].

To understand the feature of the waveform of the crosstalk signal, it is important to know its timing. Let's first estimate the travel time for the new-born photon to reach the neighbouring pixel and then compare it to the typical pixel avalanche time and recovery time.

All the photons, of any energy, travel at the same speed of light at $3x10^8$m/s in vacuum. In the material, the speed will change with its refraction index. So, in Si with a refraction index of 3.45, the speed of light is reduced to $0.87x10^8$m/s. Assume the pixel pitch (25um for S13360-1325CS SiPM) represents the typical distance for the photon to travel, the travel time to the nearest pixel is:

$$t = 25\mu m / 0.87 \times 10^8 \, m/s = 0.287 \, ps$$

This is much shorter than a typical pixel recovery time (~15ns). This means the new event(s) triggered by the cross-talk photons will stack on top of the existing event to form at least a 2p.e. high event (when one extra pixel is triggered.) This feature introduces a method to measure the cross-talk in a setup where only 1p.e. signal shall exist. A convenient source of such data is the file #1 where LED is off and only dark count presents. The study (from later section of this chapter) shows that more than 99.99% of the dark count is 1p.e. signal. So, it can be confidently claimed that those 2p.e. and above signal are the results of cross-talk. This methodology is formalized as:

$$R_{cross-talk} = \frac{N_{\geq 1.5\, p.e.}}{N_{\geq 0.5\, p.e.}} \tag{5.3}$$

Figure 6-24 shows the histogram of all the pulses peak from the data file #1 post Deconvolution-Wiener filtering. The filtered data includes all the after-pulses. Using this as the denominator might have improved the result. However, on the other hand, an after-pulse has the same probability to trigger an optical crosstalk as the genuine signal, so it is fair to count the after-pulse together.

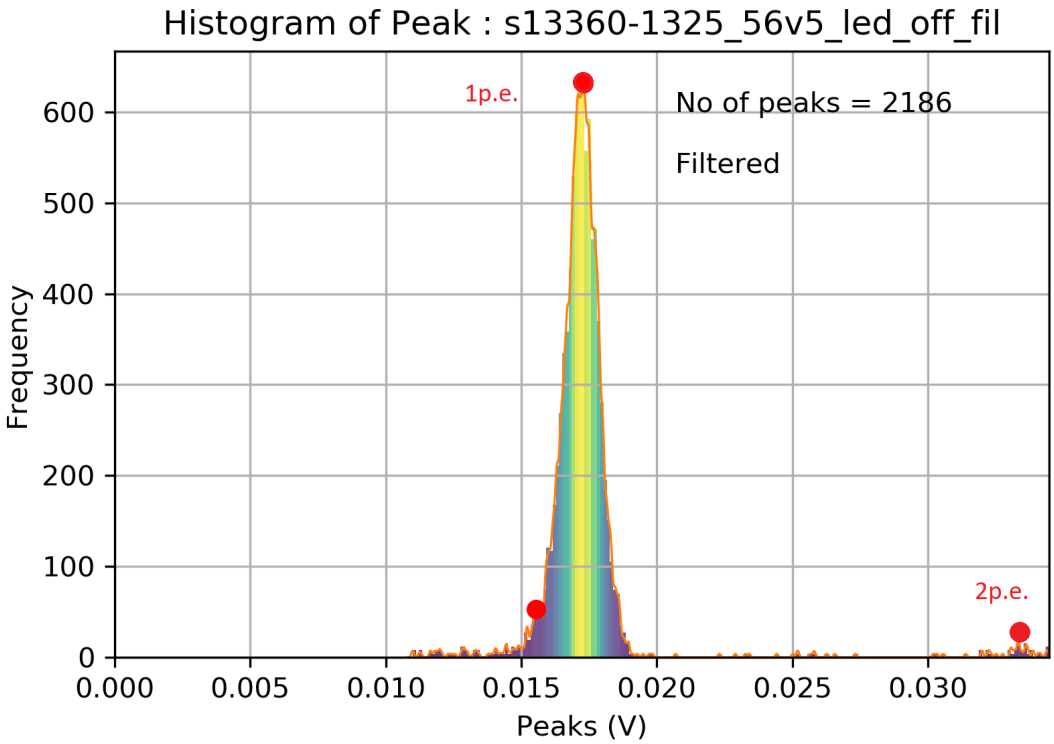


Figure 6-24 The histogram of pulse peak for the data file #1 post Wiener-Deconvolution filter

According to the plot, the 0.5p.e. threshold is set at 0.01V and 1.5p.e. threshold at 0.025V. Then run the counting statistics on the data with results shown in Table 6-3:

Table 6-3 Statistics for Optical Crosstalk calculation

Peak higher than 0.5 p.e.	2186
Peak higher than 1.5 p.e.	44
Ratio	2.01%

Compare the result to the SiPM S13360-1325CS datasheet plot shown in Figure 6-25. At 5V overvoltage, the crosstalk reads about **1.3%**. This corresponds well to our calculation result at **2.01%**.

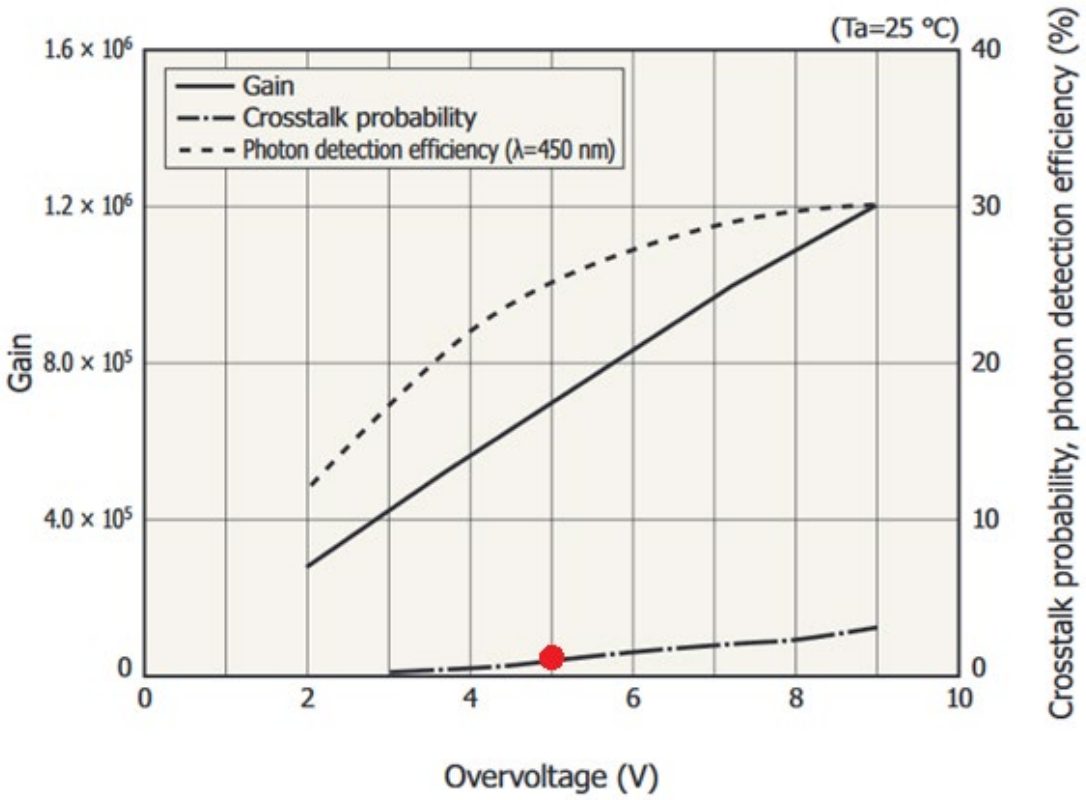


Figure 6-25 Official cross-talk plot against overvoltage from S13360-1325CS datasheet

6.4.3 SiPM photon detection efficiency

Patrick Eckert[69] has established a fine setup to measure the photon detection efficiency (PDE) of a SiPM. His setup used a range of special equipment, which is beyond the reach of this project. With the simple setup as Figure 6-9 describes, it is possible to make some analysis to estimate the photon detection efficiency.

First step is to work out how many photons arrived at the SiPM via the optical fibre. This calculation should consider how many photons are emitted from the LED, how many can fall into the acceptance angle of the optical fibre, how many is lost during the transmission over the fibre and how many is lost due to the mismatch of fibre output size to the SiPM holder acceptance hole size. These factors will be discussed in detail one by one.

Photon emission of the LED

Figure 6-26 shows the emitted photons in relation to the LED setting. This is measured by a PMT in front of the LED without the optical fibre for a single 8ns pulse. The LED driver adjusts the LED intensity by controlling its current. The plot shows the photon

emission/intensity is very linear to the LED setting between 4 and 9.5. Below 4, the photon flux is probably below the PMT's sensitivity, hence no data point is recorded. By the way, as previous section shows, our system saturates just before the LED setting of 5 and can detect signal from LED setting as low as 0.2. This shows the extreme photon sensitivity of SiPM system compared to the PMT system.

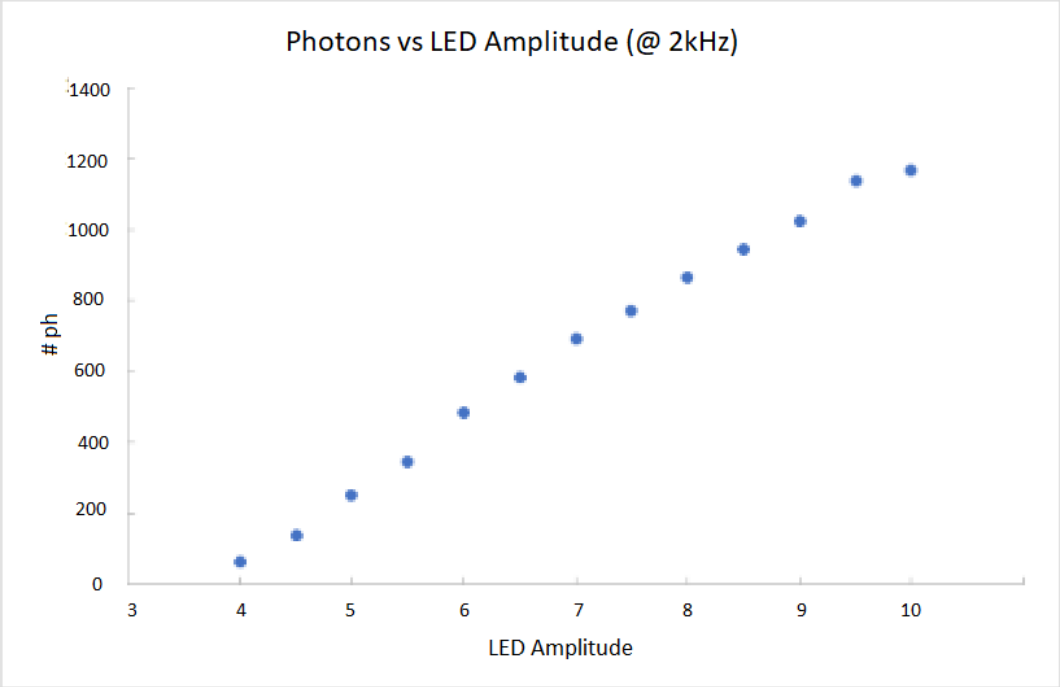


Figure 6-26 the relation between the LED setting and the output photons for SP5601 LED driver [155]

Optical fibre acceptance angle

The LED equipped in the driver is OSSV5111A from OptoSupply. Its emission directivity is shown in Figure 6-27. It shows a full cone of 15 degree for 50% of the beam power (the interaction between the beam profile and the middle half circular line which ends at the 0.5 mark on the bottom line). The full cone for 100% beam power (the interaction between the beam profile and the outer half circular line) is no more than 20 degrees.

The datasheet of the optic fibre states a full acceptance cone of **43.3** degrees, which is larger than the LED full beam angle **20** degrees. Thus, it can be assumed all the emitted photons enter the fibre.

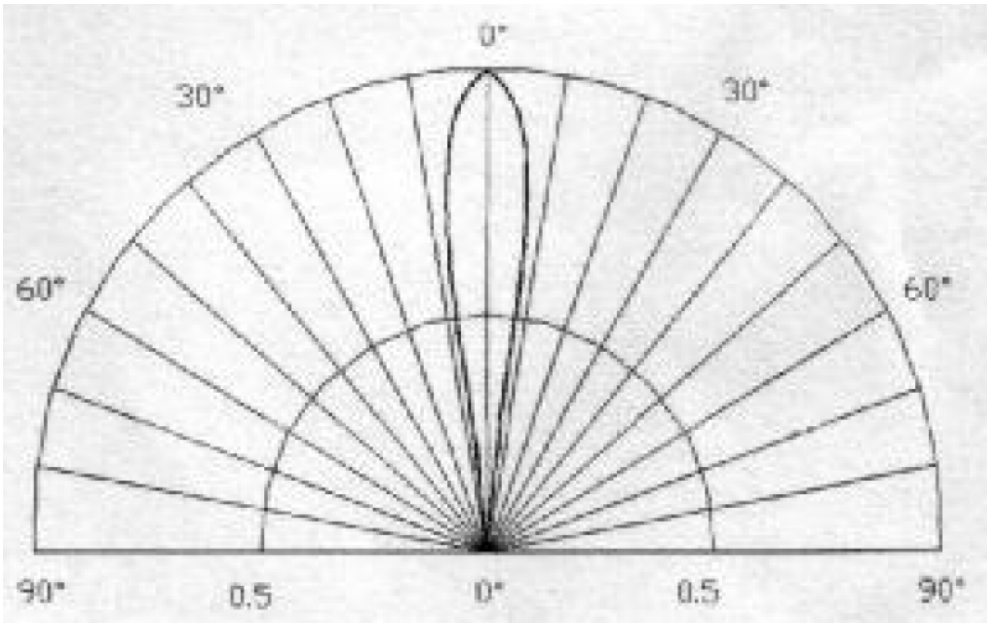


Figure 6-27 LED beam directivity [159]

Optical fibre transmission loss

Figure 6-28 is the typical spectral attenuation from the manufacturer datasheet. The principle wavelength of the LED is 400nm, which is just outside of this plot. We extend the curve by following its trend to 400nm. The estimated attenuation is roughly 60dB/km at 400nm. Our fibre length is 400mm. The attenuation is calculated as

$$A = -60dB/km \times 400mm = -0.024dB = 99.72\%$$

This means the loss is only 100%-99.72%=0.28%. Such small loss can be neglected for the simplicity of calculation. For the later calculation, the transmission efficiency is assumed to be 100%.

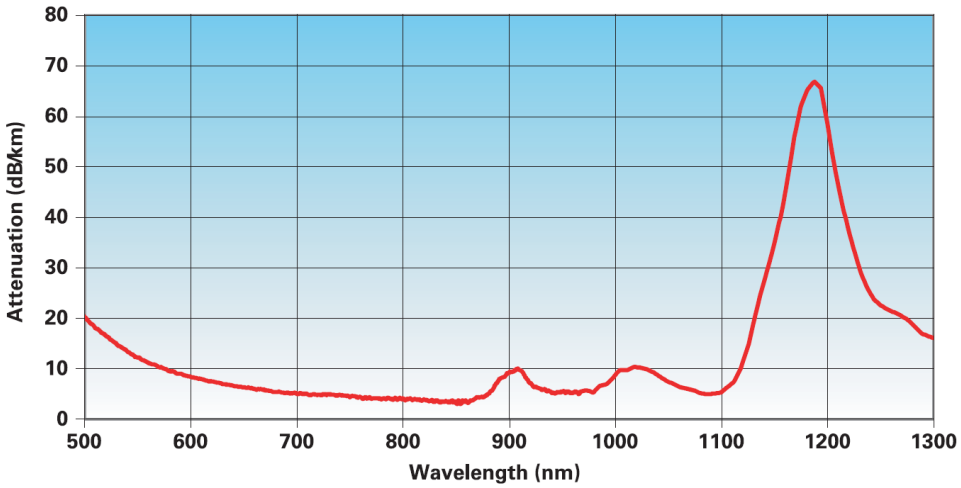
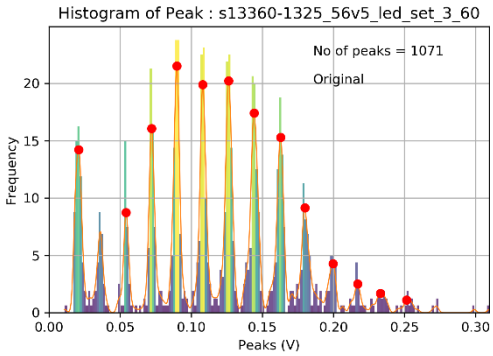


Figure 6-28 The signal (photon flux) attenuation of the fibre against the wavelength (from the datasheet of the optical fibre) [160]

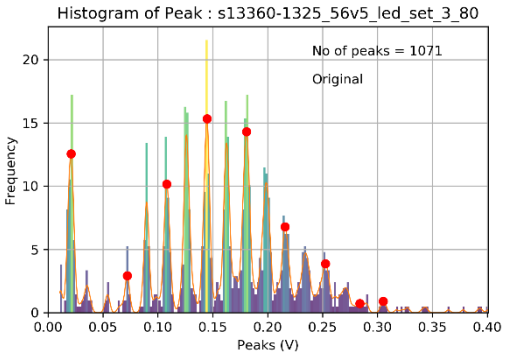
Mismatch between the fibre output core and the photon acceptance pin hole in the SiPM holder

This is a purely mechanical interface. Assume the beam is uniform across the fibre core cross-section, the dimension of the SiPM holder acceptance pin hole will determine the geometry photon loss. The diameter of the fibre core is 1mm. The diameter of the pin hole is 1.5mm. As the pin hole is bigger than the fibre core, there is no photon loss here.

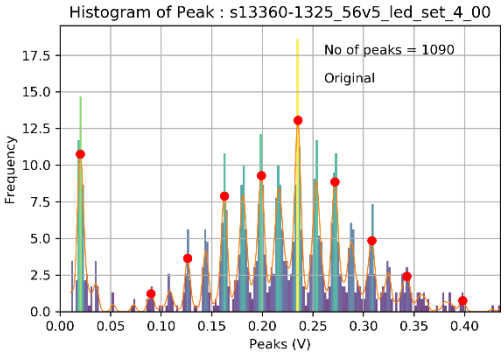
So, after this basic analysis, it is reasonable to conclude the number of photons arrives at the SiPM sensitive area is very close to the number of photons emitted by the LED. Figure 6-29 shows the histogram of the detected photons at different LED settings. The centre peak is taken as the **nominal number** of detected photons.



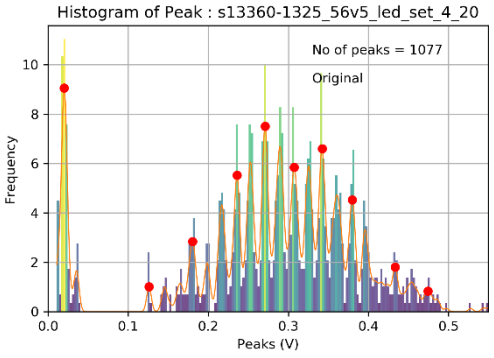
LED = 3.60



LED = 3.80



LED = 4.00



LED = 4.20

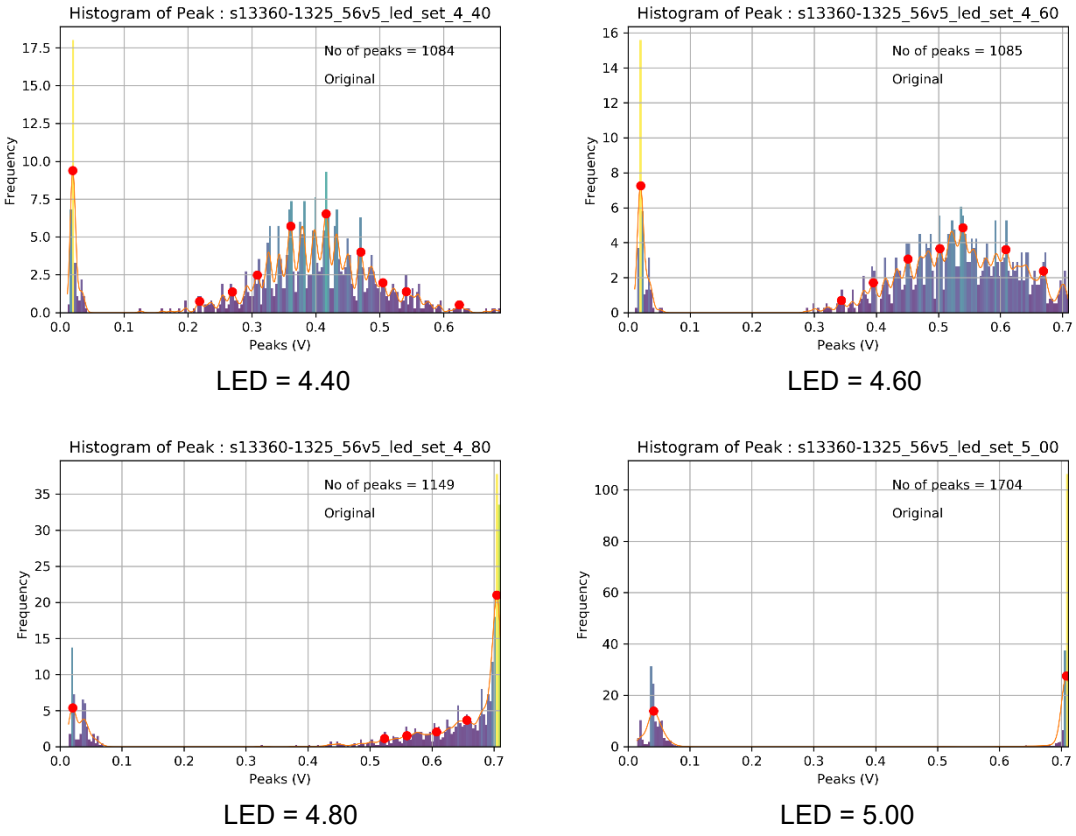


Figure 6-29 The histogram of detected signal peak at different LED settings

Through interpretation of Figure 6-26 and Figure 6-29, the photon detection efficiency is worked out as shown in Table 6-4:

Table 6-4 Calculation for photon detection efficiency

LED setting	Emitted Photon	Detected Photon	Efficiency
LED=3.60		7	
LED=3.80	35	8	22.86%
LED=4.00	74	13	17.57%
LED=4.20	113	17	15.04%
LED=4.40	152	23	15.13%
LED=4.60	191	29	15.18%
LED=4.80	230	38	16.52%
LED=5.00	269		
		Average =	15.89%

All the data are sight reading from the plots, so there is inaccuracy especially at both end of the data. (next section 6.4.4 introduced a mathematic method to calculate the

nominal SiPM system output. That should provide a better **nominal number** than the plot sight reading. However, that method also suffers the inaccuracy at both end of the LED setting.) The emitted photon at LED=3.60 is missing because that point is very difficult to interpret from the Figure 6-26. The detected photon at LED=5.00 is also missing because the signal amplitude is out of our SiPM detection system limit. So, these two lines of data are skipped. The calculated efficiency is averaged to represent the overall photon detection efficiency. Note, the line of LED=3.80 shows largely different efficiency value from the others. This probably is due to the low photon number causing excessive statistic error. So, this line is excluded from the average calculation.

The SiPM datasheet reports the photon detection efficiency is about **22%** at 400nm wavelength (Figure 6-30). This is close to the observation at **15.89%**. The difference probably is due to the inaccuracy in our data set, some omitted photon loss from LED to the SiPM and the excessive duration of the LED pulse (for the later period of the 8ns LED pulse duration, the SiPM and its readout preamplifier might be in deadtime period when light intense is high. This will result in the loss of photon detection.)

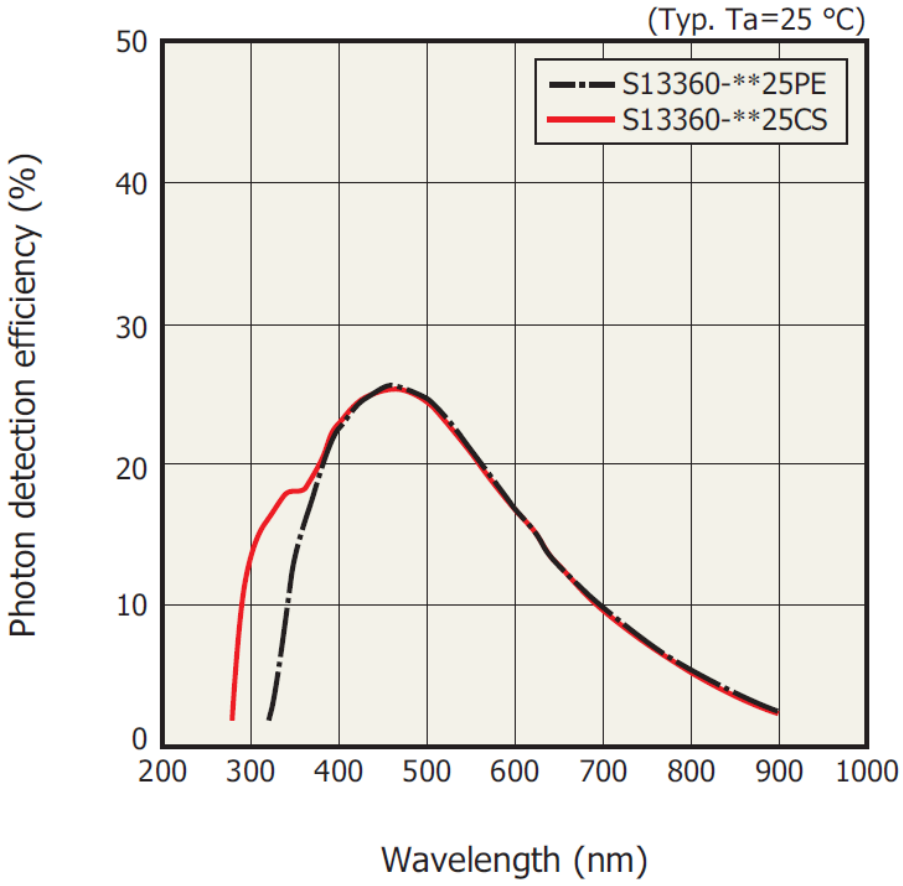
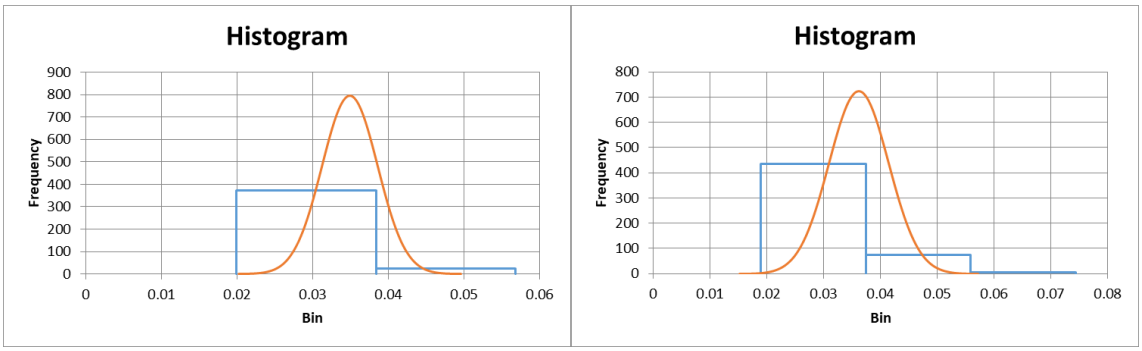


Figure 6-30 SiPM S13360-1325CS photon detection efficiency against the wavelength

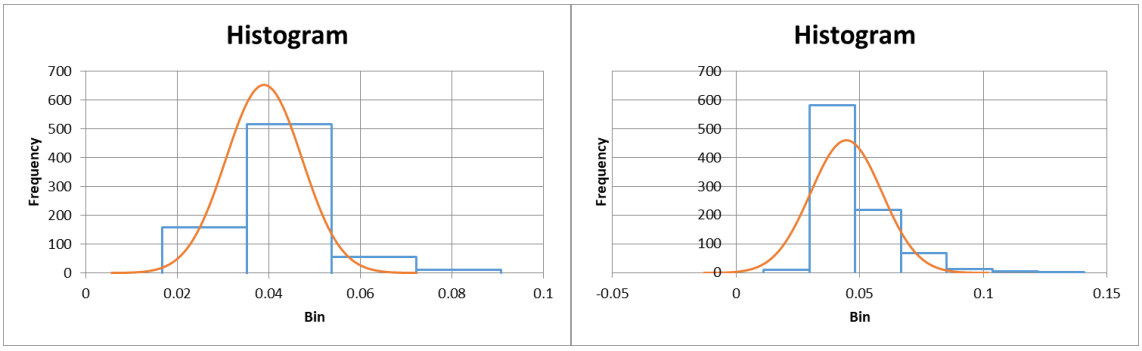
6.4.4 System linearity

The linearity is defined as the proportionality between the input number of photons and the output signal height. Figure 6-26 shows the linear response of the LED driver, but data is only available from LED setting 3.50 onwards. Our test starts from LED setting 0.0 and ends at setting 5.00 where the system is fully saturated. So, there is only a small section of overlap, which is suitable to calculate the photon detection efficiency as shown in Table 6-4.

To acquire a single number that can represent the system response to the input LED level, a method is developed to convert the distribution of response into a single number. First, for each data file, all the pulses with peak below 1.5p.e are removed. This is to remove the disturbance of dark counts. Then, the histogram is regenerated. Then, a Gaussian curve is constructed with the mean and standard deviation calculated from the data samples of the remaining pulses. The Gaussian curve is then scaled up to have its area match that of the histogram. Figure 6-31 shows some examples of the processing results. As shown, the more pulses are (corresponding to higher LED intensity), the better the Gaussian curve fits.



LED = 0.60 LED = 1.20



LED = 1.80 LED = 2.40

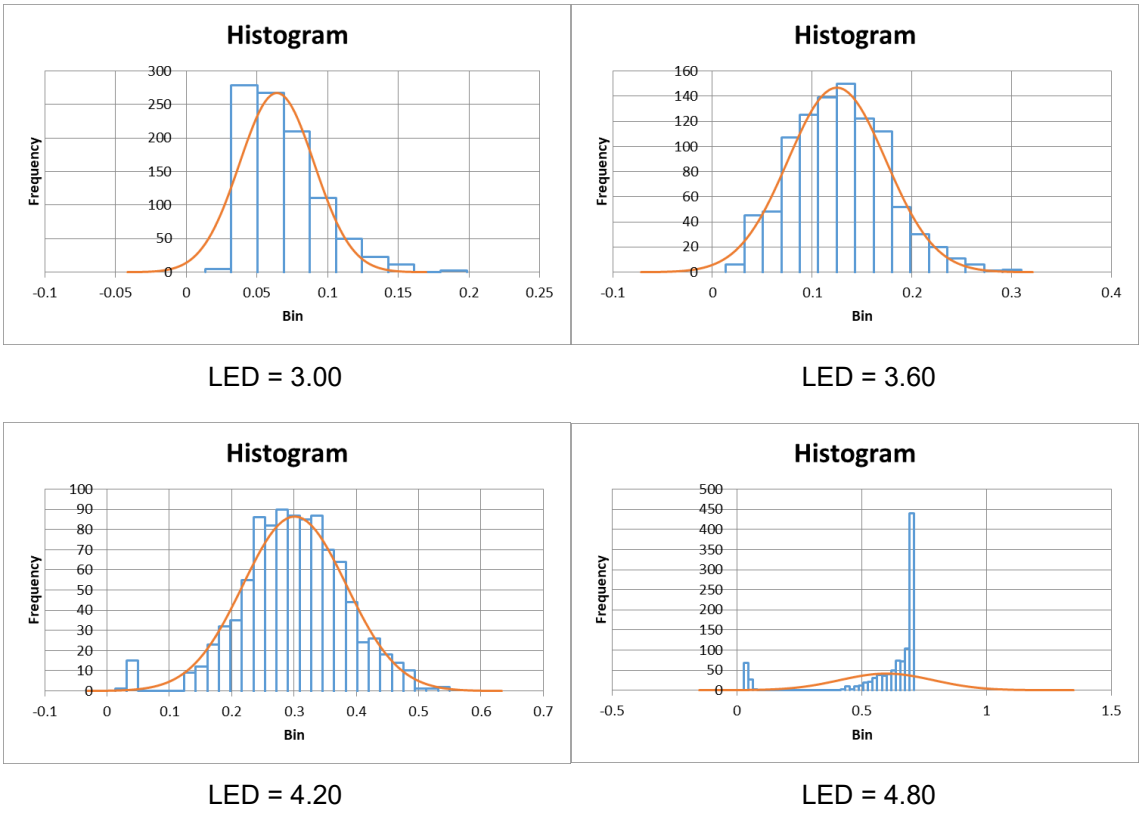


Figure 6-31 The histogram and the Gaussian curve fit for some data files at different LED setting

Then the mean of the constructed Gaussian curve is picked as the single number to represent the response of the system to the input LED level. Note that for the data file near both ends of the system dynamic range (the low end is when LED is barely on and the high end is when LED saturates the system), the fit Gaussian curve is highly distorted, so its mean can't represent the system output well. The mean represents the SiPM system output signal peak amplitude. This can be converted to the number of photons using the system gain 0.0185V/p.e. The converted photon number against the LED setting is plotted in Figure 6-32. The manufacturer calibration data is also plotted for the overlap section for comparison.

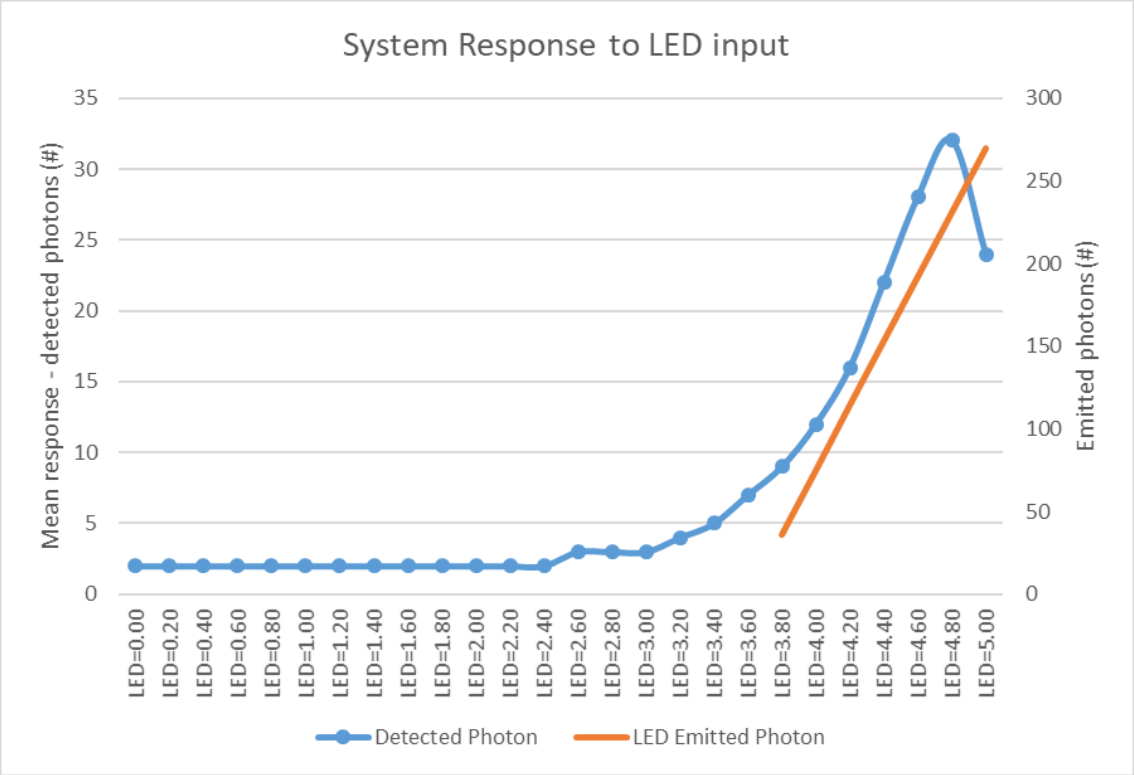


Figure 6-32 The system response to the LED input. The blue curve is the detected photons converted from the mean of the Gaussian curve. The orange curve is the replot of Figure 6-26 for the overlap section of the LED setting.

The result shows the system response to the LED input is not as linear as expected. It is more like an exponential curve. The reason is not fully clear, but it is suspected the nonlinearity is not in our detector system but in the timing of the arrival photons. When the LED driver is calibrated, probably the photons are integrated over the full 8ns gate duration. In the SiPM detector system, this 8ns gate period is too long. Only the photons generated/arrived at the initial period (although it is not clear what the exact period is) of the full gate duration will be detected. The number of photons might not be evenly distributed over the full gate duration, which results in the non-linearity behaviour of the SiPM system response.

6.4.5 System dynamic range

The dynamic range defines the minimum and maximum input signal that the system can maintain a linear response. The minimum level depends on the system noise level. A practical system shall maintain a signal to noise ratio of at least 10:1 as a figure of merit, but in theory, as long as the signal is above noise, i.e. signal to noise ratio is above 1, the signal can be discriminated from the noise. However, this doesn't allow any tolerance to uncertainty. If the signal obeys Poisson distribution, its deviation is

square root of the signal. So, for an expected nominal signal level of *A*, it is considered $A-\sqrt{A}$ as the minimum signal level. If noise also obeys Poisson distribution, for the expected nominal noise level of *N*, it is considered $N+\sqrt{N}$ as the maximum noise level. The ratio of $R=(A-\sqrt{A})/(N+\sqrt{N})$ must be maintained above 1 all the time. The minimum signal level that can satisfy this condition is defined as the **minimum input signal**. This minimum input level is also called system sensitivity.

The **maximum input signal** is defined slightly below the system saturation level for the similar reason discussed in the minimum input level. That is the signal will fluctuate around its nominal level. If the saturation level is defined as *L,* the maximum input level *A* shall be at $L-\sqrt{A}$. This way, the system linearity can be maintained, otherwise like the example of LED=4.80 in Figure 6-31, part of the distribution of the signal will be cut off, which will result in distortion in the data processing.

The **dynamic range** is defined as the range between the **minimum input signal** and **maximum input signal**. As SiPM behaves like a discrete device, the system dynamic range is expressed in the unit of p.e.

The dynamic range in p.e. term is determined by two factors of the system. Both will be discussed below. The factor that results in smaller range will determine the final system dynamic range.

The first factor is the SiPM sensor. The ideal response of a SiPM sensor is:

$$N_{p.e.}=N_{photon}\times PDE \tag{5.4}$$

As a SiPM only has a limited number of pixels. Therefore, its real response shall be adjusted as:

$$N_{p.e.}=N_{pixel}\cdot(1-e^{\frac{-N_{photon}\times PDE}{N_{pixel}}}) \tag{5.5}$$

If the LED gate duration is longer than the pixel recovery time, Eq. (5.5) needs further adjustment:

$$N_{p.e.}=\frac{N_{pixel}\times T_{Gate}}{T_{recovery}}\cdot(1-e^{\frac{-N_{photon}\times PDE}{N_{pixel}\times\frac{T_{Gate}}{T_{recovery}}}}) \tag{5.6}$$

Using the data of SiPM S13360-1325CS as N_{pixel} = 2668, PDE = 25% (official peak efficiency), Figure 6-33 plots the ideal and real response of the SiPM to the incident number of photons with ideal curve calculated using Eq. (5.6). The error between the ideal and real response increases quickly as incident photons gets more intensive. The calculation shows error reaches 5% when N_{photon} = 1120, with corresponding $N_{p.e.}$ = 266. When N_{photon} = 2280, with corresponding $N_{p.e.}$ = 513, error reaches 10%.

Thus, if aim to maintain no more than 5% linearity error, the maximum $N_{p.e.}$ is **266p.e.**

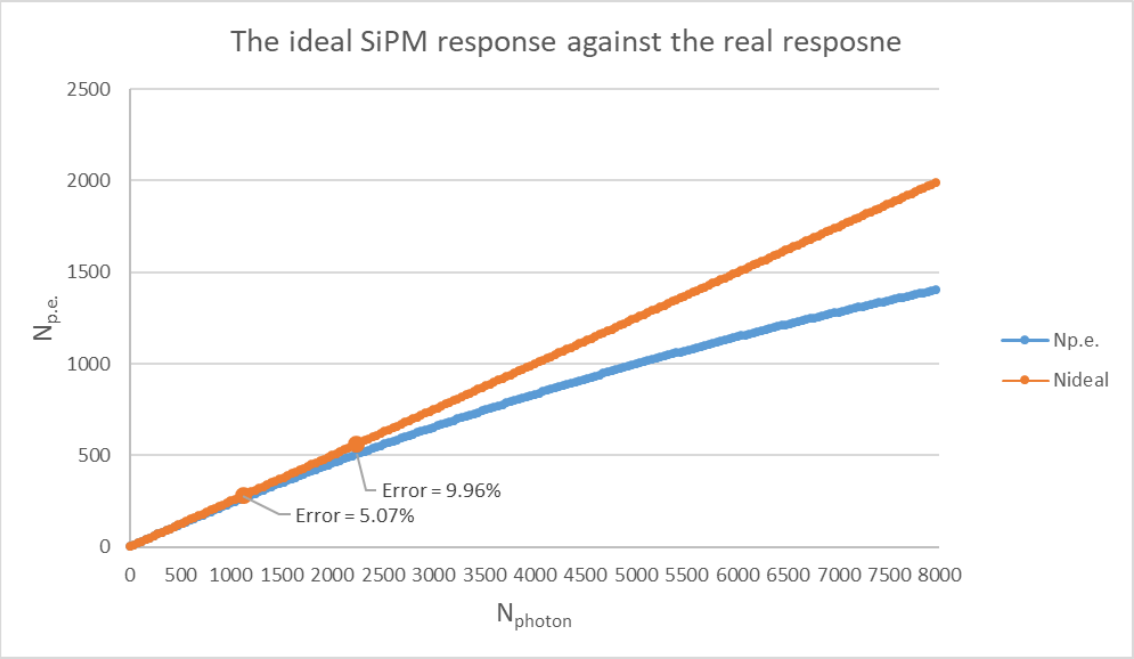


Figure 6-33 The comparison of ideal and real SiPM response. The error increases as the number of incident photon increases.

The second factor is the readout electronics (including the preamplifier, the data acquisition and digital conversion). The preamplifier circuit can output signal as high as 2.5V while the DRS4 module can only accept input signal as high as 0.95V. The real data shows the peak input signal is clamped at 0.7V (see Figure 6-16). Take the preamplifier gain of 0.0185V/p.e., the maximum signal in terms of p.e. that readout electronics can accept is $0.7 / 0.0185 \approx 40 p.e.$

Comparing the calculation outcome by the two factors, the smaller result of **40p.e.** shall be taken as the maximum input signal.

The general dark count is a 1p.e. signal. Considering the optical crosstalk impact, some dark count will appear as 2p.e. signal. In all, the noise level is no more than 2p.e. Hence, **the minimum input signal** can be defined at 3p.e. to exclude all the dark counts.

We can summarize that the system dynamic range in **p.e.** term is **3** to **40p.e.** and considering PDE at 25%, the dynamic range in **photon** term is **12** to **160** photons.

6.5 SiPM Dark Count Study

The SiPM dark count is an event triggered by electron (e-) or hole (p+) thermal activity near the depletion region of the pixel diode. At a temperature T, a fraction of charged particle (e- or p+) in the device will gain enough energy to jump from the valence band to the conduction band. In the conduction band, the charged particle will move freely. Under the influence of the electrical field applied by the external bias voltage, some of them will move into the depletion region, where extreme high electric field exists. Under the extreme high electric field, the charged particle can trigger an avalanche process, which grows into a full pixel event. This event due to its self-generated nature is called dark count.

6.5.1 Statistic Model

The dynamic process of formation of SiPM signal, particularly the dark count, can be described by a mathematical, statistical model. This model can help to understand the signal time distribution, which can be used to identify the unwanted signal (like dark count) from the wanted signal.

SiPM Signal Statistic Model

Assume each pixel is an independent trial, with status of 0 (quiet) or 1 (triggered), the output of SiPM can be described as a binomial distribution with the triggering probability of each pixel at ***p***. If define the output as a random variable ***Y***, the probability of ***k*** pixels triggered can be given by:

$$P(Y=k)=\binom{n}{k}p^k q^{n-k}, k=0,1,2...n$$
$$where \qquad (5.7)$$
$$\binom{n}{k}=\frac{n!}{k!(n-k)!}$$

Note, ***n*** is the total number of pixels and ***q*** = 1-***p***. This distribution has a mean of ***np*** and a variance of ***npq***.

Eq. (5.7) represents the SiPM output signal amplitude in terms of the number of p.e. at any instance time slice. The amplitude can range from k=0 (no pixel is triggered) to k=n (all the pixels are triggered). When counting the dark events, the number of pixels being triggered at a single time slice is not important, only the fact that if there is pixel triggered (k>0) or not (k=0) matters. Thereby, the probability of two binary cases (k=0 and k>0) can be derived from Eq. (5.7) as:

$$\begin{aligned} p_0 &= P(Y=0) = q^n = (1-p)^n \\ p_1 &= 1 - p_0 = 1 - q^n = 1 - (1-p)^n \end{aligned} \tag{5.8}$$

Then for ***m*** time slices, the total number of dark counts ***X*** can be described as another binomial distribution:

$$\begin{aligned} &P(X=k) = \binom{m}{k} p_1{}^k p_0{}^{m-k}, k = 0,1,2...m \\ &where \\ &\binom{m}{k} = \frac{m!}{k!(m-k)!} \end{aligned} \tag{5.9}$$

The mean ***μ*** is mp_1 and the variance ***σ*** is mp_0p_1. If **m** can cover the length of one second, the mean ***μ*** will equal to the dark count rate (# per second). This provides a way to calculate the single pixel triggering probability ***p***. Replace p_1 with ***p*** using eq. (5.8), the expression of ***μ*** can be expanded as:

$$\begin{aligned} \mu &= mp_1 \\ &= m(1-(1-p)^n) \end{aligned} \tag{5.10}$$

Rearrange to solve ***p*** as:

$$p = 1 - \sqrt[n]{1 - \frac{\mu}{m}} \tag{5.11}$$

Some real data like the SiPM S13360-1325 with **n** = 2668 and **μ**=70,000cps (counts per second), is used for a trial. The unit time slice period is defined at 1ns. Then for one second, the total number of time slices **m** = 1E9. Apply Eq. (5.11), then ***p*** can be worked out at 2.62378e-8.

With ***p*** worked out, ***q*** is just 1-***p***. Then a few k values of Eq. (5.7) can be calculated. The result is listed in Table 6-5:

Table 6-5 The theoretical calculation of the success rate for each k value using Eq. (5.7)

k=	P(Y=k)	Success in 1E9 Trials (m)
0	0.99993	999,930,000
1	6.99976E-05	69,998
2	2.44908E-09	2
3	5.71044E-14	0
4	9.98238E-19	0
5	1.39549E-23	0
6	1.62507E-28	0
7	1.62148E-33	0
…	…	…
2668	~0	0

The result shows the triggering probability of a single pixel is so low that at any time slice, the probability of zero pixel triggering (k=0) is 99.993%, the probability of one-pixel triggering (k=1) is 0.006999%, the probability of two or more pixels triggering (k>=2) is less than 0.00000025%. For 1E9 times of trial (see results in the third column) in one second period, the success rate for one-pixel triggering (1p.e.) is 69,998 cases, for two-pixel triggering (2p.e.) is only 2 cases and for three-pixel or more is 0. The sum (2+69,998) matches the DCR rate of 70,000 exactly. This match is no surprise as the ***p*** value used for this Binomial distribution calculation is derived from the DCR rate, which is the ***μ*** from Eq.(5.10)

This statistic model and the calculated probability table suggests Dark Count (DC) will be mostly constituted of 1p.e. signal. The chance for a DC to be 2p.e. signal is next to zero. The chance of 3p.e. signal or above can be safely assumed to be zero. However, in the real measurement, much higher rate of appearance of 2p.e. signal than what the statistic model predicts are observed. This is because the effect of cross-talk and after-pulse.

6.5.2 Simulator

Based on the statistic model developed in 6.5.1, a simulation program is developed in Python to investigate the distribution of time delay of any consecutive events. The aim is to investigate if the knowledge of delay can assist to distinguish the fully grown after-pulse event from a genuine DC event.

6.5.2.1 The simulator construction and description

The simulator program flow chart is shown in Figure 6-34. The deadtime discussed in the 6.2.3.2 is fully considered. The program is constructed with two hierarchy of loops. The outer loop iterates ***m*** times (m is the number of time slice, 1 time slice = 1ns). The inner loop iterates ***n*** times (n is the number of pixels). In each inner loop, a pixel will first be checked for status. If it is active, then a random dice will be thrown (using a random number generator obeying uniform distribution between 0 and 1). If the result is less than the pixel triggering probability ***p***, the pixel will be marked as triggered and the event is registered. Immediately, the pixel will be marked as inactive for the next 50ns (= 50 time slice). Out of the inner loop, if at least one pixel has been triggered, one DC event is recorded and for the next 20ns, the device is marked as inactive (For program convenience, simply all the pixels are marked as inactive for the next 20ns.) This 50ns inactive period for a single pixel corresponds to the dead time of the pixel. The 20ns inactive period for the device corresponds to the dead time of the detector system. 6.2.3.2 states 50ns for pixel deadtime and ~15ns for device deadtime. We add 5ns to make it 20ns to reflect the observation of real 1p.e. signal duration. The time slice number is recorded as the time stamp of the DC event. As each slice represents 1ns, the slice number simply represents the number of ns elapsed. Once the outer loop is completed, all the collected DC events are saved to a CSV file (a common text file format using comma separating the data point.)

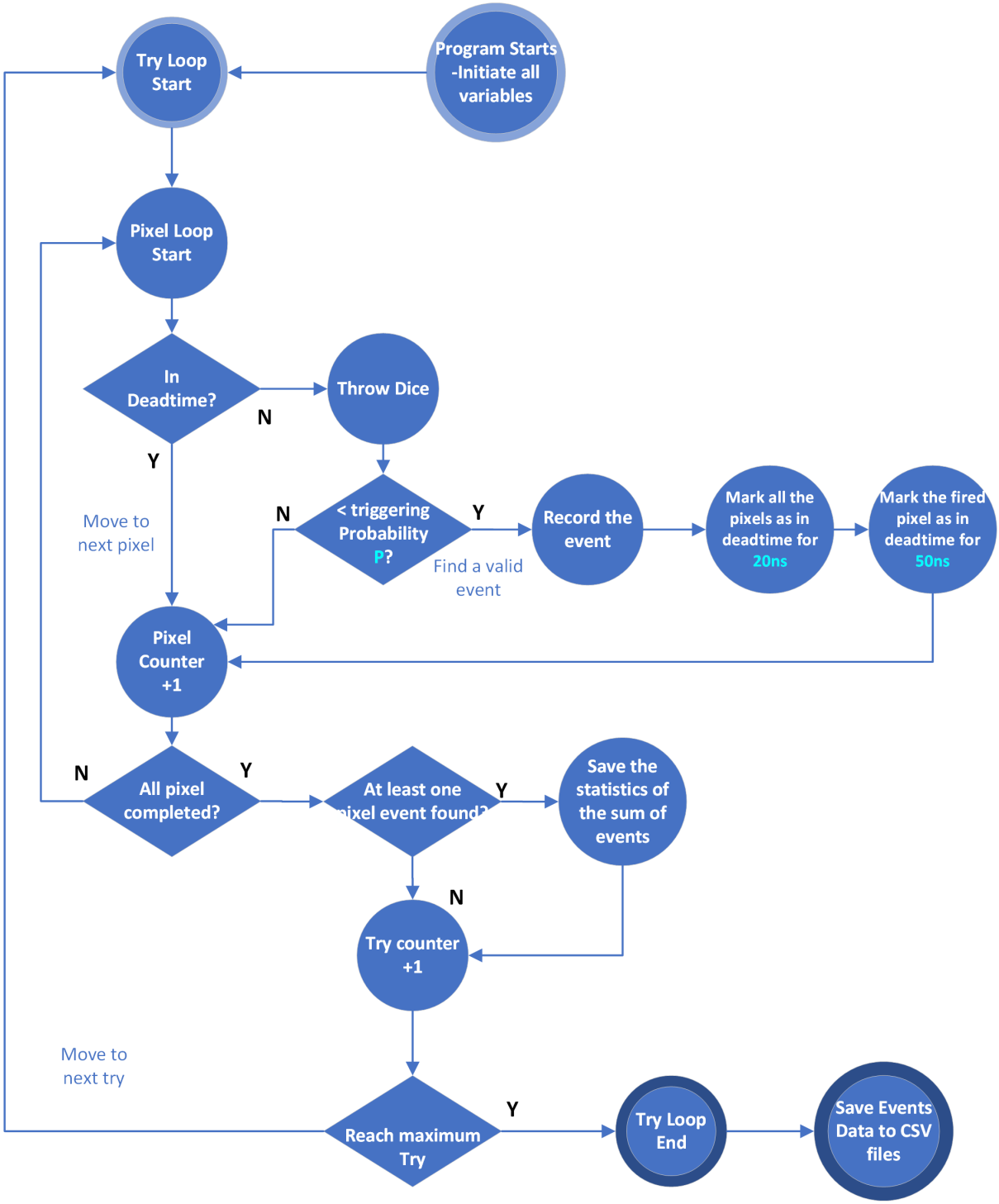


Figure 6-34 The DCR simulation algorithm flow chart

As the time slice is set to 1ns, to complete one second simulation, it requires 1E9 time slices. This make the simulation unviable big for memory consumption. So, the simulation is limited to 1ms duration, which equals to 1E6 time slices. Then the simulation

is repeated multiple times. The results can be merged together to make up a large period without distorting the results, as there is no correlation between each simulation run.

6.5.2.2 The simulation results

6.5.2.2.1 The simulated DCR

In total, 60 runs of 1ms simulation are acquired. This totals 60ms simulation time. The number of events found in each run is shown in Table 6-6.

The result ranges from minimum 46 events per 1ms simulation to maximum 89 events per 1ms, with the average level at 72.07 events. This is equivalent to **72.07kcps** DCR, if the result is scaled up to 1second. The ***p*** figure used in the simulation is derived from the real measurement of DCR at **72.93kcps** instead of the datasheet figure of 70kcps. The close match of the simulation result of 72.02kcps to the real DCR result of 72.93kcps proves the choice of randomness distribution in the simulation program matches the nature distribution of DC in the real experiment very well. The slightly lower figure of the simulation result is probably due to the impact of the deadtime and the statistic fluctuation error from the limited sample size (only 60 runs of total 60ms simulated, which represents 60ms/1000ms = 6% sample size).

The histogram of the number of events over 60 runs are constructed and shown in Figure 6-35. It is expected to see a Gaussian distribution with the peak (mean) at 72. This does not match the histogram presented. The reason is unclear and perhaps this is due to the sample size is not big enough to show up the Gaussian distribution or the 1ms duration is too short with not enough DC events.

Table 6-6 The DCR simulation results with the duration of 1ms for each run

Runs	Events	Runs	Events	Runs	Events	Average
1	74	21	70	41	76	72.07
2	69	22	66	42	80	**Min**
3	76	23	46	43	63	46
4	76	24	81	44	82	**Max**
5	77	25	71	45	49	89
6	64	26	63	46	76	
7	80	27	66	47	89	
8	84	28	64	48	62	
9	79	29	61	49	75	
10	73	30	80	50	65	

11	66	31	66	51	89
12	77	32	69	52	79
13	70	33	72	53	80
14	75	34	63	54	73
15	74	35	73	55	63
16	68	36	65	56	71
17	82	37	82	57	71
18	67	38	83	58	75
19	61	39	69	59	72
20	79	40	88	60	65

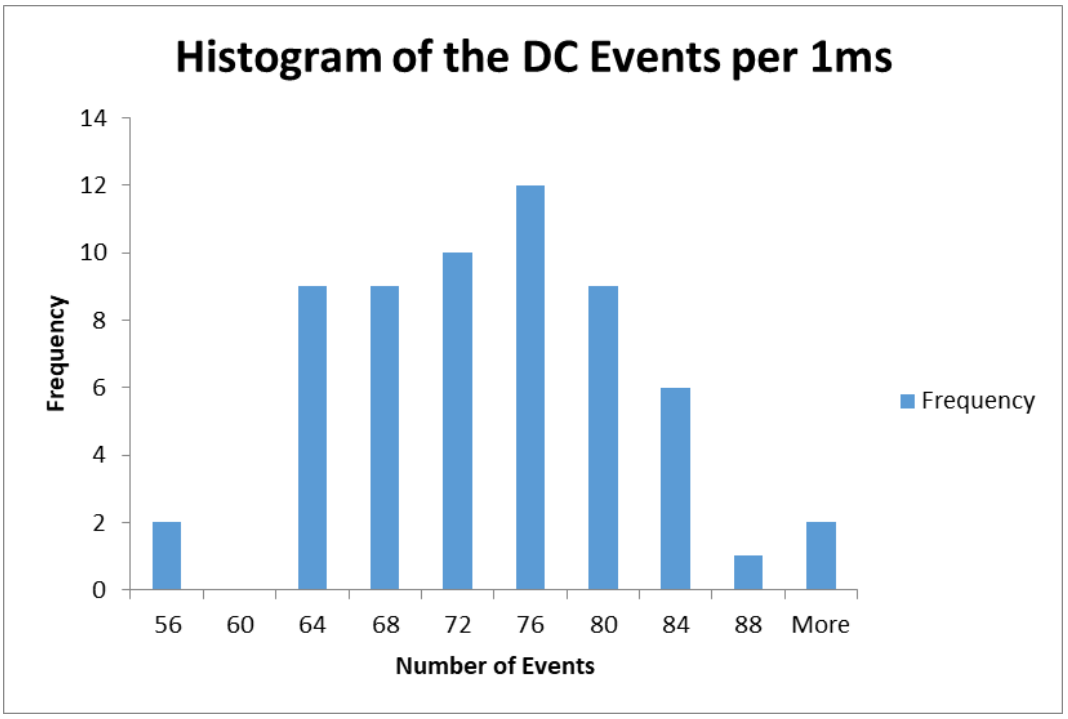


Figure 6-35 The histogram of the number of DC events in each 1ms simulation run

Next, the delay time between each consecutive event will be studied. The average gap obviously is the reciprocal of DCR. If use the figure 72.93kcps, its reciprocal is **13.7us**. There is a collection of 4324 DC events out of 60 runs. The timestamp of each DC event is used to calculate the delay. The results range from 32ns to 127028ns with an average delay (mean) at 13704ns(=**13.704us**). This matches our calculation of the reciprocal of DCR very well. With this mean, the standard variation of the delay is calculated at 13723.6ns(=**13.724us**). The coincidence of the mean and standard variation value is an important clue. And this finding will be discussed in the next section about the delay distribution.

6.5.2.2.2 The distribution of delay between events

To check out the distribution of the delay, its histogram is built as shown in Figure 6-36.

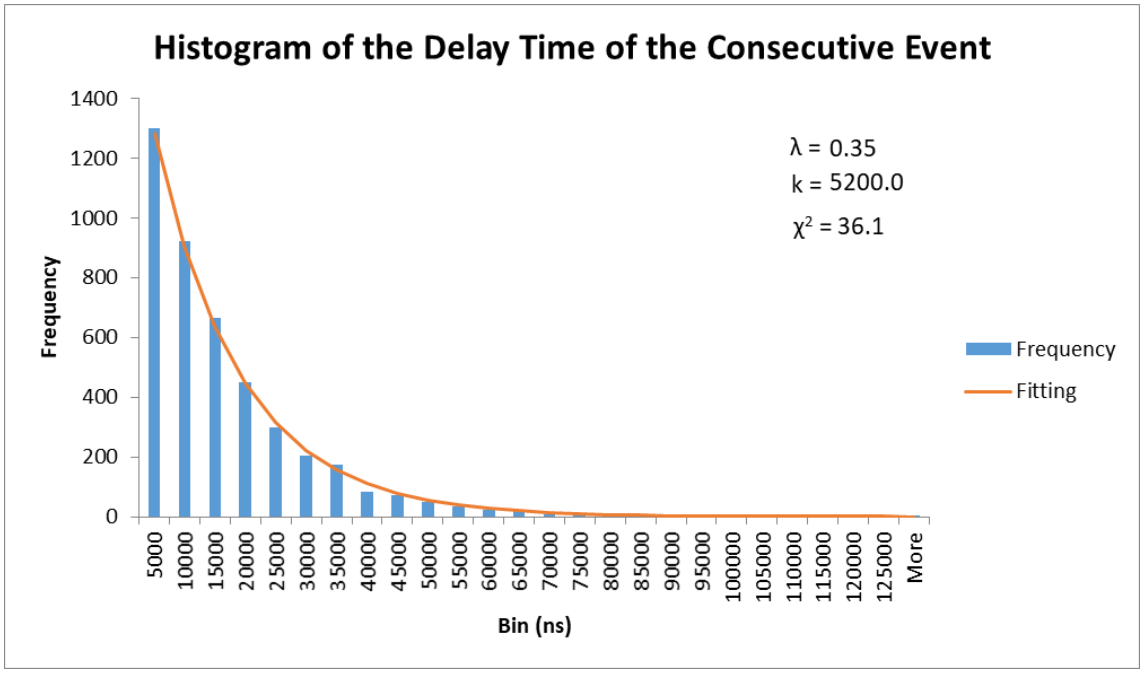


Figure 6-36 the histogram of the delay between consecutive events over the simulation data set

The plot clearly suggests an exponential distribution. The fact that its mean equals to its standard variance matches the characteristic of an exponential distribution, whose mean equals to standard variance. The probability density function (PDF) of an exponential distribution is:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, x \geq 0 \\ 0, x < 0 \end{cases} \tag{5.12}$$

Here, λ is the single parameter of the distribution, often called the rate parameter. The mean and standard variance of this distribution is both 1/λ.

The PDF provides a distribution between 0 and 1 and the integration of the PDF equals to 1. The PDF can be scaled up by a factor of k to generate the sample data between 0 and k while maintaining the distribution shape. So, the slightly modified equation (5.13) is used to fit the histogram data.

$$g(\lambda, x) = k\lambda e^{-\lambda x}, x \geq 0 \tag{5.13}$$

The fitting curve is shown as orange line in the Figure 6-36. The parameter k and λ values are experimented to get the best chi-square score.

The finding matches the following statement which links the Poisson distribution to the Exponential distribution:

"*The exponential distribution is the probability distribution of the time between events in a Poisson point process, i.e. a process in which events occur continually and independently at a constant average rate.*" [161]

6.5.2.3 Other findings

Since our DRS4 acquisition window is often defined at 200ns. It is interested to know the chance of two consecutive genuine DC events occur within the same window. From the simulation data set, a total of 53 out of 4324 cases with delay less than 200ns. This sets the probability of capturing two or more events in a single 200ns acquisition window at 53/4324=**1.23%**. As the simulation only creates genuine DC events with zero after-pulse or cross-talk, the figure 1.23% calculated from the simulation data could represent the probability of a genuine second or more DC event captured in the same 200ns window. The figure is very close to the theoretical estimation in 6.4.1.1 as 28/2000=**1.4%**, with the difference just being the fluctuation introduced by the random sampling in the simulation.

Another finding from the simulation data is that for every randomly generated DC event, there is only one pixel fired. This matches what the statistic model predicts. This provides strong support for the judgement that those events with amplitude equal to or above 2p.e. is not a DC event with multiple pixel triggering at the same time, but a DC event coupled with optical cross-talk.

6.5.3 SiPM Dark Count Measurement

6.5.3.1 Dark Count Rate

Any theoretical model can't replace the value of real measurement. So naturally, the next step following the theoretical modelling is the real measurement with the current hardware setup.

In general, the DCR can be obtained with the following simple ratio when the data is sampled with LED off:

DCR = (Number of pulses) / (Event time window)

To get good statistics, the number of pulses and event time window shall be as large as possible with minimum interruption like the dead time. Early discussion in 6.2.3 shows large system deadtime that leads to very slow event handling rate of our readout system. To overcome this, a wide bandwidth counter replaces the DRS4 module. The new system setup is shown in Figure 6-37.

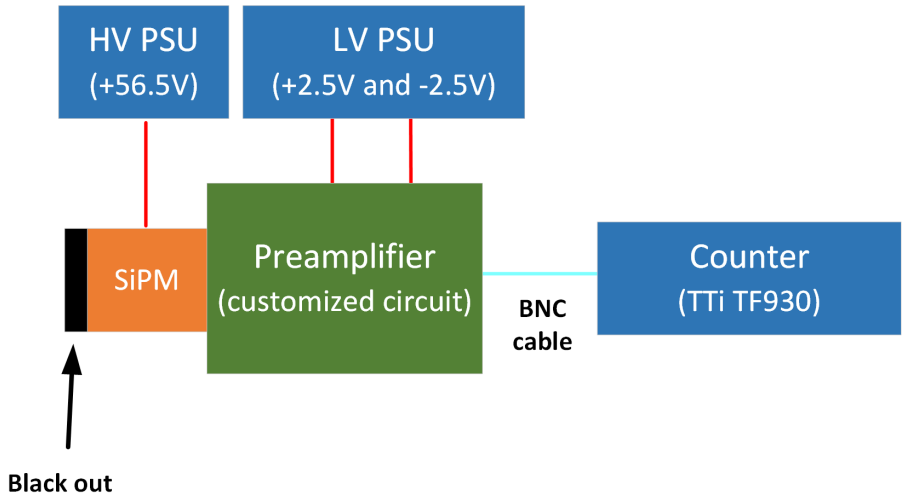


Figure 6-37 The DCR measurement system setup with a fast Counter replacing the DRS4 module

The counter has a bandwidth of 3GHz. This enables it to catch the short pulse directly out of the preamplifier. The counter has an analogue dial knob to adjust the triggering threshold. In order to set the threshold correctly, the SiPM and preamplifier board is replaced with a function generator. The generator is configured to mimic the real SiPM dark count signal (=1p.e.) at a given frequency. Then adjust the counter threshold setting until the reading is stable and match the generator signal frequency. This way the counter threshold setting is correctly calibrated and set. After this step, the SiPM and the preamp board is swapped back and ready to measure the real DCR.

The average reading of DCR acquired is about 75kcps at the bias voltage of 56.5V at the room temperature. This is close to the datasheet figure of 70kcps with 5V over-voltage. As the temperature of the SiPM is not actively controlled during the acquisition period, the real temperature of SiPM is expected to be slightly above the room temperature as the dissipated power of SiPM will heat up the device.

It is well known the DCR is highly correlated to temperature. The higher the temperature is, the higher the DCR is. DCR is also correlated to the bias voltage (strictly speaking, it is the over voltage part makes the effect.) in the similar way. The underlying

cause of the temperature influence is that statistically more electrons will gain enough energy to travel into the depletion region at higher temperature. With higher bias voltage, the depletion region expands so that the chance of a thermal electron entering the depletion region increases.

With the current setup, it is relatively easy to measure the DCR variance against the bias voltage. Figure 6-38 shows the DCR vs the bias voltage with the bias voltage sweeps from 56V to 66V.

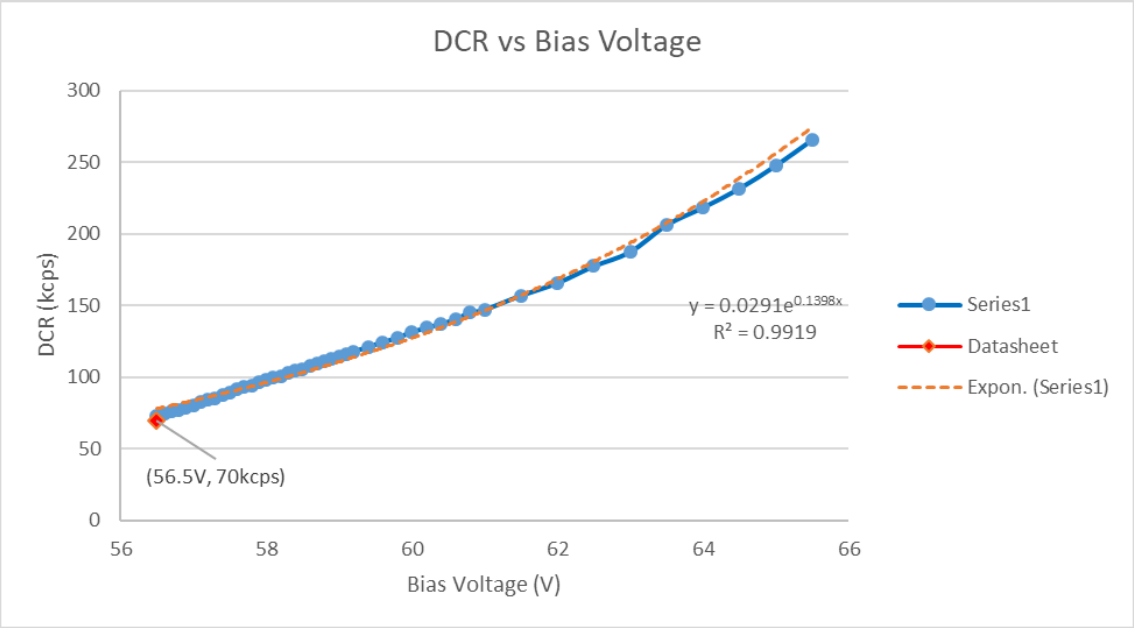


Figure 6-38 The DCR variation with the SiPM bias voltage

The plot shows a good fit with an exponential curve. But through close examination, the data is better divided into two sections. From 56.5V to 61V is the first section where data shows quite linear trend. From 61V and onwards is the second section where data shows more exponential trend.

It is not quite clear why the data shows two different trends. It could be that in the first section, the bias voltage impact is dominant. The relationship between dark count rate and bias voltage is likely linear. In the second section, the temperature impact is dominant, which might have exponential relationship to the dark count.

It is understood, the energy distribution of electrons at any thermal equilibrium with its environment is governed by Fermi-Dirac distribution function. The function has a form of:

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \tag{5.14}$$

Where, **E** is the energy level of the electron. $\mathbf{E_F}$ is the Fermi level. **k** is the Boltzmann constant and **T** is the temperature in kelvin.

At a given **T** and known $\mathbf{E_F}$, the electron energy distribution is defined. A portion of the electron above an energy threshold can move to trigger the dark count. In a semiconductor, the threshold is the difference between the bottom conduction band energy and the donor energy level. This difference is lower than the semiconductor band gap. Therefore, only a quite small energy is required to overcome this threshold. For example, for donor material Sb in silicon semiconductor, the threshold is only 39meV, a small fraction of the 1.12eV silicon bandgap.

As temperature increases, the distribution curve shifts, but the threshold doesn't change. This means more electrons gain energy above the threshold and this leads to trigger more dark counts. The increment of the number of electrons above the threshold is exponential to the increment of temperature T. Figure 6-39 shows the electron energy distribution at different temperature in relation to the threshold (this is an artificial energy threshold for illustration purpose only.) Clearly, the portion of the Fermi-Dirac curve above the threshold increases exponentially as temperature increase. Limited by the scope of this research, further theoretical study of the dark count is left for the next step.

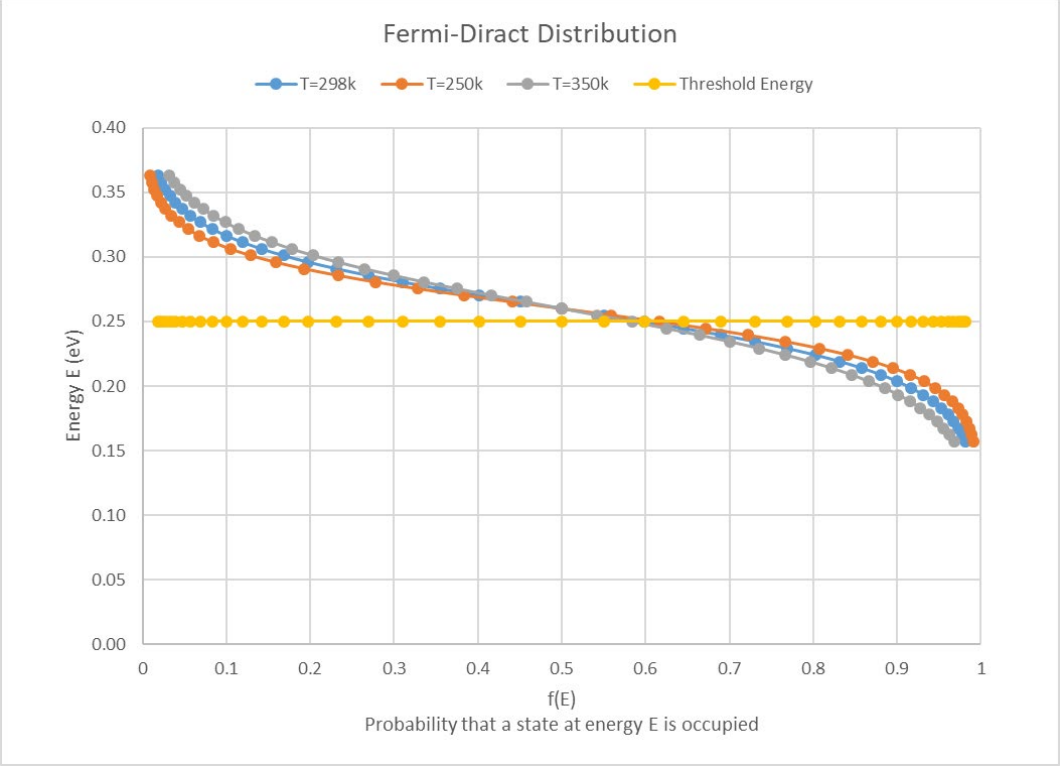


Figure 6-39 The illustration of Fermi-Dirac distribution

6.6 Conclusion and future work

This chapter introduces the SiPM detector system characterization test with a variable LED source. This serves as a preparation before the final test with the radioactive source. The Python scripts introduced in the last chapter is combined into a single master script to batch process the characterization test data.

The data set has been deeply examined to extract the SiPM detector system characteristics. They are mostly dominant by the SiPM characteristics. It is found the system total after-pulse rate is about **7.9%**, the crosstalk at 5V overvoltage is **2.01%**, the overall photon detection efficiency is about **15.89%**, the system linearity is quite good within the dynamic range, but the system effective dynamic range is merely **3 to 40p.e.**

The second part of the chapter devotes to the study of the SiPM dark count. As the most important part of the system performance, the SiPM dark count is the dominant source of noise. Comparing to MCP, CEM or PMT photon sensor, SiPM has the highest background dark count rate by a few orders. This study builds a theoretical probability model to describe the distribution of DC event occurrence. It is proved the DC event occurrence follows a Poisson distribution and the delay of the consecutive event follows an exponential distribution. It is also proven that more than **99.997%** of the total dark counts are 1p.e. signal. The likelihood of a 2p.e. signal dark count is merely 0.00286%. Therefore, the observed 2p.e. or higher signals in the dark count data set are the contribution of the optic crosstalk.

The dark count rate of our system is measured at **75kHz**, which is close to the manufacturer declared 70kHz. The rate change with bias voltage is characterized, but no characterization with temperature due to our simple test setup lack of temperature control. The underlying cause of the temperature impact on the dark count is briefly discussed with the Fermi-Dirac distribution and semiconductor band theory.

For the future work, more theoretical study about the dark count statistics derived from the semiconductor physics can be done. It is possible to establish a theoretical model to predict the dark count rate given the semiconductor parameters, such as doping level, material parameter, PN junction construction, together with bias voltage, temperature as the input parameters to such a model.

Chapter 7 Radiation Test

7.1 Introduction

After the SiPM system characterization test with LED source, it was ready for the final end to end test with the scintillator module integrated in front of a radioactive source. This is the last test for this research project, but only the first test to characterize the full scintillator+SiPM detector system. More tests with different radioactive sources or electron, proton, ion beam facility will be the future work.

The selected source is an unsealed Am-241, which emits alpha particle and a weaker gamma ray. Frank Asaro et. al.[162] has measured the alpha emission. Paul P.Day[163] has measured the electromagnetic spectrum. Their study unveils the complex alpha and electromagnetic spectrum from Am-241. For the purpose of verifying our prototype detector system, only the two prominent emissions as the 5.486MeV alpha and the 60keV gamma are studied.

The review of the scintillator module design shows that the prominent alpha will be stopped by the top layer plastic scintillator BC404 and the prominent gamma will mostly penetrate the top layer scintillator and be absorbed by the bottom layer LXSO scintillator fully.

The photon yielded by both scintillators will be collected by a single SiPM photon detector at the bottom of LXSO scintillator. This creates a problem of how to identify the origin of the photons, which is essential to calculate the energy resolution of each scintillator. This problem can be partially solved by analysing the timing of the photon. The discussion of this analysis will be one of the key parts of this chapter.

The radiation test is deemed successful in terms of demonstrating the end-to-end functionality. Yet, it also shows the improvement that is needed for this system regarding energy resolution and overall detection efficiency.

7.2 Radiation Source

7.2.1 Am-241 basic

According to Wikipedia [164] , Am-241 is a radioactive isotope of americium. It has a half-life of 432.2 years and it decays mainly via α-decay, with a weak γ-ray by-product. The decay process is shown as:

$$^{241}_{95}Am \rightarrow {}^{237}_{93}Np + {}^{4}_{2}\alpha^{2+} + \gamma 59.5409keV \tag{6.1}$$

The α-decay energies are 5.486MeV for 85% of the time, 5.443MeV for 13% of the time, and 5.388MeV for the remaining 2%. The γ-ray energy is mostly 59.5409keV. Eq. (6.1) hints a 1:1 ratio between α-decay and γ-ray for every decay. This is not true. Only about 40% of the decay will yield γ-ray.

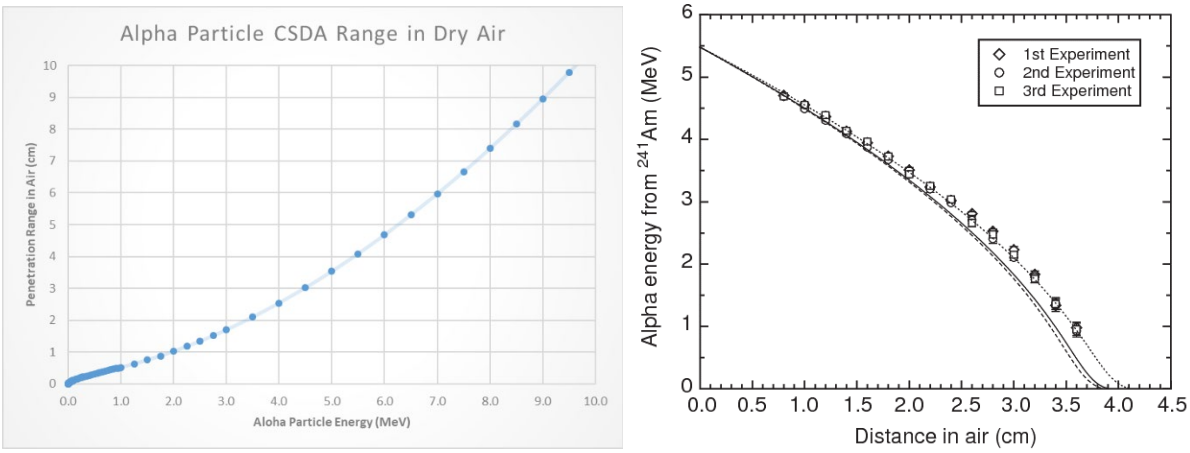
There are several Am-241 sources available at MSSL (see Appendix F: MSSL radioactive source list.) The AM7 is chosen for its easy handling and weak intensity. Figure 7-1 shows the AM7 source encapsulated in a metal coin of 25mm diameter by 0.5mm thickness. In the centre is a small exposure area of 7mm diameter. The source is manufactured in Jan 2017 with an initial activity at 0.04MBq. These figures will be used to estimate the detector efficiency in a later section.

Figure 7-1 MSSL AM7 source

7.2.2 Am-241 Range

The test is performed in air in a dark box. It will be better if the test can be done in a vacuum chamber as that can remove the attenuation by air of the source emission. However, the setup and the associated harness preparation work is much more elaborate. To save effort and time, the first test was made being in air.

As alpha particle travels through the air, it continuously loses energy due to nuclear collision and electric force. This is generally characterized as stopping power and range. The range of alpha particle in air can be calculated using NIST tool ASTAR [81]. Figure 7-2(a) is the plot of the CSDA (Continuous Slowing Down Approximation) [165] range calculated with NIST data. Experiments by Yu,K.N. et al. [166] measured Am-241 alpha particle energy loss in Air. Figure 7-2 (b) quotes their measurement result.



(a) (b)

Figure 7-2 Alpha particle energy loss in the air. (a) is the range calculated by the NIST tool (b) is the measurement result done by Yu, K. N [166]

Reading from the plot, the alpha particle energy will diminish to zero at about 4cm. This means the primary alpha particle of Am-241 source can travel by maximum 4cm in the dry air. Therefore, in the experiment, the distance between the source and the scintillator detector shall be kept within 4cm. This will be further discussed in 7.5.

7.2.3 Am-241 spectrum

Back to 1950s, the emission of Am-241 has been thoroughly studied. Frank Asaro et al [162] has measured the alpha emission with a magnetic spectrometer. Paul P.Day [163] has measured the electromagnetic spectrum with a ten-inch bent-crystal spectrometer and a sodium iodide scintillation counter.

Frank et al reported the findings of six alpha particle groups that are 5.546MeV, 0.23%; 5.535, 0.34%; 5.503,0.21%; 5.476, 84.2%; 5.433, 13.6% and 5.379,1.42% in energy and abundance. Later, they corrected the 5.546 MeV peak was not true but ascribed to the scattering from one of the baffles in the spectrometer. Figure 7-3 shows the alpha spectrum of Am241 from their report.

The experiment system photographed the track of each alpha particle to calculate the energy and abundance. Figure 7-4 is one of the photographs showing six acceptable alpha-tracks (those six in parallel). It is requoted here to give evidence that the track of an alpha particle in the matter is a straight line and the track length varies linearly with the energy.

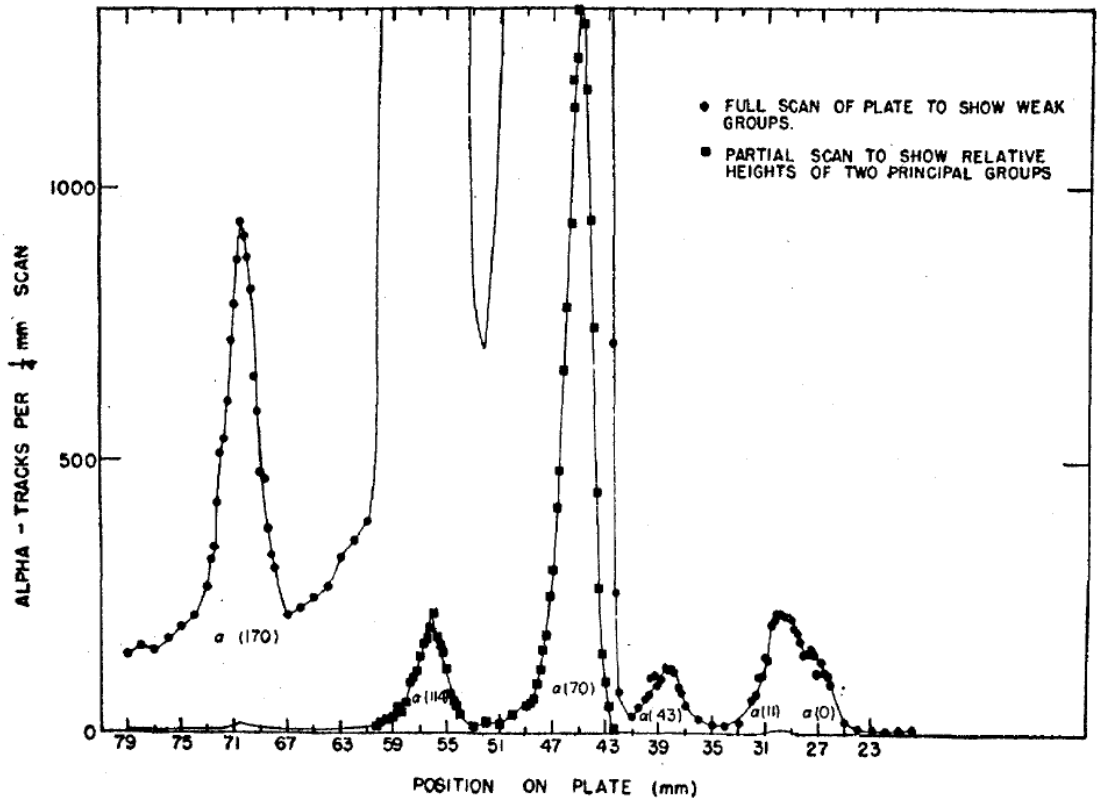


Figure 7-3 Alpha-spectrum of Am241 measured by Frank Asaro et al [162]*. The x-axis is the hit position on the measurement plate. Here 1mm represents 3.96keV difference.*

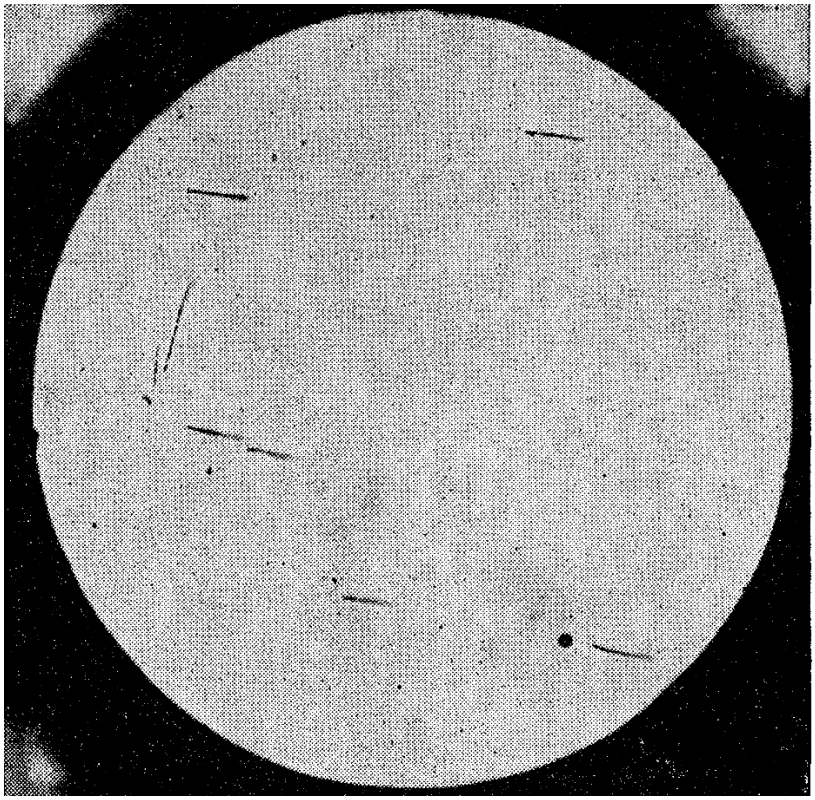


Figure 7-4 Alpha-tracks in photographic emulsion. For a 6MeV alpha particle, the track length in the emulsion is approximately 25 microns.

Paul P.Day carried on studying the electromagnetic spectrum of Am-241. The experiment used a Topaz to bend the electromagnetic wave and used a NaI scintillator as a detector. This way, the source and detector are not on a straight line, which can avoid the contamination of alpha particle from the same source. The chart of the complete spectrum is requoted in Figure 7-5. The detail of four prominent gamma rays are requoted in Table 7-1. Paul also mentioned Seaborg et al [167]'s work observed that the 60keV γ-ray occurred at a frequency of about 0.4 per alpha particle. This ratio will be used in the later analysis.

Table 7-1 The Am^{241} gamma rays observed on the bent-crystal spectrometer

Energy (kev)	Background[a] (counts/min)	Observed intensity[b] (counts/min)	Correction[c] factor	Corrected[d] intensity
26.363±0.014	40	520	1.55	8.2
33.199±0.021	39	39	1.32	0.5
43.463±0.085	65	30	2.17	0.6
59.568±0.017	70	5220	1.88	100

[a] Background computed by averaging counting rates on both sides of each peak.
[b] Background subtracted.
[c] Attenuation corrections as explained in text.
[d] Relative to the 60-kev line.

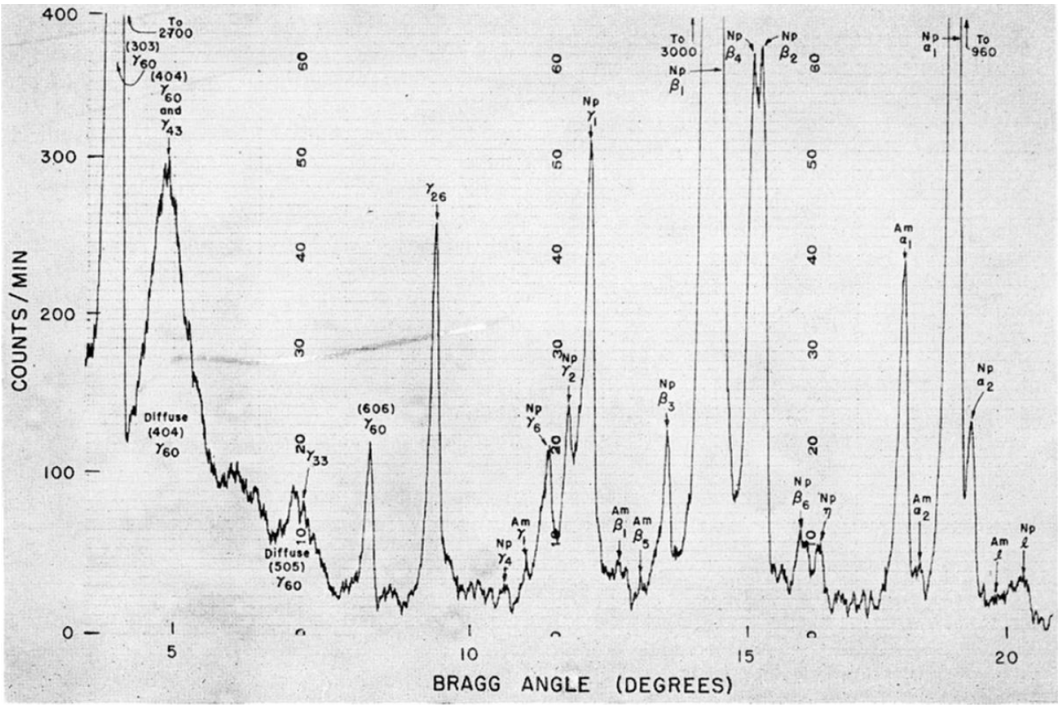


Figure 7-5 The complete electromagnetic spectrum of Am-241 source from Paul P.Day [163]

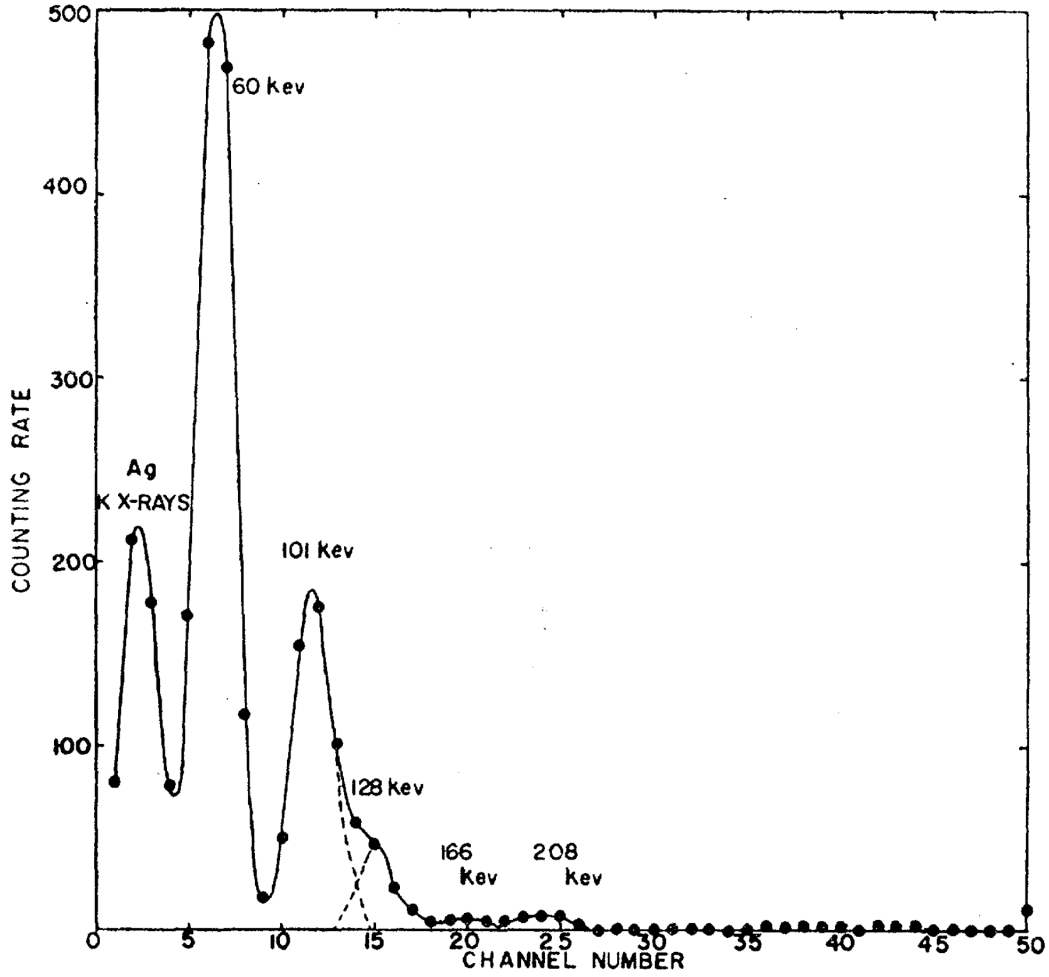


Figure 7-6 Electromagnetic spectrum of the Am-241 source measured by a NaI(Tl) scintillator and a 50 channel PHA analyser from Jaffe et al. [168]

Jaffe et al [168] did a similar measurement with the same type of bent crystal spectrometer. He used NaI(Tl) scintillator and a 50 channel pulse height analyser as the main detector and readout system. Figure 7-6 shows their spectrum result. Clearly the 60keV line is the prominent emission.

7.3 Test Setup

The separation distance between the Am-241 source and the scintillator is a key distance. In the test, this distance is varied to simulate different alpha particle energy hitting on the scintillator. The idea is that the air attenuation will absorb a fraction of the original alpha energy and the left-over energy will be the energy absorbed by the scintillator. As the attenuation is related to the air distance according to Figure 7-2, varying the distance can vary this attenuation, thus adjust the left-over energy.

The system data acquisition and the post-processing Python scripts are kept the same as that used in the SiPM system characterization. The difference is just the scintillator module is fit and the LED source is replaced with Am-241 source. Figure 7-7 shows the setup.

The front-end detector system (including scintillator module, SiPM and preamp; for simplicity, this is called detector assembly) and the radioactive source is placed inside an aluminium box. The inside surface of the box is coated with black anti-reflection material. The low voltage (LV), high voltage (HV) PSU, DRS4 module and the laptop are placed outside the box.

The source is held in a bracket secured with some yellow tape. The detector is held by two adjustable support frames. The source and detector are aligned manually. This is very coarse, and it leads to some problems in later data analysis as the exact distance between the source and the detector is unknown. Figure 7-8 shows the setup inside the dark box. All the data is captured and saved during the test. The data processing is done offline.

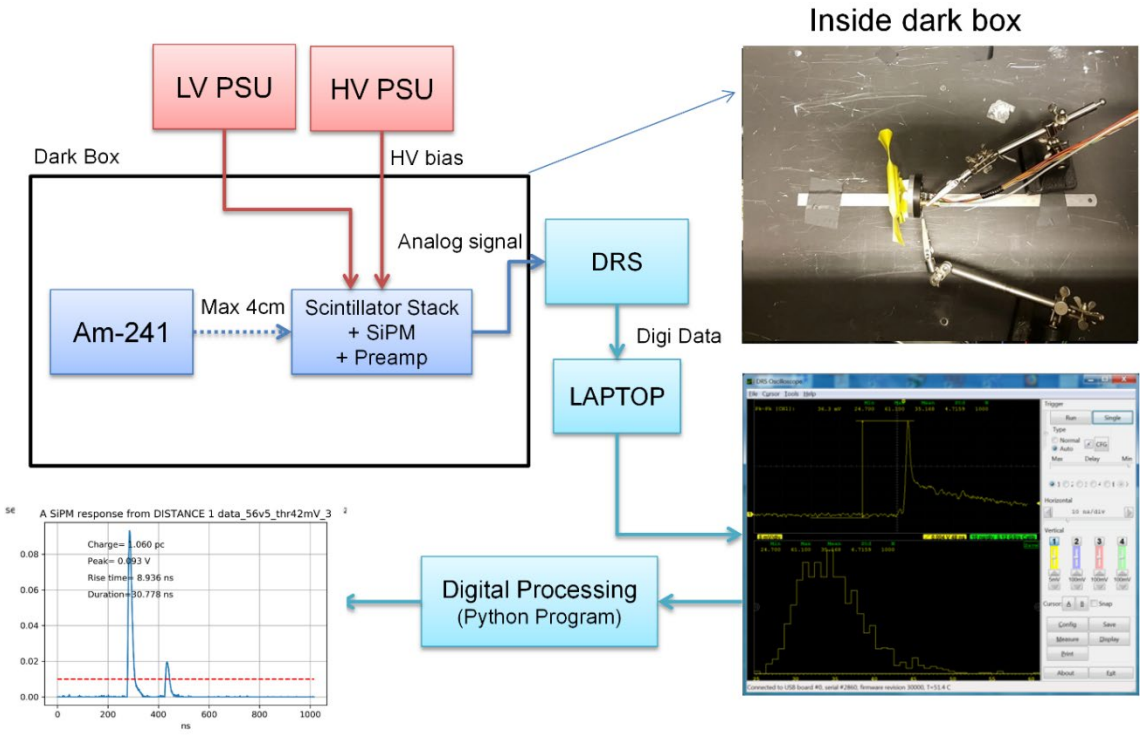


Figure 7-7 The architecture of the test setup with radioactive source Am-241

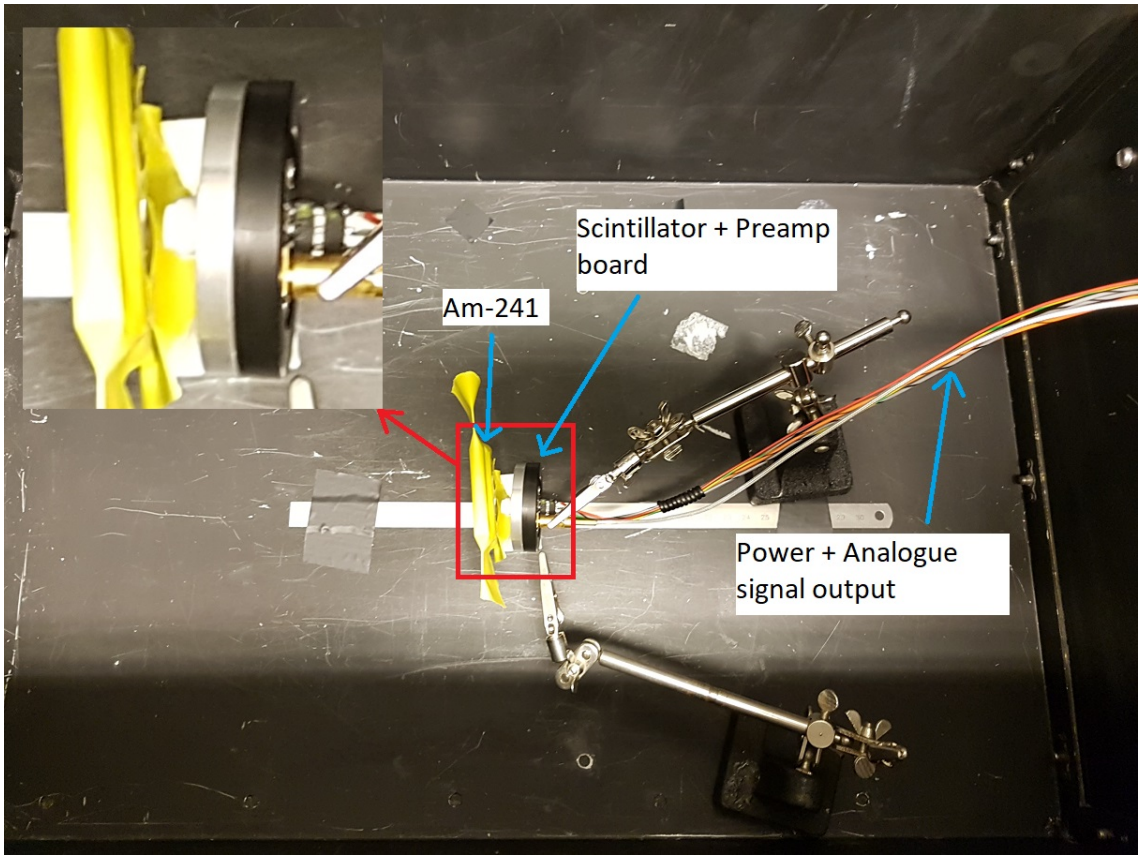


Figure 7-8 The photograph of the radiation test setup inside a dark box. The Am-241 source is held in a bracket with the active area facing the scintillator detector module top face

Figure 7-9 shows the scintillator detector module and how it is mounted in SP5650 sensor holder. The scintillator module is already coated with a thin layer of Aluminium

for containing the scintillator light within the module and rejecting the ambient light. The white Teflon wrap is mainly used to hold the scintillator module rather than light shielding.

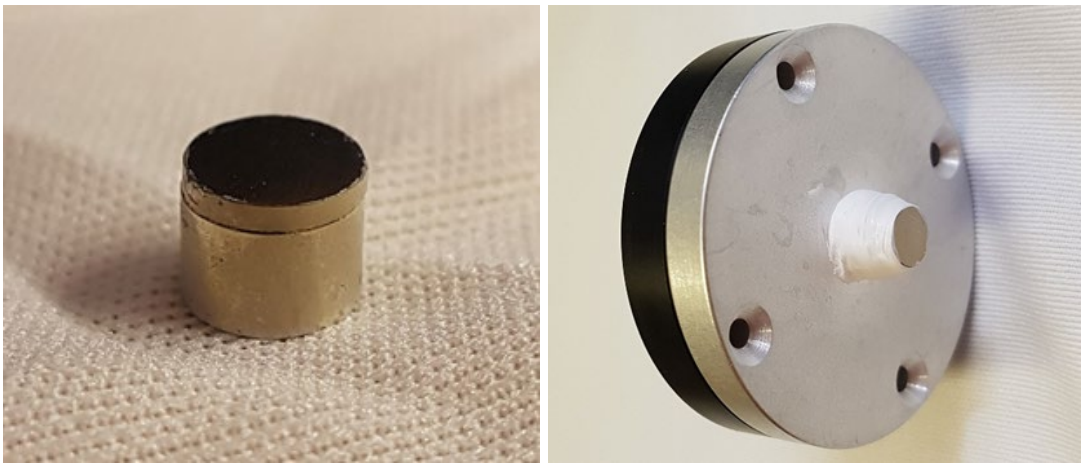


Figure 7-9 On the left is the scintillator detector module and on the right is the assembly of scintillator module with the preamplifier board holder

7.4 Scintillator module response review

Firstly, briefly revisit the scintillator design to understand what the expected response from the detector assembly shall be like for Am-241 source.

The scintillator module is constituted of two layers of scintillator stacked together. The top layer is BC404 of 1mm thickness and the second layer is LXSR of 4mm thickness. The module is coated with Aluminium reflection layer, with top coating at 50nm thickness and body coating at 100nm. The illustration of the design can be found in Figure 2-10.

Recall the discussion in 2.3.3, Table 2-6, the minimum energy of alpha particle to penetrate the BC404 scintillator is 40MeV. Therefore, the alpha decay of Am-241 at energy of ~5.5MeV will be fully stopped in BC404. According to NIST data, the absorption of 60kV γ-ray in BC404 is 1.91% and in LXSR (use CsI data to approximate) is 100%. Therefore, gamma decay of Am-241 at energy of 60keV will mostly go through BC404 and fully absorbed by the LXSR scintillator.

Before the Am-241 decays reaches the scintillator, two energy/flux losses shall be evaluated. They are the loss through air and aluminium coating. For alpha particle, the range in air is shown in 7.2.2 and the energy required to penetrate the 50nm aluminium coating is only 1.5keV as calculated from the NIST data. For γ-ray, the ab-

sorption of 4cm air is merely 0.1% at 60keV and absorption of 50nm aluminium coating is 0% according to the NIST data. Therefore, the loss through the aluminium coating can be ignored for both alpha decay and gamma decay. The loss through the air can be ignored for gamma decay, but not so for the alpha decay.

Based on the above brief calculation, the prediction is that:

a) the Am-241 alpha decay will lose some of its energy through the air (the loss is dependent on the air distance) and a neglectable small amount through the aluminium coating and will deposit the remaining energy fully in the BC404 scintillator.

b) the Am-241 gamma decay will nearly 100% penetrate through the air and the aluminium coating and lose about 1.91% flux through 1mm BC404 scintillator. The remaining flux will be fully absorbed by the LXSR scintillator.

7.5 Detector system end to end efficiency calculation

The detector system end to end efficiency calculation is made up of the following steps.

The first step is to calculate the source activity at the time of testing and the total branching ratio which represents the portion of the selected decay mode and photon/particle energy over the entire decay activity.

The second step is to calculate the effective radiation flux received by the scintillator. The emission of the source is isotropic, and the detector has a limit entry surface area placed at a distance from the source. So inevitably, most of the radiation flux won’t reach the scintillator. This calculation can be done through the analysis of the geometry made up of the source and the detector. The medium between the source and the scintillator will absorb or attenuate the emitted photon/particle. The loss through this medium shall be calculated, although the result might be so trivial that it can be neglected.

The third step is to analyse the scintillation photon loss. This loss is mainly due to a) absorption in crystal, b) reflection losses and c) interface size mismatch between scintillator exit surface and SiPM entry surface. The remaining scintillation photons will reach the SiPM sensor.

The final step is the SiPM detection efficiency, which is how much incident photon can successfully be detected. This is generally quoted as PDE (photon detection efficiency) in the manufacture datasheet. The SiPM output pulse will be amplified and digitized by the readout electronics. The electronics efficiency is assumed 100%. The pulse height will be analysed to understand how many photons are detected.

Cascade the results from the above steps, the final end to end system efficiency can be calculated. However, there are two concepts of system efficiency. One is in terms of flux as the ratio of detected flux to the source emitted flux. The other is in terms of energy as the ratio of detected energy (by counting the number of scintillator photons) to the emitted energy. Discussed in 6.2.3, this demo system has a very limit detection rate. Therefore, if the detectable incident flux is higher than the system maximum detection rate, the flux detection efficiency can’t be resolved. The demo system also has a limited dynamic range, max 40p.e. Similarly, if the output pulse height exceeds this range, the energy detection efficiency can’t be resolved.

Figure 7-10 illustrates the calculation flow. This flow mixes the flux and energy together, which will be separated in the detail calculation.

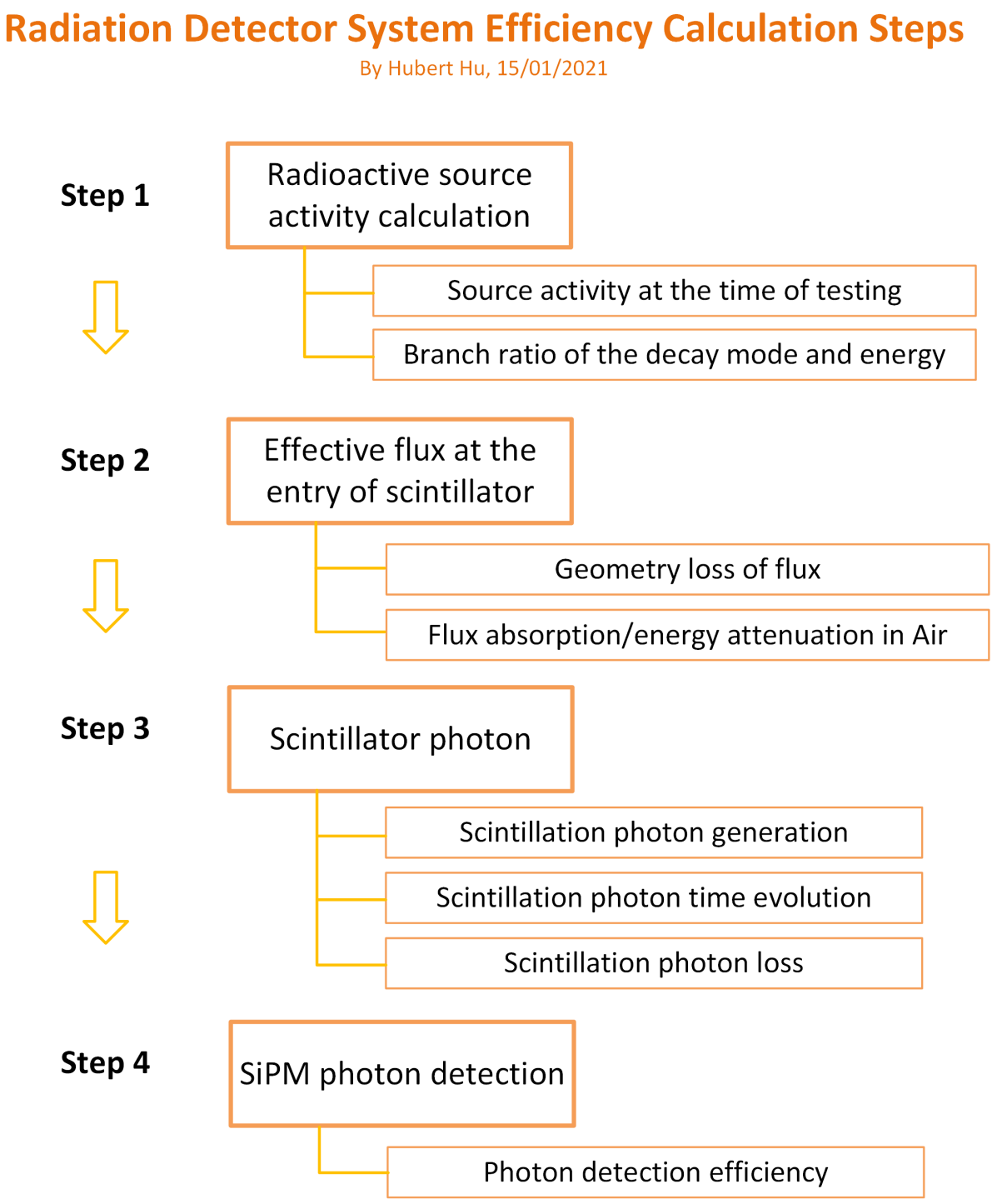


Figure 7-10 The flow of calculating the detector system end to end efficiency

7.5.1 Radioactive source activity calculation

This calculation is made up of two parts. The first part is to calculate the source strength at the time of testing if the calibration data is not current. The formula is:

$$A = A_0 \, \mathrm{e}^{-(t/\tau)} \tag{6.2}$$

Where:

A_0 is the activity of the last calibration or manufacturer data

t is the time elapsed since the last calibration or manufacturing date

τ is the mean-life (1/e life) of the source before decay. It can be related to the half-life as $\tau = T_{1/2} / \ln(2)$.

The MSSL AM7 source has an initial activity at 0.04MBq at Jan 2017. And AM-241 has a half-life of 432.2 year. The time of the test occurred at Nov 2019. So, the elapsed time is 2.8year. Apply Eq. (6.2), the source activity at the time of testing can be worked out at 0.039819Mbq.

The second part is to consider the branching fraction **BF** for that mode of decay and the branching ratio **BR** for that photon/particle energy. The product of BF and BR is sometimes defined as total branching ratio **TB**. The formula is:

$$N = BF \times BR \times A \tag{6.3}$$

Where:
BF is the branching fraction for that mode of decay
BR is the branching ratio for that photon energy
N is the flux rate
A is the source activity

Use the numbers surveyed in section 7.2.3, we can calculate the BF, BR and N for the Am-241 source. Result is listed in Table 7-2.

Table 7-2 The calculation of branch factor and branch ratio for the major Am-241 branches

Branch	Energy	Symbol	BF	BR	A	N	N
					Mbq	Mbq	Decay/s
Alpha	5.535MeV	N_{alpha_1}	1	0.34%	0.039819	0.000135385	135
Alpha	5.503MeV	N_{alpha_2}	1	0.21%	0.039819	0.000084	84
Alpha	5.476MeV	N_{alpha_3}	1	84.2%	0.039819	0.033527598	33528
Alpha	5.433MeV	N_{alpha_4}	1	13.6%	0.039819	0.005415384	5415
Alpha	5.379MeV	N_{alpha_5}	1	1.42%	0.039819	0.00056543	565
Gamma	26.363keV	N_{gamma_1}	0.4	7.50%	0.039819	0.0011949	1195
Gamma	33.199keV	N_{gamma_2}	0.4	0.46%	0.039819	0.000073	73
Gamma	43.463keV	N_{gamma_3}	0.4	0.55%	0.039819	0.000087	87
Gamma	59.568keV	N_{gamma_4}	0.4	91.49%	0.039819	0.0145724	14572

The nuclei decay is a random process. The above figures are all statistics over a long integration time. For a very short period, it is not certain that whether or exact how many times nuclei decay will occur. The statistic feature for nuclei decay over a given interval time obeys Poisson distribution.

7.5.2 Reception of the radiation flux at the scintillator entry face

7.5.2.1 Flux loss due to geometry

The source has a round shape of active area. As the source is coated onto some metal and thus forms a thin layer of active surface, we can regard it as a plane source with an isotropic emission. The scintillator is placed in a small distance to the plane source. A virtual 2π sphere can be drawn with the centre of the source as the origin and the distance between the source and the detector as the radius. This is illustrated in Figure 7-11.

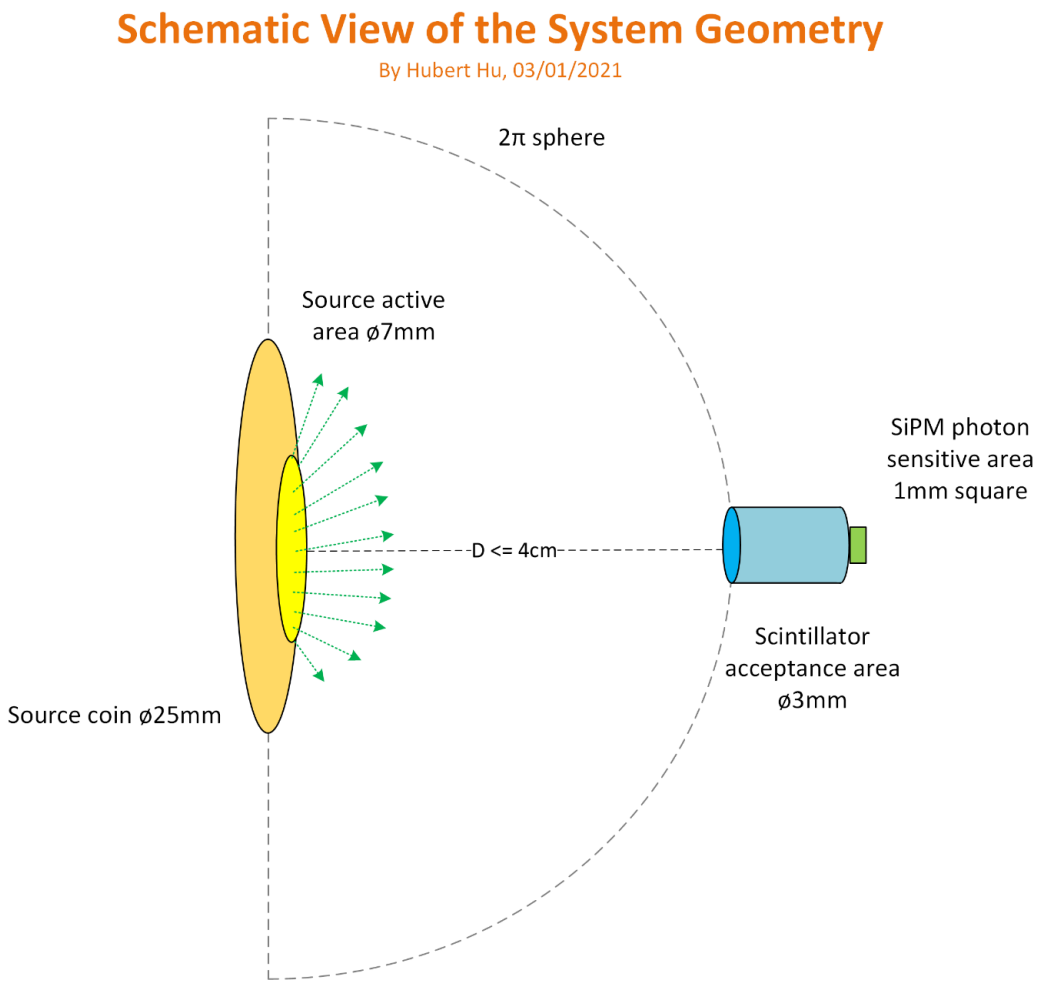


Figure 7-11 the illustration of the radioactive source and the detector system geometry loss

We can assume an even distribution of radiation flux from the source plane in the front 2π semi-sphere. The back half is blocked by the coin body. However, the source emission is isotropic. This means even the back half is blocked, we still need to count the 4π sphere when calculating the geometry loss. The distance ***D*** compared to the source active area radius ***R*** is not significant enough, so the source can't be regarded as a single point. The mitigation is to take a very small area ***dS*** in the source plane and treat it as a point source and work out its projection to the detector surface. Then integrate ***dS*** over the entire source plane to acquire the total projection.

Figure 7-12 illustrates the method. The small area ***dS*** in the source plane (marked in red) has a radial length of ***dr*** and an angular length of ***rdθ***. Its area then can be expressed as $dS=rd\theta dr$. If the source plane total activity is defined as N_0 and assume an even distribution of source activity over the entire active area, then the activity N_1 for this small area ***dS*** can be defined as the ratio of ***dS*** to the full plane size ***S***:

$$N_1 = N_0 \frac{rd\theta dr}{\pi R^2} \tag{6.4}$$

This area is small enough compared to the distance ***D***, so it is considered as a single point. The new distance ***D'*** between this point and the scintillator source can be worked out as ***D' = sqrt(r²+D²)*** from the simple right-angle triangle formed by line ***r***, ***D*** and ***D'***. Using the new origin ***O'*** and new distance ***D'*** to form the new virtual sphere, we can calculate the amount of arrival emission ***N₂*** coming from the sub-region ***dS*** of the source as:

$$N_2 = N_1 \frac{\pi r_s^2}{4\pi D'^2} \tag{6.5}$$

Here, the impact of the little tilting between the normal line of the scintillator surface and the new central line ***D'*** is neglected.

Simply replace ***N₁*** with (6.4) and then double integrate the ***N₂*** over full area of the source plain to get the total projection of emission ***N_total*** from the plain source to the scintillator surface:

$$\begin{aligned} N_{total} &= \int_0^{2\pi}\int_0^{R} N_0 \frac{\pi r_s^2}{4\pi D'^2}\frac{r}{\pi R^2} d\theta dr \\ &= \int_0^{2\pi}\int_0^{R} N_0 \frac{r_s^2}{4(\sqrt{r^2+D^2})^2}\frac{r}{\pi R^2} d\theta dr \\ &= N_0 \frac{r_s^2}{4\pi R^2}\int_0^{2\pi}\int_0^{R} \frac{r}{(r^2+D^2)} d\theta dr \\ &= N_0 \frac{r_s^2}{4\pi R^2} 2\pi \int_0^{R} \frac{r}{(r^2+D^2)} dr \end{aligned} \tag{6.6}$$

Using a variable swap technique, the solution to the integral is

$$
\begin{aligned}
N_{total} &= N_0 \frac{r_s^2}{4\pi R^2} 2\pi \int_0^R \frac{r}{(r^2 + D^2)} dr \\
&= N_0 \frac{r_s^2}{4\pi R^2} 2\pi \int_0^R \frac{1}{2(r^2 + D^2)} dr^2 \\
&= N_0 \frac{r_s^2}{4\pi R^2} 2\pi \int_0^{R^2} \frac{1}{2(x + D^2)} dx \\
&= N_0 \frac{r_s^2}{4R^2} \ln\left(x + D^2\right)\Big|_0^{R^2} \\
&= N_0 \frac{r_s^2}{4R^2} \ln \frac{R^2 + D^2}{D^2}
\end{aligned}
\tag{6.7}
$$

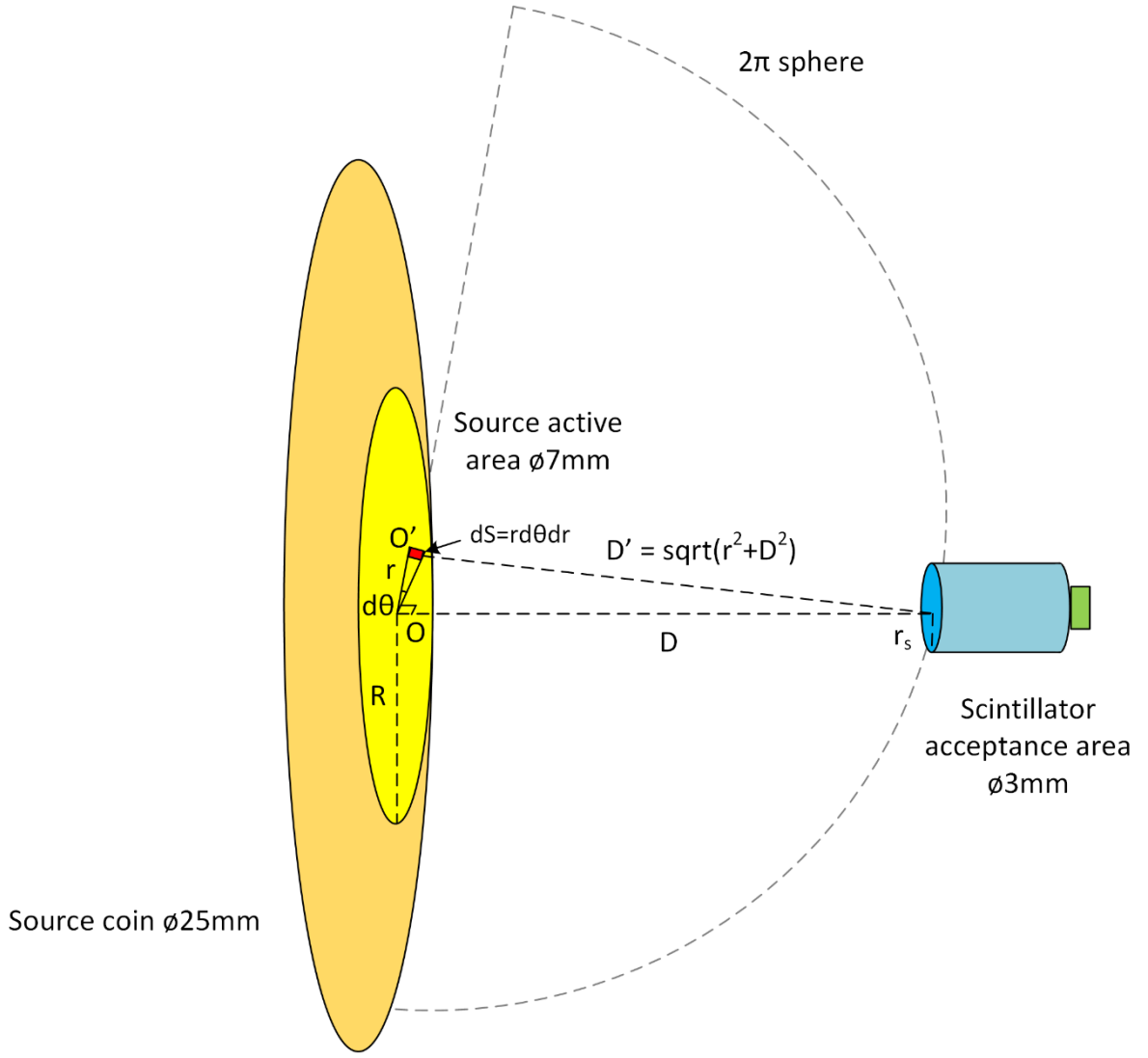


Figure 7-12 the illustration of the geometry calculation for the source emission arrived at scintillator entry surface.

Table 7-3 presents the calculation results with the parameters of our test system setup using equation (6.7).

Table 7-3 Calculation for the emission flux rate on the scintillator entry face

Parameter	Value	Unit	Note
Am-241 activity($\mathbf{N_0}$)	0.039819	Mbq	Activity calculated at the time of test
$\mathbf{r_s}$	1.5	mm	
R	3.5	mm	
D	10	mm	Test used various distances. Here, just take one typical value.
$\mathbf{N_{total}}$	0.000211	Mbq	Flux rate at the entry of scintillator
	211	s^{-1}	Equivalent to decay per second

For different distance, the receptive flux at the scintillator entry face is shown in Figure 7-13.

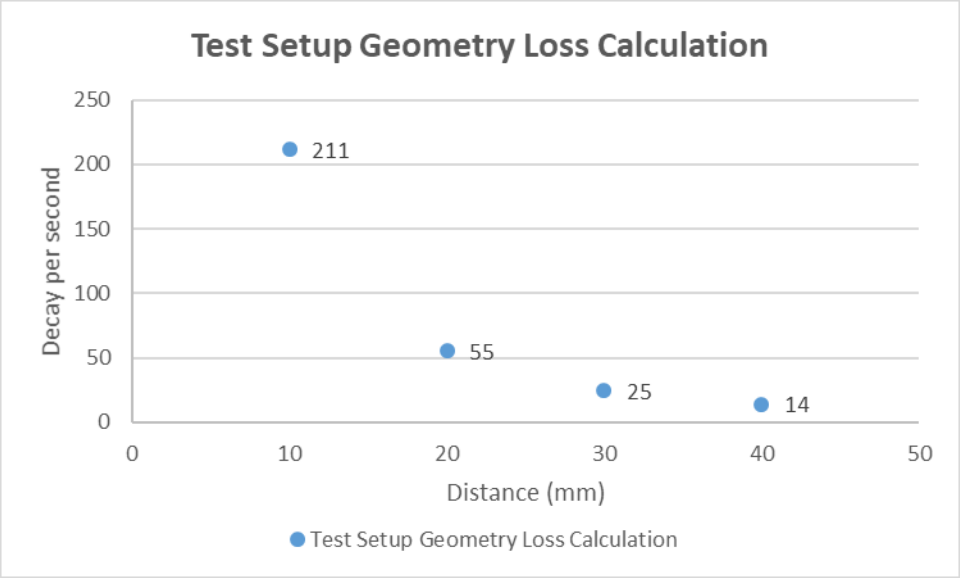


Figure 7-13 The flux level at the scintillator entrance in relation to the distance between source and detector

This flux rate is within our demo system's capability as discussed in 6.2.3.2. This means in average all the decay events from the source AM7 shall be detected. However, as the decay obeys Poisson distribution, there is a chance that some decays occurring too close to the previous one will be missed.

7.5.2.2 Flux absorption or energy loss over air

As discussed in 7.4, the flux absorption or energy loss over the aluminium coating for the AM-241 decays can be ignored. So, the discussion here will only focus on the absorption/loss over the air.

The full energy loss and flux reduction of the alpha decay through the air for the prominent alpha decay is fully calculated using Eq. (6.2), (6.3) and (6.7) with process and results shown in Table 7-4

Table 7-4 The flux and energy of the prominent alpha decay over various distance in air

Parameter	Distance				Unit
	10mm	**20mm**	**30mm**	**40mm**	
Initial activity (A_0)	0.04	0.04	0.04	0.04	Mbq
Activity @ testing (A)	0.03982	0.03982	0.03982	0.03982	Mbq
Branch factor (BF)	1	1	1	1	
Branch ratio (BR)	84.20%	84.20%	84.20%	84.20%	
Flux rate (N)	0.03353	0.03353	0.03353	0.03353	Mbq
Flux geometry loss	0.00531	0.00139	0.00062	0.00035	
Flux @ scintillator	1.78E-04	4.64E-05	2.08E-05	1.17E-05	Mbq
	178	**46**	**21**	**12**	Decay/s
Initial energy (E_0)	5.476	5.476	5.476	5.476	MeV
Energy loss through air	1	2	3.5	5.5	MeV
Energy @ scintillator	**4.476**	**3.476**	**1.976**	**0**	MeV

Using a similar method with slight modification, the prominent gamma decay is processed with results shown in Table 7-5.

Table 7-5 The flux and energy of the prominent gamma decay over various distance in air

Parameter	Distance				Unit
	10mm	**20mm**	**30mm**	**40mm**	
Initial activity (A_0)	0.04	0.04	0.04	0.04	Mbq
Activity @ testing (A)	0.03982	0.03982	0.03982	0.03982	Mbq
Branch factor (BF)	0.4	0.4	0.4	0.4	
Branch ratio (BR)	91.49%	91.49%	91.49%	91.49%	

Parameter	Distance				Unit
Flux rate (N)	0.01457	0.01457	0.01457	0.01457	Mbq
Flux geometry loss	0.00531	0.00139	0.00062	0.00035	
Flux after geometry loss	7.7E-05	2E-05	9.1E-06	5.1E-06	Mbq
Flux absorbed by air	0.03%	0.05%	0.08%	0.10%	
Final Flux @ BC404	7.7E-05	2.0E-05	9.0E-06	5.1E-06	Mbq
	77	**20**	**9**	**5**	Decay/s
Flux absorbed by BC404	1.91%	1.91%	1.91%	1.91%	Mbq
Final Flux @ LXSR	7.6E-05	2.0E-05	8.9E-06	5.0E-06	Mbq
	76	**20**	**9**	**5**	Decay/s
Initial energy (E_0)	59.568	59.568	59.568	59.568	keV
Energy loss through air	0	0	0	0	keV
Energy @ scintillator	**59.568**	**59.568**	**59.568**	**59.568**	keV

7.5.3 Scintillation photon

7.5.3.1 Scintillation photon generation

The number of photons produced in the scintillator crystal depends on the overall conversion efficiency of the energy of incident particle to secondary photon emission. The physical process is described in 2.2.3.1.

Most of the scintillation yield parameters are material dependent. Manufacturer often provides the total photon yield directly. Just note this total photon yield is the integration of all the photons over the full emission spectrum. As part of the spectrum is outside of the SiPM acceptance spectrum, photons under that area shall be excluded. However, to keep the calculation simple, this small error is tolerated.

For the scintillator used in our prototype, the photon yield is defined in Table 7-6.

Table 7-6 The summary of the scintillator photon yield and other parameters

Scintillator	Photon Yield	Decay Time Constant	Refraction Index	Density	Peak Emission Wavelength
	# per MeV	ns		g/cm3	nm
BC404	500	1.8	1.58	1.023	420
LXSR	32000	32	1.8	4.45	408

7.5.3.2 Scintillation photon time evolution

As discussed in 2.2.3.2, the time evolution of the scintillation photons N can be described by the superposition of a set of exponential formulas. To simplify, Eq. (2.12) is reduced to just one decay component:

$$N = N_0 e^{(-t/\tau)} \tag{6.8}$$

Where τ is the decay constant and N_0 is a relative amplitude at t=0.

According to the manufacture datasheet [169] , BC404 has a decay constant of 1.8ns. Figure 7-14 is the Geant4 simulation showing the distribution of photon arrival time tagged when the photon hits the SiPM sensitive area. This matches the exponential form defined in Eq. (6.8). Note there is a small difference here. Equation (6.8) describes only the scintillator photon born process while the simulation also includes the travelling time for the photon from where it is born to the detector sensitive area. However, the travelling time is so small (for a 5mm scintillator length, it only takes 16ps.) that it can be ignored. Figure 7-14 shows BC408 scintillation photon arrival time. It shows most of the photon can be collected within 8 ns. (BC408 is very similar to BC404 in performance, i.e. 2.2ns versus 1.8ns in decay time. They belong to the same family. The plot of BC408 is quoted here instead of BC404 because this Geant4 simulation is done on BC408 before the change of scintillator.)

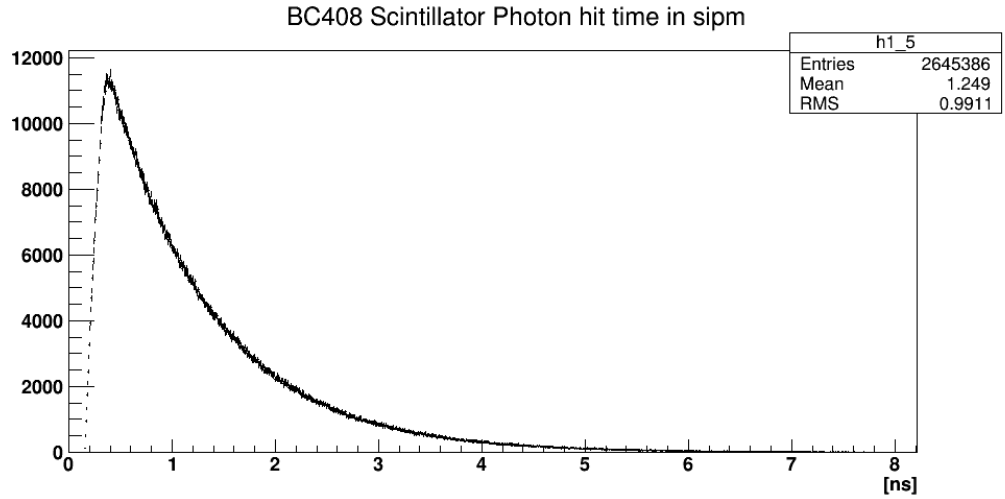


Figure 7-14 The Geant4 simulation of BC404 scintillator photon hit time at the SiPM sensitive area

The rate of photon emission can be worked out by the derivation of Eq. (6.8)

$$\frac{dN_{ph}}{dt} \equiv n_{ph}(t) = -\frac{N_0}{\tau} e^{-t/\tau} \tag{6.9}$$

Then after time T, the total number of photons emitted is:

$$N_{ph}(T) = \int_0^T n_{ph}(t)dt = N_0(1 - e^{-T/\tau}) \tag{6.10}$$

Using the scintillator parameters in Table 7-6 and Eq.(6.10), the photon emission time evolution in BC404 and LXSR are presented in Figure 7-15 and Figure 7-16.

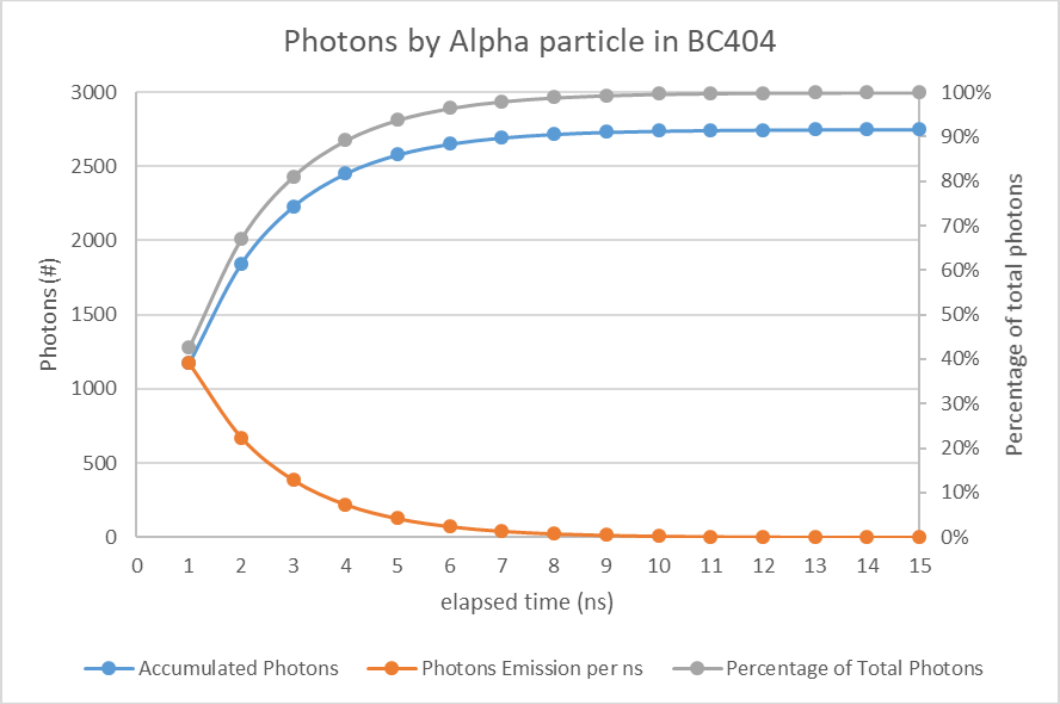


Figure 7-15 Photons yield in BC404 by the 5.5MeV alpha particle. Orange line shows the photon yield rate (#/ns) over time. Blue is the accumulated photons over time. Grey shows the percentage of total photons yielded.

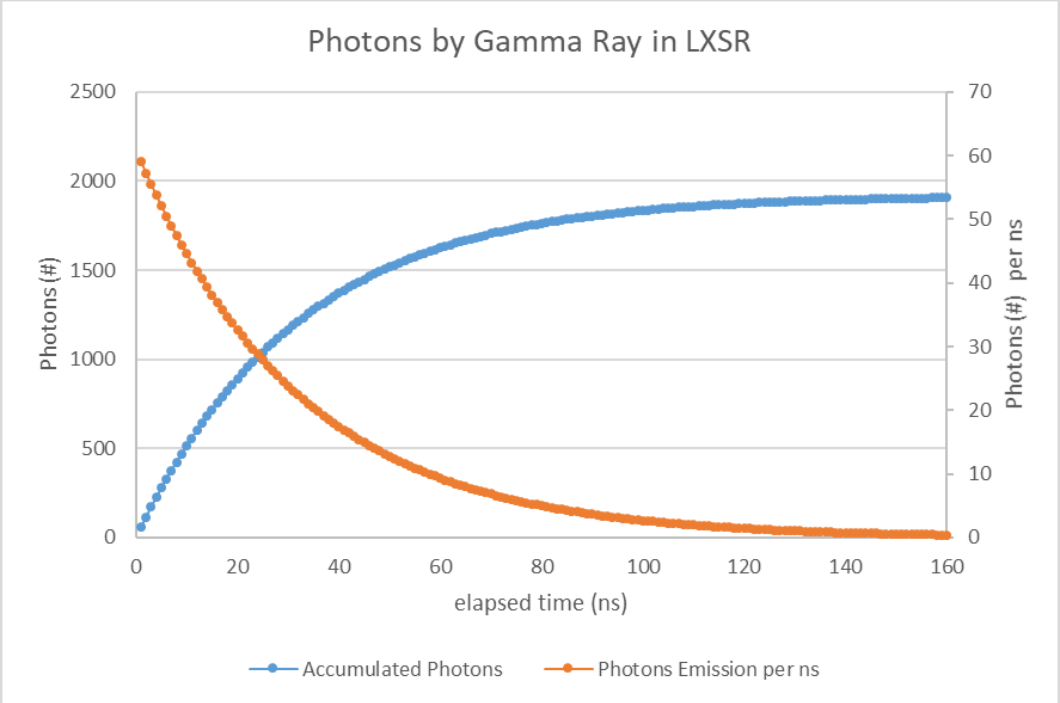


Figure 7-16 Photons yield in LXSR by the 60keV gamma ray. Orange line shows the photon yield rate over time (right axis) . Blue is the accumulated photons over time (left axis)

From the discussion of deadtime in 6.2.3.2, the SiPM has a deadtime of less than 1ns assuming the photons not all land in the same pixel. The preamplifier has a deadtime of about 15ns. Given these timings, the assumption is the SiPM can integrate all the photons within the same 1ns interval into a single output signal. Then over the next 14ns, because the preamplifier is inactive, all the successive photons will be lost. Over the following 1ns interval, the system is re-active and is able to detect new photon. The process repeats until the photon emission is over.

Apply this to the BC404 scintillator with an elapsing time of 8ns for 99% of photon emission, it means only the first 1ns photons will be registered. The remaining 7ns will be lost. For the case of LXSR scintillator with an elapsing time of 147ns for 99% of photon emission, for every 15ns, the first 1ns photon will be registered and the following 14ns will be lost. This process will repeat about 147/15 = 10 times.

This understanding of the photon detection and loss will be applied to interpret the radiation test data in 7.6.3.

7.5.3.3 Scintillation photon emission spectrum

Different scintillator emission has different spectrum. It is important this spectrum is matched by the absorption spectrum of the SiPM sensor. Figure 7-17 shows the emission spectrum of BC404 and LXSR.

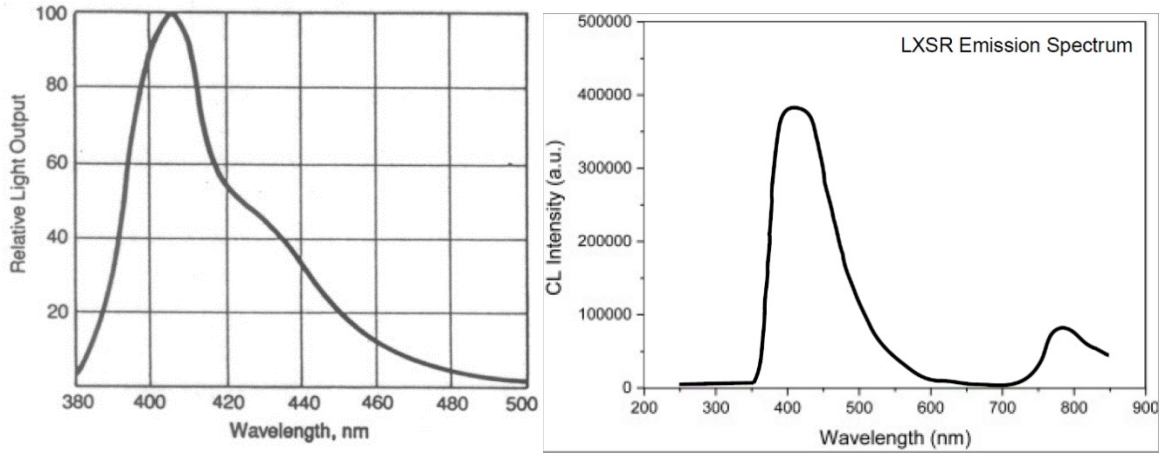


Figure 7-17 left is the emission spectrum of BC404 and right is the emission spectrum of LXSR.

It is shown that BC404 has a peak at 408nm and LXSR has a peak of 420nm. Figure 7-18 shows SiPM s13360-1325CS has a peak at 450nm in its photon detection efficiency spectrum. The detection efficiency drops slightly at the emission peaks of BC404 and LXSR. The proper detection efficiency correction due to the spectrum mismatch shall follow Eq. (2.14) in 2.2.3.3.

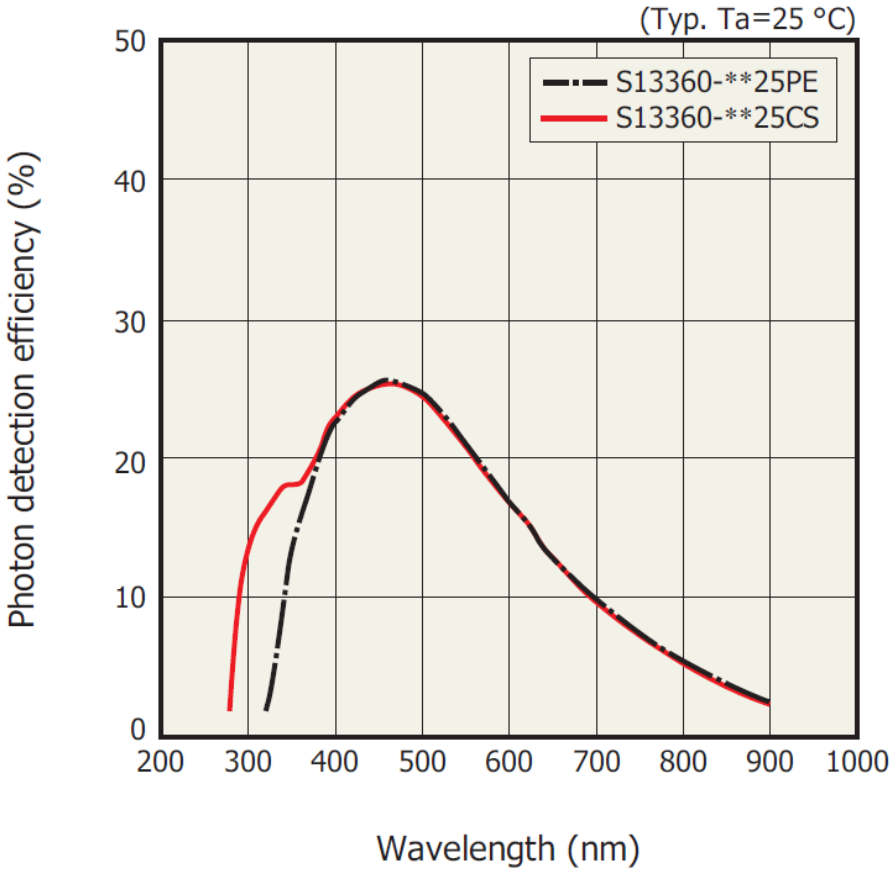


Figure 7-18 SiPM s13360-1325cs photon detection efficiency vs wavelength [170]

7.5.3.4 Photon loss due to reflection loss

Scintillation photon emitted along the track in the crystal has different directions. Only a portion of photons with the right direction towards the scintillator exit face will reach the SiPM detector. Some of the photon will hit the crystal wall. Whether a total reflection will occur or not depends on the incident angle α (the angle between the wall nominal line and the photon incident line) and the refractive index of the crystal $n_{crystal}$. The condition can be described as:

$$\sin\alpha \geq \frac{1}{n_{crystal}} \tag{6.11}$$

If the incident angle meets the condition in (6.11), a total reflection will occur. Otherwise, the photon will leave the crystal. See the simple illustration in Figure 7-19.

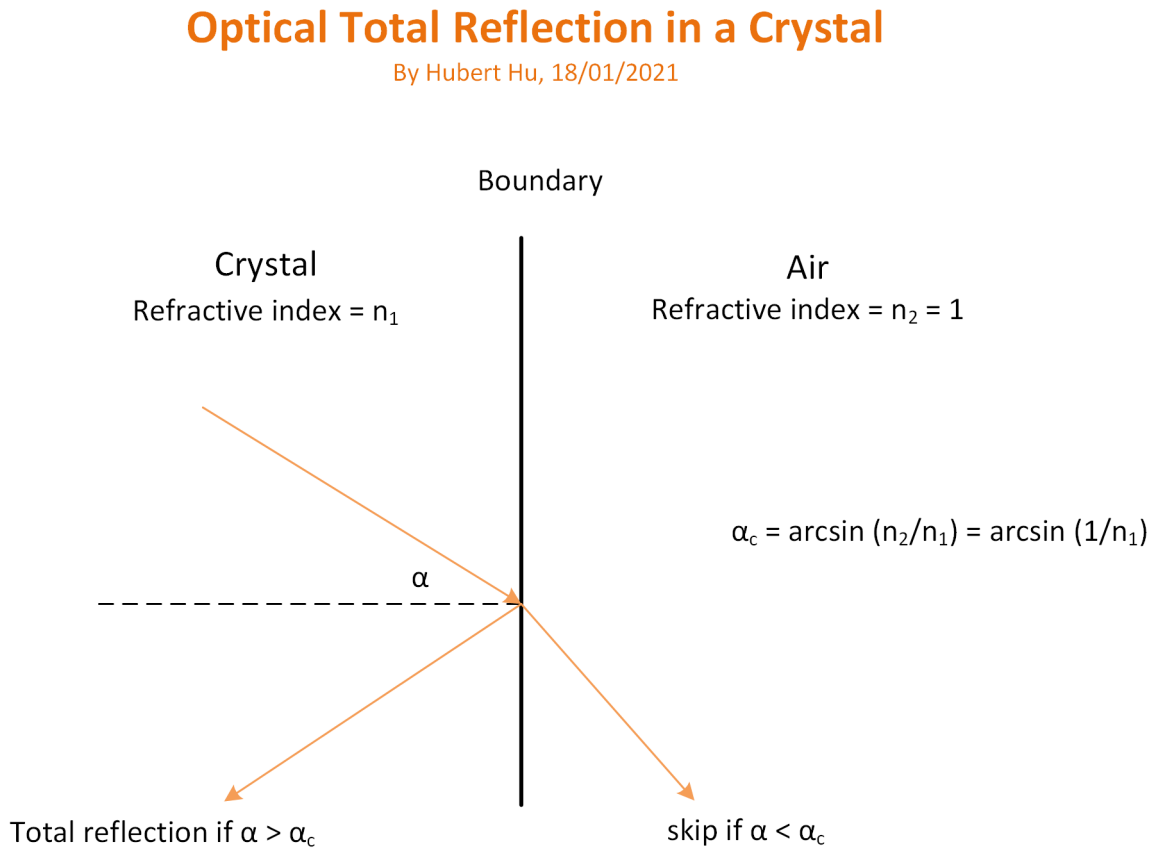


Figure 7-19 the condition for the photon total reflection on the crystal boundary

Some photons can reach the photon detector after a few reflections. Many will be reabsorbed in the crystal. One method to improve the photon collection efficiency is to wrap the crystal with a reflection layer. This layer can reflect some of those skipped photons back into the crystal.

Knowing the principle of total reflection, it is found that a certain geometry is more efficient in creating a total reflection condition than others. In 2.3.4, two geometries are compared: hemisphere and cylinder using Geant4 simulation. The study shows the photon collection ratio is almost 100% for the hemisphere geometry for either scintillator types. While for the cylinder geometry, study shows the ratio is merely 38% for the BC404 scintillator (on the top) and 68% for the LXSR scintillator (on the bottom). The LXSR scintillator is more efficient probably due to its proximity to the SiPM sensor. In comparison, the photons generated in the BC404 scintillator must go through the LXSR scintillator to reach the photon detector.

Our manufactured scintillator module is a cylinder shape. Hence, for the detector system efficiency calculation in 7.5.5, the figures **38%** for the BC404 scintillator and **68%** for the LXSR scintillator will be used as the scintillator geometry efficiency.

7.5.3.5 Interface size mismatch loss

The other photon loss comes from the interface size mismatch between the scintillator exit surface and the SiPM sensitive area. The scintillator exit surface is a ø3mm circle while the SiPM sensitive area is a 1mm square. This mismatch loss can be easily calculated as the area ratio:

$$L_{\text{int}erface} = \frac{S_{SiPM}}{S_{sc\,\text{int}illator}} = \frac{d^2}{\pi r_s^2} \tag{6.12}$$

Use the dimensions of our detector module and the chosen SiPM type, the interface mismatch loss in our case is **14.15%**.

7.5.4 SiPM PDE

For the scintillation photon arrived at the SiPM sensitive area, not all can be successfully detected. This is because the quantum efficiency of the SiPM and the dead area inside the sensitive area. In the manufacturer datasheet, these two effects are combined and reported as photon detection efficiency (PDE). This PDE is wavelength dependant. The PDE against wavelength for Hamamatsu SiPM S13360-1325CS is already shown in Figure 7-18. For the detector system efficiency calculation in 7.5.5, the PDE at the peak sensitivity wavelength as **25%** is used. The error from spectrum mismatch (7.5.3.3) is not counted here for simplicity and its small contribution.

7.5.5 Calculation for the expected system output

With all the preparation from section 7.5.1 to 7.5.4, we can calculate the expected system output in two aspects: flux rate and energy resolution. Table 7-7 and Table 7-8 lists the calculation procedure for the Alpha and Gamma decay from the Am-241 source respectively.

Table 7-7 Detection efficiency calculation for Am-241 Alpha Decay

Radioactive source						
activity at the time of testing	0.039818652	Mbq		Alpha branch energy	5.476	MeV
Alpha branch Ratio	84.20%					
Alpha Branch factor	100.00%					
Alpha branch activity	0.033527305	Mbq				
Flux loss due to geometry				**Energy loss due to geometry**		
Distance (D)	10	mm		Distance (D)	10	mm
Scintillator radius (rs)	1.5	mm				
Source radius (R)	3.5	mm				
Flux at the scintillator	0.000177904	Mbq		Residue energy at the scintillator	4.5	MeV
	178	#/s				
				Scintillation photon		
				BC404 photon yield coefficient	500	# per MeV
				Theoretical total photon emitted (N_0)	2250	#
				Decay time constant	1.8	ns
				SiPM acceptance window	1	ns
				Total photon in the acceptance window (N)	959	#
				Photon loss		
				Reduction due to the reflection	38%	
				Photon arrived at the exit surface	364	
				Reduction due to the interface size mismatch	14.15%	
				Photon arrived at the SiPM sensitive area	52	
				SiPM detector		
				Photon detector efficiency	25%	
				Final detected photon	12.89	
				Photon detection efficiency		
				Photon detection end to end efficiency	0.57%	
				system energy resolution		
				Overall system energy resolution	**0.349**	MeV/photon

Note: the photon detection end to end efficiency is defined as ratio of the SiPM final detected photons to the scintillator theoretical total yield photons. System energy resolution is defined as the ratio of incident energy to the SiPM final detected photons.

Table 7-8 Detection efficiency calculation for Am-241 Gamma decay

Radioactive source					
activity at the time of testing	0.039818652	Mbq	Gamma branch energy	0.06	MeV
Gamma branch Ratio	91.49%				
Gamma branch factor	40.00%				
Gamma branch activity	0.014572034	Mbq			
Flux loss due to geometry			**Energy loss due to geometry**		
Distance (D)	10	mm	Distance (D)	10	mm
Scintillator radius (rs)	1.5	mm			
Source radius (R)	3.5	mm			
Flux at the scintillator	0.0000773	Mbq	Residue energy at the scintillator	0.06	MeV
	77	#/s			
Flux loss due to air absorption			**Scintillation photon**		
Flux attenuation by the air	99.90%	at 10mm	LXSR photon yield coefficient	32000	# per MeV
Flux at the scintillator	0.0000772	Mbq	Theoretical total photon emitted (N_0)	1920	#
	77	#/s	Decay time constant	32	ns
			SiPM acceptance window	1	ns
			Total photon in the acceptance window (N)	59	#
			Photon loss		
			Reduction due to the reflection	68%	
			Photon arrived at the exit surface	40	
			Reduction due to the interface size mismatch	14.15%	
			Photon arrived at the SiPM sensitive area	6	
			SiPM detector		
			Photon detector efficiency	25%	
			Final detected photon	1.421	
			Photon detection efficiency		
			Photon detection end to end efficiency	0.07%	
			system energy resolution		
			Overall system energy resolution	**0.042**	MeV/photon

The calculation shows the expected alpha flux rate is **178** per second and the gamma flux rate is **77** per second. The expected alpha output signal is **13**p.e. and the gamma's is **~1**p.e. These numbers will be compared with the experiment data in 7.6.

7.6 Test Results

In the experiment, four different distances between the source and the detector are tried. The original plan sets the four distances at 4cm, 3cm, 2cm and 1cm. But in the real experiment operation, the exact distance is not measured so the knowledge of the exact distance is not known. Hence, we use the label Distance 1, 2, 3, 4 to denote four

different distances from far to near as explained in Table 7-9. As the exact real distance is unknown, we can only explain them with some unprecise text.

Table 7-9 The definition of distance label in relationship to the planned distance

Label	Planned Distance	Real Distance
Distance 1	4cm	Furthest from source
Distance 2	3cm	Closer
Distance 3	2cm	Much closer
Distance 4	1cm	Nearly touch the source

This is a mistake in operation and it leads to some data analysis inaccuracy, particularly for the alpha particle. But we can see the data histogram distribution changes as distance varies. So, we are convinced the system is working and responsive.

The data is acquired in the trigger mode. First a set of background run is done to establish the preamplifier electric noise and the SiPM dark count 1p.e. amplitude. Then the triggering threshold is set between 2p.e. and 3p.e. This aims to remove all the background signal up to 2p.e. amplitude. As discussed in 6.5, most of the thermal generated dark count is 1p.e. signal. Some will present as 2p.e. signal due to the optical cross-talk. However, not all the background signal can be completely removed at this threshold, because with the scintillator attached, very occasionally there will be background cosmic ray passing through the scintillator and trigger multiple p.e. signal. Also, some optical cross-talk can trigger signal above 3p.e. Moreover, the distribution span of the signal amplitude means there is no absolute boundary between 2p.e. and 3p.e. signal. Some 2p.e. signal with high enough amplitude will present over the threshold and be captured. In short summary, the background signal cannot be completely eliminated, but can only be reduced to statistically non-significant. What this means is that when the radiation source is presented, most of the triggered signal shall be genuine events from the source with a non-significant portion of signal from the background. However, quantify the background signal requires efforts beyond the reach of this simple setup and author's available time. So, this will be left for further study.

7.6.1 Overview

Table 7-10 lists all the data recorded during the experiment. Upon triggering, the full frame will be saved as an event. Each run contains 2000 events and is saved as an individual data file. The frame length is configured at 1000ns which is slightly different to the 200ns set in the previous system characterization test. This is achieved by reducing the sampling frequency from 5GHz to 1GHz. So, the same buffer can hold 5 times longer period. This arrangement is to capture as much signal as possible.

Each event has a timestamp added by the DRS4Osc software with the accuracy of 1ms. The timestamp of the first event and the last event are listed here to allow the calculation of the duration of the run. One immediate finding is that the duration is reduced as the distance between the source and the scintillator detector is shortened. This means it takes less time to accumulate 2000 events, in another word, more signal is detected when source is moved closer to the detector.

Figure 7-20 calculates the average event rates from the event timestamp. From 6.2.3.2, it is established that our detection system max frame rate is about 354. Data in Table 7-10 shows the highest observed frame rate (or trigger rate) at Distance 4 is still within the system capability.

Table 7-10 List of data recorded during the radioactive test with Am-241

Test Data	Description	Events	First Event Time Stamp	Last Event Time Stamp	Duration (ms)	Frame Rate (s^{-1})
Background 1	electronics noise	n/a	n/a	n/a	n/a	n/a
Background 2	Without source, biased close to active level	2000	13:58:24.810	13:59:12.921	48111	41.57
Distance 1	Test with source Run 2	2000	14:22:43.183	14:23:00.948	17765	112.58
	Test with source Run 3	2000	14:23:34.340	14:23:51.612	17272	115.79

Test Data	Description	Events	First Event Time Stamp	Last Event Time Stamp	Dura-tion (ms)	Frame Rate (s^{-1})
	Test with source Run 4	2000	14:24:22.505	14:24:39.535	17030	117.44
Distance 2	Test with source Run 1	2000	14:25:38.566	14:25:50.837	12271	162.99
	Test with source Run 2	2000	14:26:03.139	14:26:15.687	12548	159.39
	Test with source Run 3	2000	14:26:26.82	14:26:38.520	11700	170.94
Distance 3	Test with source Run 1	2000	14:29:30.672	14:29:37.895	7223	276.89
	Test with source Run 2	2000	14:29:55.392	14:30:02.580	7188	278.24
	Test with source Run 3	2000	14:30:15.427	14:30:22.698	7271	275.07
	Test with source Run 4	2000	14:30:42.602	14:30:49.790	7188	278.24
	Test with source Run 5	2000	14:31:02.564	14:31:09.780	7216	277.16
	Test with source Run 6	2000	14:31:25.841	14:31:33.70	7859	254.49
Distance 4	Test with source Run 1	2000	14:33:23.969	14:33:30.470	6501	307.64
	Test with source Run 2	2000	14:33:42.602	14:33:49.200	6598	303.12
	Test with source Run 3	2000	14:34:03.267	14:34:09.811	6544	305.62
	Test with source Run 4	2000	14:34:24.804	14:34:31.290	6486	308.36

Test Data	Description	Events	First Event Time Stamp	Last Event Time Stamp	Dura-tion (ms)	Frame Rate (s^{-1})
	Test with source Run 5	2000	14:34:53.767	14:35:00.316	6549	305.39

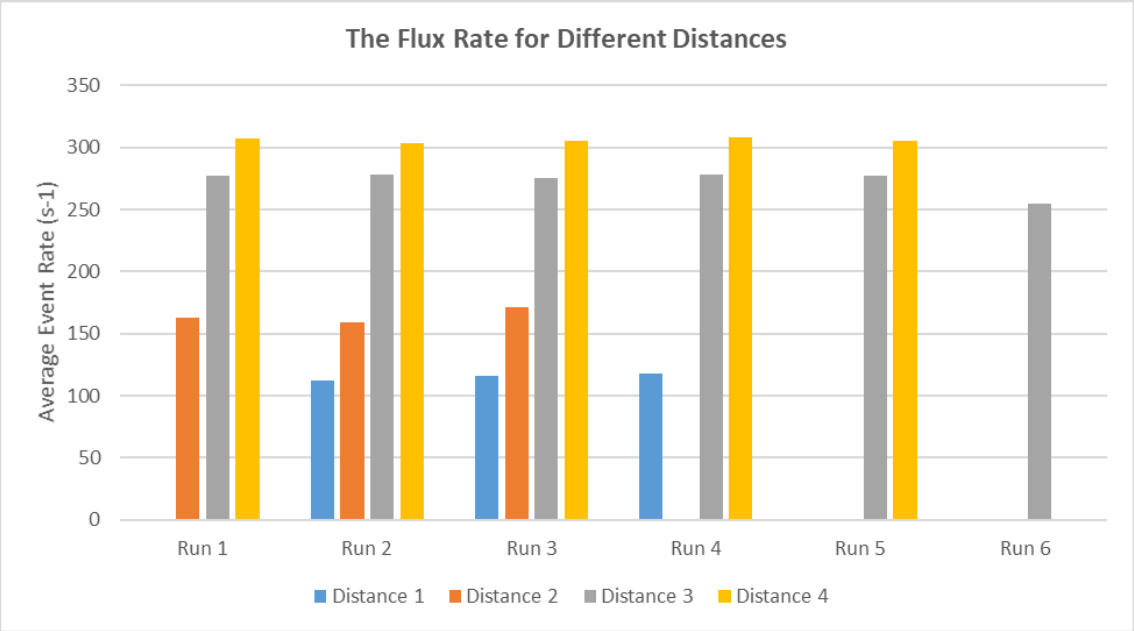


Figure 7-20 The average flux rate for the four different distances.

Figure 7-20 indicates Distance 3 and 4 are quite close and Distance 1 and 2 are relatively close. The average event rate across different Run is quite consistent.

The captured flux rate is compared to the calculated values (assuming distance at 1cm, 2cm, 3cm and 4cm) as shown in Figure 7-21. This comparison is not quite meaningful as the real distance is unknown. Interesting is that the real observed flux rate is generally higher than the theoretical calculation.

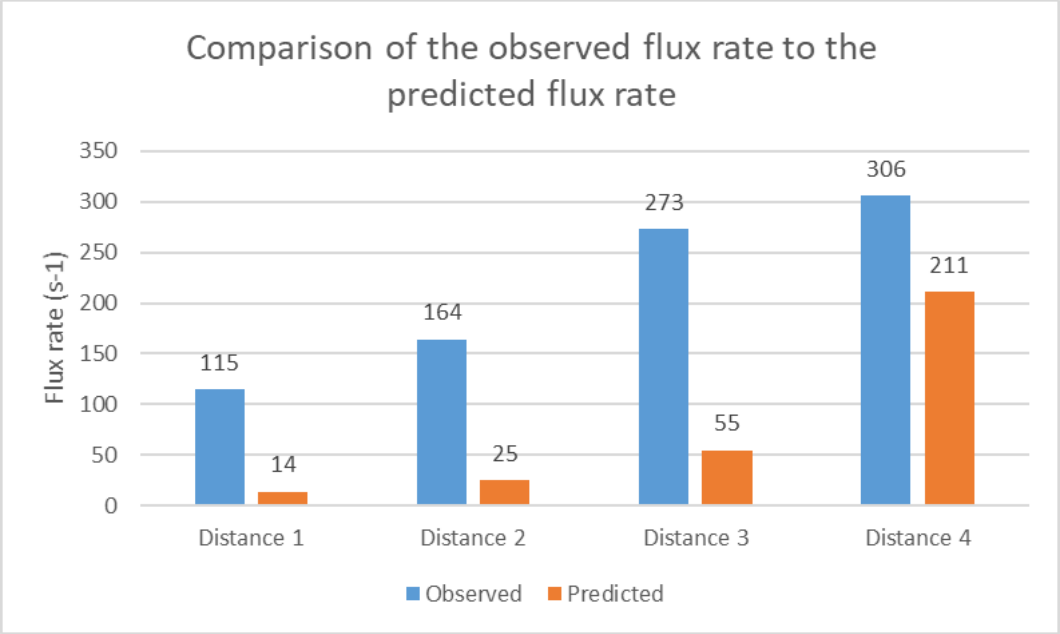
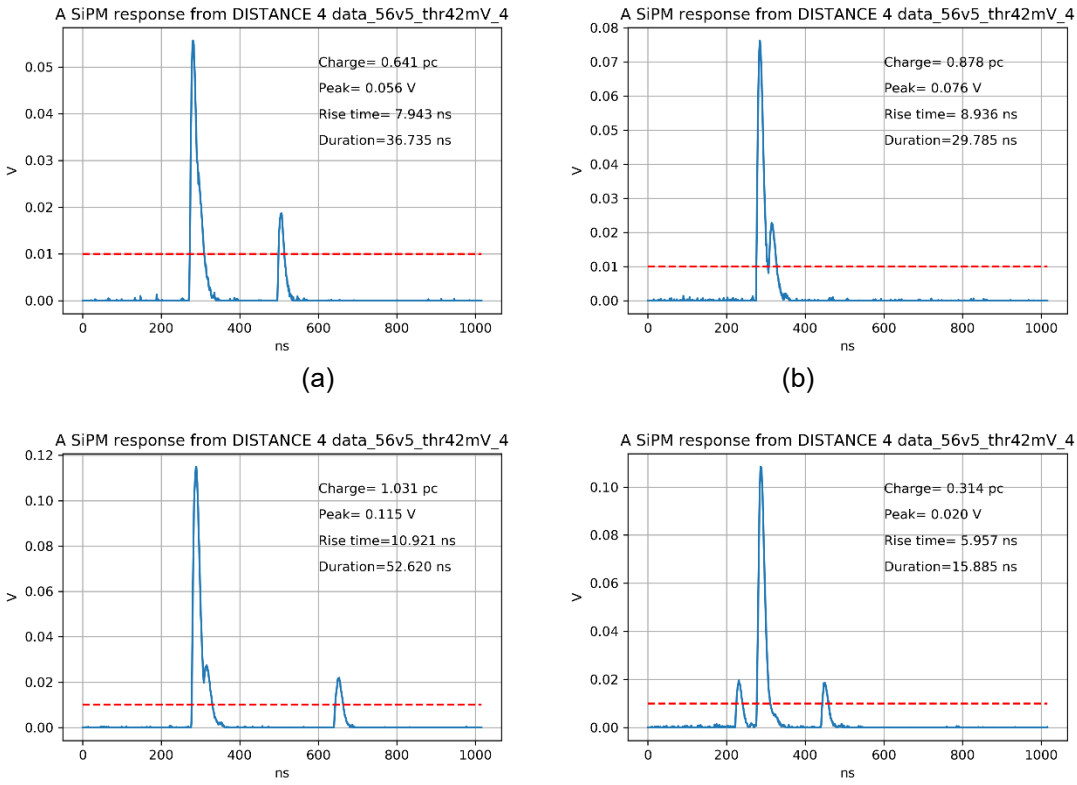


Figure 7-21 Comparison of the observed flux rate and the predicted flux rate.

7.6.2 Data

Each event is recorded like an oscilloscope screenshot with a length of 1000ns at the resolution of 1GHz sampling rate. Figure 7-22 shows some typical event plots with more than one signal captured.



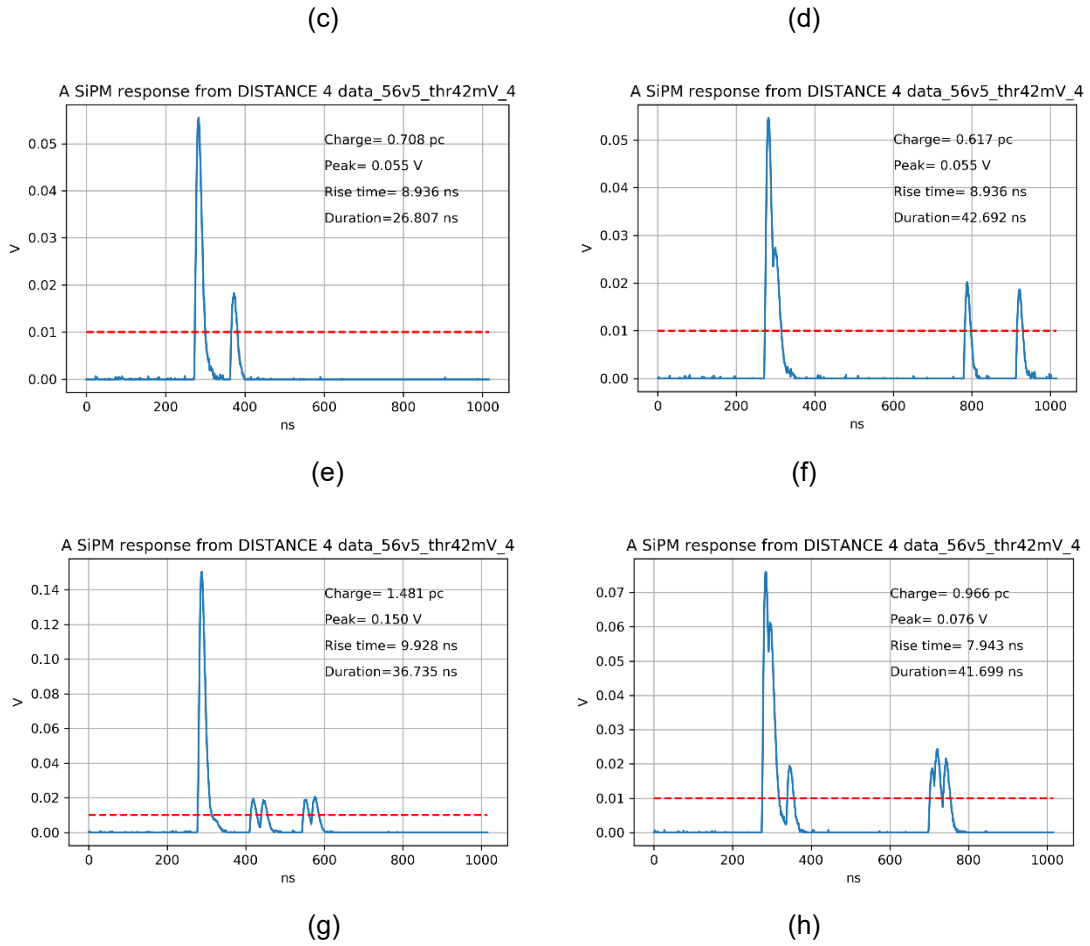
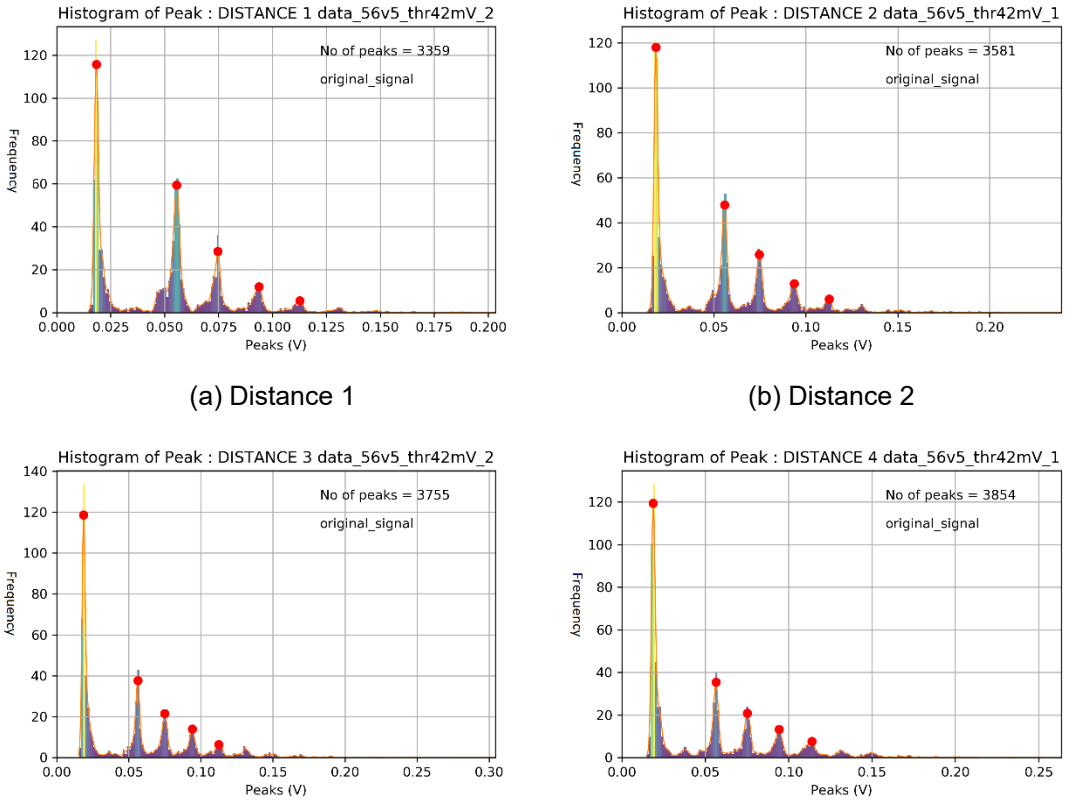


Figure 7-22 Typical DRS4 module capture of the preamp output signal for those events with more than one signal.

The DRSOsc software is set to trigger at 42mV which is between 2p.e. and 3p.e. and the trigger point is set at about 282ns of the 1000ns long event frame. So, the leading triggering signal will be always equal to or above 3p.e. and appear near 282ns time. There is occasionally a 1p.e. signal before the triggering signal like plot (d). That probably is a dark count signal falling into the event frame at that time by chance. There are many smaller signal (mostly 1p.e. signal) trailing the triggering signal. To be convenient for discussion, the triggering signal is defined as primary, and those trailing signals as secondary. To be clear, these are just names not suggesting the secondary signals are descendants of the primary signal.

Using the same Python processing software, the histogram of signal peak for each run data at different distance is presented in Figure 7-23.



(a) Distance 1 (b) Distance 2

(c) Distance 3 (d) Distance 4

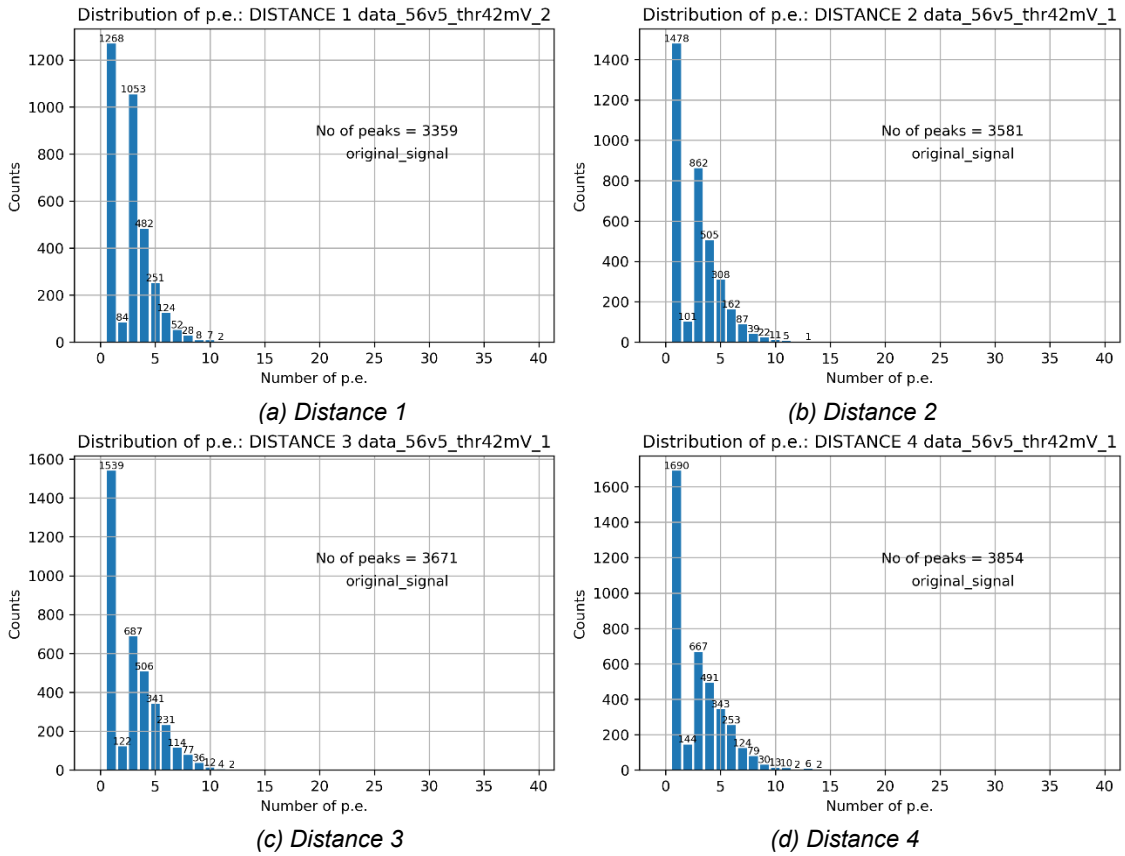
Figure 7-23 Histogram of signal peak for four various distances

The leading pulse is the 1p.e. signal. The 2p.e. signal is missing because the threshold is set between 2p.e. and 3p.e. So, the 2p.e. is filtered as expected. The excessive 1p.e. signal is an unexpected result. More analysis will be done to explore the source of this 1p.e. signal in 7.6.3.

Due to the excessive high number of the 1p.e. signal, the structure of the other p.e. signals are not obvious for different distance. If examined very closely, as distance shortens, more higher p.e. signal appears. This broadly matches the expectation.

7.6.3 Analysis

To make a quantitative analysis, the analogue plots in Figure 7-23 are first converted to digital plots using the method developed in the 5.4.4, with result shown in Figure 7-24.



(a) Distance 1 *(b) Distance 2*

(c) Distance 3 *(d) Distance 4*

Figure 7-24 The histogram of system output signal peak in digital p.e. number

Using the data in the plot, the frequency of 1p.e. signal in respect to total signal is calculated in the Table 7-11. The ratio is very close to the 40% occurrence rate of γ-ray by-product of the Am-241 emission. More in-depth analysis of the 1p.e. signal will follow.

Table 7-11 Summary of the frequency of 1p.e. signal for one run at every distance

	1p.e. frequency	Total (inc. 1p.e.)	Ratio
Distance 1 run 2	1268	3359	37.7%
Distance 2 run 1	1478	3581	41.3%
Distance 3 run 1	1539	3671	41.92%
Distance 4 run 1	1690	3854	47.15%

7.6.3.1 The source of 1p.e. signal

There are a few potential sources for the 1p.e. signal. Each will be evaluated to identify their contribution.

7.6.3.1.1 Dark Counts

The dark count rate (DCR) for SiPM S13360-1325CS at room temperature and at standard bias voltage is about 70kHz. Each dark count (DC) occurs randomly, and the timing obeys the Poisson distribution. The mean of the distribution is the dark count rate. Over a long period, the total number of dark count events can be expressed as DC = DCR x Period.

In this experiment, the data of each run is composed of 2000 events. Each event is 1000ns long. So, the expected total DCs over total period is

$$DC = 2000events \times 1000ns \times 70kHz = 140counts \tag{6.13}$$

Most of the dark count is 1p.e. signal. Compared to the frequency of 1p.e. signals recorded in Table 7-11, which is between 1000 and 2000 counts, the estimated total DC at 140 counts is just a small portion.

7.6.3.1.2 After-pulse

The second possible source is the after-pulse of the primary signal. As calculated in 6.4.1.1, the after-pulse rate is about 7.9% for the S13360-1325CS. Here, only the after-pulse related to the primary signal is considered for simplicity. Those related to the secondary signal is deemed rare and probably fall out of the frame window, so they are ignored. Then, with 2000 events, there are 2000 primary signal captured. The simple calculation is:

$$After - pulse = 2000events \times 7.9\% = 158counts \tag{6.14}$$

This is another important source but still not enough to count for the total 1p.e. signal.

7.6.3.1.3 Second hit from the radioactive source in the same 1000ns capture window

The third source can be a second hit from the radioactive source within the 1000ns capture window. We can evaluate the likelihood of such case based on the data calculated in section 7.5.5:

Received primary alpha particle flux rate = 178 per second

Received gamma photon flux rate = 77 per second

So, the chance of receiving a second alpha hit or gamma hit is:

$$P_{2nd\_alpha} = \frac{178}{10^9 ns} \times 1000 ns = 0.0178\% \quad (6.15)$$

$$P_{2nd\_gamma} = \frac{77}{10^9 ns} \times 1000 ns = 0.0077\% \quad (6.16)$$

For 2000 events, the combined number of possible second hit is:

$2000 \times (0.0178\% + 0.0084\%) = 0.524$, less than 1 count.

This calculation assumes the decay occurrence obeys the uniform distribution over the observation period. While this is not quite true, the decay obeys Poisson distribution and the gap between consecutive decay obeys exponential distribution. However, as long as the observation period is long enough, the uniform distribution or average is a good approximate for the estimation here.

The calculation shows the chance of second hit is so small that it can ignored.

7.6.3.1.4 Delayed scintillation photon

The fourth source can be the delayed scintillation photon. From the discussion in 7.5.3.2, it is learned that the scintillator emission timing obeys an exponential distribution. At 4.6 times of the time constant, the accumulated emission will reach 99% of the total. So, the emission span of BC404 can be closely estimated at $1.8ns \times 4.6 = 8.28ns$ and that of LXSR can be around $32ns \times 4.6 = 147.2ns$. Considering the preamplifier circuit will prolong the output signal, slightly longer decay from the experiment data is expected.

Figure 7-25 shows the histogram of delay of all the secondary signal with respect to the primary signal. This generally matches the expected exponential distribution with two exceptions. One is that the scale of delay is much larger than the theoretical calculation plotted in the Figure 7-15 and Figure 7-16. The other is that the tail of the data plot is higher than a typical exponential distribution shown as a fit function in orange.

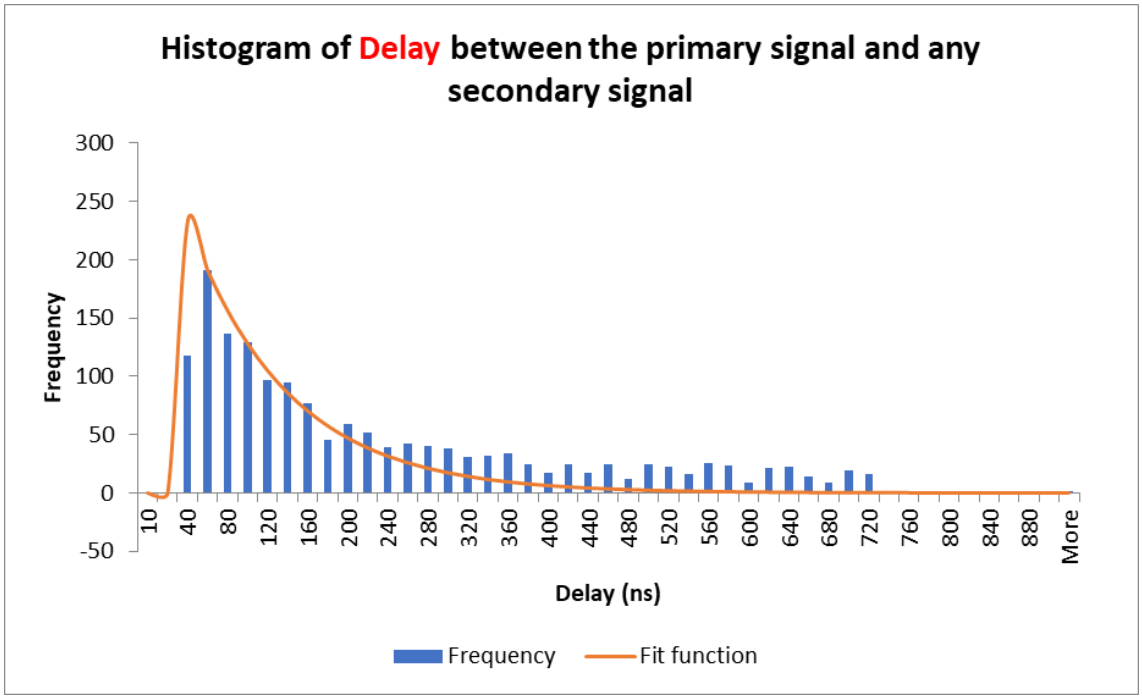


Figure 7-25 the histogram of delay between the primary signal and any secondary signal within the same 1000ns capture window. The plot is based on one dataset at Distance 4.

This suggests the delayed scintillation photon forms part of the secondary signal (most are 1p.e.) but it is not quantitively known how big this contribution is. To understand this more, some experimental data interpretations are tried.

7.6.3.1.5 First try – sum all the secondary onto the primary signal

The first experiment of interpretation assumes all the secondary signals are originated from those delayed scintillation photons. Here, the contribution of dark count, after-pulse or the second hit are ignored as they only form a very small part of the secondary and also for each individual event, it is almost impossible to separate the signals of the "delayed" scintillation photon from these unwanted sources. Based on this assumption, the calculation is simple. All the secondary signals within the same event frame are summed and added to the primary signal. For example, if the primary signal is 3p.e., and the event frame contains another two 1p.e. signal, then the total signal for this event is 5p.e. The effect can be seen in Figure 7-26, where green is the peak histogram of the original data and blue is the new histogram after summing up all the signals in the same frame.

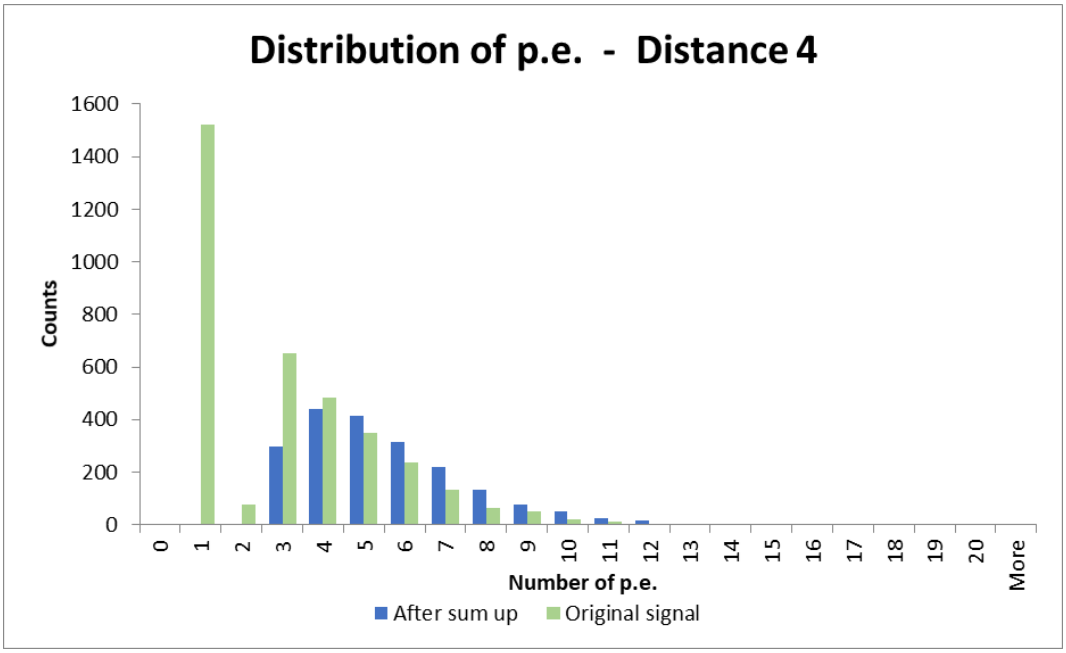
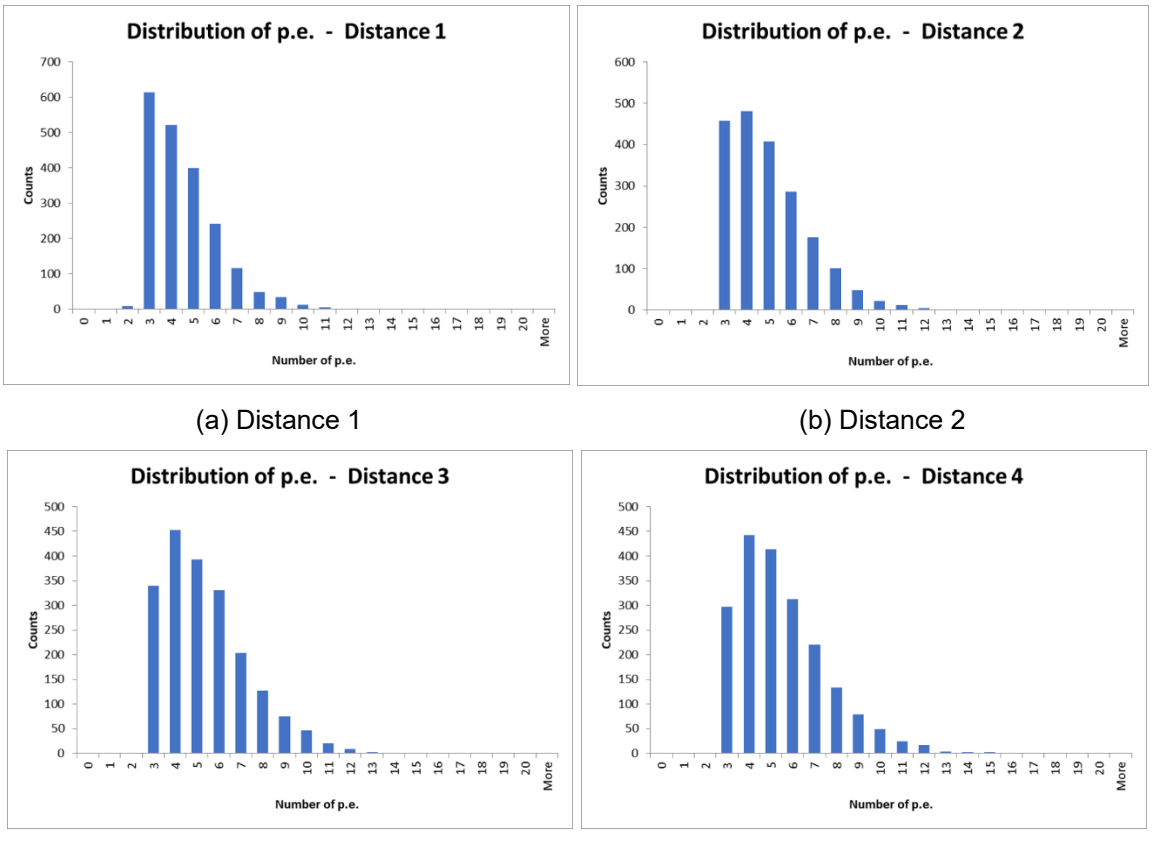


Figure 7-26 Comparison of the histogram of signal peak in number of p.e.: green is the original data and blue is the data after summing up all the primary and secondary signals

In the new histogram, the standing out 1p.e. signal has been absorbed. The profile stretches out and becomes smoother. The process is repeated for one run data in each distance with results shown in Figure 7-27.



(a) Distance 1 (b) Distance 2

(c) Distance 3 (d) Distance 4

Figure 7-27 Rearranged p.e. distribution for four different Distance data set using data processing method 1

It is clear the histogram profile central point is shifting towards the higher p.e. number as the distance shortens. This simple method seems to give us a reasonable result. However, it is not entirely satisfactory as this method doesn't separate the response from the two different scintillators. Therefore, a more complex method is tried.

7.6.3.1.6 Second try – two separate responses

In this second experiment of data interpretation, the event frame is divided into two parts. The first part is the primary signal and any immediate secondary signal within 50ns. Signals within this part is summed to represent the response of BC404 scintillator. 50ns is chosen as the cut off threshold because it is found 106 out of 131 total after-pulse signal occurred within 50ns to the primary signal from the SiPM characterization data set. The second part is all the secondary signals after 50ns. They are combined to represent the response from the LXSR scintillator. This method is illustrated in Figure 7-28 on an example:

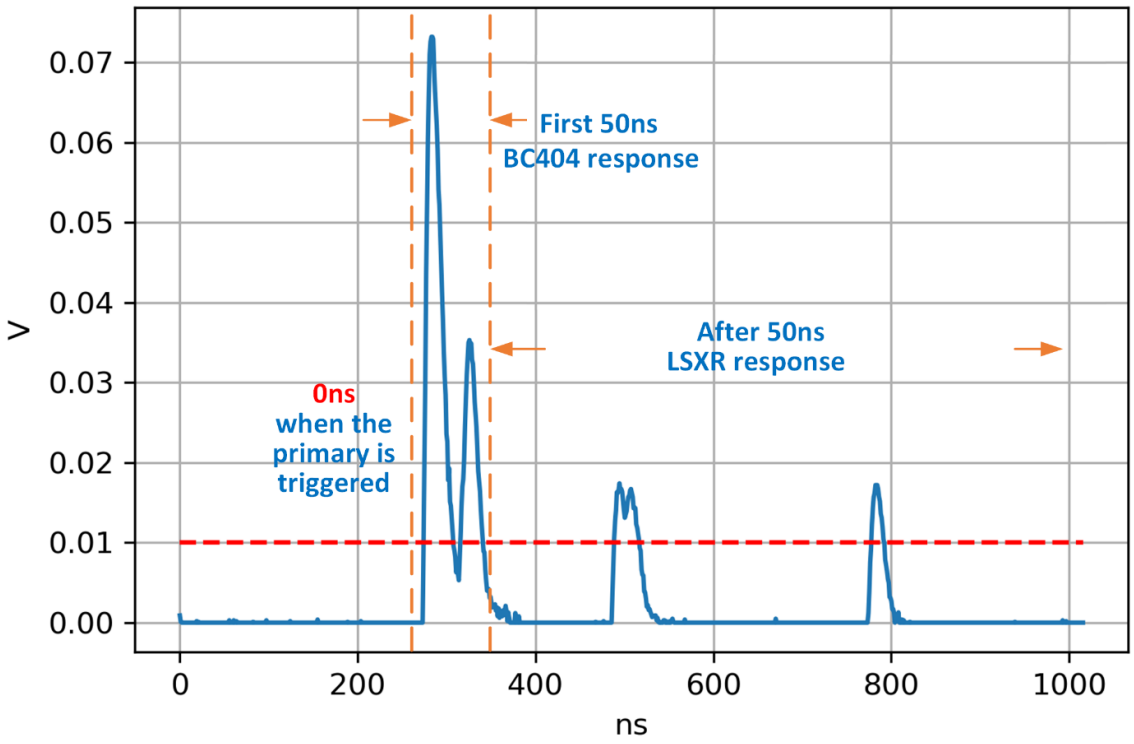
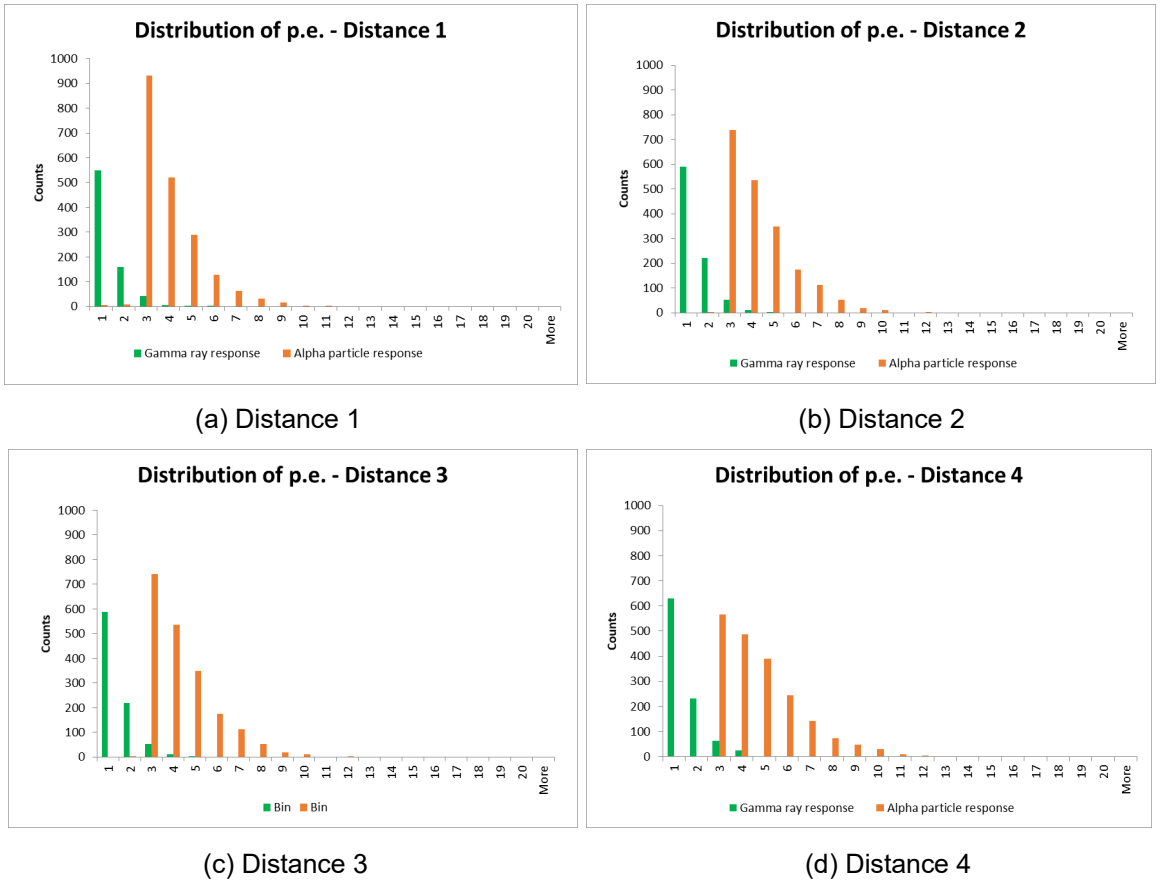


Figure 7-28 Illustration of separation of the response of two scintillators by the timing

The re-processed data is shown in Figure 7-29 for four different distances. The orange data is the response of BC404 and the green data is that of LXSR.



(a) Distance 1

(b) Distance 2

(c) Distance 3

(d) Distance 4

Figure 7-29 Breakup of two sources for the four different distance using data processing method 2

Statistically sum the frequency of counts (regardless of the amplitude, each combined signal is regarded as one count) from BC404 and LXSR. The results are listed in Table 7-12.

Table 7-12 Statistical summary of counts from BC404 and LXSR

	Counts from BC404	Counts from LXSR	Ratio
Distance 1 run 4	2000	760	38%
Distance 2 run 3	2000	873	43.65%
Distance 3 run 5	2000	924	46.2%
Distance 4 run 4	2000	957	47.85%

The data is triggered on the response of BC404 so that its total counts are always 2000, the same as the number of captured events. The counts of LXSR is the collection of those remaining secondary signals after 50ns. The data shows the counts of LXSR is increasing as the distance is decreasing. The ratio is calculated as the number of counts from LXSR to BC404. The discussion in 7.4 suggest the response of

LXSR is from the gamma ray and the response of BC404 is from the alpha. Also, the review of Am-241 decay spectrum in 7.2.3 states the ratio of gamma decay to alpha decay is about 0.4:1 (=40%). The ratio presented in Table 7-12 confirms the 40% ratio. The flux loss due to the distance between the source and the detector discussed in 7.5.2.1 is reflected by the increasing gamma to alpha ratio as distance decreases.

The data quality can be further improved by removing the dark counts statistically. Though at individual base, the dark count can't be distinguished from a normal 1p.e. signal, statistically it is reasonable to remove it over a long enough of period of acquisition using the knowledge of average dark count rate. The portion of each scintillator effective emission period over the full frame period is used to work out the portion of the dark counts. For LXSR, it roughly equals to 1000ns – 282ns (triggering point) – 50ns = 668ns. For BC404, it is 50ns. Quick calculation based on the result from 7.6.3.1.1 shows:

$$\begin{aligned} BC404\_DC_{total} &= \frac{50ns}{1000ns} \times 140counts = 7counts \\ LXSR\_DC_{total} &= \frac{668ns}{1000ns} \times 140counts = 93.52counts \end{aligned} \tag{6.17}$$

As those dark counts (1 p.e.) occur during the BC404 response period is mixed with the delayed scintillator photon from BC404. So, in our treatment, they have been combined with the primary signal together (illustrated in Figure 7-29). Hence, there is no way to remove the dark counts for the BC404 response. However, most of the LXSR response is made up of just one 1p.e. signal and is scattered around in the time line. Some of the 1p.e. signal can be the dark count. Table 7-13 shows the distribution of the composition for the LXSR response (after summing up all the secondary after 50ns) for one data set as an example.

Table 7-13 The composition data of the LXSR response with all the secondary summed up from one data set

Bin (p.e.)	Frequency
0	0
1	621
2	210
3	72
4	16

Bin (p.e.)	Frequency
5	4
6	1
7	0
More	0
Total	924

As Table 7-13 is constructed from the sum of all the secondary assigned to the LXSR response period, there will be dark count included in each bin. For example, a 2p.e. bin might be composed of two 1p.e. signal, of which, one is a dark count and the other is a valid signal. Once summed, this detail information is lost, and it becomes a single 2p.e. signal. Removing dark count from such bin will lead to loss of valid signal. Therefore, dark count number cannot be subtracted from such bin. This means it is only possible to remove the dark count from the 1p.e. bin, which only contains one 1p.e. signal, that will be either a dark count or valid signal. Statistically, over the full samples, it is reasonable to remove a portion of the 1p.e. signal as the dark count. However, the average dark counts calculated in (6.17) can't be used directly for the deduction of 1p.e. bin, because part of that total 1p.e. signal will enter those higher than 1p.e. bins. Hence, it is necessary to make some adjustment using the assumption that the portion of dark count within the 1p.e. bin is proportional to the size of the 1p.e. bin over the total sample size as:

$$\begin{aligned} DC_{to\_be\_removed} &= \frac{1p.e.\_counts}{total\_counts} \times LXSR\_DC_{total} \\ &= \frac{621}{924} \times 93.52 \\ &= 62.85 counts \end{aligned} \tag{6.18}$$

Apply the same treatment for all the four data sets in Table 7-12 and re-calculate the ratio. The new ratio is closer to the theoretical 40% value as shown in Table 7-14.

Table 7-14 Revised counts from LXSR after removing the dark counts using method in (6.18)

	Counts from BC404	Counts from LXSR	Counts from LXSR (after DC adjustment)	New Ratio
Distance 1 run 4	2000	760	692	34.62%
Distance 2 run 3	2000	873	810	40.5%
Distance 3 run 5	2000	924	861	43.06%
Distance 4 run 4	2000	957	895	44.77%

7.6.3.1.7 Data fit

Knowing Figure 7-29 is made up of two independent scintillator response, it is proper to use the linear superposition of two exponential functions to fit the plot. Each exponential function represents the relevant scintillator response to the relevant incident particle/photon. The expression of the exponential function can be written as:

$$N = \begin{cases} Ae^{-(x-B)/C} & x \geq B \\ 0 & x < B \end{cases} \tag{6.19}$$

Where,

A is a relative number to indicate the amplitude of distribution.

B is the shift of peak point on the x-axis relative to the origin.

C defines the shape of the exponential curve.

Both exponential functions can share the same format as described in (6.19) with just different coefficients. Their linear superposition shall be like:

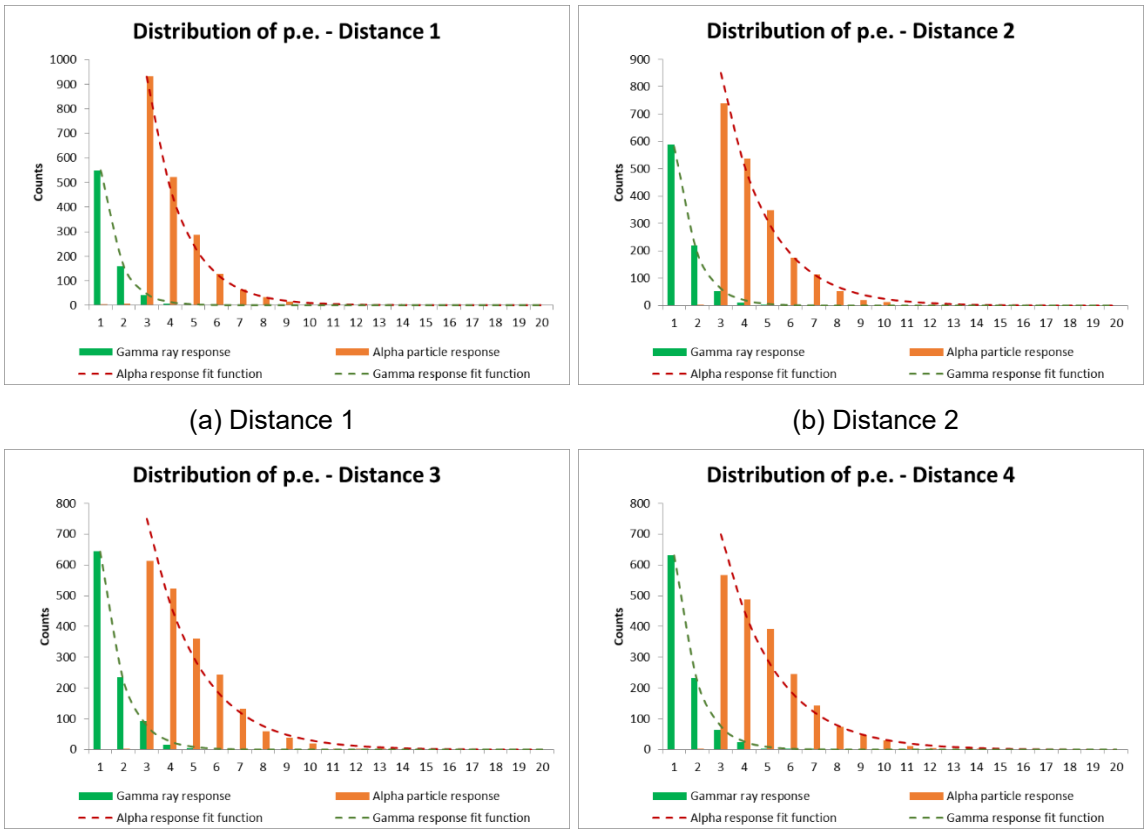
$$\begin{aligned} N_{all} &= N_1 + N_2 \\ &= A_1 e^{-(x-B_1)/C_1}\Big|_{x\geq B_1} + A_2 e^{-(x-B_2)/C_2}\Big|_{x\geq B_2} \end{aligned} \tag{6.20}$$

The data in Figure 7-29 is fit with the set of coefficients that can score the lowest in Chi-squared test by trial and error. The result is presented in Table 7-15 and Figure 7-30.

Table 7-15 The optimized fit function parameters and the associated Chi-squared score

Data	BC404 Response				LXSR Response			
	A_1	B_1	C_1	X^2	A_2	B_2	C_2	X^2
Distance 1	932	3	1.52	22.23	550	1	0.8	4.36
Distance 2	850	3	2	80.99	580	1	0.9	19.28
Distance 3	750	3	2.1	86.69	621	1	0.88	5.186
Distance 4	700	3	2.3	117.27	631	1	0.95	18.8*

* There is one data point for the gamma ray response at 14p.e., which contributes to enormous χ^2 error. That data point is temporally removed to make the result comparable to the other dataset. It is possible that this abnormal data point is caused by a background cosmic ray. This seems to occur only once within all the data samples.



(a) Distance 1

(b) Distance 2

(c) Distance 3

(d) Distance 4

Figure 7-30 Add the two exponential fit functions

From Table 7-15 and Figure 7-30, it is observed that the alpha response is flattening (spread more towards high p.e. number as the distance of the source is closer.) This can also be indicated as the parameter C is increasing.

7.6.3.1.8 Short summary

Now, with reasonable confidence, it can be concluded that the excessive 1p.e. signals shown in Figure 7-23 is originated from the LXSR scintillator in response to the gamma decay. This also echoes the calculation in section 7.5.5 that shows the nominal response of the LXSR scintillator from the Am-241 60keV Gamma ray is about 1 photon.

7.7 System performance summary

7.7.1 System energy resolution

Energy resolution is an important indication to evaluate the detector system performance. Generally, the energy resolution is defined as full width at half maximum (FWHM) for the distribution of signal amplitude. This can be simply expressed as:

$$Energy\_resolution = \frac{\Delta E}{E} \tag{6.21}$$

Where ***ΔE*** is the FWHM and ***E*** is the energy corresponding to the distribution peak. Most observed signal distribution is like a Gaussian shape. For a Gaussian distribution with standard deviation ***σ*** the FWHM can be defined as:

$$FWHM \approx 2.35\sigma \tag{6.22}$$

Then, Eq. (6.21) can be rewritten as:

$$Energy\_resolution = \frac{2.35\sigma}{E} \tag{6.23}$$

When the only source of fluctuations in the signal is the statistic fluctuation of the number N of scintillation photons and their occurrence follows the Poisson statistics, for N>20, the response can be well approximated as a Gaussian distribution. This prompts a **relative energy resolution** as

$$\frac{\Delta E}{E} \propto \frac{\sqrt{E}}{E} = \frac{1}{\sqrt{E}} \tag{6.24}$$

In our system, the number of p.e. is proportional to the energy. So we can alter equation (6.24) as

$$\frac{\Delta E}{E} \propto \frac{1}{\sqrt{N_{p.e.}}} \tag{6.25}$$

However, our system detection efficiency is too low to reach >20p.e. for Am-241 alpha decay or gamma decay. Therefore, we can't work out a meaningful energy resolution or relative energy resolution from the existing data. This will be left for the next stage of work.

7.7.2 System count rate

The system count rate is determined by a number of factors. The first factor is the scintillator module handling rate. The fast scintillator BC404 has a decay time of 1.8ns and the slow scintillator LXSR has a decay time of 32ns. According to the calculation in 7.5.3.2, this corresponds to 8ns and 147ns for 99% of photon yield. We can define this time as the scintillator recovery time. To be conservative, we opt for the slow scintillator recovery time as the module recovery time. Hence, the scintillator module handling rate can be worked out as 1/147ns = 6.8MHz.

The second factor is the SiPM sensor and its readout circuit handling rate. This is thoroughly discussed in 6.2.3.2. Here, we just take the conclusion of 15.5ns as the sensor and readout circuit combined recovery time. This corresponds to 1/15.5ns = 64.5MHz.

The third factor is the DRS4 module data acquisition system output rate. As discussed in 6.2.3.1, this DRS4 module has a waveform processing rate of 354 fps (frame per second). We simplify the calculation by assuming one particle event per frame. This frame rate then equals to the event rate.

As the three factors form a serial chain, the slowest factor will determine the final system count rate, that is 354 evert per second. This is too slow for any practical application, but there are rooms to improve it.

7.7.3 System detection efficiency and noise floor

The system detection efficiency is made up of the scintillator photon yield efficiency as introduced in 2.2.3.1 and the photon detection efficiency as calculated in 7.5.5.

The photon yield efficiency is generally expressed as 500 photon/MeV for BC404 and 32000 photon/MeV for LXSR. These are just nominal values measured by the manufacturer using a calibrated PMT. The yield fluctuation shall follow a typical Gaussian distribution.

The photon detection efficiency is calculated as merely 0.57% for BC404 emission and 0.07% for LXSR emission.

Product the photon yield and detection efficiency, the overall system detection efficiency can be worked out as: 500*0.57%=2.85 photon/MeV for BC404 and 22.4 photon/MeV for LXSR. From our experiment data, the system detection efficiency is even lower. We merely get 3 p.e. (1p.e.=1photon) from the BC404 absorbing the alpha decay and just 1 p.e. from the LXSR absorbing the gamma decay. Given the alpha after air attenuation at 4.5MeV and gamma at 60keV, this equals to 0.67photon/MeV and 16.7photon/MeV. Hence, the real system detection efficiency is even poorer than the theoretical calculation.

Our system noise floor is determined by the SiPM dark count, which is typically 1p.e. With the cross-talk effect, many 1p.e. dark count will become 2p.e. or even more but very rare. So, the system can safely define the noise floor at max 2p.e.

Clearly, for Am-241 source, the current system detection efficiency is not sufficient to produce a good quality (i.e. good signal to noise ratio) measurement.

7.8 Conclusion and future work

We can conclude that our prototype system works from end-to-end after this radioactive source Am-241 test. By setting the triggering threshold at 2.5p.e. level, all the system background noise is discriminated. This in theory will leave only the signal from the radioactive source. However, the real experiment data shows excessive 1p.e. signal captured after the leading triggering signal which is discriminated at and above 3p.e. This led to the thorough investigation on the source of this 1p.e. signal. The conclusion is that majority of this 1p.e. signal is from the LXSR scintillator in response to the gamma decay and minority is from the dark count and the after-pulse of the triggering signal.

The full system detection efficiency was studied. All the potential loss is counted and calculated. This yields the theoretical system response. It is found the real efficiency from the experiment data is even lower. This is the area that requires further work.

The theoretical analysis predicts the alpha decay will only be detected by the BC404 scintillator and the gamma decay will mostly be detected by the LXSR detector. The scintillator photons are separated by their timing. By examining the experiment data, it is found the ratio of the occurrence of the two scintillator emission closely match the theoretical ratio of the two decays from the Am-241 source as shown in Table 7-14.

While studying the experiment data, an exponential photon decay model is proposed to fit the data. In general, the fit has a good chi-square score. However, a Gaussian fit with the mean representing the nominal system response and the standard deviation representing the fluctuation may be a better approach. This will be left to the next study.

The next step shall be testing the system in the vacuum environment to remove the impact of air attenuation, which is sensitive to distance. More radioactive sources shall be tried to generate more data and have wider energy coverage. The very low system efficiency is affecting the data quality. So, the detail of the loss will be further examined, and efforts will be put to reduce the loss.

Chapter 8 Conclusion

8.1 Summary

This thesis has developed a scintillator + SiPM based detector system as a pathfinder for a potential full-scale space instrument. The prime target application is an in-situ radiation monitor to provide real time warning for the satellite or human crew in space. Moreover, this type of detector system can find many other applications beyond space astronomy. It is already widely used in nuclear high energy particle experiments, medical instruments and so on. Hence, this research outcome has the potential for very wide impact with many beneficiaries.

This thesis structure is designed to follow the standard development flow of a scientific instrument, which is:

1. start with a science requirement definition,
2. then to carry out the detector design,
3. then to verify the design with simulation,
4. then to develop the electronics system to drive the detector,
5. then to develop the software to drive the electronics and analyse the data,
6. then to test the system with artificial stimulating signal (e.g. LED) to test the electronics system and its driving software,
7. then to carry out the final end to end system test with a real source.

Thus, each step forms a chapter of this thesis. A full instrument development requires much more efforts than a PhD research project can cover. Therefore, we have only picked up one or two topics from each step to focus our work on.

The thesis starts with the study of the space environment for the near-Earth orbit. A set of particle species and flux for a typical low Earth orbit are derived using the SPENVIS tool. This information forms the basic performance requirements for the detector system, particularly for the scintillator design.

The second chapter has discussed the scintillator design. The background study about scintillator material and principle helps to establish the ground of the material choice. The concept of fast and slow sandwiched structure has been demonstrated very successful in identifying ion species by Geant4 simulation. The same algorithm is not very successful in separating gamma, or electron from the ion species. A different approach is required for identifying gammas and electrons. A cylindric detector module with two layers of scintillators (BC404+LXSR) was manufactured. Based on the scintillator module design, an instrument concept is proposed.

The third chapter continues to demonstrate the performance of the scintillator module using Geant4 simulation. The simulation programming and the detailed simulation are all done on the initial hemisphere scintillator module as the module change came later. The simulation examines the energy response, angle response of the scintillator module under the bombarding of e-, p+, alpha and gamma. For the angle response test, a sectioned global source was developed. All the simulation data is processed with ROOT software. The results show good angle resolution and provide a detail energy response range, which is in broad agreement with the requirements. For the next stage of research, more simulation on the current cylindric module and more details of the photon detection efficiency in the SiPM should be done. SPENVIS tool has very close link to the Geant4 tool. With SPENVIS, it is possible to generate the orbit flux profile as the input macro script for the Geant4 simulation program. This feature can be explored in the future.

The fourth chapter has discussed about the SiPM readout electronic system. The chapter starts with the exploration of SiPM SPICE model for circuit simulation and the approach to acquire all the necessary SiPM electric parameters, some of which are not provided by the manufacture datasheet. Then the chapter focuses on the preamplifier circuit design and analysis. Two popular architectures: current preamplifier and voltage preamplifier are thoroughly compared on bandwidth, small signal and stability. The current preamplifier is deemed better with wider bandwidth and better stability. The circuit phase compensation as a method to improve the stability is introduced. A very detail noise model for the current preamplifier circuit was constructed and analysed. This helps to understand the source of the circuit electronic noise and the role of the SiPM capacitance. Besides frequency domain, a small section is dedicated to the time domain and a time-domain signal model is derived. The chapter concludes

with the engineering realization of a preamp PCB. The initial electric test shows good single photon detection capability of the SiPM + preamplifier circuit. The remaining of the readout electronic system is made up of the discrete and off the shelf electronics. Thus, a complete SiPM readout electronic system is fully constructed.

The fifth chapter has discussed the data processing software. To address the signal pile-up issues, the Deconvolution + Wiener filter combination is introduced. This combination successfully separates the pile-up signal while maintaining a good signal to noise ratio. The key parameters of both filters have been discussed. The further optimization of the filter implementation, parameter determination and possible dynamic parameter adjustment can be the topic of the next stage. The chapter also introduces a series of software developed in Python, for data extraction, waveform analysis, histogram analysis and digitization.

The sixth chapter has discussed the SiPM system characterization with an adjustable LED source. This is based on the hardware built in Chapter 4 and the software developed in Chapter 5. This simple and low-cost system provides reasonably accurate measurements of the SiPM performance with many results matching the manufacturer's data closely. High dark count rate is a major hurdle to the SiPM based system. Much efforts have been spent on studying the dark count. A mathematic probability model has been constructed to simulate the occurrence of the dark count. The model proves that more than 99.997% of the dark count is 1p.e. signal. Therefore, the excessive 2p.e. or above signal that appeared in the dark count measurement experiment is from the photon mediated crosstalk.

The last chapter has discussed the final end-to-end test with a radioactive source Am-241. The study of the Am-241 spectrum unveils the dominant ~5.5MeV alpha decay and a weaker 60keV gamma decay. The coexistence of the two components is perfect to test our sandwiched scintillator module as alpha at 5.5MeV will only deposit in the BC404 scintillator (fast) and gamma at 60keV will only deposit in the LXSR scintillator (slow). We can clearly see the system response to the radioactive source. However, due to the overall very low detection efficiency and air attenuation of the source, the registered signal in p.e. format is very low. The signal is too low to form any meaningful energy resolution calculation. And only from statistic calculation, we can extract the part of the signal that is attributed from the gamma ray detection by the LXSR scintillator. The test is inconclusive, and it points out the direction that the system needs to

be improved. And we plan to repeat this test under vacuum condition as next step. To fully characterize the stacked scintillator module, test with more radioactive source or particle beam are necessary. And these could well be our next step.

8.2 Future plan

This is the end of this thesis, but it is just the beginning of a new era. With the SiPM sensor and the new type of scintillator material, this scintillator + SiPM technology platform opens a new world for the particle instrument, particularly at high energy region. Next step, we shall further characterize the scintillator detector module, improve the SiPM readout circuit on speed and dynamic range. Then, we can seek some path-finding mission opportunity to develop a full-scale instrument. The future is only limited by imagination and determination.

8.3 Author contribution

All the engineering works including both hardware and software development presented in this thesis are originated from the author himself. All the survey work and many chapter introduction sections quote the existing research from publication.

In Chapter 1, the survey of space environment, the instrument and detector technology quote many existing research outcomes. The detector system requirement derived from a target orbit using Spenvis tool is the author's original work.

In Chapter 2, the scintillator principle introduction is based on the publication survey. The scintillator module is designed by the author. In the third iteration, the new material is recommended by the manufacturer, but the author has done analysis to confirm the choice. The instrument concept is the author's own idea.

In Chapter 3, the simulation with Geant4 software is the author's original work. The Author learned the Geant4 software from scratch and developed the simulation software.

In Chapter 4, the author updated Corsi and Marano's electrical model for SiPM with a small practical innovation. The author used Corsi's method and other common methods and conducted the SiPM parameter measurements work. The Author studied the preamplifier circuit theory and conducted the in-depth analysis of the circuit performance covering bandwidth, stability, noise source and derived the formula for

the time domain signal waveform. The author then developed the preamplifier circuit board by designing the schematic, PCB layout , procurement and final PCB soldering and testing all by himself. The author experimented with many different circuit configurations with different components, particularly the op-amps.

In Chapter 5, the author got inspiration from the Hamamatsu technical note[139] on SiPM about the Deconvolution + Wiener filter to address the signal pile-up issue. The author developed the Python code to realize the filters and experimented the filter on the real data. The Author developed the full chain software to process the waveform data from the DRS4 module and generate all the histograms.

In Chapter 6, the author built the full data acquisition system with self-built hardware and software, and some third-party module and instruments. The author studied the methods about characterizing the SiPM from publication and followed these methods to characterize the SiPM sensor performance. The author developed a relatively simple statistical model for the dark count occurrence using Binomial distribution. Based on this statistical model, the author developed a simple simulation program to test the theory. The model proved the chance for a 2p.e. dark count is next to zero.

In Chapter 7, the author carried out the first radioactive test for the detector system with the support from colleague on the test setup and conduction. The Author carried out the test plan, data collection and data analysis. The author also conducted a thorough system performance analysis in order to check the real data.

Reference

[1] E. R. Benton and E. V Benton, “Space radiation dosimetry in low-Earth orbit and beyond.,” *Nucl. Instrum. Methods Phys. Res. B.*, vol. 184, no. 1–2, pp. 255–94, Sep. 2001, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/11863032.

[2] E. R. Benton and E. V. Benton, “Space radiation dosimetry in low-Earth orbit and beyond,” *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms*, vol. 184, no. 1–2, pp. 255–294, Sep. 2001, doi: 10.1016/S0168-583X(01)00748-0.

[3] J. G. Roederer, *Dynamics of Geomagnetically Trapped Radiation*, vol. 2. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970.

[4] M. Walt, *Introduction to Geomagnetically Trapped Radiation*. Cambridge University Press, 1994.

[5] J. F. Lemaire, D. Heynderickx, and D. N. Baker, Eds., *Radiation Belts: Models and Standards*. Washington, D. C.: American Geophysical Union, 1996.

[6] D. M. Sawyer and J. I. Vette, “Ap-8 trapped proton environment for solar maximum and solar minimum.” 1976, Accessed: May 15, 2021. [Online]. Available: http://inis.iaea.org/Search/search.aspx?orig_q=RN:9351076.

[7] J. I. Vette, “The AE-8 Trapped Electron Model Environment,” 1991. doi: Astrophysics.

[8] “Background: Trapped particle radiation models.” https://www.spenvis.oma.be/help/background/traprad/traprad.html#APAE (accessed May 15, 2021).

[9] D. Heynderickx, “SPENVIS Tutorial : Radiation models in SPENVIS and their accuracy,” 2013.

[10] D. J. Williams, “The Earth’s ring current: Causes, generation, and decay - Tutorial lecture,” *Space Sci. Rev.*, vol. 34, no. 3, pp. 223–234, Mar. 1983, doi: 10.1007/BF00175279.

[11] I. A. Daglis, R. M. Thorne, W. Baumjohann, and S. Orsini, “THE TERRESTRIAL RING CURRENT’ ORIGIN, FORMATION, AND DECAY,” 1999. doi: 10.1029/1999RG900009.

[12] L. J. Lanzerotti, “Space weather effects on technologies,” in *Geophysical Monograph Series*, vol. 125, Blackwell Publishing Ltd, 2001, pp. 11–22.

[13] “Earth’s Magnetosphere | NASA.” https://www.nasa.gov/mission_pages/sunearth/science/inner-mag-mos.html (accessed May 12, 2021).

[14] D. V. Reames, *Solar Energetic Particles*, vol. 932. Cham: Springer International Publishing, 2017.

[15] D. H. Brooks and S. L. Yardley, “The source of the major solar energetic particle events from super active region 11944,” *Sci. Adv.*, vol. 7, no. 10, Mar. 2021, doi: 10.1126/sciadv.abf0068.

[16] E. N. Parker, “THE PASSAGE OF ENERGETIC CHARGED PARTICLES

THROUGH INTERPLANETARY SPACE*,” 1965.

[17] D. Reames, “Solar energetic particles: is there time to hide?,” *Radiat. Meas.*, vol. 30, no. 3, pp. 297–308, 1999, doi: 10.1016/S1350-4487(99)00066-9.

[18] “Solar wind - Wikipedia.” https://en.wikipedia.org/wiki/Solar_wind (accessed May 16, 2021).

[19] R. M. Skoug *et al.*, “Extremely high speed solar wind: 29-30 October 2003,” *J. Geophys. Res. Sp. Phys.*, vol. 109, no. A9, p. 9102, Sep. 2004, doi: 10.1029/2004JA010494.

[20] J. H. King, “Solar proton fluences for 1977-1983 space missions,” *J. Spacecr. Rockets*, vol. 11, no. 6, pp. 401–408, May 1974, doi: 10.2514/3.62088.

[21] J. Feynman, G. Spitale, J. Wang, and S. Gabriel, “Interplanetary proton fluence model: JPL 1991,” *J. Geophys. Res. Sp. Phys.*, vol. 98, no. A8, pp. 13281–13294, Aug. 1993, doi: 10.1029/92ja02670.

[22] M. A. Xapsos and G. P. Summers, “Probability model for cumulative solar proton event fluences,” *IEEE Trans. Nucl. Sci.*, vol. 47, no. 3 PART 1, pp. 486–490, 2000, doi: 10.1109/23.856469.

[23] J. N. Kapur, *Maximum-entropy models in science and engineering / J. N. Kapur*. .

[24] A. J. Tylka *et al.*, “CREME96 : A Revision of the Cosmic Ray E ects on Micro-Electronics Code,” vol. 44, no. 6, 1997.

[25] P. Blasi, “The Origin of Galactic Cosmic Rays,” *Astron. Astrophys. Rev.*, vol. 21, no. 1, Nov. 2013, doi: 10.1007/s00159-013-0070-7.

[26] I. A. Grenier, J. H. Black, and A. W. Strong, “The nine lives of cosmic rays in galaxies,” *Annual Review of Astronomy and Astrophysics*, vol. 53, no. 1. Annual Reviews Inc., pp. 199–246, Aug. 18, 2015, doi: 10.1146/annurev-astro-082214-122457.

[27] L. J. Gleeson and W. I. Axford, “Solar Modulation of Galactic Cosmic Rays,” *Astrophys. J.*, vol. 154, p. 1011, Dec. 1968, doi: 10.1086/149822.

[28] A. J. Tylka *et al.*, “CREME96: A Revision of the Cosmic Ray E ects on Micro-Electronics Code,” 1997.

[29] R. A. Nymmik, M. I. Panasyuk, T. I. Pervaja, and A. A. Suslov, “A model of galactic cosmic ray fluxes,” *Int. J. Radiat. Appl. Instrumentation. Part*, vol. 20, no. 3, pp. 427–429, Jul. 1992, doi: 10.1016/1359-0189(92)90028-T.

[30] “Help: GCR particle models.” https://www.spenvis.oma.be/help/background/gcr/gcr.html (accessed May 17, 2021).

[31] “TechDemoSat - eoPortal Directory - Satellite Missions.” https://earth.esa.int/web/eoportal/satellite-missions/t/techdemosat-1 (accessed May 18, 2021).

[32] “Small Satellite supplier | Surrey Satellite Technology Ltd | SSTL | SSTL.” https://www.sstl.co.uk/ (accessed May 18, 2021).

[33] D. Hastings and H. Garrett, *Spacecraft-Environment Interactions*. Cambridge University Press, 2004.

[34] F. Corporation, “State of the Satellite Industry Report,” 2012.

[35] M. Wüest, D. S. Evans, and R. Von Steiger, “Calibration of Particle Instruments in Space Physics.” Accessed: May 19, 2021. [Online]. Available: www.issibern.

[36] R. F. Pfaff, J. . Borovsky, and D. T. Young, Eds., *Measurement Techniques in Space Plasmas: Particles*. Washington, D. C.: American Geophysical Union, 1998.

[37] D. T. Y. Robert F. Pfaff, Joseph.E Borovsky, *Measurement Techniques in Space Plasmas: Fields*, vol. 103. Washington, D. C.: American Geophysical Union, 1998.

[38] L. H. Brace, “Langmuir probe measurements in the ionosphere,” in *Geophysical Monograph Series*, vol. 102, 1998, pp. 23–35.

[39] R. L. Merlino, “Understanding Langmuir probe current-voltage characteristics,” 2007, doi: 10.1119/1.2772282.

[40] L. Conde, “An introduction to Langmuir probe diagnostics of plasmas,” 2011.

[41] H. M. Mott-Smith and I. Langmuir, “The theory of collectors in gaseous discharges,” *Phys. Rev.*, vol. 28, no. 4, pp. 727–763, Oct. 1926, doi: 10.1103/PhysRev.28.727.

[42] F. F. Chen *et al.*, “Electric Probes,” *PLASMA DIAGNOSTIC Tech. Ed. BY HUDDLESTONE, R.H. LEORNARD, S.L*, 1965, Accessed: May 23, 2021. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.259.8626.

[43] X. Wang, H. W. Hsu, and M. Horányi, “Identification of when a Langmuir probe is in the sheath of a spacecraft: The effects of secondary electron emission from the probe,” *J. Geophys. Res. Sp. Phys.*, vol. 120, no. 4, pp. 2428–2437, Apr. 2015, doi: 10.1002/2014JA020624.

[44] F. F. Chen, “Langmuir Probe Diagnostics,” 2003.

[45] K. W. Ogilvie *et al.*, “SWE, a comprehensive plasma instrument for the WIND spacecraft,” *Space Sci. Rev.*, vol. 71, no. 1–4, pp. 55–77, Feb. 1995, doi: 10.1007/BF00751326.

[46] M. Neugebauer and C. W. Snyder, “Solar plasma experiment,” *Science (80-.).*, vol. 138, no. 3545, pp. 1095–1097, 1962, doi: 10.1126/science.138.3545.1095-a.

[47] F. R. Paolini and G. C. Theodoridis, “Charged particle transmission through spherical plate electrostatic analyzers,” *Rev. Sci. Instrum.*, vol. 38, no. 5, pp. 579–588, 1967, doi: 10.1063/1.1720771.

[48] C. W. Carlson, D. W. Curtis, G. Paschmann, and W. Michel, “An instrument for rapidly measuring plasma distribution functions with high resolution,” *Adv. Sp. Res.*, vol. 2, no. 7, pp. 67–70, Jan. 1982, doi: 10.1016/0273-1177(82)90151-X.

[49] “UCL Department of Space and Climate Physics - UCL – University College London.” https://www.ucl.ac.uk/mssl/ (accessed May 24, 2021).

[50] D. T. Young *et al.*, “Hot Plasma Composition Analyzer for the Magnetospheric Multiscale Mission,” *Sp. Sci Rev*, vol. 199, pp. 407–470, 2016, doi:

10.1007/s11214-014-0119-6.

[51] D. M. Hassler *et al.*, “The Radiation Assessment Detector (RAD) Investigation,” *Space Sci. Rev.*, vol. 170, no. 1–4, pp. 503–558, Jul. 2012, doi: 10.1007/s11214-012-9913-1.

[52] A. Mohammadzadeh *et al.*, “Monitor Program First Results From,” vol. 50, no. 6, pp. 2272–2277, 2003.

[53] E. F. Mitchell *et al.*, “The Highly Miniaturised Radiation Monitor,” vol. 1.

[54] “LUCID.” http://www.thelangtonstarcentre.org/index.php/what-is-lucid.

[55] LIP, “RADFET,” [Online]. Available: http://www.lip.pt/~space/.

[56] A. Holmes-siedle, L. Adams, and A. H. Iedle, “RADFET : A REVIEW OF THE USE OF METAL-OXIDE-SILICON DEVICES AS INTEGRATING DOSIMETERS,” vol. 28, no. 2, pp. 235–244, 1986.

[57] V. Angelopoulos, “Energetic Particle Detector.” http://elfin-lomo.igpp.ucla.edu/?epd.shtml.

[58] F. Chastellain, “Nos creare scientia hodie ad cras Radiation Hard Electron Monitor Radiation Hard Electron Monitor - RADEM Electron detector,” no. May, pp. 9–10, 2012.

[59] M. P. Seah, “Channel electron multipliers: quantitative intensity measurement-efficiency, gain, linearity and bias effects,” *J. Electron Spectros. Relat. Phenomena*, vol. 50, no. 1, pp. 137–157, Jan. 1990, doi: 10.1016/0368-2048(90)80015-3.

[60] J. H. Moore, “BUILDING SCIENTIFIC APPARATUS Fourth Edition-energy electron scattering, and the design and fabrication of instruments for use in the laboratory and on spacecraft.”

[61] M. Krems, J. Zirbel, M. Thomason, and R. D. DuBois, “Channel electron multiplier and channelplate efficiencies for detecting positive ions,” *Rev. Sci. Instrum.*, vol. 76, no. 9, p. 093305, Sep. 2005, doi: 10.1063/1.2052052.

[62] “Dr. Sjuts Optotechnik GmbH - Channel electron multipliers and complete measurement systems.” http://www.sjuts.com/Quality_efficiencies.html (accessed May 25, 2021).

[63] J. L. Wiza, “MICROCHANNEL PLATE DETECTORS,” 1979.

[64] T. Information, “TECHNICAL INFORMATION MCP ASSEMBLY.”

[65] G. Lutz and R. Klanner, “Solid State Detectors,” in *Particle Physics Reference Library*, Cham: Springer International Publishing, 2020, pp. 137–200.

[66] K. H. Photonics, “Photomultiplier tubes: Basics and applications,” *Ed. 3a*, p. 323, 2006.

[67] B. Dolgoshein *et al.*, “Status report on silicon photomultiplier development and its applications,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 563, no. 2, pp. 368–376, Jul. 2006, doi: 10.1016/j.nima.2006.02.193.

[68] P. Buzhan *et al.*, “ICFA Instrumentation Bulletin AN ADVANCED STUDY OF SILICON SiPM Description and Performance SiPM topology.”

[69] P. Eckert, H.-C. Schultz-Coulon, W. Shen, R. Stamen, and A. Tadday, “Characterisation studies of silicon photomultipliers,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 620, no. 2–3, pp. 217–226, Aug. 2010, doi: 10.1016/j.nima.2010.03.169.

[70] S. R. and D. V. Giancarlo Barbarino, Riccardo de Asmundis, Gianfranca De Rosa, Carlos Maximiliano Mollo, “Silicon Photo Multipliers Detectors Operating in Geiger Regime: an Unlimited Device for Future Applications,” in *Photodiodes - World Activities in 2011*, P. J. W. Park, Ed. InTech, 2011, p. 410.

[71] S. Tudisco, “Towards a new concept of photomultiplier,” no. February 2008, pp. 1–23.

[72] G. Collazuol, “The SiPM Physics and Technology - a Review,” in *PhotoDet 2012*, 2012, no. June, [Online]. Available: http://indico.cern.ch/event/164917/contribution/72/material/slides/0.pdf.

[73] H. Kanda, Y. Kasai, K. Maeda, and T. Nishizawa, “Development of a scintillation counter with MPPC readout for the internal tagging system.”

[74] E. M. Becker, A. T. Farsoni, A. M. Alhawsawi, and B. Alemayehu, “Small Prototype Gamma Spectrometer Using CsI(Tl) Scintillator Coupled to a Solid-State Photomultiplier,” *IEEE Trans. Nucl. Sci.*, vol. 60, no. 2, pp. 968–972, Apr. 2013, doi: 10.1109/TNS.2012.2228236.

[75] S. Majewski, J. Proffitt, J. Mckisson, R. Raylman, A. Stolin, and A. G. Weisenberger, “Initial Tests of a Compact Imaging Photomultiplier Made From Array of 3x3mm 2 Hamamatsu MPPC- SMD Modules.”

[76] D. Ginzburg *et al.*, “Optimizing the design of a silicon photomultiplier-based radiation detector,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 652, no. 1, pp. 474–478, Oct. 2011, doi: 10.1016/j.nima.2011.01.022.

[77] P. Pd, S. Gomi, T. Nakaya, M. Yokoyama, and H. Kawamuko, “Research and development of MPPC for T2K experiment,” 2009.

[78] A. Vacheret, M. Noy, M. Raymond, and A. Weber, “First results of the Trip-t based T2K front end electronics performance with GM-APD,” 2007.

[79] G. Blasse;B.C.Grabmaier, *Luminescent Materials*. Springer, Berlin, Heidelberg, 1994.

[80] D. Carbone *et al.*, “Mini-phoswich and SiPM for heavy ion detection,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 912, pp. 128–131, Dec. 2018, doi: 10.1016/J.NIMA.2017.10.095.

[81] M. A. Z. and J. C. M.J. Berger, J.S. Coursey, “Stopping-Power & Range Tables for Electrons, Protons, and Helium Ions,” 2017. doi: https://dx.doi.org/10.18434/T4NC7P.

[82] J.H.Hubbell and S.M.Seltzer, “X-Ray Mass Attenuation Coefficients | NIST,” *NIST*, 2004. https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients (accessed May 07, 2021).

[83] James Ziegler, “James Ziegler - SRIM & TRIM.” http://www.srim.org/ (accessed May 07, 2021).

[84] S. Agostinelli *et al.*, “GEANT4 - A simulation toolkit,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 506, no. 3, pp. 250–303, Jul. 2003, doi: 10.1016/S0168-9002(03)01368-8.

[85] G. G.Guilbault, *Practical Fluorescence, Second Edition*, 2nd editio. CRC Press;, 1990.

[86] M. J. Weber, “Inorganic scintillators: Today and tomorrow,” *J. Lumin.*, vol. 100, no. 1–4, pp. 35–45, Dec. 2002, doi: 10.1016/S0022-2313(02)00423-4.

[87] S. E. Derenzo, M. J. Weber, E. Bourret-Courchesne, and M. K. Klintenberg, “The quest for the ideal inorganic scintillator $,” *Nucl. Instruments Methods Phys. Res. A*, vol. 505, pp. 111–117, 2003, doi: 10.1016/S0168-9002(03)01031-3.

[88] T. Yanagida, “Inorganic scintillating materials and scintillation detectors,” *Proceedings of the Japan Academy Series B: Physical and Biological Sciences*, vol. 94, no. 2. Japan Academy, pp. 75–97, 2018, doi: 10.2183/pjab.94.007.

[89] M. Nikl and A. Yoshikawa, “Recent R&D Trends in Inorganic Single-Crystal Scintillator Materials for Radiation Detection,” *Adv. Opt. Mater.*, vol. 3, no. 4, pp. 463–481, Apr. 2015, doi: 10.1002/adom.201400571.

[90] W. H. Bragg and R. Kleeman, “ XXXIX. On the α particles of radium, and their loss of range in passing through various atoms and molecules ,” *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 10, no. 57, pp. 318–340, Sep. 1905, doi: 10.1080/14786440509463378.

[91] E. P. Particle Data Group , P A Zyla, R M Barnett, J Beringer, O Dahl, D A Dwyer, D E Groom, C -J Lin, K S Lugovsky, “Review of Particle Physics,” *Prog. Theor. Exp. Phys.*, vol. 2020, no. 8, doi: https://doi.org/10.1093/ptep/ptaa104.

[92] Y. S. Tsai, “Pair production and bremsstrahlung of charged leptons,” *Rev. Mod. Phys.*, vol. 46, no. 4, pp. 815–851, Oct. 1974, doi: 10.1103/RevModPhys.46.815.

[93] E. D. Aluker, D. Y. Lusis, and S. A. Chernov, “Electron excitations and radioluminescence of alkali halide crystals.” 1979, Accessed: Apr. 07, 2021. [Online]. Available: http://inis.iaea.org/Search/search.aspx?orig_q=RN:12636289.

[94] P. Lecoq *et al.*, “Lead tungstate (PbWO4) scintillators for LHC EM calorimetry,” *Nucl. Inst. Methods Phys. Res. A*, vol. 365, no. 2–3, pp. 291–298, Nov. 1995, doi: 10.1016/0168-9002(95)00589-7.

[95] T. Yanagida, G. Okada, T. Kato, D. Nakauchi, and S. Yanagida, “Fast and high light yield scintillation in the Ga2O3 semiconductor material,” *Appl. Phys. Express*, vol. 9, no. 4, p. 042601, Apr. 2016, doi: 10.7567/APEX.9.042601.

[96] F. D. Brooks, “Development of organic scintillators,” *Nucl. Instruments Methods*, vol. 162, no. 1–3, pp. 477–505, Jun. 1979, doi: 10.1016/0029-554X(79)90729-8.

[97] R. Voltz, H. Dupont, and G. Laustriat, “Radioluminescence des milieux organiques. II. Vérification expérimentale de l’étude cinétique,” *J. Phys.*, vol.

29, no. 4, pp. 297–305, Apr. 1968, doi: 10.1051/jphys:01968002904029700.

[98] J. B. Birks, “The fluorescence and scintillation decay times of crystalline anthracene,” *Proc. Phys. Soc.*, vol. 79, no. 3, pp. 494–496, Mar. 1962, doi: 10.1088/0370-1328/79/3/306.

[99] G. Blasse, “Reviews: Scintillator Materials,” *Chemistry of Materials*, vol. 6, no. 9. American Chemical Society, pp. 1465–1475, Sep. 01, 1994, doi: 10.1021/cm00045a002.

[100] W. S. Choong, “The timing resolution of scintillation-detector systems: Monte Carlo analysis,” *Phys. Med. Biol.*, vol. 54, no. 21, pp. 6495–6513, 2009, doi: 10.1088/0031-9155/54/21/004.

[101] P. Lecoq *et al.*, “Factors Influencing Time Resolution of Scintillators and Ways to Improve Them,” *IEEE Trans. Nucl. Sci.*, vol. 57, no. 5, 2010, doi: 10.1109/TNS.2010.2049860.

[102] Y. Evangelista *et al.*, “The scientific payload on-board the HERMES-TP and HERMES-SP CubeSat missions,” p. 168, Jan. 2021, doi: 10.1117/12.2561018.

[103] N. Werner *et al.*, “CAMELOT: Cubesats Applied for MEasuring and LOcalising Transients - Mission Overview,” *arXiv*, Jun. 2018, Accessed: Apr. 30, 2021. [Online]. Available: http://arxiv.org/abs/1806.03681.

[104] J. R. Smith *et al.*, “BurstCube: Concept, Performance, and Status,” Jul. 2019, Accessed: Apr. 30, 2021. [Online]. Available: http://arxiv.org/abs/1907.11069.

[105] “M.E. Taylor Engineering, Inc. | SEMicro Division.” https://semicro.org/# (accessed Jun. 10, 2021).

[106] Norland, “Norland Optical Adhesive 61.” https://www.norlandprod.com/adhesives/NOA 61.html.

[107] F. Benrachi *et al.*, “INVESTIGATION OF THE PERFORMANCE OF CsI(TI) FOR CHARGED PARTICLE IDENTIFICATION BY PULSE-SHAPE ANALYSIS 137,” 1989.

[108] J. Alarja *et al.*, “Charged particles identification with a CsI(Tl) scintillator.” 1985, Accessed: May 08, 2021. [Online]. Available: http://inis.iaea.org/Search/search.aspx?orig_q=RN:17074036.

[109] Y. M. Protopopov and V. G. Vasil’chenko, “Radiation damage in plastic scintillators and optical fibers,” *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms*, vol. 95, no. 4, pp. 496–500, Apr. 1995, doi: 10.1016/0168-583X(94)00599-0.

[110] J. Chen, R. Mao, L. Zhang, and R. Y. Zhu, “Gamma-ray induced radiation damage in large size LSO and LYSO crystal samples,” *IEEE Trans. Nucl. Sci.*, vol. 54, no. 4, pp. 1319–1326, Aug. 2007, doi: 10.1109/TNS.2007.902370.

[111] R. Brun and F. Rademakers, “ROOT - An object oriented data analysis framework,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 389, no. 1–2, pp. 81–86, Apr. 1997, doi: 10.1016/S0168-9002(97)00048-X.

[112] “Introduction to Geant4 — IntroductionToGeant4 10.7 documentation.” https://geant4-userdoc.web.cern.ch/UsersGuides/IntroductionToGeant4/html/index.html

(accessed May 31, 2021).

[113] R. Chytracek, J. McCormick, W. Pokorski, and G. Santin, “Geometry description markup language for physics simulation and analysis applications,” *IEEE Trans. Nucl. Sci.*, vol. 53, no. 5, pp. 2892–2896, Oct. 2006, doi: 10.1109/TNS.2006.881062.

[114] M. Pinto and P. Gonçalves, “GUIMesh: a tool to import STEP geometries into Geant4 via GDML,” *Comput. Phys. Commun.*, vol. 239, pp. 150–156, Jul. 2018, doi: 10.1016/j.cpc.2019.01.024.

[115] Christopherpoole, “CADMesh.” https://github.com/christopherpoole/cadmesh.

[116] “FastRAD.” http://www.fastrad.net/.

[117] “About NIST SRMs | NIST.” https://www.nist.gov/srm/about-nist-srms (accessed May 31, 2021).

[118] M. J. Berger, “Monte Carlo Calculation of the Penetration and Diffusion of Fast Charged Particles*.”

[119] J. C. Butcher and H. Messel, “Electron number distribution in electron-photon showers in air and aluminium absorbers,” *Nucl. Phys.*, vol. 20, no. C, pp. 15–128, 1960, doi: 10.1016/0029-5582(60)90162-0.

[120] H. Messel and D. F. Crawford, *Electron-photon shower distribution function : Tables for lead, copper and air absorbers*. Braunschweig : Pergamon Pr., 1970, 1970.

[121] W. R. Nelson, D. W. O. Rogers, and H. Hirayama, “The EGS4 Code System, Report SLAC-265,” Stanford University, Stanford, 1985.

[122] CERN, “Geant4 class architecture.” .

[123] L. Landau, “On the energy loss of fast particles by ionization,” *J.Phys.(USSR)*, vol. 8, pp. 201–205, 1944.

[124] T. Niggemann, E. Dietz-Laursonn, T. Hebbeker, A. Künsken, M. Lauscher, and M. Merschmeyer, “G4SiPM: A novel silicon photomultiplier simulation package for Geant4,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 787, pp. 344–347, Aug. 2015, doi: 10.1016/j.nima.2015.01.067.

[125] E. C. Gil, E. M. Albarrán, and E. Minucci, “GODDeSS: a Geant4 extension for easy modelling of optical detector components Related content The beam and detector of the NA62 experiment at CERN,” *J. Instrum.*, vol. 12, no. 04, p. P04026, Apr. 2017, doi: 10.1088/1748-0221/12/04/P04026.

[126] F. Corsi *et al.*, “Modelling a silicon photomultiplier (SiPM) as a signal source for optimum front-end design,” *Nucl. Instruments Methods Phys. Res. A*, vol. 572, pp. 416–418, 2007, doi: 10.1016/j.nima.2006.10.219.

[127] D. Marano *et al.*, “Improved SPICE electrical model of silicon photomultipliers,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 726, pp. 1–7, Oct. 2013, doi: 10.1016/j.nima.2013.05.127.

[128] L. W. Nagel, “SPICE2: A Computer Program to Simulate Semiconductor Circuits | EECS at UC Berkeley,” 1975. Accessed: Jun. 10, 2021. [Online]. Available: https://www2.eecs.berkeley.edu/Pubs/TechRpts/1975/9602.html.

[129] R. H. Haitz, A. Goetzberger, R. M. Scarlett, and W. Shockley, “Avalanche Effects in Silicon p - n Junctions. I. Localized Photomultiplication Studies on Microplasmas,” *J. Appl. Phys.*, vol. 34, no. 6, pp. 1581–1590, 1963, doi: 10.1063/1.1702639.

[130] R. J. Mcintyre, “Theory of Microplasma Instability in Silicon,” *J. Appl. Phys.*, vol. 32, p. 983, 1961, doi: 10.1063/1.1736199.

[131] “What is an SiPM and how does it work? | Hamamatsu Photonics.” https://hub.hamamatsu.com/jp/en/technical-note/how-sipm-works/index.html (accessed Jan. 02, 2022).

[132] F. Acerbi and S. Gundacker, “Understanding and simulating SiPMs,” *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 926. Elsevier B.V., pp. 16–35, May 11, 2019, doi: 10.1016/j.nima.2018.11.118.

[133] G. Collazuol, “The Silicon PhotoMultiplier Status and Perspectives - VCI 2019 - G Collazuol | Enhanced Reader,” 2019. .

[134] W. Shen, “Development of high performance readout ASICs for silicon photomultipliers (SiPMs) - heiDOK,” uni-heidelberg, 2012.

[135] V. Radeka, “LOW-NOISE TECHNIQUES in detectors,” 1988.

[136] J. Graeme, *Photodiode Amplifiers: OP AMP Solutions*, 1st editio. McGraw-Hill Education, 1995.

[137] H. Hashemi, “Transimpedance Amplifiers (TIA): Choosing the Best Amplifier for the Job,” 2015. Accessed: Jun. 12, 2021. [Online]. Available: www.ti.com.

[138] “DesignSpark.” https://www.rs-online.com/designspark/home (accessed Jun. 13, 2021).

[139] K. Sato, Y. Enomoto, and Y. Adachi, “MPPC,” 2017.

[140] Y. Qiang, C. Zorn, F. Barbosa, and E. Smith, “Radiation hardness tests of SiPMs for the JLab Hall D Barrel calorimeter,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 698, pp. 234–241, Jan. 2013, doi: 10.1016/J.NIMA.2012.10.015.

[141] B. Doyon *et al.*, “SiPM proton irradiation for application in cosmic space,” doi: 10.1088/1748-0221/15/03/P03002.

[142] R. Pagano *et al.*, “Radiation hardness of silicon photomultipliers under 60Co γ-ray irradiation,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 767, pp. 347–352, Dec. 2014, doi: 10.1016/J.NIMA.2014.08.028.

[143] “Evaluation Board | DRS Chip | Paul Scherrer Institut (PSI).” https://www.psi.ch/en/drs/evaluation-board (accessed Jan. 03, 2022).

[144] IAEA, “IAEA-TECDOC-363 SELECTED TOPICS IN NUCLEAR ELECTRONICS,” 1986.

[145] “Welcome to Python.org.” https://www.python.org/ (accessed Jun. 14, 2021).

[146] “NumPy.” https://numpy.org/ (accessed Jun. 14, 2021).

[147] “Project Jupyter | Home.” https://jupyter.org/ (accessed Jun. 14, 2021).

[148] M. Gabbouj, E. J. Coyle, and N. C. Gallagher, “AN OVERVIEW OF MEDIAN AND STACK FILTERING*,” 1992.

[149] “Matplotlib: Python plotting — Matplotlib 3.4.2 documentation.” https://matplotlib.org/ (accessed Jun. 15, 2021).

[150] D. W. Scott, “A Service of zbw Multivariate Density Estimation and Visualization Standard-Nutzungsbedingungen: Multivariate Density Estimation and Visualization.” Accessed: Jun. 15, 2021. [Online]. Available: www.econstor.euhttp://www.stat.rice.edu/~scottdw.

[151] D. Freedman and P. Diaconis, “On the Histogram as a Density Estimator: L 2 Theory,” 1981.

[152] S. J. Sheather and M. C. Jones, “A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation,” *J. R. Stat. Soc. Ser. B*, vol. 53, no. 3, pp. 683–690, Jul. 1991, doi: 10.1111/j.2517-6161.1991.tb01857.x.

[153] “SciPy.org — SciPy.org.” https://www.scipy.org/ (accessed Jun. 15, 2021).

[154] “PeakUtils · PyPI.” https://pypi.org/project/PeakUtils/ (accessed Jun. 15, 2021).

[155] CAEN, “SP5601-LED Driver datasheet,” 2011.

[156] S. Ritt, “DRS4 Evaluation Board User’s Manual,” 2018.

[157] J. Rosado and S. Hidalgo, “Characterization and modeling of crosstalk and afterpulsing in Hamamatsu silicon photomultipliers,” 2015.

[158] R. Newman, “Visible light from a silicon p-n junction,” *Phys. Rev.*, vol. 100, no. 2, pp. 700–703, Oct. 1955, doi: 10.1103/PhysRev.100.700.

[159] OptoSupply, “OSSV5111A datasheet.”

[160] Polymicro, “SILICA/HARD POLYMER CLAD Optical Fiber JTFLH.”

[161] “Exponential distribution - Wikipedia.” https://en.wikipedia.org/wiki/Exponential_distribution (accessed Jun. 18, 2021).

[162] F. Asaro, F. L. Reynolds, and I. Perlman, “The complex alpha-spectra of Am241 and Cm242,” *Phys. Rev.*, vol. 87, no. 2, pp. 277–285, Jul. 1952, doi: 10.1103/PhysRev.87.277.

[163] P. P. Day, “Electromagnetic spectrum of Am241,” *Phys. Rev.*, vol. 97, no. 3, pp. 689–700, Feb. 1955, doi: 10.1103/PhysRev.97.689.

[164] Wikipedia, “Americium-241,” *Wikipedia*. 2020, Accessed: Dec. 14, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Americium-241.

[165] H. W. Lewis, “Multiple scattering in an infinite medium,” *Phys. Rev.*, vol. 78, no. 5, pp. 526–529, Jun. 1950, doi: 10.1103/PhysRev.78.526.

[166] K. N. Yu, C. W. Y. Yip, D. Nikezic, J. P. Y. Ho, and V. S. Y. Koo, “Comparison among alpha-particle energy losses in air obtained from data of SRIM, ICRU and experiments,” *Appl. Radiat. Isot.*, vol. 59, no. 5–6, pp. 363–366, Nov. 2003, doi: 10.1016/S0969-8043(03)00201-X.

[167] L. O. M. G.T.Seaborg, R.A.James, *The new element americium (atomic number 95)*, vol. 14B, no. 22.1. New York: McGraw-Hill Book Co., Inc., 1949.

[168] H. Jaffe, T. O. Passell, C. I. Browne, and I. Perlman, “Gamma and X-radiation

in the decay of Am241," *Phys. Rev.*, vol. 97, no. 1, pp. 142–150, Jan. 1955, doi: 10.1103/PhysRev.97.142.

[169] "BC-400 BC-404 | Products | Saint-Gobain Crystals." https://www.crystals.saint-gobain.com/products/bc400-bc404 (accessed Jan. 28, 2021).

[170] K. H. Photonics, "S13360-1325CS datasheet."

Appendix A: Survey of scintillator detector application in space instrument

Mission	Instrument	Detector	Target	Energy Range	Feature
Compton Gamma Ray Observatory (CGRO) 1991 to 2000, Earth orbit	Burst and Transient Source Experiment (BATSE)	NaI(Tl)	Gamma-ray burst	20 to >600keV	Surround the core detector with plastic scintillator for active anti-coincidence to veto the background
	Oriented Scintillation Spectrometer Experiment (OSSE)	NaI(Tl)	Gamma ray	0.05 to 10 MeV	Operate as a phoswich utilizing slow and fast signal
	Energetic Gamma Ray Experiment Telescope (EGRET)	NaI(Tl)	Energetic Gamma-ray	30MeV to 30GeV	with a plastic scintillator anti-coincidence dome
High Energy Astronomy Observatories (HEAO) 1977	Hard X-Ray/ Low-Energy Gamma-ray experiment	NaI	Hard X-ray or low Gamma-ray	120 keV to 10MeV	
Rossi X-Ray Timing Explorer (RXTE) 1995 - 2012	The High Energy X-ray Timing Experiment (HEXTE)	NaI/CsI	X-ray	15-250keV	
INTErnational Gamma-Ray Astrophysics Laboratory (INTEGRAL) 2002- Earth orbit	IBIS	CdTe tiles (front) and CsI tiles (back)	Hard X-ray to Gamma-ray	15 keV to 10 MeV	
	JEM-X	Xenon plus methane (gas scintillator)	X-ray	3 to 35 keV	

Mission	Instrument	Detector	Target	Energy Range	Feature
Swift Observatory 2004 Earth orbit	Burst Alert Telescope (BAT)	CdZnTe tiles	X-ray	15 - 150keV	
Fermi Gamma-ray Space Telescope (FGST) 2008 – Earth orbit	Gamma-ray Burst Monitor (GBM)	NaI BGO	Gamma	8keV to 1MeV 150keV to 30 MeV	
Alpha Magnetic Spectrometer (AMS-02)	Time of Flight	Plastic scintillator EJ-200 or BC-408	Time of flight		

Appendix B: Survey of popular scintillators

	Basic Material Info				Physics Parameters										
No	Manufacturer	Material Type	Catelogue	Material	Density	Dielectric Constant	Melting point	Thermal expansion coefficient	Thermal conductivity	Cleavage plane	Hardness	Soluble in Alcohol	Hygroscopic	Structure	Vapor Pressure
					g.cm-3		°C	(/C) x 10-6	$W.m^{-1}k^{-1}$						
1	Hilger Crystals	Inorganic crystal	Alkali halide crystals	CsI(Tl)	4.51	5.65@1Mhz(298k)	621		1.13	None		Yes	Slightly	BCC	
2	Hilger Crystals	Inorganic crystal		NaI(Tl)	3.7		651		3.47	(100)		Yes	Very	BCC	
3	Hilger Crystals	Inorganic crystal	Non- Alkali crystals	CaF_2(Eu)	3.18		1418			(111)			no	Cubic	
4	Hilger Crystals	Inorganic crystal	Newly develop	LYSO(Ce)	7.1		2050			None		Not in water	no	Cubic	
5	Hilger Crystals	Inorganic crystal	Non- Alkali crystals	$ZnWO_4$	7.62		1220			(010)				Monoclinic	
6	Hilger Crystals	Inorganic crystal	Newly develop	$Bi_4Ge_3O_{12}$ (BGO)	7.13		1050			None			no	Cubic	
5	Hilger Crystals	Inorganic crystal		$CdWO_4$	7.9		1320			<010>		stable		Monoclinic	
5	Hilger Crystals	Inorganic crystal		Cs2LiYCl6:Ce	3.31		640			None		water soluble	yes	Cubic	
7	SAINT-GOBAIN	Inorganic crystal		BrilLanCe380- LaBr3(Ce)	5.08			8		<100>			yes		
8	SAINT-GOBAIN	Inorganic crystal		NaI(Tl)	3.67			47.4		<100>			yes		
9	SAINT-GOBAIN	Inorganic crystal		Lu1.8Y2SiO5(Ce)	7.1			--		none			no		
10	SAINT-GOBAIN	Plastic scintillators	Base: Polyvinyltoluene	BC-444	1.032		70	0.78				soluable in aromatic solvents,chlorinated solvents, acetone		sheet	May be used in vacuum
11	SAINT-GOBAIN	Plastic scintillators	Base: Polyvinyltoluene	BC-400, BC-408,BC-412,BC-416	1.032		70	0.78				soluable in aromatic solvents,chlorinated solvents, acetone		sheet	May be used in vacuum
12	SAINT-GOBAIN	Plastic scintillators		BC-700 series											
13	RMD	Inorganic crystal		CsI(Tl) - microcolumnar											
14	RMD	Inorganic crystal	Alkaline earth	CLYC - (Cs2LiYCl6:Ce)	3.31		640	34.34	0.67@50degC	None			Yes	cubic	
15	Toshiba material	Ceramic scintillator		Gd2O2S:Pr	7.34			8.8					No		
16	ELJEN TECHONOLOGY	Plastic scintillators		EJ-228/BC-418											

Continue column from last page

	Photon output																	
No	Wavelength of emission max.	Wavelength min	Wavelength max	Refractive Index at peak emission	Primary decay time	After glow	Pulse width (FWHM)	Energy resolution Cs137	Light yield	Photoelectron yield	Temperature coefficient of light yield	Light Escape Cone to Air	Light Escape Cone to Glass	Optical Transmission Range	Radiation Length	Gamma and X-ray absorption coefficients	Comments
	nm	nm	nm		us					Photons/MeV	%.°C^{-1}	°	°	nm	cm	cm-1	
1	565	375	725	1.78	1	Apx 2.8% at 1ms, Apx 2.0% at 2ms.				52,000	0.32	34.1	57.2	240-70,000	1.86	0.48 at 660keV 10.00 at 100keV	Highest light yield, perfect for gamma/x-ray detection. But slow response
2	410	325	525	1.8	0.23	Apx 8.0% at 1ms, Apx 4.0% at 2ms.				40,000	0.08	44.7	56.5	250-35000	2.59	0.29 at 660keV 5.5 at 100keV	Most popular in industry
3	435	400	500	1.43	1					23,650	-0.28				3.5		
4	420	380	480	1.81	0.04	N/A				32,000	-0.2				1.15		
5	490	310	500	2.32	20					9,500	-1.2				1.1		Too slow response time
6	480	375	650	2.15	0.3					8,500	-1.6				1.13		Light yield is too low
5	520	330	540	2.25	20					13,000	-0.1				1		Cd is a hazard material
5				1.81@405nm, 1.67@532nm, 1.61@633nm													
7	380			1.9	0.016					63,000	0					1.8 at 662keV 50%	General purpose, best energy resolution, rate of change of light output w/temperature is small
8	415			1.85	0.25					38,000	-0.3					2.5 at 662keV 50%	
9	420			1.81	0.041					32,000	-0.28					1.1 at 662keV 50%	Bright, high Z, fast ,dense, background from ^{176}Lu activity
10	428	250	525	1.58	0.285										180		Long timing property. Good for phoswitch, particle identification, low background counting, active shieldings. Can be used in conjunction with a fast scintillator such as BC-400
11	423			1.58	0.0024					65% of Anthracene; typical 25-30% of NaI(Tl)	Indpendent of temperature between -60 to +20. At +60degC = 95% of that at +20degC				160		General purpose, fast
12																	for the efficient detection of neutrons in the presence of gamma radiation.
13																	Columnized, range from 250nm to 10um. Films are 10 to 700um thick. Sctructures are up to 3mm thickness
14	370	275	450	1.81	0.001,0.05,1					20,000					3.42	0.251 at 660keV 3.97 at 100keV	Gamma-neutron scintillation, better gamma-ray resolution than NaI or CsI. It uses pulse shape discrimination for neutron detection
15	512			2.2	3	<0.1% after 3ms			35% (2mm thickness) not fully transparent	200% of CdWO4	-0.1						Radiation degration 1% per 7Gy. Similar light output to CsI but not fully transparent. Can be made into array
16	391				1.4		1.2			13000/67% of Anthracene							

Appendix C: Initial Simulation Data

No	Particle Energy	Energy deposit in the Scintillator(mean)	Photons detected in SiPM(mean)	Photon hit time(mean)	Particles fired	Particle triggered	Photons per keV	Detection ratio	Energy deposit ratio
	keV	keV	#	ns	#	#			
1	5	5	282.3	1093	100	4	56.460	4.00%	100.00%
2	10	10	535.6	992.8	100	85	53.560	85.00%	100.00%
3	20	19.8	1066	1003	100	97	53.838	97.00%	99.00%
4	30	30	1607	1001	100	98	53.567	98.00%	100.00%
5	40	40	2152	997.1	100	98	53.800	98.00%	100.00%
6	50	49.97	2682	1002	100	99	53.672	99.00%	99.94%
7	60	60	3224	1003	100	99	53.733	99.00%	100.00%
8	70	70	3759	998.6	100	93	53.700	93.00%	100.00%
9	80	80	4297	1001	100	87	53.713	87.00%	100.00%
10	90	87.14	4671	998.1	100	73	53.603	73.00%	96.82%
11	100	98.46	5268	1001	100	64	53.504	64.00%	98.46%
12	150	132.5	7122	1001	100	31	53.751	31.00%	88.33%
13	160	144.3	7758	998.8	100	30	53.763	30.00%	90.19%
14	170	146.6	7862	1004	100	27	53.629	27.00%	86.24%
15	180	158.2	8480	995.7	100	24	53.603	24.00%	87.89%
16	190	163.1	8772	1001	100	23	53.783	23.00%	85.84%
17	200	167.7	9023	998.3	100	21	53.804	21.00%	83.85%
18	210	115.8	6225	1002	100	19	53.756	19.00%	55.14%
19	220	182	9732	1003	100	12	53.473	12.00%	82.73%
20	230	182.9	9803	1002	100	22	53.598	22.00%	79.52%
21	240	151.1	8047	1004	100	12	53.256	12.00%	62.96%
22	250	147.3	7892	1001	100	13	53.578	13.00%	58.92%
23	260	190.7	10270	999.3	100	13	53.854	13.00%	73.35%
24	270	160.5	8682	1002	100	9	54.093	9.00%	59.44%
25	280	186.7	10040	1002	100	11	53.776	11.00%	66.68%
26	290	187.9	10110	999.1	100	4	53.805	4.00%	64.79%
27	300	219.9	11830	996.9	100	9	53.797	9.00%	73.30%
28	400	136.6	7344	996.1	100	7	53.763	7.00%	34.15%
29	500	359.7	19380	1001	100	9	53.878	9.00%	71.94%
30	600	231.5	12510	1000	100	9	54.039	9.00%	38.58%
31	700	285	15320	996	100	9	53.754	9.00%	40.71%
32	800	379	20400	998.4	100	2	53.826	2.00%	47.38%
33	900	364.3	19580	997.1	100	6	53.747	6.00%	40.48%
34	1000	641	34400	996.6	100	2	53.666	2.00%	64.10%

Appendix D: Work out the directional source

The position and orientation of the circular plane source is set by two direction vectors lying in the plane, rot1 and rot2. They can be derived from the position vector of the circular plane.

The position vector of the centre of the circular plane source is:

$$\vec{c} = \begin{pmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{pmatrix} \tag{0.1}$$

Vector rot1 and rot2 are obtained by differentiation with respect to θ and ø and subsequent normalization:

$$rot1 = \frac{d\vec{c}}{d\theta} = \begin{pmatrix} \cos(\theta)\cos(\phi) \\ \cos(\theta)\sin(\phi) \\ -\sin(\theta) \end{pmatrix} \tag{0.2}$$

$$rot2 = \frac{d\vec{c}}{d\phi} = \begin{pmatrix} -\sin(\phi) \\ \cos(\phi) \\ 0 \end{pmatrix} \tag{0.3}$$

Now rot1 and rot2 can be directly used in the Geant4 General Particle Source. The following is an example macro script for a circular plane pointing at θ=45° and φ=45°.

/gps/particle proton

/gps/ene/type Mono

/gps/ene/mono 100 MeV

/gps/pos/type Plane

/gps/pos/shape Circle

/gps/pos/radius 14.0 cm

/gps/pos/centre 7.0 7.0 9.89 cm

/gps/pos/rot1 0.5 0.5 -0.71

/gps/pos/rot2 -0.71 0.71 0

/gps/direction -7.0 -7.0 -9.89 cm

/gps/ang/type planar

The C++ utility runs the equation (0.1), (0.2) and (0.3) automatically for each θ and φ pair and generates the macro script for each circular plane source.

Appendix F: MSSL radioactive source list

MSSL code	Isotope	Emission (energy, particle)	Original Activity (MBq)	Arrival date (y)	Half life (y)
Sm1	Samarium 151	80 keV b	925	1987	101.3005
Sr1	Strontium 90	0.55 MeV b	0.479964	2001	28.78
C1	Carbon 14	49 keV mean <156 keV b	1.85	1980	5730
Tl2	Thalium 204	760 keV b; Hg-k 71 keV X	0.386613	2001	3.78
Cs1	Caesium 137	0.512 Mev b, 1.174 Mev b; 662 keV g; Ba-k 32-38 keV X	0.36963	1997	30.17
Am4	Americium 241	5,44 5,48 MeV a; 26,33, 60 keV g	74	1996	432.2
Am3	Americium 241	5,44 5,48 MeV a; 26,33, 60 keV g; 8-50 keV X (Cu, Rb, Mo, Ag, Ba, Tb)-k	370	1979	432.2
Am2	Americium 241	5,44 5,48 MeV a; 26,33, 60 keV g; 8-50 keV X (Cu, Rb, Mo, Ag, Ba, Tb)-k	370	1974	432.2
Am1	Americium 241	5,44 5,48 MeV a; 26,33, 60 keV g; 8-50 keV X (Cu, Rb, Mo, Ag, Ba, Tb)-k	370	1973	432.2
Fe57	Iron 55	5.9 keV X	3700	1991	2.73
H8	Tritium(Hydrogen 3)	18 keV b	1850	1974	12.35
H7	Tritium(Hydrogen 3)	18 keV b	1850	1974	12.35
Am5	Americium 241	5.5 MeV alpha	0.00577	2015	432.2
Am6	Americium 241	5.5 MeV alpha	0.04	2017	432.2
Am7	Americium 241	5.5 MeV alpha	0.04	2017	432.2

Appendix G: Survey of some SiPM sensor performance

Manufacturer	Type	Pkg	Eff. A	Pixels	Pix. Size	FF	Max λ	Peak λ	QE	PDE
			mm		um	%	nm	nm	%	%
Hamamatsu, Japan	S10362-11-100U	metal	1*1	100	100 * 100	78.5	900	440	65	72
Hamamatsu, Japan	S10362-33-100c	ceramic	3*3	900	100*100	78.5	900	440		
NDL, China	large dynamic range		1*1	10000	10*10					>10%@500nm
Philips Aarchen, Germany	DPC6400-22-44		3.9*3.2	6396	59.4*32	54	700	420		30%
Philips Aarchen, Germany	dpc3200-22-44		3.9*3.3	3200	59.4*64	74	700	420		40%
SENSL, Ireland	MicroSL family	10020	1*1	1296		48%	400-1000	500		8%
SENSL, Ireland	MicroSL family	10100	1*1	90		85%	400-1000	500		28%
Zecotek	MAPD-3A		3*3	135000			800	450-500		20%
KETEK, Germany	pm3350		3*3	3600	50*50	70%	300-1000	420		>50%

Manufacturer	Type	DC	DC	TR	BV	Gain	RT	X-talk	Comments
		kcps	pixel	ns	V		ns		
Hamamatsu, Japan	S10362-11-100U	600	6		70±10	2.4*10^6			
Hamamatsu, Japan	S10362-33-100c	8000	8.9						
NDL, China	large dynamic range	1000	0.1			>10^5	5		
Philips Aarchen, Germany	DPC6400-22-44	5000	0.78	0.044	27±0.5				Active quenching, full digital,
Philips Aarchen, Germany	dpc3200-22-44	7000	2.19	0.044	27±0.6				
SENSL, Ireland	MicroSL family	600	0.463		29.5	1.2*10^6	30	17%	
SENSL, Ireland			0		29.5	1*10^7	1800	49%	
Zecotek	MAPD-3A	9000000	66.667		65-70	60,000	80-100		high pexel density
KETEK, Germany	pm3350	200000	55.56		27	2*10^6		35%	

FF = fill factor, QE = quantum efficiency, PDE = photon detection efficiency, DC = dark count, TR = time resolution, BV = bias voltage, RT = recovery time

Appendix H: Review of Scintillator+SiPM based detector system performance

#	Title	Description	Scintllator		SiPM	Bond	Wrapping	Performance		Note	Yr	Inst.	Auth
			Type	Geometry	Type			ER	RT				
1	On the limited amplitude resolution of multipixel Geiger-mode APDs	Discuss about the amplitude resolution(energy resolution) limited by the number of cells;								Intrinsic energy resolution can be calculated using formula: Rmin = 2.92/sqrt(m), m is the number of cells.	2007	PSI	Alexey Stoykov
2	A time-resolution study with a plastic scintillator read out by a Geiger-mode Avalanche Photodiode	Explore the time resolution of scintillator + MPPC	BC-422	2x3x3 mm^3	MPPC S10362-33-050	optical grease				Best time resolution for scintillator + PMT is: ssqrt(E) = 19ps*$MeV^{0.5}$; Best time resolution for scintillator + MPPC is: ssqrt(E) = 18ps*$MeV^{0.5}$;	2011	PSI	Alexey Stoykov
3	Optimizing the design of a silicon photomultiplier-based radiation detector		CsI(Tl)	12.5mm x 12.5mm x 12.5mm	SPM Plus 144mm2		Balck opaques tape optical poly reflector white Teflon film Enhanced Specular Reflector (ESR)	7% for 137Cs (emitting 662keV gamma)			2011	Rotem Industries Ltd, Israel	D.Ginzburg
4	Initial Tests of a Compact Imaging Photomultiplier Made From Array of 3x3mm2 Hamamatsu MPPC-SMD Modules	Compare three scintillators LYSO, NaI(Tl) and CsI(Tl)	LYSO (from Proteus)	3mmx3mmx10mm	MPPC array, s10362-33-050SMD			12.4% for N^{22}, 511keV gamma			2011	Radiation Detector & Medical Imaging Group	S.Majewski
5	Development of High Precision Timing Counter Based on Plastic Scintillator with SiPM Readout		BC422	5mmx5mmx5mm	HPK SiPM S10362-33-050C	optical greaseOKEN6262A				Read out by DRS4 chip from PSI.	2014	INFN	PW Cattaneo

#	Title	Description	Scintllator		SiPM	Bond	Wrapping	Performance		Note	Yr	Inst.	Auth
6	Small Prototype Gamma Spectrometer Using CsI(Tl) Scintillator Coupled to a Solid-State Photomultiplier		CsI(Tl)	6mmx6mmx10mm		BC630 silicon grease	Teflon	Cs137, 622keV, 7.88%			2012	Oregon state university	Erik M Becker
7	Energy resolution of scintillation detectors with SiPM light readout	comapre the performance with scintillator + PMT, PMT has better energy resolution 6.7% vs 7.1%	CsI(Tl) LFS-3	3mmx3mmx3mm				622keV, 7.1% 622keV, 9.6%		shaping time 6us; shaping time 0.25us;	2010	National Centre for Nuclear Research, Poland	M. Grodzicka

ER = energy resolution, RT = response time

Appendix I: The hardware configuration for each chapter

Chapter	Scintillator Module	SiPM	Preamplifier	Comments
Chapter 2	BC408 + CsI (v1) BC404 + LXSR (v3)	N/A	N/A	Scintillator module v3 is used for the instrument concept
Chapter 3	BC408 + CsI (v1)	N/A	N/A	Used for Geant4 simulation
Chapter 4	N/A	S13360-1325CS	V1((FR4 version)	Used for preamplifier development
Chapter 5	N/A	S13360-1325CS	V2 (RF version)	Used to acquire the dark count data
Chapter 6	N/A	S13360-1325CS	V2 (RF version)	Used to test with LED source
Chapter 7	BC404 + LXSR (v3)	S13360-1325CS	V2 (RF version)	Used for the radioactive test